

Lecture 6: Basic Functional Analysis

Ebbighausen

July 14, 2026

1 Introduction

Up to this point, our main solution method has been to use some inherent symmetry or behavior to reduce the PDE to an ODE. If we want to start to push beyond these cases, what we need to study is the partial differential operator itself. If we view a PDE as $Lu = f$, whether we can solve the PDE amounts to whether we can invert L for $u = L^{-1}f$. The solution only exists when f is in the image of L , and it is only unique when L is injective. Whether the solution is continuous on the data has to do with whether L^{-1} is continuous and some other caveats.

During the development of many of the solutions we have looked at, there was significant development in a field called functional analysis that studied operators like L in great detail. The goal of functional analysis was to understand infinite-dimensional linear algebra, by way of operators on spaces like $C^m(\mathbb{R})$ and $L^2([0, 1])$. We are studying a special type of operator on spaces of continuous functions, so we reduce well-posedness to studying the operator L such that we can use these powerful tools.

Learning Objectives:

- Give examples of infinite-dimensional vector spaces.
- Describe a norm and an inner product by definition; prove that a function is either a norm or an inner product
- Define and differentiate L^p spaces.
- Describe the convergence of sequences of functions.
- Describe the properties of a basis and tools to check that it is a basis.
- Construct a basis given a space.
- Connect a eigenvectors of a self-adjoint operator to a basis.

2 Function Spaces

We will consider a vector space V , usually over $\mathbb{C}$ or $\mathbb{R}$. Our goal here is to look at generalizing linear algebra, so we first need to see some more general spaces. While there are many types of spaces to which our theory will apply, the spaces studied historically were those of continuous or integrable functions. We display why the continuous functions are interesting.

Recall that on such a space, a norm is a function $V \rightarrow \mathbb{C}$ such that

- 1.) Positive Definiteness: $\|u\| \geq 0$ with equality iff. $u = 0$
- 2.) Homogeneity: $\|\lambda u\| = |\lambda| \|u\|$ for $\lambda \in \mathbb{C}$
- 3.) Triangle Inequality: $\|u + v\| \leq \|u\| + \|v\|$

Example: $C^0(\mathbb{R})$ is infinite-dimensional

In linear algebra, we had quite a few theorems that relied on the space being finite-dimensional, such as the rank-nullity theorem. However, many of the spaces we care about are not finite-dimensional. To see that

$C^0(\mathbb{R})$ is not finite dimensional, consider the "sup norm" $\|f\| = \sup_{\Omega} |f(x)|$. Let ψ be a nonzero smooth bump supported in $[0, 1]$. Then, the set $\{\psi(x - n)\}_{n \in \mathbb{N}}$ is linearly independent.

Example:

We may place another norm on $C_c^0((0, 1))$ by creating an *inner product* (which we focus on more later)

$$\langle f, g \rangle = \int_0^1 f g dx$$

and one interesting facet of this inner product is that

$$\langle \Delta f, g \rangle = \langle f, \Delta g \rangle$$

which makes treating PDEs very easy.

Our next goal is to investigate more spaces like this to get an idea of the right choice of space to set our operators on. We often run into integrals, especially of weird functions, when doing PDEs. It turns out that the main tool we need to find the right spaces and right theory is to strengthen these integrals first (we strengthen derivatives later, for weak solutions). In particular, we investigate Lebesgue integration.

2.1 Integration and Measures

Bernhard Riemann introduced the classical Riemann integral in 1854 (though you probably learned Darboux's integral in analysis because it is slightly better-posed). In the late 1800s, Caratheodory, Hausdorff, and some other figures were developing the idea of *measures*, which allowed for the study of the size of sets and for the reconstruction of probability theory with rigorous mathematics. Henri Lebesgue used these measures in the early 20th century to construct an integral that worked on functions that were not Riemann integrable. For example, the function

$$f(x) = \begin{cases} 1 & x \in \mathbb{Q} \\ 0 & x \notin \mathbb{Q} \end{cases}$$

is not Riemann integrable, but its Lebesgue integral is defined (and is 0). This integral also has amazing properties when it comes to sequences of functions (Fatou's lemma, Dominated Convergence Theorem). Furthermore, when a function is Riemann integrable, the two integrals agree, so it is truly a generalization of classical integrals. For these reasons, we will need the stronger integration techniques to solve higher PDEs. Now, we could easily spend an entire course (or actually several) focused on simply measures and integration theory, but we will just quickly take some highlights for our purposes.

In the most basic sense, a measure is a way to generalize the notion of length or volume. We will only be talking about the *Lebesgue measure*, though there are many, many types of measures that are important even in Euclidean spaces. As a generalization of volume, it agrees with volume on basic sets. For example, an interval $[0, 5] \subset \mathbb{R}$ has measure $\mu([0, 5]) = 5$. In $\mathbb{R}^n$, a rectangle $R = \prod [a_i, b_i]$ has measure $\mu(R) = \prod (b_i - a_i)$, the formula you know from elementary school. When we encounter other sets, we measure them by covering them with rectangles. For example, consider $B(0, 1) \subset \mathbb{R}^2$. We can place it inside a rectangle of side-length 2, so $\mu(B(0, 1)) \leq 4$. If you do the same thing with rectangles of side-length $1/5$, it is possible to cover the circle with 88 of them, giving an upper bound of 3.52. It gets tricky to count, but using a tool like Desmos, you may plot the circle and notice that the area approximated decreases to exactly π , the expected area. In general, we use this type of process to measure weird sets, defining a function $\mu : P(\mathbb{R}^n) \rightarrow [0, \infty]$ by

$$\mu(A) = \inf \left\{ \sum_{j=1}^{\infty} \text{vol}(R_j) \mid A \subset \bigcup_{j=1}^{\infty} R_j \right\}$$

[You will not be asked to use this definition]. Technically, this is called an *outer measure*, but we will not care about the distinction here because we will not worry much about the details of measurable sets. We will add the assumption that sets and functions are measurable onto our theorems because some wacky fringe cases.

The *characteristic function* of a set A is

$$1_A(x) = \begin{cases} 1 & x \in A \\ 0 & x \notin A \end{cases}$$

And such a set allows us to define basic integrals, provided A is measurable, as

$$\int_{\mathbb{R}^n} 1_A d\mu = \mu(A)$$

to agree with the normal definition of volume. To get to an integral of a function f , we first extend by linearity

$$\int \sum_{A \in \mathcal{A}} c_A 1_A d\mu = \sum c_A \mu(A)$$

then, we use linear combinations like this to approximate functions (again, only measurable functions). For example, consider $\sin(x)$ for $x \in [0, \pi]$. Since $\sin(x) \in [-1, 1]$, let $A_k = \sin^{-1}([\frac{k-1}{n}, \frac{k}{n}])$. Then,

$$\psi_n(x) = \sum \frac{k}{n} 1_{A_k}$$

sits just under the sine curve, like a stack of coins. As $n \rightarrow \infty$, $\int_0^\pi \psi_n(x) d\mu = \sum \frac{k}{n} \mu(A_n)$ gives $\int_0^\pi \sin(x) dx$.

Beyond this concept, we ask the reader to accept without justification some facts following from the definition. First, changing a function on a set of measure 0 doesn't change the integral, so when working with integration, we actually work with equivalence classes where $f \equiv g$ iff. $f = g$ except on a set of measure 0. This is precisely the same as saying

$$\int_{\Omega} |f - g| d\mu = 0$$

This does occasionally cause issues, since a continuous function may be equivalent to a non-continuous one. For this purpose, we say f is C^m if it is equivalent to a C^m function, and we always assume we are working with such a representative of the equivalence class.

Second, this integral agrees with the Riemann integral on functions which are Riemann integrable, so we swap $d\mu$ for dx .

Then, we say a function $f : \Omega \rightarrow \mathbb{C}$ is *integrable* if

$$\int_{\Omega} |f| dx < \infty$$

For $p \geq 1$, we define the space of p -integrable functions by

$$L^p(\Omega) = \{f : \Omega \rightarrow \mathbb{C}; \int_{\Omega} |f|^p dx < \infty\}$$

(the spaces with $p \leq 1$ are studied, but have very different properties). These are the vector spaces we are going to care most about. To show that it is truly a vector space, notice that it is closed under scalar multiplication. Then, since $x \mapsto |x|^p$ is convex when $p \geq 1$,

$$|\frac{f+g}{2}|^p \leq \frac{|f|^p + |g|^p}{2}$$

so we are closed under addition. The other axioms may be checked easily due to the linearity of the integral. We define the L^p norm

$$\|f\|_p = \left(\int_{\Omega} |f|^p dx \right)^{1/p}$$

where this is truly a norm due to the *Minkowski Inequality*

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p$$

(we omit the proof). In the case $p = 1$, this follows from the triangle inequality for absolute value, and in the case $p = 2$, the norm arises from an inner product

$$\langle f, g \rangle = \int_{\Omega} f \bar{g} dx$$

and the Cauchy-Schwartz inequality gives the triangle inequality (see the next section).

What happens when we change p ? If we look at the basic case of $f = k1_A$, $\|k1_A\|_p = k\mu(A)^{1/p}$. As $p \rightarrow \infty$, this approaches k as long as A has nonzero, finite measure. Thus, we are instead measuring the height of the function, much like a supremum. This leads us to define the *essential supremum* of a set $A \subset \mathbb{R}$ and the *infinity norm* for a function $f : \Omega \rightarrow \mathbb{C}$ by

$$\text{ess-sup}(A) = \int \{a \in \mathbb{R} \mid \mu(A \cap [a, \infty) = 0\} \|f\|_{\infty} = \inf\{a \in \mathbb{R} \mid \{|f| > a\} \text{ has measure } 0\}$$

Question: What is the infinity norm of $\sin(x)$, $\frac{1}{x}$, and of $1_{\mathbb{Q}}$?

We also define $L^{\infty}(\Omega)$ to be the set of functions with finite infinity norm.

2.2 Convergence, Completeness

Convergence in a normed vector space V means $\{v_n\} \rightarrow v$ when

$$\lim_{n \rightarrow \infty} \|v_n - v\| = 0$$

Recall from early in the class that, when a limit may not be known, we also talk about a sequence being *Cauchy*, which is to say that given any $\epsilon > 0$, there is some N so for $k, m \geq N$,

$$\|v_k - v_m\| < \epsilon$$

and that a *complete* space is one in which every Cauchy sequence converges.

We will take as given two facts about L^p spaces.

- [2.1]** 1.) Assume $1 \leq p < \infty$. For any $f \in L^p(\Omega)$, there exists some $\{\psi_k\} \subset C_c^{\infty}(\Omega)$ so $\psi_k \rightarrow f$.
 2.) Assume $1 \leq p \leq \infty$ and Ω is a domain. Then, $L^p(\Omega)$ is complete.

Question: A subspace $W \subset V$ is closed if it contains the limit of every sequence in W that converges in V . Show that a closed subspace of a complete space is complete.

Example 1: The sequence in $C^0([-1, 1])$ given by

$$f_n(x) = \begin{cases} -1 & x < -\frac{1}{n} \\ nx & -\frac{1}{n} \leq x \leq \frac{1}{n} \\ 1 & x > \frac{1}{n} \end{cases}$$

is Cauchy since (using the area of two triangles)

$$\|f_k - f_m\|_1 = \left| \frac{1}{k} - \frac{1}{m} \right|$$

and so there is a limit in L^1 , which is precisely $f(x) = 1_{[0,1]} - 1_{[-1,0]}$. Notice that we fail to converge in the space of continuous functions.

Example 2:

We may also define $l^p(\mathbb{C})$ to be the space of sequences $\{a_i\}$ such that

$$\|a\|_{l^p} = \left[\sum_{j=1}^{\infty} |a_j|^p \right]^{1/p} < \infty$$

For example, $\{\frac{1}{n^2}\}$ is in l^1 , but $\{\frac{1}{n}\}$ is not (though it is in l^2). This space is often used to come up with examples in the more difficult L^p spaces. From the previous sequence, we notice

$$f(x) = \sum \frac{1}{n} 1_{[n, n+1)}$$

is in L^2 but not L^1 , whereas

$$g(x) = \sum n 1_{[n, n+\frac{1}{n^3}]}$$

is in L^1 but not L^2 .

3 Linear Algebra Recap: Inner Products

Recall that on a vector space V over $\mathbb{C}$, an *inner product* is a function $u, v \mapsto \langle u, v \rangle \in \mathbb{C}$ such that

- 1.) $\langle v, v \rangle \geq 0$, with $\langle u, v \rangle = 0$ if and only if $v = 0$
- 2.) $\langle u, v \rangle = \overline{\langle v, u \rangle}$
- 3.) For $c_1, c_2 \in \mathbb{C}$ and $v_1, v_2, w \in V$,

$$\langle c_1 v_1 + c_2 v_2, w \rangle = c_1 \langle v_1, w \rangle + c_2 \langle v_2, w \rangle$$

The inner product measures the angle between the two vectors, so we call u, v *orthogonal* if $\langle u, v \rangle = 0$. This generalizes the fact that, for $u, v \in \mathbb{R}^n$, $u \cdot v = |u||v|\cos(\theta)$ where θ is precisely that angle between u and v .

On an inner product space, we define $\|v\| = \sqrt{\langle v, v \rangle}$. Our most commonly seen inner product space is $\mathbb{R}^n$. In general, a norm $\|\cdot\|$ need not arise from an inner product.

Example: The space $C^m(\Omega)$ is a vector space for $\Omega \subset \mathbb{R}^n$. The standard norm on $C^m(\Omega)$ is

$$\|f\| = \sum_{i=0}^m \sup_{\Omega} |f^{(i)}(x)|$$

This norm does not arise from an inner product (we'll show this on the HW).

Similarly, $C_c^m(\Omega)$ is an inner product space with standard inner product

$$\langle f, g \rangle = \int_{\Omega} f \bar{g} dx$$

In the norm case, the triangle inequality is extremely useful. In the inner product case, we gain a further inequality

[3.1] (Cauchy-Schwartz Inequality) For an inner product space V with $\|\cdot\|$ defined by the inner product,

$$|\langle v, w \rangle| \leq \|v\| \|w\|$$

for all $v, w \in V$.

Note: In the Euclidean case, this is very easy to see because $|\cos(\theta)| \leq 1$, but it can also be derived directly from sum properties.

Proof:

For $t \in \mathbb{R}$, consider

$$q(t) = \|v + t\langle v, w \rangle w\|^2$$

if $w = 0$, the inequality is immediate, and in the case $w \neq 0$,

$$\begin{aligned} q(t) &= \langle v + t\langle v, w \rangle w, v + t\langle v, w \rangle w \rangle \\ &= \|v\|^2 + 2t\|\langle v, w \rangle\|^2 + t^2|\langle v, w \rangle|^2\|w\|^2 \end{aligned}$$

which is a quadratic in t . We use calculus to say that the minimum is at $t_0 = -\|w\|^{-2}$, and

$$0 \leq q(t_0) = \|v\|^2 - \frac{|\langle v, w \rangle|^2}{\|w\|^2}$$

which is the desired inequality.

Q.E.D.

Remark: The Cauchy-Schwartz inequality also implies the triangle inequality by

$$\begin{aligned} \|u + v\|^2 &= \langle u + v, u + v \rangle \\ &= \|u\|^2 + 2\operatorname{Re}\langle u, v \rangle + \|v\|^2 \\ &\leq \|u\|^2 + 2|\langle u, v \rangle| + \|v\|^2 \\ &\leq \|u\|^2 + 2\|v\|\|u\| + \|v\|^2 = \|u + v\|^2 \end{aligned}$$

3.1 Hilbert Spaces, Orthonormal Bases

A *Hilbert Space* is a complete inner-product space (such as $\mathbb{R}^n$). We will consider an infinite-dimensional complex Hilbert space (such as $L^2([0, 1])$). A sequence of vectors $\{e_n\}$ in this space is orthonormal if the vectors are pairwise orthonormal, or $\langle e_n, e_m \rangle = 0$ if $n \neq m$ and each vector has norm 1.

An *orthonormal basis* for H is a an orthonormal sequence so each $v \in H$ admits a unique representation as a convergent series

$$v = \sum_{j=1}^{\infty} c_j e_j$$

Notice that convergence here means absolute convergences of the sum $\sum |c_j|$. This allows us to compute c_k , since

$$\left\langle \sum_{j=1}^n c_j e_j, e_k \right\rangle = c_k$$

for $n \geq k$ and since Cauchy-Schwartz implies the inner product is continuous in each entry, taking the limit gives $c_k = \langle v, e_k \rangle$. Therefore, we often write $c_j[v] = \langle v, e_j \rangle$ and also sometimes talk about sums

$$S_n[v] = \sum_{j=1}^n c_j[v] e_j$$

Therefore, $S_n[v] \rightarrow v$ for every v if and only if $\{e_j\}$ is a basis.

We tend to think of a basis as a way to break down a vector into its components. Then, each component should be smaller than the original, the following result clarifies just this in our case.

[3.2] Bessel's Inequality Assume that $\{e_j\}$ is an orthonormal sequence in Hilbert space H . For $v \in H$, the series

$$\sum |c_j[v]|^2$$

converges, and

$$\sum_{j=1}^{\infty} |c_j[v]|^2 \leq \|v\|^2$$

with equality iff. $S_n[v] \rightarrow v$ in H .

Note: From Bessel's inequality, we know that $S_n[v]$ always converges to some vector, but if $\{e_j\}$ is not a basis, this vector becomes the projection of v onto the span of this sequence.

Proof:

First,

$$\|v - S_n[v]\|^2 = \langle v, v \rangle - 2\operatorname{Re}(\langle v, S_n[v] \rangle) + \langle S_n[v], S_n[v] \rangle$$

Since the e_j are orthonormal, $\langle v, S_n[v] \rangle = \langle S_n[v], S_n[v] \rangle = \sum_{j=1}^n |c_j|^2$. Therefore,

$$\|v - S_n[v]\|^2 = \|v\|^2 - \sum_{j=1}^n |c_j|^2$$

Since the left hand side is nonnegative, we have established the inequality

$$\sum_{j=1}^n |c_j|^2 \leq \|v\|^2$$

where the infinite sum reaches equality precisely when $\|v - S_n[v]\|^2 \rightarrow 0$.

Q.E.D.

We end with a characterization of a basis that will be useful in our search to compute bases.

[3.3] Suppose H is an infinite-dimensional Hilbert space. An orthonormal sequence in H forms a basis iff. 0 is the only element orthogonal to every vector in the sequence.

Proof:

If $\{e_j\}$ forms a basis, a vector v being orthogonal to every term means $c_j[v] = 0$ for all j . Since this is a basis, $\sum 0$ converges to v , so $v = 0$.

Next, let $\{e_j\}$ simply be orthonormal and assume that 0 is the only vector orthogonal to every e_j . Let $S_n[v] \rightarrow \tilde{v}$ and notice that

$$\langle v - S_n[v], e_j \rangle = 0$$

for all $n \geq j$. Taking the limit on n , $\langle v - \tilde{v}, e_j \rangle = 0$ for all j , so that $v = \tilde{v}$ by assumption. We then have a basis.

Q.E.D.

3.2 Self-Adjoint Operators

When we investigated the separation of variables, a key computation was to find the eigenvalues of the Laplacian. This motivates us to understand how we may simplify operators like the Laplacian in similar ways to what we do with matrices when we diagonalize, in particular for orthogonal diagonalization. Recall that a matrix may be orthogonally diagonalized when it is symmetric. This meant that $Ax \cdot y = x \cdot Ay$. If we generalize to

$$\langle Av, w \rangle = \langle v, Aw \rangle$$

for all $u, v \in H$ a Hilbert space, this property is called being self-adjoint (the adjoint of A is the operator so $\langle Av, w \rangle = \langle v, A^*w \rangle$). In particular, this property leads to a much more technical version of the Spectral theorem that we will not use. When we look at L^2 , the Laplacian is a self-adjoint operator with some technicalities to work out non-differentiable functions. To avoid those technicalities, we reduce to C^2 :

[3.4] Suppose that $\Omega \subset \mathbb{R}^n$ is a bounded domain with C^1 boundary. If $u, v \in C^2(\Omega)$ both satisfy either Dirichlet or Neumann boundary conditions then

$$\langle \Delta u, v \rangle = \langle u, \Delta v \rangle$$

Proof:

$$\int_{\Omega} u \overline{\Delta v} - \overline{v} \Delta u dx = \int_{\partial\Omega} u \partial_{\nu} \overline{v} - \overline{v} \partial_{\nu} u dS = 0$$

by the boundary conditions.

Q.E.D.

We call boundary conditions where the above property holds *self-adjoint boundary conditions*.

Example: On $[0, \pi]$, recall that the eigenvalue equation with Dirichlet boundary conditions is

$$\begin{aligned} \phi'' &= -\lambda \phi \\ \phi(0) &= \phi(\pi) = 0 \end{aligned}$$

and has eigenvalue solutions $\lambda = n^2$ for $n \in \mathbb{N}$ with eigenfunctions $\sin(nx)$. Notice, in particular, that

$$\int_0^{\pi} \sin(nx) \sin(mx) dx = \int_0^{\pi} \cos((m+n)x) - \cos((m-n)x) dx = 0$$

so that this sequence is orthogonal. If we allow Neumann boundary conditions, we instead of eigenfunctions $\{\cos(nx)\}$, which are also orthogonal. Furthermore, we can form convergent Fourier series for any C^2 functions on $[0, \pi]$ using these eigenfunctions, so they are a basis in some sense (in fact a basis for L^2).

[3.5] Suppose $\{\lambda_j\}$ is a sequence of eigenvalues of $-\Delta$ on a bounded domain $\Omega \subset \mathbb{R}^n$ with eigenvectors in $C^2(\overline{\Omega})$ subject to self-adjoint boundary conditions. Then, $\lambda_j \in \mathbb{R}$ and, after some possible rearrangement, the eigenvectors form an orthonormal sequence in $L^2(\Omega)$. Furthermore, $\lambda_j > 0$ for Dirichlet conditions and $\lambda_j \geq 0$ for Neumann boundary conditions.

Proof:

Suppose we have a sequence of eigenvectors $\{\phi_j\}$ corresponding to eigenvalues $\{\lambda_j\}$. The self-adjoint condition implies

$$-\lambda_j \langle \phi_j, \phi_k \rangle = \langle \Delta \phi_j, \phi_k \rangle = \langle \phi_j, \Delta \phi_k \rangle = -\overline{\lambda_k} \langle \phi_j, \phi_k \rangle$$

In the case $j = k$, the inner product is nonzero and $\lambda_j = \overline{\lambda_j}$ implies the eigenvalues are real. If $\lambda_j \neq \overline{\lambda_k}$, the functions must be orthogonal.

If there are multiple eigenfunctions for one eigenvalue, we may apply Gram-schmidt to create an orthogonal sequence of eigenfunctions for that eigenvalue. Further, for all eigenfunctions, we may divide by their magnitude to make an orthonormal sequence.

Lastly, given an orthonormal sequence

$$\lambda_j = \langle -\Delta\phi_j, \phi_j \rangle = \int_{\Omega} |\nabla\phi_j|^2 dx - \int_{\partial\Omega} \overline{\phi_j} \partial_{\nu}\phi_j dS$$

so that, in the case of Dirichlet and Neumann boundary conditions, we have $\lambda_j \geq 0$.

If $\lambda_j = 0$, this also shows the eigenfunction is constant, so that in the Dirichlet case the only such solution is 0 and $\lambda_j > 0$.

Q.E.D.