18.102 — Introduction to Functional Analysis

Class by Marjorie Drake

Notes by Sanjana Das

Spring 2024

Notes for the MIT class **18.102** (Introduction to Functional Analysis), taught by Marjorie Drake. All errors are my responsibility.

Contents

§1 February 5, 2024

The textbook for this course is *Introductory Functional Analysis* by Erwin Kreyszig. (There's several other recommended textbooks on the syllabus, which will be posted on Canvas later today.)

§1.1 Course introduction

First, we'll give a quick introduction to functional analysis.

Functional analysis originated in the 1900s with people like Hilbert. At the time, there was a fashionable trend to axiomize mathematics. In particular, we've all taken linear algebra. In linear algebra, if we have a map $A\colon \mathbb{R}^d \to \mathbb{R}^m$ and we know it's linear, then there's a clean, elegant theory — we can talk about the eigenvalues and eigenvectors of A, and we understand lots about it.

Marjorie works in a field called *extension theory*; this has to do with the theoretical side of interpolation of data. Extension theory is contained in *harmonic analysis*, which is contained in *analysis*. We've taken a real analysis course by now, so we know analysis is like calculus, but broader. In calculus, we're looking at functions $f\colon \mathbb{R}^n \to \mathbb{R}$ which are *really nice* — by *nice* we mean smooth in some sense, e.g., differentiable. Then *Taylor's theorem* says that at a point, a differentiable function looks approximately like a line; if it's twice differentiable then it looks like a parabola; if it's three times differentiable, then it looks like a cubic; and so on. So in particular, up to some error, at this point the function is *linear*. And that means we can use some of the theory of linear algebra to understand what's happening.

But in calculus, we only really deal with functions that are way too nice — they're super smooth and friendly. In analysis, we're trying to think about more interesting classes of functions and types of functions. Then the question is, can we still use linear algebra and this clean elegant theory to understand what's going on?

This brings us to functional analysis. In functional analysis, we'll begin by studying *function spaces* — usually, these are *normed linear spaces*. (We'll define a norm precisely later, but it's like a distance — it's related to metric spaces, but we require a bit more structure.) We'll talk about *complete normed linear spaces* — *complete* means that every Cauchy sequence converges. (These are also called *Banach spaces*.) And then we'll talk about *Hilbert spaces*, which have an *inner product structure* — an inner product is the same as a dot product. These spaces have structure that looks most like the structure from linear algebra we've worked with before, so these will be in some sense the most familiar.

We'll introduce these objects. Then the quintessential theorems of functional analysis are the *Hahn–Banach theorem* (about the dual space of a function space — linear maps from one of these spaces to $\mathbb{R}$) and the *uniform boundedness theorem* (about when operators are bounded or linear or continuous, and what happens with sequences of operators), and the *open mapping theorem*.

The final big section of the course is *spectral theory*. In linear algebra, we talk about eigenvalues and eigenvectors. If we want to generalize to operators on function spaces, that generalization is called *spectral theory*. When we finish this, assuming we're on schedule, we'll cover a bit of topics that these tools show up in — we'll see a bit of Fourier spaces, distributions, a taste of extension theory, Sobolev spaces (these are some spaces that show up in the study of advanced PDEs).

§1.2 Measure theory

We'll now give a review of measure theory. (There's a textbook by Royden that takes a similar approach, which we can use as a reference.)

We'll do this abstractly first.

Definition 1.1. Let S be a set. A family of subsets $\mathcal{S}$ is called an *algebra* if:

(1) $\emptyset \in \mathcal{S}$.

(2) If $A \in \mathcal{S}$, then $A^c \in \mathcal{S}$.

(3) $\mathcal{S}$ is closed under unions — if $A, B \in \mathcal{S}$, then $A \cup B \in \mathcal{S}$.

If $\mathcal{S}$ is also closed under *countable unions* — i.e., for every family $(A_i)_{i \in \mathbb{N}}$ of sets with $A_i \in \mathcal{S}$ for all i, we have $\bigcup_{i \in \mathbb{N}} A_i \in \mathcal{S}$ — then we say $\mathcal{S}$ is a *σ-algebra*.

Example 1.2 (Borel σ-algebra)

If (S, τ) is a topological space (here, τ is the collection of open sets), then $\sigma(\tau)$ is the *Borel σ-algebra* generated by all open sets — the smallest σ-algebra containing all the open sets. We write $\mathcal{B}(\mathbb{R})$ to represent the Borel σ-algebra on $\mathbb{R}$ generated by open sets defined by the Euclidean metric.

Definition 1.3. A pair $(S, \mathcal{S})$ where S is a set and $\mathcal{S}$ is a σ-algebra is a *measurable space*.

Definition 1.4. Given a measurable space $(S, \mathcal{S})$ and a function $f: S \to \mathbb{R}$, we say f is *$\mathcal{S}$-measurable* if for every Borel set $B \in \mathcal{B}(\mathbb{R})$, we have $f^{-1}(B) \in \mathcal{S}$.

Remark 1.5. What's an example of a family of subsets that's an algebra but not a σ-algebra?

We can take our generating set to be $\{(n, n+1) \mid n \in \mathbb{N}\}$. Then $\{(2n+1, 2n+2) \mid n \in \mathbb{N}\}$ should be in the σ-algebra generated by this set, but not the algebra — it can be written as a countable union of these sets $(n, n+1)$, but it should be possible to show that we can't get this set by taking just complements and finite unions. This gives an algebra that's not a σ-algebra.

We've now defined what it means for a function to be $\mathcal{S}$-measurable; now we're going to define a measure.

Definition 1.6. Let $(S, \mathcal{S})$ be a measurable space, and let μ be a map $\mu: \mathcal{S} \to [0, \infty]$. We say μ is a *(positive) measure* if:

(1) $\mu(\emptyset) = 0$.

(2) μ is countably additive — this means given any collection of *pairwise disjoint* sets $\{A_i\}_{i \in \mathbb{N}}$ (i.e., sets such that $A_i \cap A_j = \emptyset$ for all i and j), we have $\mu(\bigcup_{i \in \mathbb{N}} A_i) = \sum_{i \in \mathbb{N}} \mu(A_i)$.

If this is true, we say $(S, \mathcal{S}, \mu)$ is a *measure space*.

Usually, when we talk about measures, we only mean positive measures (in particular, this will be true for this class).

§1.2.1 The Lebesgue measure

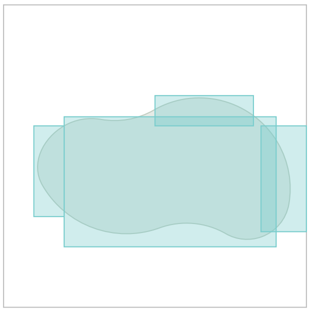
Next, we're going to define the *Lebesgue outer measure*.

Definition 1.7. Let $A \subseteq \mathbb{R}$ be *any* set. We define the *Lebesgue outer measure* $m^*(A)$ as

$$m^*(A) = \inf \left\{ \sum_{k=1}^{\infty} \ell(I_k) \mid A \subseteq \bigcup_{k=1}^{\infty} I_k \right\},$$

where each I_k is an interval.

In other words, we're looking at all the ways for A to be contained in a countable union of *intervals*, and we're adding up the *lengths* of these intervals; and we're taking the infimum of this sum of lengths of intervals (over all such ways to do so). (We use $\ell(I)$ to denote the length of an interval I.)



(This diagram is an illustration for two dimensions; in one dimension we'll have intervals instead of boxes.)

Lemma 1.8

The outer measure m^* satisfies the following properties:

- $m^*(I) = \ell(I)$ for an interval I.
- m^* is *countably subadditive* — this means that for any pairwise disjoint sets $\{A_i\}_{i \in \mathbb{N}}$, we have $\mu(\bigcup_{i=1}^\infty A_i) \le \sum_{i=1}^\infty \mu(A_i)$.

Importantly, m^* is *not* countably additive; there exists a collection of sets $\{A_i\}_{i \in \mathbb{N}}$ for which this inequality is strict; one example is the *Vitali set*.

Example 1.9 (Vitali set)

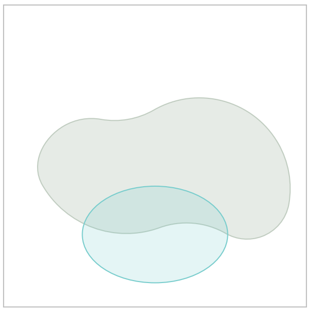
Consider the set $[0,1]$, and say $x \sim y$ if $x = y + r$ for some $r \in \mathbb{Q}$. Let V be a set consisting of one representative of each equivalence class, and for each $p \in \mathbb{Q}$, let $V_p = V + p = \{v + p \mid v \in V\}$. (We may need to intersect with $[0,1]$.) Then we have $\bigcup_{p \in \mathbb{Q}} V_p = [0,1]$. But the Lebesgue outer measure of $[0,1]$ is 1, while $m^*(V_p) = 0$ for each V_p. This means $m^*(\bigcup_p V_p) = 1$, while $\sum m^*(V_p) = 0$.

(The inequality points the wrong way here, so something has gone wrong; I'm not sure what, but it's probably that $m^*(V_p)$ is not 0. We'll fix this next class.)

So the Lebesgue measure has a problem, that it's not countably additive. But we have a criterion due to Caratheodory to fix this.

Definition 1.10 (Caratheodory criterion)**.** Given a set S and outer measure m^*, we say that $A \subseteq S$ is *measurable* if for all $B \subseteq S$, we have $m^*(B) = m^*(A \cap B) + m^*(A^\mathsf{c} \cap B)$.

Remark 1.11. This is kind of about splitting sets — you only want to include sets A in your measure if they split other sets the way you'd expect them to. This criterion essentially says that A splits all sets B the way you'd intuitively expect.



This criterion is fantastic because the collection of measurable sets is a σ-algebra, and m^* is countably additive on such sets — and this collection contains the Borel σ-algebra. So then if we consider $(S, \mathcal{S}, m^*)$ where S is some set, $\mathcal{S}$ is the σ-algebra of *measurable* sets, and m^* is the Lebesgue outer measure, then $(S, \mathcal{S}, m^*)$ is a measure space. In particular, when m^* is restricted to just $\mathcal{S}$ (the σ-algebra of measurable sets), we write m in place of m^* (and we call it the *Lebesgue measure*).

Remark 1.12. In other words, this criterion says that a set A is measurable if for *every* subset B of S, the measure of that subset B can be written as $m^*(B) = m^*(A \cap B) + m^*(A^c \cap B)$. We won't prove that the set of such A forms a σ-algebra, but this is true. And when we restrict m^* to this σ-algebra, we define a measure, which we call the *Lebesgue measure*.

We can define the Lebesgue measure in $\mathbb{R}^n$ by extending this construction from intervals $I = (a, b)$ to rectangles $R = \prod_{i=1}^n (a_i \times b_i)$. (We won't do this in more detail right now; but the idea is that the Caratheodory criterion isn't restricted to just $\mathbb{R}$, and we can talk about the same concept in $\mathbb{R}^n$.)

§1.2.2 Functions

Definition 1.13. We use $\mathcal{L}^0(S, \mathcal{S})$ (which we may write as $\mathcal{L}^0(S)$ when $\mathcal{S}$ is evident, or just $\mathcal{L}^0$ when S is also evident) to denote the vector space of measurable functions on S.

Example 1.14

Any continuous function is Lebesgue measurable, because the pre-image of an open set is open.

(Most functions you'll want end up being Lebesgue measurable, but not all end up being Lebesgue *integrable*, which we'll see next class.)

Definition 1.15 (Simple functions)**.** We say $f \in \mathcal{L}^{0,\text{simple}}(S, \mathcal{S})$ if f takes on finitely many values — i.e., $f = \sum_{k=1}^n \alpha_k \chi_{A_k}$ for some disjoint sets $A_1, \ldots, A_k$, where

$$\chi_{A_k} = \begin{cases} 1 & \text{if } x \in A_k \\ 0 & \text{if } x \notin A_k \end{cases}$$

is the characteristic function (or indicator function) of A_k.

§2 February 7, 2023

§2.1 Review

Last time, we defined the Lebesgue measure space — we have a set S (e.g., $\mathbb{R}$ or $\mathbb{R}^n$), the σ-algebra $\mathcal{S}$ of measurable sets, and a measure μ (the Lebesgue outer measure restricted to $\mathcal{S}$). To define $\mathcal{S}$, we use Caratheodory's criterion — we say $A \subseteq S$ is *Lebesgue measurable* if for all $B \subseteq S$, we have

$$\mu^*(B) = \mu^*(A \cap B) + \mu^*(A^c \cap B).$$

Remark 2.1. We didn't prove that sets satisfying this property form a σ-algebra. They actually don't; they form an *algebra*. We then use this algebra to generate a σ-algebra; and that set is $\mathcal{S}$.

§2.2 The Lebesgue integral

Definition 2.2. We say a function $f\colon S \to \mathbb{R}$ is *Lebesgue measurable* if for all Borel sets $B \in \mathcal{B}(\mathbb{R})$, we have $f^{-1}(B) \in \mathcal{S}$.

We'll now define the vector space of measurable functions.

Definition 2.3. Let $\mathcal{L}^0(S, \mathcal{S}, \mu)$ be the vector space of measurable functions.

We'll sometimes abbreviate this as $\mathcal{L}^0(S, \mathcal{S})$, $\mathcal{L}^0(S)$, or just $\mathcal{L}^0$, if the remaining parameters are self-evident.

Remark 2.4. We haven't said why $\mathcal{L}^0$ is a vector space; this is worth thinking about.

Definition 2.5. We say a function $f\colon S \to \mathbb{R}$ is a *simple* function, written $f \in \mathcal{L}^{0,\text{simp}}$, if it takes on finitely many values — equivalently, we can write $f(x) = \sum_{i=1}^n \alpha_i \chi_{A_i}$ for a collection of measurable sets $\{A_i\}_{i=1}^n$ (i.e., sets $A_i \in \mathcal{S}$), where $\chi_{A_i}(x)$ (or equivalently $\mathbf{1}_{A_i}(x)$) is the indicator function of A_i.

Definition 2.6. We let $\mathcal{L}_+^{0,\text{simp}}$ be the space of nonnegative simple functions.

For $f \in \mathcal{L}_+^{0,\text{simp}}$, what do we think $\int f$ should be? We'd expect to have $\int_S f\, d\mu = \sum_{i=1}^n \alpha_i \mu(A_i)$. And that's exactly how we define the Lebesgue integral.

Definition 2.7. For $f \in \mathcal{L}_+^{0,\text{simp}}$, we define $\int_S f\, d\mu = \sum_i \alpha_i \mu(A_i)$.

Definition 2.8. Let $\mathcal{L}_+^0$ be the space of nonnegative measurable functions. Then we define

$$\int_S f\, d\mu = \sup\left\{ \int_S g\, d\mu \mid g \in \mathcal{L}_+^{0,\text{simp}},\ g \le f \right\}.$$

In other words, we look at all ways to approximate f from below by a simple (nonnegative) function g; and we define $\int_S f$ based on these approximations.

Finally, we'll define the integral for a general function $f \in \mathcal{L}^0$. To do so, note that we can write $f(x) = f^+(x) - f^-(x)$ where

$$f^+(x) = \max\{f(x), 0\} \text{ and } f^-(x) = \max\{-f(x), 0\}.$$

Then f^+ and f^- are both elements of $\mathcal{L}_0^+$. We'd expect $\int_S f\, d\mu = \int_S f^+\, d\mu - \int_S f^-\, d\mu$, and so this is how we define it.

Definition 2.9. For a general function $f \in \mathcal{L}^0$, we define $\int_S f\, d\mu = \int_S f^+\, d\mu - \int_S f^-\, d\mu$.

Definition 2.10. We say a function $f \in \mathcal{L}^0$ is *Lebesgue integrable* if $\int_S f^+\, d\mu$ and $\int_S f^-\, d\mu$ are both finite. We use $\mathcal{L}^1$ to denote the set of Lebesgue integrable functions.

Example 2.11

The function $f(x) = x$ on $\mathbb{R}$ is not in $\mathcal{L}^1$.

Definition 2.12. For $f, g \in \mathcal{L}^0$, we say that $f = g$ *almost everywhere* (abbreviated a.e.) if

$$\mu(\{x \in S \mid f(x) \neq g(x)\}) = 0.$$

Lemma 2.13

If $f = g$ almost everywhere, then $\int f\, d\mu = \int g\, d\mu$.

So we've defined the Lebesgue integral for measurable functions, and we've defined what it means for a function to be Lebesgue integrable.

§2.3 Integrals of sequences

One reason the Lebesgue integral is nice is that the Riemann integral doesn't behave well with respect to limits — the Riemann integral involves partitioning the domain, and this requires that our functions be very smooth in some sense. (If the function isn't continuous, you can imagine building an example where the value of your concept of an integral changes as you shrink the partition size.) On the other hand, the Lebesgue integral behaves very well with respect to limits.

We'll now look at sequences of functions; we want our integrals to behave well with respect to taking limits.

Question 2.14. When is it true that $\lim_{n\to\infty} \int f_n(x)\, d\mu = \int \lim_{n\to\infty} f_n(x)\, d\mu$?

For the Riemann integral this is 'rarely' true; this is one of the big advantages of the Lebesgue integral.

The first statement of this form is the monotone convergence theorem.

Theorem 2.15 (Monotone convergence theorem)

Let $\{f_n\}_{n\in\mathbb{N}}$ be a sequence in $\mathcal{L}^0_+(\mathbb{R})$ satisfying $f_1(x) \leq f_2(x) \leq \cdots$ for all $x \in \mathbb{R}$. Then

$$\lim_{n\to\infty} \int f_n(x)\, d\mu = \int f(x)\, d\mu,$$

where $\lim_{n\to\infty} f_n(x) = f(x)$ for all $x \in \mathbb{R}$.

So here we're looking at nonnegative functions on $\mathbb{R}$ (still with the Lebesgue measure).

Remark 2.16. Right now, we'll write $d\mu$ instead of dx to emphasize that we're working with the Lebesgue measure; at some point we will probably switch to dx (because people generally don't write $d\mu$ unless working in abstract measure theory).

Remark 2.17. For *any* function $f \in \mathcal{L}^0_+(\mathbb{R})$, there exist simple functions $\{f_i\}_{i\in\mathbb{N}} \subseteq \mathcal{L}^{0,\text{simp}}_+$ satisfying $f_1 \leq f_2 \leq \cdots$ and $\lim_{n\to\infty} f_n(x) = f(x)$ for all x; in particular, we can choose our sequence f_i such that this convergence is uniform on the set $\{x \in \mathbb{R} \mid f(x) \leq m\}$ for all $m \geq 0$. So in particular, if f is bounded, then the convergence is uniform.

This result says we can build a sequence of increasing simple functions whose limit is any nonnegative function f we like. The same idea can be used to show that simple functions are dense in $\mathcal{L}^1(\mathbb{R})$.

This is an important concept because often in analysis, we want to prove a property holds for all functions in $\mathcal{L}^1$, and we start by proving it for a simpler class of functions. (A simpler class of functions can mean many things — e.g., continuity; in this case that class might be simple functions.)

Here's the second result on swapping integrals and limits.

Theorem 2.18 (Lebesgue's dominated convergence theorem)
Let $\{f_n\}_{n\in\mathbb{N}} \subseteq \mathcal{L}^0(\mathbb{R})$ satisfy $|f_n(x)| \leq g(x)$ almost everywhere, for some function $g \in \mathcal{L}^1(\mathbb{R})$, and suppose that $\lim_{n\to\infty} f_n(x) = f(x)$ for each x. Then $\lim_{n\to\infty} \int_{\mathbb{R}} f_n(x)\, d\mu = \int f(x)\, d\mu$.

This is one of the most general results about the convergence of sequences of functions.

§2.4 Metric spaces

Now we'll review metric spaces.

Definition 2.19. A *metric space* is a pair (X, d), where X is a set and d is a map $d\colon X \times X \to \mathbb{R}_{\geq 0}$ satisfying the following properties (for all $x, y, z \in X$):

- We have $d(x, y) = 0$ if and only if $x = y$.
- We have $d(x, y) = d(y, x)$.
- The triangle inequality — $d(x, y) \leq d(x, z) + d(y, z)$.

(Note that the fact d is a map to $\mathbb{R}_{\geq 0}$ means that we should have $0 \leq d(x, y) < \infty$ for all $x, y \in X$ — i.e., d is always nonnegative and finite.)

§2.4.1 Some examples

We'll now look at a few examples.

Example 2.20
Some examples of metric spaces:

- $\mathbb{R}$ is a metric space with $d(x, y) = |x - y|$;
- $\mathbb{R}^2$ is a metric space with $d(x, y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}$ (i.e., the Euclidean metric).
- More generally, $\mathbb{R}^n$ is a metric space with

$$d(x, y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \cdots + (x_n - y_n)^2}.$$

One very important example for us is *sequence* spaces.

Definition 2.21. We define ℓ^∞ as the space of *bounded* sequences $x = (x_1, x_2, \ldots) = (x_i)_{i\in\mathbb{N}}$, where

$$d(x, y) = \sup_i |x_i - y_i| .$$

Definition 2.22. We define $\mathcal{C}[a, b]$ as the space of continuous functions $[a, b] \to \mathbb{R}$, where for two functions $f, g \in \mathcal{C}[a, b]$, we define

$$d(f, g) = \max_{x\in[a,b]} |f(x) - g(x)| .$$

Remark 2.23. At this point we've talked about some metric spaces that were *finite*-dimensional ($\mathbb{R}$, $\mathbb{R}^2$, and $\mathbb{R}^n$), and one that was *countable* in dimension (ℓ^∞); but $\mathcal{C}[a, b]$ is neither. (We don't yet have the setup needed to define dimension; but we'll do this soon.)

Definition 2.24. Let X be a set. Then the *discrete metric space* on X is the metric space with distance defined as

$$d(x, y) = \begin{cases} 1 & \text{if } x \neq y \\ 0 & \text{if } x = y. \end{cases}$$

The final example, which we'll also talk a lot about, is ℓ^p spaces.

Definition 2.25. For $p \in [1, \infty)$, we define the ℓ^p *space* as the metric space on the set

$$X = \{x = (x_1, x_2, \ldots) \mid \|x\|_{\ell^p} < \infty\},$$

where $\|x\|_{\ell^p}$ is defined as

$$\|x\|_{\ell^p}^p = \sum_{i=1}^{\infty} |x_i|^p ,$$

with distance

$$d(x, y) = \|x - y\|_{\ell^p} = \left(\sum_{i=1}^{\infty} \|x_i - y_i\|^p\right)^{1/p} .$$

Remark 2.26. Today we talked about $\mathcal{L}^1$, which can be equivalently defined as the space of functions with $\int |f| \; d\mu < \infty$. We'll later talk about $\mathcal{L}^p$ — the space of functions with $\int |f|^p \; d\mu < \infty$. This is an analog of the ℓ^p spaces.

§2.4.2 Convergence

We'll now briefly talk about convergence, Cauchy sequences, and completeness. This should be review; there's more detail in Kreyszig if we don't remember this.

Definition 2.27. Let (X, d) be a metric space, and let $(x_n)_{n\in\mathbb{N}}$ be a sequence in X. We say $(x_n)_{n\in\mathbb{N}}$ is *convergent* if there exists $x \in X$ such that $\lim_{n\to\infty} d(x_n, x) = 0$. (We also write this as $\lim_{n\to\infty} x_n = x$, or as $x_n \to x$ as $n \to \infty$.)

Remark 2.28. As we continue in the course, we'll have to be careful about the notation $x_n \to x$ as $n \to \infty$. For example, we say f_n converges to f *pointwise* if $f_n(x) \to f(x)$ for all $x \in \mathbb{R}$. But $\mathcal{L}^1$ is also a metric space (which we will talk about later), and there's a notion of $f_n \to f$ in the $\mathcal{L}^1$ metric that's different from pointwise convergence (we say $f_n \to f$ in $\mathcal{L}^1$ if $\int_{\mathbb{R}} |f_n - f| \, d\mu \to 0$.). So we should be careful what type of convergence we mean.

Definition 2.29. We say a sequence $(x_n)_{n\in\mathbb{N}}$ is *Cauchy* if for all $\varepsilon > 0$, there exists some N (depending on ε) such that we have $d(x_n, x_m) < \varepsilon$ for all $n, m > N$.

Definition 2.30. We say a metric space is *complete* if every Cauchy sequence converges.

In other words, what it means for a metric space to be complete is that whenever $(x_n)_{n\in\mathbb{N}}$ is Cauchy, there exists some $x \in X$ such that $x_n \to x$.

§2.5 Remark on the Vitali set

Last time, we considered the interval $[0, 1]$; for $x, y \in [0, 1]$, we said $x \sim y$ if $x - y \in \mathbb{Q}$. (This is an equivalence relation.) We defined V as a set of one representative from each equivalence class (so $V \subseteq [0, 1]$). Then for each $p \in \mathbb{Q} \cap [0, 1]$, we define

$$V_p = \{x \in [0, 1] \mid x = y + p \text{ for } y \in V\}.$$

(In other words, we imagine translating the set V by p.)

Last time, we incorrectly said that $\mu^*(V_p) = 0$. It's actually that $\mu^*(V_p) = 1 - p$ (there might be a way to do the construction so that it ends up being just 1, but this way we get $1 - p$ because $V_p \subseteq [p, 1]$). (This is not really obvious; you have to use the definition of measure.)

Now the point is that $\mu^*(\bigcup_p V_p) = \mu^*([0, 1]) = 1$. But the sets V_p are disjoint, and we have $\sum_p \mu^*(V_p) = \infty$. In particular, this means μ^* is not countably additive (if it were, we would need to have $\sum_p \mu^*(V_p) = 1$).

So this shows that the outer measure is not countably additive, and V is an example of a set that is not measurable.

Remark 2.31. The fact that the sets V_p are disjoint follows from the fact that V is a set of representatives. Explicitly, suppose $x \in V_p \cap V_q$; then we can write $x = v + p = w + q$ for some $v, w \in V$. But then $v \sim w$, which is a contradiction.

§3 February 12, 2024

§3.1 Logistics

We have office hours today at 11-12 adn 4-5 in Room 2-255.

Midterm 1 is scheduled for March 4; Midterm 2 was originally scheduled for April 8, but because this is the date of a total solar eclipse which several students want to go to, we are pushing it back to April 10 or 17.

Marjie received feedback that this pset was long and not particularly deep. There will be some more interesting problems as we progress, but we have to use a lot of different tools before we can think about some of those problems; but she'll try to put more interesting problems on future psets. She will also shorten the assignments for the future.

Today we'll continue presenting new ideas, which we're going to start using them as we go further into the class.

§3.2 Vector spaces

Definition 3.1. A *vector space* over a field $\mathbb{K}$ is a nonempty set of elements X and two operations, *vector addition* and *scalar multiplication*, such that:

- Vector addition is commutative ($x + y = y + x$ for all $x, y \in X$) and associative ($x + (y + z) = (x + y) + z$ for all $x, y, z \in X$).
- X has a zero element — there is a vector $0 \in X$ such that $0 + x = x$ for all $x \in X$.
- Every element $x \in X$ has an additive inverse $-x$ (such that $x + (-x) = 0$).
- Scalar multiplication is distributive — for all $\alpha, \beta \in \mathbb{K}$ and $x, y \in X$, we have $\alpha(x+y) = \alpha x + \alpha y$ and $(\alpha + \beta)x = \alpha x + \beta x$.

Intuitively, a vector space is a generalization of $\mathbb{R}^n$. We'll think of our field $\mathbb{K}$ as either $\mathbb{R}$ or $\mathbb{C}$.

Example 3.2

The set $\mathbb{R}^n$ (the set of n-tuples of complex numbers) is a vector space, with addition componentwise. Similarly $\mathbb{C}^n$ is a vector space.

Example 3.3

The space $\mathcal{C}[a, b]$ of continuous functions on $[a, b]$ is a vector space, with $f + g$ defined as $(f + g)(x) = f(x) + g(x)$. (Usually we're discussing real-valued functions, so $\mathbb{K}$ is $\mathbb{R}$.)

Example 3.4

For any $p \in [1, \infty]$, the sequence space ℓ^p is a vector space, again with coordinate-wise addition.

Definition 3.5. A *subspace* of a vector space X is a nonempty subset $Y \subseteq X$ which is closed under vector addition and scalar multiplication — for all $\alpha \in \mathbb{K}$ and all $x, y \in Y$, we have $\alpha(x + y) \in Y$.

Note that the subspace Y is a vector space.

Example 3.6

If $X = \mathbb{R}^2$, the set $\mathbb{R} \subseteq \mathbb{R}^2$ is a subspace (and in fact, we can rotate it to get that *any* line through the origin is a subspace).

Definition 3.7. We say a set of vectors $\{x_1, \ldots, x_n\}$ is *linearly independent* if the equation $\alpha_1 x_1 + \cdots + \alpha_n x_n = 0$ (with $\alpha_i \in \mathbb{K}$) has only the trivial solution $\alpha_1 = \cdots = \alpha_n = 0$. If this is not the case, we say $\{x_1, \ldots, x_n\}$ is *linearly dependent.*

Definition 3.8. We define the *span* of a set of vectors $\{x_1, \ldots, x_n\}$ as the set of all linear combinations of $\{x_1, \ldots, x_n\}$ — i.e.,

$$\operatorname{Span}(x_1, \ldots, x_n) = \{\alpha_1 x_1 + \cdots + \alpha_n x_n \mid \alpha_i \in \mathbb{K}\}.$$

(These are all generalizations of the usual notions in $\mathbb{R}^n$.)

Definition 3.9. A vector space X is *finite-dimensional* if there exists some $n \in \mathbb{N}$ such that X contains a linearly independent set of n vectors, but *any* set of $n+1$ vectors is linearly dependent. In this case, we say X has dimension n.

In other words, the dimension is the maximum size of a linearly independent set.

Example 3.10

The space $\mathbb{R}^n$ has dimension n.

Example 3.11

The space $\mathcal{C}[a, b]$ is infinite-dimensional.

Looking back at some of the definitions from linear algebra, we can wonder how they carry over to infinite dimensions — for example, can we have a basis in an infinite-dimensional vector space? We'll discuss this later.

Definition 3.12. If a vector space X has dimension $\dim X = n$, a linearly independent set of n vectors is called a *basis* for X.

Remark 3.13. *Every* vector space has a basis (not just finite-dimensional ones), though this requires the axiom of choice (we'll later discuss Zorn's lemma and then prove this).

Remark 3.14. What we're doing today is trying to generalize what we've seen from finite-dimensional linear algebra to more interesting spaces like $\mathcal{C}[a, b]$; we can ask if *everything* generalizes; for bases, the answer is yes, but we need the axiom of choice.

§3.3 Normed spaces

Definition 3.15. A *normed (linear) space* is a vector space with a norm defined on it.

(We'll define norms soon.)

Definition 3.16. A *Banach space* is a complete normed space.

This means that Cauchy sequences (in the given norm) converge to elements in our space.

Now we'll talk about what a norm is.

Definition 3.17. A *norm* on a vector space X is a $\mathbb{R}$-valued function on X with the following properties:

(i) $\|x\| \geq 0$ for all $x \in X$.

(ii) $\|x\| = 0$ if and only if $x = 0$.

(iii) $\|\alpha x\| = |\alpha| \cdot \|x\|$ for all $\alpha \in \mathbb{K}$ and $x \in X$.

(iv) The triangle inequality — $\|x + y\| \leq \|x\| + \|y\|$ for all $x, y \in X$.

What $|\alpha|$ means depends on whether we're working over $\mathbb{R}$ or $\mathbb{C}$.

Fact 3.18 — A norm on X induces a metric on X, where $d(x, y) = \|x - y\|$.

Example 3.19

In $\mathbb{R}^n$, one norm is the Euclidean distance $\|x\| = \sqrt{x_1^2 + \cdots + x_n^2}$. This norm induces the Euclidean metric that we're familiar with.

Example 3.20

We can similarly define a norm on $\mathbb{C}^n$ as

$$\|x\| = \sqrt{|x_1|^2 + \cdots + |x_n|^2}.$$

Example 3.21

In ℓ^∞ (the space of bounded sequences), we can define a norm $\|\cdot\|_\infty$ as

$$\|x\|_\infty = \sup_{i \in \mathbb{N}} |x_i|.$$

This induces the metric we discussed on ℓ^∞ as well.

Example 3.22

For the space $\mathcal{C}[a,b]$, we define the *supremum norm* $\|\cdot\|_{\mathcal{C}[a,b]}$ as

$$\|f\|_{\mathcal{C}[a,b]} = \sup_{t \in [a,b]} |f(t)|.$$

Claim 3.23 — The space $\mathcal{C}[a,b]$ with the supremum norm is complete.

Proof. A Cauchy sequence in the supremum norm is *uniformly* convergent to some function f (obtained as a pointwise limit), and a uniform limit of continuous functions is continuous. □

So far, we've only discussed norms that induce a metric which is complete — i.e., if we take a sequence of points that's Cauchy in any of these vector spaces, then the Cauchy sequence is convergent to an element of the vector space. But we can also have normed spaces that are *not* complete; here is an example.

Example 3.24 (Incomplete normed space)

For the space $\mathcal{C}[a,b]$, we can define a norm $\|f\| = \int_a^b |f(t)|\, dt$. But there exists a sequence of functions f_n which is Cauchy — meaning that $\int_a^b |f_n - f_m| \to 0$ — but which does not have a limit in $\mathcal{C}[a,b]$, i.e., $\lim_{n\to\infty} f_n \notin \mathcal{C}[a,b]$.

Proof. As an example of such a sequence, we can take the interval $[0,1]$, with $f_n(x) = x^n$. Then the pointwise limit of f_n is

$$f(x) = \begin{cases} 0 & \text{if } x \in [0,1) \\ 1 & \text{if } x = 1. \end{cases}$$

(We can check that the sequence f_n is Cauchy — for $n \le m$ we have

$$\int_0^1 |f_n(x) - f_m(x)|\, dx = \int_0^1 (x^n - x^m)\, dx = \frac{1}{n} - \frac{1}{m},$$

which goes to 0 for large n and m.) □

Remark 3.25. Note that this sequence f_n isn't Cauchy in the supremum norm.

Remark 3.26. Doesn't the sequence f_n converge to the 0 function in this norm? The answer is yes — we have $\lim_{n\to\infty} \int_0^1 f_n(x)\,dx = 0$.

§3.4 Completions of normed spaces

In real analysis, we saw the following theorem about completing a *metric* space.

Theorem 3.27

For a metric space (X, d), there exists a complete metric space $(\widehat{X}, \widehat{d})$ which has a subspace W that is isometric to X (i.e., there exists a map $T\colon X \to W$ such that $d(x, y) = \widehat{d}(T(x), T(y))$ for all $x, y \in X$) and is dense in $\widehat{X}$ (i.e., $\overline{W} = \widehat{X}$). Furthermore, $\widehat{X}$ is unique up to isometry.

The detailed proof is in the text; we'll just give an overview, so that we can see how to extend it to show that a normed linear space can be completed.

Proof sketch. First, we'll construct $(\widehat{X}, \widehat{d})$ by taking $\widehat{X}$ to be the set of all equivalence classes of Cauchy sequences in X, where we say $x \sim y$ (for Cauchy sequences $x = (x_i)$ and $y = (y_i)$ in X) if $\lim_{i\to\infty} d(x_i, y_i) = 0$, and defining our metric as

$$\widehat{d}(x, y) = \lim_{i\to\infty} d(x_i, y_i).$$

(We've defined an equivalence relation on Cauchy sequences, where we say two Cauchy sequences (x_i) and (y_i) are equivalent if $\lim_{i\to\infty} d(x_i, y_i) = 0$. We then define $\widehat{X}$ as the set of equivalence classes — each class has some representative, so we can imagine choosing one for each class and placing it in $\widehat{X}$.)

We have to show that this limit exists and that $\widehat{d}$ is a metric; we'll omit these steps.

We now need to define the map T. We define $T\colon X \to W$ as

$$T(b) = (b, b, b, \ldots).$$

(So $T(b)$ is a Cauchy sequence which is just constant at the point b; this is clearly Cauchy, as all distances between terms are 0.)

The next step is to show that $\overline{W} = \widehat{X}$, and once we've done this, we need to show that $\widehat{X}$ is complete. To do this, let x^n be a Cauchy sequence in $\widehat{X}$ (where we use superscripts to denote indices in the sequence; each x^n is of the form $(x_i^n)_{i\in\mathbb{N}}$). We use the density of W to choose $z^n \in W$ which is 'close' to x^n for each $n \in \mathbb{N}$. Then $T^{-1}(z^n) = z^n \in X$. Then (z^n) is Cauchy in (X, d), and we define $\lim_{n\to\infty} x^n$ as the sequence $(z^n)_{n\in\mathbb{N}}$ (or rather, the equivalence class associated with this Cauchy sequence).

The final step is to show that $\widehat{X}$ is unique up to isometry, but we'll skip that for now. □

We now want to extend this to normed linear spaces.

Theorem 3.28

Let $(X, \|\cdot\|)$ be a normed space. Then there exists a Banach space $\widehat{X}$ and an isometry $T\colon X \to W$ for a subspace W of $\widehat{X}$ with $\overline{W} = \widehat{X}$. Furthermore, $\widehat{X}$ is unique up to isometry.

Proof. We've defined $\widehat{X}$ above (in our proof for metric spaces), but we now need to turn it into a vector space. To do this, we need to define vector addition — given two Cauchy sequences $x = (x_i)$ and $y = (y_i)$, we define their sum componentwise — we define $x + y = (x_i + y_i)$. (Note that if (x_i) and (y_i) are both Cauchy, then $(x_i + y_i)$ is Cauchy as well.)

We also have to define scalar multiplication — for $\alpha \in \mathbb{K}$, we define $\alpha x = (\alpha x_i)$. Again, because (x_i) is Cauchy, so is (αx_i).

To be more precise, we need to be working with representatives of the equivalence classes of Cauchy sequences. To show that this is well-defined, we need to show that if we choose different representatives, we still get the same sums and scalar products. For the first part, suppose that $x \sim x'$ and $y \sim y'$; we then need to show $(x + y) \sim (x' + y')$, and for the second, we need to show that $\alpha x \sim \alpha x'$. But this is pretty immediate.

We also need to show that $0 \in \widehat{X}$, which is pretty clear.

So we have shown that $\widehat{X}$ is a vector space. The last step is to define a norm on this vector space; for $x \in \widehat{X}$, we define

$$\|x\|_{\widehat{X}} = \widehat{d}(0, \widehat{x}) = \lim_{i\to\infty} d(0, x_i) = \lim_{i\to\infty} \|x_i\|$$

(we use the subscript to distinguish from the original norm $\|\cdot\|$ on X). ☐

So we've just shown that we can always complete a vector space. This results in the following theorem.

Theorem 3.29

A subspace Y of a Banach space X is complete if and only if it is closed.

(This is immediate from what we just discussed.)

§3.5 Preview

Next class, we'll prove the following:

Theorem 3.30

Every finite-dimensional vector space is complete.

Theorem 3.31

Every norm on a finite-dimensional vector space is equivalent.

This is a super-powerful theorem in analysis — it means it doesn't matter what norm we take on a finite-dimensional vector space.

Definition 3.32. We say two norms $\|\cdot\|_1$ and $\|\cdot\|_2$ are *equivalent* if there exist constants c_1 and c_2 such that $c_1 \|x\|_1 \le \|x\|_2 \le c_2 \|x\|_1$ for all $x \in X$.

The idea is that the two norms give us comparable notions of size in our space. This is interesting because it means that in $\mathbb{R}^n$, the Euclidean distance is equivalent to the taxicab distance. And the ℓ^p norms on *finite* sequences are all equivalent.

This is *not* true in infinite dimensions. That's the other theorem we'll talk about next time — in infinite dimensions, the unit ball is not compact. (In $\mathbb{R}^n$ the unit ball is compact, but this isn't true in infinite dimensions.)

Remark 3.33. We just talked about two different norms on $\mathcal{C}[a, b]$. These aren't equivalent — in fact, $\mathcal{C}[a, b]$ is infinite-dimensional. We don't yet have the tools to analyze it — we'll do so once we talk about the axiom of choice and Zorn's lemma — but once we have those tools, we'll see that these theorems don't generalize to infinite-dimensional vector spaces.

§4 February 14, 2024

§4.1 Logistics

In the future, we'll have office hours Monday 12–2 and 4–5, Tuesday 11–12 and 4–5, Wednesday 4–5, and Thursday 11–12 and 4–5. We're also moving the exams to March 6 and April 17.

Right now, the sections we're going through match up with Kreyszig pretty well, and the syllabs keeps track of which sections of Kreyszig the lectures correspond to. (We're slightly off the syllabus schedule right now, though.)

Marjie will post the next pset today or tomorrow. It'll be shorter than the last one.

§4.2 Completeness of finite-dimensional subspaces

First, here's a remark on what we saw last time.

Fact 4.1 — In a normed space $(X, \|\cdot\|)$, the map $\|\cdot\|: X \to \mathbb{R}$ is continuous.

Proof. This follows from the triangle inequality — the triangle inequality tells us that $\|x + y\| \le \|x\| + \|y\|$, so $\|x + y\| - \|x\| \le \|y\|$. Now suppose we have a sequence $x_i \to x$ (so to prove continuity, we want to show $\|x_i\| \to \|x\|$). Then we get that $\|x_i\| - \|x\| \le \|x_i - x\|$, and similarly $\|x\| - \|x_i\| \le \|x_i - x\|$; so $|\|x_i\| - \|x\|| \le \|x_i - x\| \to 0$, as desired. □

Last class, we mentioned that this class, we'll prove that in an infinite-dimensional space, the closed unit ball is not compact — that's our big goal for today.

First, here's a lemma that we'll use in the proof.

Lemma 4.2

Let $\{x_1, \dots, x_n\}$ be a linearly independent set of vectors in the normed space X. Then there exists a constant $c > 0$ such that for all $\alpha_1, \dots, \alpha_n \in \mathbb{K}$, we have

$$\|\alpha_1 x_1 + \cdots + \alpha_n x_n\| \ge c(|\alpha_1| + \cdots + |\alpha_n|).$$

(Here $\mathbb{K}$ is the field that X is a vector space over — it's either $\mathbb{R}$ or $\mathbb{C}$.)

Proof. First note that we can assume $\sum |\alpha_i| = 1$, by dividing by $(|\alpha_1| + \cdots + |\alpha_n|)$ (i.e., replacing each α_i by α_i divided by the sum). Then we want to show that $|\alpha_1 x_1 + \cdots + \alpha_n x_n| \ge c$.

Assume for contradiction that this does not hold for any c. Then there exists a sequence $(\alpha^j)_{j \in \mathbb{N}}$ (where each α^j is a vector with n components — i.e., $\alpha^j = (\alpha_1^j, \dots, \alpha_n^j)$) such that

$$\|\alpha_1^j x_1 + \cdots + \alpha_n^j x_n\| \to 0$$

as $j \to \infty$. Now notice that in particular, the sequence $(\alpha_1^j)_{j\in\mathbb{N}}$ is a bounded subset of $\mathbb{R}$; so by the Bolzano–Weirstrass theorem, it has a convergent subsequence — there exists a subsequence $(\alpha_1^{j_k})_{k\in\mathbb{N}}$ converging to some value β_1.

(The reason we divided through by $\sum|\alpha_i|$ is to ensure that each term is bounded.)

Now we can take the sequence $(\alpha_2^{j_k})_{k\in\mathbb{N}}$ (which is a subsequence of our original sequence $(\alpha_2^j)_{j\in\mathbb{N}}$); this is also a bounded sequence in $\mathbb{R}$, so again by Bolzano–Weirstrass there exists a convergent subsequence $(\alpha_2^{j_{k_\ell}})_{\ell\in\mathbb{N}}$ converging to some β_2. And we can repeat this n times — taking a subsequence each time — to eventually find that there exists a subsequence $(\alpha^{j_k})_{k\in\mathbb{N}}$ converging to a vector $\beta = (\beta_1, \dots, \beta_n)$, which satisfies that $|\beta_1| + \cdots + |\beta_n| = 1$.

Since $x_1, \dots, x_n$ are linearly independent, we know that $y = \beta_1 x_1 + \cdots + \beta_n x_n$ cannot be 0. But since $\|\cdot\|$ is continuous and we assumed $\|\alpha_1^j x_1 + \cdots + \alpha_n^j x_n\| \to 0$, we must have $\|\beta_1 x_1 + \cdots + \beta_n x_n\| = 0$; this is a contradiction (since $\|z\| = 0$ if and only if $z = 0$). □

Remark 4.3. Here we assumed that $\mathbb{K} = \mathbb{R}$. But it doesn't really matter — the same proof would work in $\mathbb{C}$ (the Bolzano–Weirstrass theorem is true in $\mathbb{C}$ as well).

Theorem 4.4

Every finite-dimensional subspace Y of a normed space X is complete.

This immediately implies that every finite-dimensional normed space is complete.

Proof. Let $(y^m)_{m\in\mathbb{N}}$ be a Cauchy sequence in Y. The idea is that we're going to take a basis for Y and write out the elements of our Cauchy sequence in this basis. Then we'll use our lemma to show that (y^m) has a limit in Y.

Let $\{e_1, \dots, e_n\}$ be a basis for Y. Then there exists $(\alpha_1^m, \dots, \alpha_n^m)$ such that $y^m = \alpha_1^m e_1 + \cdots + \alpha_n^m e_n$ (so each α^m is the coordinate representation of y^m in our chosen basis).

Since our sequence (y^m) is Cauchy, we know that for every $\varepsilon > 0$, there exists N such that for all $m, n > N$, we have $\|y^m - y^n\| < \varepsilon$. And by definition we have

$$\|y^m - y^n\| = \left\|\sum_{i=1}^n (\alpha_i^m - \alpha_i^n) e_i\right\|.$$

By the previous lemma, we then have

$$\varepsilon \geq \|y^m - y^n\| = \left\|\sum_{i=1}^n (\alpha_i^m - \alpha_i^n) e_i\right\| \geq c \sum_{i=1}^n |\alpha_i^m - \alpha_i^n|$$

for some constant c.

And so now $(\alpha_i^m)_{m\in\mathbb{N}}$ is a Cauchy sequence. So we can perform the same trick to see that the vectors α^m converge to some vector $\beta = (\beta_1, \dots, \beta_n)$ — using the fact that our sequences $(\alpha_i^m)_{m\in\mathbb{N}}$ are Cauchy in $\mathbb{R}$, and $\mathbb{R}$ is complete.

Now let $y = \beta_1 e_1 + \cdots + \beta_n e_n$. Then

$$\|y - y^m\| = \left\|\sum_{i=1}^n (\beta_i - \alpha_i^m) e_i\right\| \leq \sum_{i=1}^n |\beta_i - \alpha_i^m| \, \|e_i\|$$

(using properties of the norm — specifically the triangle inequality and scalar multiplication). And the values of $\|e_i\|$ are bounded, and we know $|\beta_i - a_i^m| \to 0$; so the right-hand side converges to 0, and therefore we have $y^m \to y$. □

Remark 4.5. The nice thing about the norm is that even though our spaces don't have to be complete, because the norm is a map onto $\mathbb{R}$, we can use the completeness of $\mathbb{R}$ to understand what's happening with points in our normed space.

Here's a quick corollary of the above result — recall that a subspace M of a complete metric space X is complete if and only if M is closed in X.

Corollary 4.6

Every finite-dimensional subspace of a normed space is closed.

This is *not* true for infinite-dimensional subspaces — there exist infinite-dimensional subspaces that are *not* closed.

Example 4.7

Let $X = \mathcal{C}[0,1]$ (with the sup norm), and let $M = \text{Span}\{1, t, t^2, \dots\}$. Then M is not closed in X — to see this, we can note that

$$\lim_{n\to\infty} \sum_{j=0}^{n} \frac{t^j}{j!} = e^t$$

is not in M (there's no way to represent e^t as a polynomial). So we have a sequence that has a limit in $\mathcal{C}[0,1]$ (and is therefore Cauchy), but does not have a limit in M.

Remark 4.8. When we talk about the span of an infinite set, we're only referring to the set of vectors we can get from *finite* sums. (This is important here.)

§4.3 Equivalent norms

Definition 4.9. We say two norms $\|\cdot\|_1$ and $\|\cdot\|_2$ on a space X are *equivalent* if there exist constants $c_1, c_2 > 0$ such that for all $x \in X$, we have

$$c_1 \|x\|_1 \le \|x\|_2 \le c_2 \|x\|_1 .$$

Remark 4.10. Equivalent norms generate the same topology.

Theorem 4.11

On a finite-dimensional vector space, any two norms are equivalent.

Remark 4.12. This is really useful — it's a very nice idea that it doesn't matter what norm you use when looking at a finite-dimensional vector space. There can be very clever ways to solve problems where you just change to a norm appropriate to the problem — sometimes the norm you start with will be very messy algebraically, but you can get a much cleaner solution by switching to a different norm.

Proof. Let $\{e_1, \dots, e_n\}$ be a basis for the vector space X, and let $x \in X$. Then there exist $\alpha_1, \dots, \alpha_n \in \mathbb{K}$ such that

$$x = \alpha_1 x_1 + \cdots + \alpha_n x_n.$$

From the earlier lemma, for any norm $\|\cdot\|$ we then have

$$\|x\| \geq c \sum_{i=1}^{n} |\alpha_i| \, .$$

On the other hand, by the triangle inequality we have

$$\|x\| \leq \sum_{i=1}^{n} |\alpha_i| \, \|e_i\| \, .$$

Letting $c' = \max \|e_i\|$, this gives $\|x\| \leq c' \sum |\alpha_i|$. (The constants c and c' depend on the norm.)

This implies the desired result — applying the first inequality to $\|\cdot\|_1$ and the second to $\|\cdot\|_2$, we get

$$\|x\|_1 \geq \frac{c}{c'} \|x\|_2 \, .$$

We can get a bound in the other direction (with different values of c and c') in the same way, by switching the roles of $\|\cdot\|_1$ and $\|\cdot\|_2$. ☐

§4.4 Compactness of unit balls

Definition 4.13. A topological space X is *compact* if every open cover has a finite subcover.

For a metric space, compactness is equivalent to sequential compactness — that every sequence has a convergent subsequence.

Fact 4.14 — If a set M is compact, then M is closed and bounded.

However, the converse is false, as seen by the following example.

Example 4.15

Consider the space ℓ^∞ (with distance defined by $d(x, y) = \sup_i |x_i - y_i|$), and consider the sequence $(1, 0, 0, \ldots)$, $(0, 1, 0, \ldots)$, $(0, 0, 1, \ldots)$, $\ldots$; let M be the set consisting of its elements. Then M is closed and bounded, but this sequence has no convergent subsequence (since the 1 keeps moving to the right forever).

Theorem 4.16

In a finite-dimensional normed space, if $M \subseteq X$ is closed and bounded, then M is compact.

This makes sense — in the above counterexample, we crucially used the fact that our space was infinite-dimensional.

Proof. Let $(x^m)_{m\in\mathbb{N}}$ be a sequence in M. Because X is finite-dimensional, we can take a basis $\{e_1, \ldots, e_n\}$ for X, and we can write each x^m in this basis as $x^m = \sum_{i=1}^{n} \alpha_i^m e_i$.

Since M is closed and bounded, the sequence α^m must be bounded. So we can again use the Bolzano–Weirstrass theorem to choose a convergent subsequence of α^m; suppose that this subsequence converges to β. Now because M is closed, we must have $\sum \beta_i e_i \in M$ as well, and therefore M is sequentially compact. ☐

Now we can get to the theorem we stated at the beginning of this class.

Theorem 4.17

If a normed space X is infinite-dimensional, then the closed unit ball is *not* compact.

Proof. We'll use $\mathbb{B}$ to denote the closed unit ball — so $\mathbb{B} = \{x \in X \mid \|x\| \le 1\}$.

We'll first use the following lemma.

Lemma 4.18 (Riesz)

Let Y and Z be subspaces of X. If $Y \subsetneq Z$ and Y is closed, then for all $\theta \in (0, 1)$, there exists a point $z \in Z$ such that $\|z\| = 1$ and $\|z - y\| \ge \theta$ for all $y \in Y$.

As some intuition for why this is true, imagine we have the plane — here we're taking $X = Z = \mathbb{R}^2$ — and suppose we have some one-dimensional subspace Y (which is a line). Then we can find a point z on the unit circle which is a distance of 1 from the line. Here the intuition is right angles; we can't do this here because we don't have an inner product structure, but we can still make something similar work.

Now we'll use this to prove the theorem. Suppose that X is infinite-dimensional. Let $x_1 \in \mathbb{B}$ satisfy $\|x_1\| = 1$. Then $X_1 = \operatorname{Span}\{x_1\}$ is a closed 1-dimensional subspace satisfying $X_1 \notin X$. So we can apply the lemma to choose $x_2 \in X \setminus X_1$ satisfying $\|x_2\| = 1$ and $\|x_2 - x\| > \frac{1}{2}$ for all $x \in X_1$. We now let $X_2 = \operatorname{Span}\{x_1, x_2\}$.

Now we can repeat — by the lemma, we can find $x_3 \in X \setminus X_2$ satisfying $\|x_3\| = 1$ and $\|x_3 - x\| > \frac{1}{2}$ for all $x \in X_2$. We now let $X_3 = \operatorname{Span}\{x_1, x_2, x_3\}$, and so on.

Repeating this, we produce a sequence of points $x_1, x_2, \ldots \in \mathbb{B}$ satisfying that $\|x_m - x_n\| > \frac{1}{2}$ for all $m \ne n$. This sequence therefore *cannot* contain a convergent subsequence, implying that $\mathbb{B}$ is not compact. □

§5 February 20, 2024

Today we'll talk about L^p spaces (which will be new, but will be an application of analysis and a few things we've talked about so far).

§5.1 L^p spaces

Recall that we've defined the space L^1, and seen that $f \in L^1$ if and only if $\int |f| < \infty$. (Here we're integrating over some set $E \subseteq \mathbb{R}^n$ which is Lebesgue measurable, using the Lebesgue measure.)

Definition 5.1. We define the space $L^p(E)$ as the space of functions $f \colon E \to \mathbb{R}$ such that $\int_E |f|^p < \infty$.

To be more precise, L^p actually consists of *equivalence classes* of functions — where we say $f \sim g$ if $\mu(\{x \in E \mid f(x) \ne g(x)\}) = 0$ (i.e., $f = g$ almost everywhere). The space L^p actually consists of equivalence classes, but we'll abuse notation and just write f instead of $[f]$ when referring to its elements.

Remark 5.2. Note that if $f \sim g$, then $\int_E |f| = \int_E |g|$; by the same logic, $\int_E |f|^p = \int_E |g|^p$.

§5.2 The L^p norm

Today we're going to prove that $L^p(E)$ is a Banach space (i.e., it is a complete normed space).

Definition 5.3. We define the L^p-norm of f as $\|f\|_{L^p(E)} = (\int_E |f|^p)^{1/p}$.

We'll first need to show that this is a norm.

- We have $\|f\|_{L^p(E)} = 0$ if and only if $f = 0$ — here it's important that we're only considering equivalence classes up to equality almost everywhere. (We showed on the problem set that a function has integral 0 if and only if it is 0 almost everywhere.)
- It's easy to see that $\|\alpha f\|_{L^p(E)} = |\alpha| \, \|f\|_{L^p(E)}$, i.e., the norm is homogeneous.
- It is clear that $\|f\|_{L^p(E)} \geq 0$.

The hard part is proving the triangle inequality. We'll need some intermediate inequalities for this.

Lemma 5.4 (Young's inequality)
For $a, b > 0$ and $p, q \in (1, \infty)$ with $\frac{1}{p} + \frac{1}{q} = 1$, we have

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}.$$

(We sometimes call such p and q *(harmonic) conjugate exponents.*)

Proof. This follows from the concavity of log — given two points x_1 and x_2, p and q determine a convex combination of points between x_1 and x_2 — so we can consider the point $x = \frac{x_1}{p} + \frac{x_2}{q}$. Because log is concave, we can imagine drawing the secant line between $(x_1, \log x_1)$ and $(x_2, \log x_2)$, and log lies above this line; this gives

$$\frac{1}{p} \log x_1 + \frac{1}{q} \log x_2 \leq \log\left(\frac{x_1}{p} + \frac{x_2}{q}\right).$$

Taking $x_1 = a^p$ and $x_2 = b^q$ gives

$$\log a + \log b \leq \log\left(\frac{a^p}{p} + \frac{b^q}{q}\right),$$

and raising e to both sides gives $ab \leq \frac{a^p}{p} + \frac{b^q}{q}$. □

Theorem 5.5 (Hölder's inequality)
For $p, q \in (1, \infty)$, let $f \in L^p(E)$ and $g \in L^q(E)$ where $\frac{1}{p} + \frac{1}{q} = 1$. Then we have

$$\int |fg| \leq \|f\|_{L^p(E)} \cdot \|g\|_{L^q(E)}.$$

The proof follows from Young's inequality.

Proof. First, we normalize f and g to the functions $\widehat{f} = f/\|f\|_{L^p(E)}$ and $\widehat{g} = g/\|g\|_{L^q(E)}$. Then it suffices to show that $\int |\widehat{f}\widehat{g}| \leq 1$. And now we're prepared to apply Young's inequality — we have

$$\int |\widehat{f}\widehat{g}| = \int |\widehat{f}||\widehat{g}| \leq \int \left(\frac{|\widehat{f}|^p}{p} + \frac{|\widehat{g}|^q}{q}\right) = \frac{\|\widehat{f}\|_{L^p(E)}^p}{p} + \frac{\|\widehat{g}\|_{L^q(E)}^q}{q}.$$

But we normalized f and g, so the respective norms of $\widehat{f}$ and $\widehat{g}$ are 1; this gives a bound of $\frac{1}{p} + \frac{1}{q} = 1$. □

Remark 5.6. This is a really useful inequality.

This has a useful corollary.

Definition 5.7. For $f \in L^p$ with $f \neq 0$, let $f^* = \|f\|_{L^p}^{1-p} \operatorname{sgn}(f) |f|^{p-1}$.

Corollary 5.8

We have $\int_E f f^* = \|f\|_{L^p(E)}$ and $\|f^*\|_{L^q(E)} = 1$, where $\frac{1}{p} + \frac{1}{q} = 1$.

Proof. For the first statement, we can simply plug in the definition of f^* — we get

$$\int_E f f^* = \|f\|_{L^p}^{1-p} \int_E f \operatorname{sgn}(f) |f|^{p-1} = \|f\|_{L^p}^{1-p} \int_E |f|^p = \|f\|_{L^p}.$$

For the second, we have $q = \frac{p}{p-1}$, so we get

$$\|f\|_{L^p}^{1-p} \left(\int_E (\operatorname{sgn} f) |f|^{(p-1)q} \right)^{1/q},$$

and plugging in $(p-1)q = p$ and $1/q = (p-1)/p$, we get that the second term is exactly $\|f\|_{L^p}^{p-1}$. □

We're now ready to prove the triangle inequality.

Theorem 5.9 (Minkowski's inequality)

For all $f, g \in L^p(E)$ and $p \in (1, \infty)$, we have $\|f + g\|_{L^p(E)} \leq \|f\|_{L^p(E)} + \|g\|_{L^p(E)}$.

Proof. By the above corollary, we have $\|f+g\|_{L^p(E)} = \int_E (f+g)(f+g)^* \leq \int |f| \, |(f+g)^*| + \int |g| \, |(f+g)^*|$ (here we're applying the triangle inequality on $\mathbb{R}$ to create a pointwise inequality inside the integral). Now we can apply Hölder's inequality to get that

$$\int_E |f| \, |(f+g)^*| \leq \|f\|_{L^p(E)} \, \|(f+g)^*\|_{L^q(E)} = \|f\|_{L^p(E)}$$

(since $\|(f+g)^*\|_{L^q(E)} = 1$ — where q is the harmonic conjugate of p), and we can do the same for the second term; this gives an upper bound of $\|f\|_{L^p(E)} + \|g\|_{L^p(E)}$, as desired. □

Remark 5.10. This may look a bit magical, but we actually did the main work in the earlier corollary.

Remark 5.11. What's the intuition behind f^*? It'll make more sense later — it's sort of an element of the *dual space*, which gives another way of defining the norm.

§5.3 The L^∞ norm

Definition 5.12. We define the *essential supremum* as

$$\operatorname{esssup}(f) = \inf\{\alpha \mid \mu(\{x \in E \mid |f(x)| > \alpha\}) = 0\}.$$

Fact 5.13 — The function esssup is a norm, written as $\|\cdot\|_{L^\infty(E)}$.

In words, this is an upper bound on $|f|$ outside a set of measure 0. (A function f can take on value ∞, but only on a measure-0 set; in that case, we say the integral of this part is 0.)

Proof. First, from the definitions it's clear that $\|f\|_{L^\infty(E)} = 0$ if and only if $f = 0$ (since we're looking at equivalence classes of functions). It's also clear that $\|f\|_{L^\infty(E)} \geq 0$ and that $\|\beta f\|_{L^\infty(E)} = |\beta| \, \|f\|_{L^\infty(E)}$. The final thing to check is the triangle inequality — that $\|f+g\|_{L^\infty(E)} \leq \|f\|_{L^\infty(E)} + \|g\|_{L^\infty(E)}$, which is not hard to check by contradiction. □

This allows us to consider the harmonic conjugate pair $(1, \infty)$ — when we have $\frac{1}{p} + \frac{1}{q} = 1$, we can take $p = 1$ and $q = \infty$ (or vice versa). Hölder's inequality for this conjugate pair is immediate — for $p = 1$ and $q = \infty$, we have

$$\int |fg| \leq \int |f| \, \text{essup}(g) = \|f\|_{L^1(E)} \, \|g\|_{L^\infty(E)} .$$

So we can actually replace the conditions $p \in (1, \infty)$ with $p \in [1, \infty]$ in the previous section.

§5.4 Completeness of L^p

Now we'll get to the hard part, proving that L^p is a Banach space.

Theorem 5.14

The space L^p is a Banach space (for any $p \in [1, \infty]$).

We'll prove this for $p \in [1, \infty)$; the proof for $p = \infty$ is not hard.

First, we'll need the following important analysis lemma.

Lemma 5.15

Let (f_n) be a Cauchy sequence in a normed space X, and suppose that there exists a convergent subsequence $(f_{n_k}) \to f$. Then the entire sequence (f_n) also converges to f.

(This is true in any metric space, not just a normed space; it's just an analysis result.)

Remark 5.16. Recall that in general a Cauchy sequence is bounded, but in infinite dimensions a bounded sequence might not have a convergent subsequence. However, it *is* true that if it *does* have a convergent subsequence, then we get convergence of the entire sequence. The proof is left to us.

We'll now prove that L^p is a Banach space, i.e., that it is complete.

Proof of theorem. Fix $p \in [1, \infty)$, and let (f_n) be a Cauchy sequence, and let (f_{n_k}) be a subsequence satisfying that

$$\|f_{n_{k+1}} - f_{n_k}\|_{L^p} \leq 2^{-k}$$

for all $k \in \mathbb{N}$. (Here we're fixing E and writing L^p as shorthand for $L^p(E)$.) Note that we can write

$$f_{n_k} = f_{n_1} + \sum_{i=1}^{k-1} (f_{n_{i+1}} - f_{n_i}).$$

Then let $(g_{n_k})_{k\in\mathbb{N}}$ be the sequence of functions in $\mathcal{L}_0^+$ (the space of measurable nonnegative functions) defined as

$$g_{n_k} = |f_{n_1}| + \sum_{i=1}^{k} |f_{n_{i+1}} - f_{n_i}|.$$

Now g_{n_k} is nonnegative and measurable. So by the monotone convergence theorem, letting $g = \lim_{k\to\infty} g_k$ (note that the sequence $g_k(x)$ is monotone for each fixed x), we have

$$\lim_{k\to\infty} \int g_k^p = \int g^p.$$

Then by Minkowski's inequality (the triangle inequality), we have

$$\int g_k^p = \|g_k\|_{L^p}^p \le \left(\|f_{n_1}\| + \sum_{i=1}^{k-1} \|f_{n_{i+1}} - f_{n_k}\|_{L^p} \right)^p.$$

(Right now, we haven't yet shown that $\|g_k\|_{L^p}$ and $\|g\|_{L^p}$ are finite; we will do so soon, but for now we allow the possibility that it's equal to ∞.)

To see that $\|g_k\|_{L^p}^p$ is bounded, we can note that by the above equation it's at most $(\|f_{n_1}\|_{L^p} + 1)^p$, which is some finite number (not depending on k). So in particular, this shows that $g_k \in L^p$ for all k and that $g \in L^p$. In particular, this means $g < \infty$ almost everywhere.

So then for almost all x, the sequence $f_{n_k}(x)$ converges in $\mathbb{R}$ — this is because of the absolute convergence of its consecutive differences. Then we can define $f(x) = \liminf_{k\to\infty} f_{n_k}$. Because $|f| \le g$ almost everywhere and $g \in L^p$, we have $f \in L^p$. In particular, $|f(x) - f_{n_k}(x)|^p \le 2\,|g(x)|^p$ almost everywhere, so by Lebesgue's dominated convergence theorem we have $\int |f - f_{n_k}|^p \to 0$ (the point is that the dominated convergence theorem lets you put the limit inside the integral).

So now we have convergence of a subsequence in L^p — this is the same as saying that $f_{n_k} \to f$ in L^p norm, where f is as we defined. □

Remark 5.17. What we've done is taken our subsequence, and then found a subsequence that converges in norm very quickly. We then know that the right thing to do is take its pointwise limit. The reason we define g and do all this is to show that this actually converges in the way we want.

Corollary 5.18

If we have a sequence of functions $(f_n) \to f$ in L^p, then there is a subsequence (f_{n_k}) converging to f pointwise almost everywhere.

Remark 5.19. The converse is false — pointwise convergence doesn't imply L^p convergence, which is why we had to do all this work to build an object where we have L^p convergence.

Remark 5.20. We only proved this for $p < \infty$, but L^∞ is also complete; these proofs can also be used to prove that ℓ^p is complete for $p \in [1, \infty)$.

§6 February 21, 2024

Today we'll talk about linear operators (or bounded linear functionals).

§6.1 Linear operators

Definition 6.1. A *linear operator* T is a map whose domain $\mathcal{D}(T)$ is a vector space and whose range $\mathcal{R}(T)$ is contained in a vector space (over the same field), such that T is a vector space homomorphism — i.e.,

$$T(\alpha x + y) = \alpha Tx + Ty$$

for all $x, y \in \mathcal{D}(T)$ and α in the field.

Note that (by the definition of a range) $T\colon \mathcal{D}(T) \to \mathcal{R}(T)$ is *onto*, meaning that for all $y \in \mathcal{R}(T)$ there exists $x \in \mathcal{D}(T)$ such that $Tx = y$. If $\mathcal{D}(T) = X$ and $\mathcal{R}(T) \subseteq Y$ (for vector spaces X and Y), then we write $T\colon X \to Y$.

Definition 6.2. We use $\mathcal{N}(T)$ to denote the *null space* of T — i.e., if $T\colon \mathcal{D}(T) \to Y$, then

$$\mathcal{N}(T) = \{x \in \mathcal{D}(T) \mid Tx = 0\}.$$

Remark 6.3. Some books call this the *kernel*, but we'll use the kernel for a different purpose later in the text (regarding integral operators), so we'll use this terminology instead.

Example 6.4

- The map $\text{Id}\colon X \to X$ given by $Ix = x$ is a linear operator.
- The 0 map $X \to Y$ (for any vector spaces) given by $0x = 0$ is a linear operator.
- Differentiation is a linear operator — for example, if we let X be the vector space of polynomials on $[0, 1]$, then the differentiation map $T\colon X \to X$ sending $Tx(t) = x'(t)$ is a linear map.
- Let $X = \mathcal{C}[a, b]$. Then multiplication by t (the variable of our function) — i.e., the linear map $T\colon X \to X$ sending $x(t) \mapsto tx(t)$ — is a linear map.
- Matrix multiplication is a linear map $\mathbb{R}^m \to \mathbb{R}^n$.

Theorem 6.5

Let T be a linear operator.

(a) The range $\mathcal{R}(T)$ is a vector sapce.

(b) If $\dim \mathcal{D}(T) = n < \infty$, then $\dim \mathcal{R}(T) \leq n$.

(c) The null space $\mathcal{N}(T)$ is also a vector space.

Proof. For (a), from the definition, we know $\mathcal{R}(T)$ is contained in some vector space Y, so we just need to show that it's closed under addition and scalar multiplication — i.e., that for all $x, y \in \mathcal{R}(T)$ and $\alpha \in \mathbb{R}$, we have $\alpha x + y \in \mathcal{R}(T)$. (This works for any field; but we'll just assume our field is $\mathbb{R}$ for convenience.)

If $x, y \in \mathcal{R}(T)$, then there exist $x_1, y_1 \in \mathcal{D}(T)$ such that $Tx_1 = x$ and $Ty_1 = y$. Then

$$T(\alpha x_1 + y_1) = \alpha T(x_1) + T(y_1) = \alpha x + y.$$

And because the domain $\mathcal{D}(T)$ is a vector space, then $\alpha x_1 + y_1$ is an element of this domain; so $\alpha x + y$ is an element of the range.

For (b), we'll take a collection of $n+1$ elements of the range, and show that they can't be linearly independent (given that we can't have $n + 1$ linearly independent elements of the domain) — let $\{y_1, \ldots, y_{n+1}\} \subseteq \mathcal{R}(T)$ be $n + 1$ elements of the range of T. Then for each i, we can define $x_i \in \mathcal{D}(T)$ such that $Tx_i = y_i$.

Since $\dim \mathcal{D}(T) = n < n+1$, we can find an equation $\sum_{i=1}^{n+1} \alpha_i x_i = 0$ where $\alpha_1, \ldots, \alpha_{n+1}$ are not all zero (because the x_i can't be linearly independent). Now we can take T of both sides, and we end up with

$$\sum_{i=1}^{n+1} \alpha_i T(x_i) = \sum_{i=1}^{n+1} \alpha_i y_i = 0.$$

So this means the y_i are *also* not linearly independent.

This indicates that we can't have $n+1$ linearly independent vectors in $\mathcal{R}(T)$, and therefore $\dim \mathcal{R}(T) < n+1$.

Finally, to prove (c), we'll do the same thing as for (a) — let $x, y \in \mathcal{N}(T)$; we need to show that $\alpha x + y \in \mathcal{N}(T)$ as well. But we have

$$T(\alpha x + y) = \alpha T(x) + T(y) = 0,$$

which implies that $\alpha x + y \in \mathcal{N}(T)$, as desired. □

Theorem 6.6

Let X and Y be vector spaces, and let $T\colon \mathcal{D}(T) \to \mathcal{R}(T)$ for some $\mathcal{D}(T) \subseteq X$ and $\mathcal{R}(T) \subseteq Y$.

(a) T has an inverse $T^{-1}\colon \mathcal{R}(T) \to \mathcal{D}(T)$ if and only if $\mathcal{N}(T) = \{0\}$.

(b) If T^{-1} exists, then it is linear.

(c) If $\dim \mathcal{D}(T) = n < \infty$ and T^{-1} exists, then $\dim \mathcal{D}(T^{-1}) = n$ as well.

These are straightforwards, so we're not going to go through the proofs (e.g., for (a), $\mathcal{N}(T) = \{0\}$ is equivalent to injectivity).

§6.2 Bounded linear operators

Definition 6.7. Let X and Y be normed spaces, and let $T\colon \mathcal{D}(T) \to Y$ be a linear operator (with $\mathcal{D}(T) \subseteq X$). We say T is *bounded* if there exists $c > 0$ such that $\|Tx\| \le C \|x\|$ for all $x \in \mathcal{D}(T)$.

Remark 6.8. Note that these two norms are in different spaces — $\|Tx\|$ uses the norm associated with Y, and $\|x\|$ uses the norm associated with X.

Remark 6.9. The idea is that a bounded linear operator maps bounded sets to bounded sets. It's important to note that this does *not* mean that the entire range of T is bounded.

§6.2.1 Norm of a linear operator

Definition 6.10. Let $T\colon X \to Y$ be any operator (i.e., a map on vector spaces or normed spaces, which may or may not be linear). Then the *norm* of T is

$$\|T\| = \sup_{x \neq 0} \frac{\|Tx\|}{\|x\|}.$$

This is a bit questionable since we haven't defined what space of operators we're defining a norm on yet, but we'll fix that soon. Note that again $\|Tx\|$ uses the norm on Y, while $\|x\|$ uses the norm on X.

Note that if T is linear, then using linearity we can rewrite

$$\|T\| = \sup_{\|x\|=1} \|Tx\|.$$

(This is because in the earlier definition, we can simply scale x to have norm 1 without affecting the ratio.)

Before we define the space where this norm is defined, we'll check that it operates the way we want (i.e., satisfies the axioms of a norm).

- The norm is nonnegative — $\|Tx\|$ is always nonnegative, and $\|x\|$ is always positive.
- We have $\|T\| = 0$ if and only if T is identically 0 — if the supremum is 0 then we must have $\|Tx\| = 0$ for all x.
- To see homogeneity, if we scale T by α, then by the linearity of the norm we have
$$\|\alpha T\| = \sup_{x \neq 0} \frac{\|\alpha T x\|}{\|x\|} = \sup_{x \neq 0} \alpha \frac{\|Tx\|}{\|x\|}.$$
- The triangle inequality is immediate — consider two operators T and S (so that $(T+S)(x) = Tx+Sx$). Then we have
$$\|T+S\| = \sup_{x \neq 0} \frac{\|Tx + Sx\|}{\|x\|} \leq \sup_{x \neq 0} \frac{\|Tx\| + \|Sx\|}{\|x\|}.$$

Remark 6.11. What about an operator T defined as
$$T(x) = \begin{cases} 0 & \text{if } x \neq 0 \\ 1 & x = 0? \end{cases}$$

This seems to have norm 0 but to not be the zero map.

The point should be that there's some rigidity about what it means to be an operator that we haven't expressed; or we might want to say that our operator has infinite norm if it doesn't map 0 to 0. For example, we could define
$$\|T\| = \inf\{C \mid \|Tx\| \leq C \|x\| \text{ for all } x\}.$$

(This is probably the right definition.)

§6.2.2 Some examples

We'll now see some examples of operators, and discuss whether they're bounded and linear.

Example 6.12

The operator $\text{Id}\colon X \to X$ (defined as $\text{Id}\, x = x$) is bounded, and $\|\text{Id}\| = 1$.

Example 6.13

The zero operator $0\colon X \to Y$ (defined as $0x = 0$) is bounded, and its norm is 0.

Example 6.14

Consider the map $T\colon X \to X$, where X is the space of polynomials on $[0, 1]$ with the sup norm
$$\|x(t)\| = \sup_{t \in [0,1]} |x(t)|,$$

and T is the differentiation map $Tx(t) = x'(t)$. Is T bounded?

The answer is no — to see this, we need to come up with a sequence of functions whose norm (when we apply T) s unbounded. Take $x_n(t) = t^n$; then $Tx_n(t) = nt^{n-1}$. We have $\|x_n\| = 1$, while $\|Tx_n\| = n$; this means

$$\frac{\|Tx_n\|}{\|x_n\|} = n.$$

So we have a sequence going to ∞, which means the operator norm is unbounded.

Remark 6.15. Since differentiation is a pretty important operator, we can already see that there'll be operators we care about that are unbounded.

Example 6.16

Fix $g \in \mathcal{L}^q(\mathbb{R}^n)$, and define a map $T\colon L^p(\mathbb{R}^n) \to \mathbb{R}$ as $Tf = \int_{\mathbb{R}^n} fg$. Then T is linear and bounded.

Proof. To see T is bounded, we can use Hölder's inequality — we have

$$\left|\int_{\mathbb{R}^n} fg\right| \le \int_{\mathbb{R}^n} |fg| \le \|f\|_{L^p} \|g\|_{L^q}.$$

Now by definition (since T is linear), we have

$$\|T\| = \sup_{\|f\|=1} \|Tf\| \le \|g\|_{L^q}.$$

□

§6.3 Some theorems

Theorem 6.17

If a normed space X is finite-dimensional, then every linear operator on X is bounded.

Proof. We have a finite-dimensional normed space X, so let's start by taking a basis $\{e_1, \dots, e_n\}$ for X. Let $T\colon X \to Y$ be a linear operator. Consider any $x \in X$, and write $x = \sum \alpha_i e_i$ as a linear combination of our basis vectors; then we have

$$Tx = \sum_{i=1}^n \alpha_i Te_i.$$

Now we want to compare $\|Tx\|$ and $\|x\|$. By the triangle inequality, we have

$$\|Tx\| \le \sum |\alpha_i| \|Te_i\|.$$

We've fixed a basis in advance, so we can pull out the terms $\|Te_i\|$ — we have

$$\|Tx\| \le \max_i \|Te_i\| \cdot \sum_{i=1}^n |\alpha_i|.$$

Now recall that all norms are equivalent on a finite-dimensional vector space (we proved this in an earlier class), and in particular, the map $x \mapsto \sum |\alpha_i|$ is a norm (we proved this statement by showing that every norm is equivalent to this one). So in particular, $\sum_{i=1}^n |\alpha_i|$ is equivalent to the norm $\|\cdot\|$ we're using on our vector space X; this means there exists a constant c such that

$$\|Tx\| \le \max_i \|Te_i\| \cdot \sum_i |\alpha_i| \le \max_i \|Te_i\| \cdot c \cdot \|x\|.$$

This means T is bounded, as desired. □

Theorem 6.18

Let $T\colon \mathcal{D}(T) \to Y$ be a linear operator, where $\mathcal{D}(T) \subseteq X$ and X and Y are normed spaces. Then:

(a) T is continuous if and only if T is bounded.

(b) If T is continuous at a single point, then T is continuous.

Proof. This is a cool result; the scheme of the proof is as follows: first, we'll prove that bounded implies continuous. Then we'll prove that continuous at a point implies bounded. This gives the desired statements.

If $T = 0$ the result is immediate; in what follows we assume $T \neq 0$.

First suppose that T is bounded; we'll show that then T is continuous. Consider any $x_0 \in \mathcal{D}(T)$, and let $\varepsilon > 0$. Define $\delta = \varepsilon/\|T\|$ (we're assuming T is bounded, so $\|T\|$ is finite). Now take any $\|x - x_0\| < \delta$; then we have

$$\|Tx - Tx_0\| = \|T(x - x_0)\| \leq \|T\| \, \|x - x_0\| < \varepsilon.$$

So T is continuous — we've shown that given ε, there exists δ such that if $\|x - x_0\| < \delta$ then $\|Tx - Tx_0\| < \varepsilon$.

Now we'll prove that if T is continuous at a point, then it's bounded. Suppose that T is continuous at a point $x_0 \in \mathcal{D}(T)$, so that for every ε, there exists δ such that whenever $\|x - x_0\| \leq \delta$ we have $\|Tx - Tx_0\| \leq \varepsilon$. Fix any such ε and δ.

Let y be any nonzero element of $\mathcal{D}(T)$, and let

$$x = x_0 + \frac{\delta}{\|y\|} \cdot y.$$

In words, we're looking at a point in the y-direction from x_0. Then $\|x - x_0\| = \delta$, which means $\|Tx - Tx_0\| \leq \varepsilon$. But we can write

$$Tx - Tx_0 = T(x - x_0) = \frac{\delta}{\|y\|} \cdot Ty$$

(since $\delta/\|y\|$ is a constant, which we can pull out). So we get that

$$\frac{\|Ty\| \, \delta}{\|y\|} < \varepsilon,$$

and this gives that

$$\frac{\|Ty\|}{\|y\|} < \frac{\varepsilon}{\delta}.$$

This is true for all y, so T is bounded. (Then by the previous statement we proved, this implies that T is continuous.) □

Corollary 6.19

Let T be a bounded linear operator. Then:

(a) If we have a sequence $x_n \to x$, then $Tx_n \to Tx$.

(b) The null space $\mathcal{N}(T)$ is closed.

(The first statement follows directly from continuity, and (b) follows from (a).)

§6.4 Functionals

Definition 6.20. Let X be a vector space over a field $\mathbb{K}$ (which is either $\mathbb{R}$ or $\mathbb{C}$).

- A *functional* is an operator $f\colon X \to \mathbb{K}$.
- A *linear functional* is a linear operator $f\colon X \to \mathbb{K}$.
- If X is a normed space, then a *bounded linear functional* is a bounded linear operator $f\colon X \to \mathbb{K}$.

In general, we've been using T for maps; it's convention to use the letter T when the vector space $\mathcal{R}(T)$ lies is not simply the field $\mathbb{K}$, and f when it is.

Definition 6.21. The *algebraic dual* of a vector space X, denoted X^*, is the set of all linear functionals on X.

Note that X^* is a vector space. Also note that X embeds into its second algebraic dual, written $X \hookrightarrow (X^*)^*$.

Definition 6.22. If the canonical embedding $X \hookrightarrow (X^*)^*$ surjective (or *onto*), we say X is *reflexive*.

(We'll talk about this more later in the class.)

Definition 6.23. For normed spaces X and Y, we define $B(X, Y)$ as the space of bounded linear operators from X to Y.

We defined a norm on $B(X, Y)$ earlier, when we defined the norm on an operator $T\colon X \to Y$. Under this norm

$$\|T\| = \sup_{\|x\|=1} \|Tx\|$$

(we can use this definition because T is linear), the space $B(X, Y)$ is a normed space.

Here are some facts, which we'll discuss next class.

Theorem 6.24

If Y is Banach, then $B(X, Y)$ is also Banach.

Definition 6.25. We define the *dual* of X as $X' = B(X, \mathbb{K})$ (the space of *bounded* linear functionals).

Because $\mathbb{K}$ is either $\mathbb{R}$ or $\mathbb{C}$, the above theorem means that X' is *always* Banach (i.e., complete).

Remark 6.26. It's important that the algebraic dual and the dual space are not in general the same object — the definition of the algebraic dual does not require that our functionals be bounded (it doesn't even require that our space is a normed space, so that wouldn't mean anything). There do exist normed spaces whose algebraic dual is distinct from their dual space; we'll see examples next time.

§7 February 26, 2024

Today we'll talk more about dual spaces — the algebraic dual and the dual space. (We'll go to inner product spaces on Wednesday and next Monday.)

§7.1 The algebraic dual

Definition 7.1. Let X be a normed space. Then its *algebraic dual* X^* is defined as the space of linear functionals on X.

Example 7.2

The map $\|\cdot\|_X$ is *not* linear, though it is a functional on X (i.e., a map $X \to \mathbb{R}$).

Example 7.3

If we fix some $(a, b) \in \mathbb{R}^2$, then the map $T\colon \mathbb{R}^2 \to \mathbb{R}$ defined by $T((x, y)) = ax + by$ (e call this map the *dot product* of (a, b) and (x, y)) is a linear functional, and it is bounded — we have $|T(x, y)| \leq \|(x, y)\|_2 \|(a, b)\|_2$.

Example 7.4

Let $T\colon \mathcal{C}[a, b] \to \mathbb{R}$ be defined as $Tf = \int_a^b f(t)\, dt$. Then T is linear and bounded — we have

$$|Tf| \leq \int_a^b |f(t)|\, dt \leq \int_a^b \|f\|_{\mathcal{C}[a,b]}\, dt = \|f\|_{\mathcal{C}[a,b]} \cdot (b - a)$$

(using the sup norm).

Example 7.5

Let $T\colon \mathcal{C}[a, b] \to \mathbb{R}$ be defined as $Tf = f(x_0)$ for some fixed $x_0 \in [a, b]$. Then T is linear (e.g., we have $T(\alpha f + g) = (\alpha f + g)(x_0) = \alpha f(x_0) + g(x_0)$ for $\alpha \in \mathbb{R}$ and $f, g \in \mathcal{C}[a, b]$) and bounded — we have $|Tf| = |f(x_0)| \leq \sup_t |f(t)| = \|f\|_{\mathcal{C}[a,b]}$.

§7.1.1 The canonical embedding

This last example sets us up for the canonical embedding of X into X^{**} (i.e., an isomorhpism from X onto a subspace of X^{**}).

Remark 7.6. For vector spaces, an isomorphism is a bijective homomorphism.

Fact 7.7 — There is a canonical embedding $X \hookrightarrow X^{**}$.

Proof. Consider some $f \in X^*$, so f is a map $X \to \mathbb{K}$. Now if we fix $x \in X$, then $f(x)$ is an element of $\mathbb{K}$ for all $f \in X^*$. So the idea is that we can think of x as an element of X^{**}, a functional acting on X^* by evaluation at x. This is linear — we have

$$f(\alpha x + y) = \alpha f(x) + f(y)$$

because f is linear for all $f \in X^*$.

So explicitly, we associate to each $x \in X$ the map $g_x \in X^{**}$ defined as $g_x\colon f \mapsto f(x)$ for $f \in X^*$. Then the canonical embedding $C\colon X \to X^{**}$ is defined as $C\colon x \mapsto g_x$. We proved above that it's a homomorphism, and it's injective; so $X \hookrightarrow X^{**}$ (i.e., X embeds in X^{**}). □

Last time, we gave the following definition:

Definition 7.8. If $C(X) = X^{**}$ (i.e., C is onto), we say X is *algebraically reflexive.*

§7.1.2 Finite-dimensional vector spaces

Consider finite-dimensional vector spaces X and Y, and let $T: X \to Y$ a linear operator. Let $\{e_1, \dots, e_n\}$ be a basis for X, and let $\{b_1, \dots, b_r\}$ be a basis for Y. Then for any $x \in X$, we can uniquely write $x = \sum_{i=1}^n \alpha_i e_i$ for some coefficients α_i (this is the unique representation of x with respect to our basis). Then its image $Tx = \sum_{i=1}^n \alpha_i T(e_i)$ is uniquely determined by the values of T on the basis vectors e_i (for $i = 1, \dots, n$).

So when we have finite-dimensional vector spaces, everything is like in linear algebra — we can write each $T(e_i)$ uniquely in the form $\sum_{j=1}^r \beta_j^i b_j$ (in terms of our basis for Y). Then we can substitute these into our formula for Tx, and get that

$$Tx = \sum_{i=1}^n \alpha_i \sum_{j=1}^r \beta_j^i b_j = \sum_{j=1}^r b_j \sum_{i=1}^n \alpha_i \beta_j^i.$$

Here each b_j is a basis vector for Y, and $\sum_{i=1}^n \alpha_i \beta_j^i$ is its coefficient. So we end up with a unique matrix representing this operator — we see that given our linear operator T, there's a unique matrix representing T with respect to the given bases (whose entries are given by the β_j^i, and such that applying T corresponds to matrix multiplication).

Definition 7.9. Let X be a finite-dimensional vector space, and let $\{e_1, \dots, e_n\}$ be a basis for X. Then we define the *dual basis* in the following way — for $k = 1, \dots, n$, we define

$$f_k(e_j) = \delta_{jk} = \begin{cases} 1 & \text{if } j = k \\ 0 & \text{otherwise.} \end{cases}$$

(Once we've defined f_k on each basis vector, there's a unique way to extend it to X.)

Theorem 7.10

Let X be a finite-dimensional vector space. Then given a basis $\{e_1, \dots, e_n\}$, its dual basisis $\{f_1, \dots, f_n\}$ is indeed a basis for X^*, and $\dim X^* = \dim X$.

To prove $\{f_1, \dots, f_n\}$ is a basis, we need to prove two things — that every element $f \in X^*$ can be written in the form $\sum_{i=1}^n \alpha_i f_i$, and that $\{f_i\}$ are linearly independent. Both of these are straightforward.

Lemma 7.11

Let X be a finite-dimensional vector space, and suppose there exists $x_0 \in X$ such that $f(x_0) = 0$ for all $f \in X^*$. Then $x_0 = 0$.

Proof. Let $\{e_1, \dots, e_n\}$ be a basis for X, so that $x_0 = \sum \alpha_i e_i$ (for some $\alpha_i \in \mathbb{K}$). Then $f(x_0) = \sum \alpha_i f(e_i)$ is supposed to be 0 for all $f \in X^*$. But we can simply take $f = f_i$ for each i; then $f_i(x_0) = \alpha_i$, so this is only possible if $\alpha_i = 0$ for each i. □

Note that in particular since $\dim X = \dim X^*$, we have $\dim X = \dim X^* = \dim X^{**}$. And since C is an injective map $X \to X^{**}$ (the above lemma implies injectivity), we must have $\dim \mathcal{R}(C) = \dim(X)$; this means $\mathcal{R}(C) = X^{**}$. So this means we actually have $X = X^{**}$, giving the following statement.

Theorem 7.12
Any finite-dimensional vector space is algebraically reflexive.

§7.2 Completeness of bounded linear operators

Now we'll show the following theorem, which we mentioned last time.

Theorem 7.13
If Y is a Banach space and X is *any* normed space, then $B(X, Y)$ is complete.

Recall that $B(X, Y)$ is the space of bounded linear operators $X \to Y$ (as defined last class).

Proof. Let (T_n) be a Cauchy sequence in $B(X, Y)$; we just need to show that there exists an operator T that this sequence converges to.

First let's unwind what the definitions mean — we know that for every $\varepsilon > 0$, there exists N such that $\|T_n - T_m\| < \varepsilon$ for all $n, m > N$. And recall that $\|T_n - T_m\|$ is the operator norm, which (since we're working with linear operators) we can write as

$$\|T_n - T_m\|_{B(X,Y)} = \sup_{\|x\|=1} \|T_n(x) - T_m(x)\|_Y .$$

Now there's a natural way to construct the limit, using the following claim.

Claim 7.14 — For each fixed $x \in X$, the sequence $(T_n x)_{n\in\mathbb{N}}$ is Cauchy in Y.

Proof. To prove this, we need to show that $\|T_n x - T_m x\|_Y = \|(T_n - T_m)(x)\|_y \leq \|T_n - T_m\|_{B(X,Y)} \|x\|_X$. We can make $\|T_n - T_m\|_{B(X,Y)}$ smaller than ε for any ε, so then for each x this sequence is Cauchy. □

Now we're just going to define our limiting operator pointwise, the way we'd expect to — let $T \colon X \to Y$ be defined as

$$T(x) = \lim_{n\to\infty} T_n(x)$$

for each x. Note that because Y is complete and $T_n(x)$ is Cauchy, then $T_n(x)$ converges to some element of Y — which we're using to define our operator.

Now we need to show that T is really an element of $B(X, Y)$.

Claim 7.15 — The operator T is linear.

Proof. Let $x_1, x_2 \in X$ and $\alpha \in \mathbb{K}$. Then we have

$$T(\alpha x_1 + x_2) = \lim_{n\to\infty} T_n(\alpha x_1 + x_2) = \lim_{n\to\infty} (\alpha T_n(x_1) + T_n(x_2)) = \alpha \lim_{n\to\infty} T_n(x_1) + \lim_{n\to\infty} T_n(x_2)$$

(using the fact that T_n is linear for each n), which is $\alpha T(x_1) + T(x_2)$. □

Claim 7.16 — The operator T is bounded, and $\lim_{n\to\infty} \|T_n - T\| = 0$.

This in particular means $T_n \to T$ in $B(X, Y)$, so this finishes the proof.

To prove this, for every (fixed) $n \in \mathbb{N}$ we have that

$$\|T_n x - Tx\| = \left\| T_n x - \lim_{m\to\infty} T_m x \right\| = \lim_{m\to\infty} \|T_n x - T_m x\|$$

(using the definition of T and the continuity of the norm), which we can bound by

$$\|T_n x - Tx\| \le \lim_{m\to\infty} \|T_n - T_m\|_{B(X,Y)} \|x\| .$$

In particular, this means $T_n - T$ is bounded; and since T_n is bounded, this means T is bounded.

And given any $\varepsilon > 0$, we can make $\lim_{m\to\infty} \|T_n - T_m\| < \varepsilon$ by taking n large enough; and we showed that $\|T_n - T\| \le \lim_{m\to\infty} \|T_n - T_m\|$, so this shows that $\|T_n - T\| \to 0$. □

Here's an immediate corollary, which we mentioned last time:

Definition 7.17. For a normed space X, we define its *dual space* X' as the space of bounded linear functionals on X — equivalently, $X' = B(X, \mathbb{K})$.

(This is different from the algebraic dual if X isn't finite-dimensional, which we'll see soon.)

Corollary 7.18

For any normed space X, the space X' is complete.

This is because $\mathbb{K}$ (which is either $\mathbb{R}$ or $\mathbb{C}$ is complete).

§7.3 Some examples of dual spaces

Example 7.19

The dual space of $\mathbb{R}^n$ is $\mathbb{R}^n$.

Example 7.20

The dual space of ℓ^1 is ℓ^∞.

Proof. The proof has three components — first we're going to show that every element of the dual space is in ℓ^∞ (i.e., that if $f \in (\ell^1)'$ then $f \in \ell^\infty$), then that every element of ℓ^∞ is an element of the dual space, and finally we'll show that the norms on ℓ^∞ and the dual space are the same (i.e., that $\|f\|_{(\ell^1)'} = \|f\|_{\ell^\infty}$).

First we'll show that if $f \in (\ell^1)'$, then $f \in \ell^\infty$. Consider the Schauder basisf or ℓ^1 — the basis $(e_i)_{i\in\mathbb{N}}$ where e_i has a 1 in the ith coordinate and 0's everywhere else. Now let $f \in (\ell^1)'$. Then for each $x \in \ell^1$, we can write $x = \sum_{i\in\mathbb{N}} \alpha_i e_i$, and using continuity we can show that $f(x) = \sum_{i\in\mathbb{N}} \alpha_i f(e_i)$, which means f is entirely determined by its values of $f(e_i)$. So we'll define $\gamma_i = f(e_i)$. Note that $|\gamma_i| \le \|f\|_{(\ell^1)'} \sup_i \|e_i\| = \|f\|_{(\ell^1)'}$ for each i, and therefore $\sup_i |\gamma_i| \le \|f\|$ is finite; this means $\gamma = (\gamma_i)$ is indeed an element of ℓ^∞.

Now for the second step, consider some $b = (\beta_k)_{k\in\mathbb{N}} \in \ell^\infty$; we need to define a corresponding linear functional. To do so, we define $g\colon \ell^1 \to \mathbb{R}$ by $g(x) = \sum \beta_k \alpha_k$ (where $x = \sum \alpha_k e_k$). We need to show that g is bounded; to do so, we have

$$|g(x)| \le \sum |\beta_k \alpha_k| \le \|b\|_{\ell^\infty} \sum |\alpha_k| = \|b\|_{\ell^\infty} \|x\|_{\ell^1} .$$

So now we've shown that $b \in \ell^\infty$ corresponds to an element of $(\ell^1)'$, as desired.

The final step is to show that the norms are the same. To do so, we have

$$|f(x)| = \left|\sum \alpha_i f(e_i)\right| \leq \|\alpha\|_{\ell^1} \sup_i |f(e_i)|,$$

and this supremum is the ℓ^∞ norm of the sequence $(f(e_1), f(e_2), \dots)$ (which is the element of ℓ^∞ corresponding to f).

On the other hand, we have $\|f\|_{(\ell^1)'} = \sup_{\|x\|} |f(x)|$, and we just showed that this is bounded by $\|\gamma\|_{\ell^\infty}$ (where γ is the corresponding element of ℓ^∞). So we have shown that $\|f\|_{(\ell_1')} \leq \|\gamma\|_{\ell^\infty}$.

So here we've shown a bound in one direction, and at the beginning of the solution we showed the reverse inequality; this gives us the equality. □

§8 February 28, 2024

Last time, we proved that $(\ell^1)' = \ell^\infty$ — specifically, that there is an isomorphism of normed spaces between the two.

§8.1 Dual of L^p spaces

Today, we'll start by talking about L^p spaces.

Theorem 8.1

Let $\Omega \subseteq \mathbb{R}^n$. Then the dual of $L^p(\Omega)$ is $L^q(\Omega)$, where $1/p + 1/q = 1$ and $p \in [1, \infty)$.

What does this actually mean? This means that the normed space $(L^p(\Omega))'$ is *isomorphic* to $L^q(\Omega)$ — i.e., there exists a map $T\colon L^q(\Omega) \to (L^p(\Omega))'$ such that T is a bijective, norm-preserving homomorphism. (*Bijective* means one-to-one and onto; *norm-preserving* means that for all $f \in L^q(\Omega)$, we have $\|f\|_{L^q(\Omega)} = \|Tf\|_{B(L^p(\Omega),\mathbb{R})}$. (We'll just work with $\mathbb{R}$ for now.) A *homomorphism* is a map preserving the linear structure of the space — meaning $T(\alpha f + g) = \alpha T(f) + T(g)$.)

Remark 8.2. These three conditions are what it means to be an isomorphism of normed spaces.

There's some component of the proof requiring measure theory we didn't cover; this won't be tested on, but it's satisfying to learn.

Proof. Let μ be the Lebesgue measure. We'll assume that $\mu(\Omega) < \infty$ (the adaptation to the case where Ω is σ-finite — a countable union of finite-measure sets — is not too hard, but we won't go into it). We'll also only consider the case $p > 1$.

Our first step is to define the map T — let $T\colon L^q(\Omega) \to (L^p(\Omega))'$ be the map $g \mapsto T_g$, where

$$T_g(f) = \int_\Omega fg$$

for all $f \in L^p(\Omega)$.

First let's check that this map is indeed a homomorphism — we can see that $T_{\alpha g_1 + g_2} f = \int_\Omega f(\alpha g_1 + g_2) = \alpha \int_\Omega f g_1 + \int_\Omega f g_2 = \alpha T_{g_1} f + T_{g_2} f$ by the linearity of the integral.

The next thing we want to show is that T_g really is in $(L^p(\Omega))'$ — i.e., that it is a bounded linear functional on $L^p(\Omega)$.

Claim 8.3 — The map T is a bounded linear functional on $L^p(\Omega)$.

Proof. We need to show that $|T_g f|$ is bounded by a constant times $\|f\|_{L^p(\Omega)}$. To do this, by Hölder's inequality we have

$$|T_g f| = \left|\int_\Omega fg\right| \le \int_\Omega |fg| \le \|f\|_{L^p(\Omega)} \|g\|_{L^q(\Omega)}. \qquad \square$$

The fact that T is injective is not hard — if $T_{g_1} f = T_{g_2} f$ for all $f \in L^p(\Omega)$, then $\int fg_1 = \int fg_2$ for all $f \in L^p(\Omega)$, which means $\int f(g_1 - g_2) = 0$ for all $f \in L^p(\Omega)$. This is only possible if $g_1 - g_2 = 0$ almost everywhere, which is exactly what it means for g_1 and g_2 to be equal in $L^q(\Omega)$ (which is really a collection of *equivalence classes*). So this implies $g_1 = g_2$.

The next thing we'll need to do is prove that T is onto.

This requires measure theory that we may not be familiar with and won't be tested on.

In order to prove this, we'll need a few lemmas.

Lemma 8.4

Simple functions are dense in $L^p(\Omega)$ for any p.

For the second lemma, we've earlier defined what it means for something to be a measure — a measure is a function on (measurable) sets which is positive and countably additive. We're now going to work with *signed* measures, which are now allowed to take negative values (but must still be countably additive).

Definition 8.5. We say a measure ν is *absolutely continuous* with respect to the Lebesgue measure if for all measurable sets A with $\mu(A) = 0$, we have $\nu(A) = 0$.

Theorem 8.6 (Lebesgue decomposition theorem)

If a signed measure is absolutely continuous with respect to the Lebesgue measure, then it can be written as an integral — there exists $f \in L^1(\Omega)$ such that $\nu(A) = \int_A f\, d\mu$ for all measurable sets A.

Now to prove that T is onto, let $S \in (L^p(\Omega))'$; our goal is to show that there is some $g \in L^q(\Omega)$ with $T_g = S$. To see this, define the function $\nu(E) = S(1_E)$ (where $E \subseteq \Omega$ is a measurable set, and 1_E denotes the characteristic function of E — so ν is a function on measurable sets).

It's possible to show ν is a signed measure (i.e., it's countably additive). And this implies that $S(1_E) = \int_E f\, d\mu = \int_\Omega 1_E \cdot f$ for some function f.

Now let φ be a simple function. Then $S(\varphi) = \int_\Omega \varphi \cdot f\, d\mu$ (this follows by writing φ as a sum of finitely many indicator functions 1_E). This means

$$|S(\varphi)| = \left|\int_\Omega \varphi f\, d\mu\right| \le \|S\|_{B(L^p(\Omega),\mathbb{R})} \|\varphi\|_{L^p(\Omega)}$$

(because S is a bounded linear functional).

This is true for *any* simple function φ. Now we can take a sequence of functions φ_i converging to $\operatorname{sgn} f \cdot |f|^{q-1}$. Why are we doing this? There's two things to notice. The first is that if we plug this into the integral, then we end up with $|f|^q$; and we want to make the claim that that's bounded. The term $\|S\|_{B(L^p(\Omega),\mathbb{R})}$ is of course bounded, so we need to show $\|\varphi\|_{L^p(\Omega)}$ is bounded. To see this, it suffices to show that

$$\int_\Omega |\varphi|^p = \int_\Omega |f|^{(q-1)p} = \int_\Omega |f|^q.$$

So we've shown that $f \in L^q(\Omega)$; and then we can use the density of simple functions to prove that for *all* $\varphi \in L^p(\Omega)$, we have $S(\varphi) = \int_\Omega \varphi \cdot f$. So we've proven that φ is onto.

Finally, we just need to show that T is norm-preserving. This relates back to a corollary of Hölder's inequality that we saw earlier:

Corollary 8.7

If $f \in L^p(\Omega)$, then the function $f^* = \|f\|_{L^p(\Omega)}^{1-p} \operatorname{sgn} f\, |f|^{p-1}$ is an element of $L^q(\Omega)$, and we have $\int f f^* = \|f\|_{L^p(\Omega)}$ and $\|f^*\|_{L^q(\Omega)} = 1$.

Now we're going to use this to show that T is norm-preserving — we have

$$\|T_g\|_{B(L^p(\Omega),\mathbb{R})} = \sup_{\|f\|_{L^p(\Omega)}} |T_g f| \,.$$

We showed earlier that $|T_g f| \le \|f\|_{L^p(\Omega)} \|g\|_{L^q(\Omega)}$ (when proving T_g was bounded), so we immediately get that

$$\|T_g\|_{B(\dots)} \le \|g\|_{L^q(\Omega)} \,.$$

To prove the reverse inequality, we apply the corollary to $g \in L^q(\Omega)$ to produce $g^* \in L^p(\Omega)$ (the roles of p and q from the corollary are switched). Now take f to equal g^*; then we get $\|g^*\|_{L^p(\Omega)} = 1$, and

$$|T_g g^*| = \int_\Omega g g^* = \|g\|_{L^q(\Omega)} \,.$$

So we get equality — we've chosen an appropriate function f (namely, g^*) to show that the operator norm of T_g is the L^q-norm of g. □

Remark 8.8. Why does this imply T is onto? We've shown that $|S(\varphi)| \le \|S\|_{B(L^p(\Omega),\mathbb{R})} \|\varphi\|_{L^p(\Omega)}$. Now we choose a sequence of functions converging to $\operatorname{sgn} f \cdot |f|^{q-1}$. Then $S(\varphi) = \int_\Omega |f|^q$. And the right-hand side has an operator norm, and so the inequality we get is that

$$\int_\Omega |f|^q \le \|S\| \cdot \left(\int |f|^q\right)^{1/p} .$$

And if we divide through, then we end up with

$$\left(\int_\Omega |f|^q\right)^{1/q} \le \|S\| \,.$$

(This is important, and we left it out.)

So once we take $\varphi_i \to \operatorname{sgn}(f)\, |f|^{q-1}$, the left-hand side becomes $\int_\Omega |f|^q$; and this is at most $\|S\| \, \|\varphi\|_{L^p} = \|S\| \left(\int |f|^q\right)^{1/p}$. But now we divide everything by the right-hand side.

Remark 8.9. Justifying that you can take simple functions converging to $\operatorname{sgn} f \cdot |f|^{q-1}$ is nontrivial. What we did was we chose a *pointwise* limit to start with, and then we showed the convergence was in L^q. But this was handwavy.

§8.2 Inner product spaces

Definition 8.10. Let X be a vector space over a field $\mathbb{K}$ (either $\mathbb{R}$ or $\mathbb{C}$). An *inner product* is a mapping $X \times X \to \mathbb{K}$, written $\langle -, - \rangle$, satisfying the following properties (for all $x, y, z \in X$ and $\alpha \in \mathbb{K}$):

(i) $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$.

(ii) $\langle \alpha x, y \rangle = \alpha \langle x, y \rangle$.

(iii) $\langle x, y \rangle = \overline{\langle y, x \rangle}$.

(iv) $\langle x, x \rangle \geq 0$, and $\langle x, x \rangle = 0$ if and only if $x = 0$.

Here $\overline{z}$ denotes complex conjugation — note that (ii) and (iii) imply $\langle x, \alpha y \rangle = \overline{\alpha} \langle x, y \rangle$. So $\langle -, - \rangle$ is linear in the first element and conjugate-linear in the second (this property is called *sesquilinearity*). If the field is $\mathbb{R}$, then complex conjugation doesn't do anything, and the inner product is completely linear.

An inner product defines a norm:

Fact 8.11 — Given an inner product $\langle -, - \rangle$, then $\|x\| = (\langle x, x \rangle)^{1/2}$ is a norm.

(We'll prove this in a moment.)

Definition 8.12. A *Hilbert space* is a complete inner product space (with respect to this norm, or the induced metric).

Now we'll prove that this is really a norm. First, (iv) implies that $\|x\| \geq 0$, with equality if and only if $x = 0$. It's also clear from (ii) and (iii) that the norm is homogeneous — we have

$$\|\alpha x\|^2 = \langle \alpha x, \alpha x \rangle = \alpha \overline{\alpha} \langle x, x \rangle = |\alpha|^2 \|x\|^2,$$

which gives that $\|\alpha x\| = |\alpha| \|x\|$.

So now it remains to prove the triangle inequality, which is the challenging step.

Lemma 8.13 (Schwarz inequality)

We have $|\langle x, y \rangle| \leq \|x\| \|y\|$.

Proof. If $y = 0$ then this is clear immediately (both sides are 0), so we assume $y \neq 0$. Then we have

$$0 \leq \|x - \alpha y\|^2 = \langle x - \alpha y, x - \alpha y \rangle = \|x\|^2 - \overline{\alpha} \langle x, y \rangle - \alpha \langle y, x \rangle + |\alpha|^2 \|y\|^2.$$

We'll choose α such that $\langle y, x \rangle - \overline{\alpha} \langle y, y \rangle = 0$, by letting

$$\overline{\alpha} = \frac{\langle y, x \rangle}{\langle y, y \rangle}.$$

Then the second two terms disappear, and plugging in $\overline{\alpha}$ into the rest of our expression gives that

$$0 \leq \|x\|^2 - \frac{\langle y, x \rangle}{\|y\|^2} \cdot \langle x, y \rangle.$$

Finally, $\langle y, x \rangle \langle x, y \rangle = \overline{\langle x, y \rangle} \langle x, y \rangle = |\langle x, y \rangle|^2$, so we immediately get that

$$|\langle x, y \rangle|^2 \leq \|x\|^2 \|y\|^2. \qquad \square$$

Remark 8.14. When does equality hold? Equality holds if $y = 0$, or if $x - \alpha y = 0$, i.e., $x = \alpha y$ — meaning that x and y are linearly dependent.

Now we'll prove the triangle inequality — we have

$$\|x+y\|^2 = \langle x+y, x+y\rangle = \|x\|^2 + \langle x, y\rangle + \langle y, x\rangle + \|y\|^2.$$

And the Schwarz inequality tells us that this is bounded by

$$\|x\|^2 + \|y\|^2 + 2\|x\|\|y\| = (\|x\| + \|y\|)^2.$$

So that gives us the triangle inequality.

Remark 8.15. As before, we can comment on when equality holds. Equality holds if and only if $\langle x, y\rangle + \langle y, x\rangle = 2\|x\|\|y\|$. But the left-hand side is exactly $2\operatorname{Re}\langle x, y\rangle$. So we must have $\operatorname{Re}\langle x, y\rangle = \|x\|\|y\|$. But we know by the Schwarz inequality that $|\langle x, y\rangle| \leq \|x\|\|y\|$. Of course we can't have $\operatorname{Re} z > |z|$, so equality must hold everywhere. From the fact that equality holds in the Schwarz inequality, either $y = 0$ or $x = \alpha y$ for some α. If $x = \alpha y$, then $\langle x, y\rangle = \alpha\|y\|^2$; so we need α to be a nonnegative real number.

So equality holds in the triangle inequality if and only if $y = 0$ or $x = \alpha y$ for some *real* $\alpha \geq 0$.

Now we'll move to the parallelogram equality (which can be proven by direct computation).

Fact 8.16 (Parallelogram equality) — We have $\|x+y\|^2 + \|x-y\|^2 = 2(\|x\|^2 + \|y\|^2)$.

This is a geometric fact about parallelograms (we can imagine a parallelogram whose sides are x and y, and whose diagonals are $x+y$ and $x-y$). But it's also true for *all* inner products.

Remark 8.17. There exist norms which don't satisfy the parallogram equality; such norms can't be obtained from an inner product.

Definition 8.18. We say x is *orthogonal* to y (denoted $x \perp y$) if $\langle x, y\rangle = 0$.

§8.2.1 Some examples

Example 8.19

$\mathbb{R}^n$ with the usual dot product is an inner product space; teh norm this induces is the Euclidean metric. The same is true for $\mathbb{C}^n$.

Example 8.20

$L^2([a, b])$ is an inner product space with

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

and the norm this induces is the familiar L^2 norm. (The same holds for general Ω.)

Remark 8.21. The book only considers $L^2([a, b])$, which it defines by considering continuous real valued functions on $[a, b]$ with finite $\int |f|^2$ and then completing the metric space. That's equivalent to the definition we saw in the course.

Example 8.22

ℓ^2 is a metric space, with $\langle x, y\rangle = \sum_j x_j \overline{y_j}$; this is a *complete* inner product space (and the historic quintessential example of a Hilbert space).

Example 8.23

ℓ^p (for $p \neq 2$) is *not* an inner product space (with the ℓ^p norm) — we can see this by finding two elements which don't satisfy the parallelogram inequality. Consider $x = (1, 1, 0, 0, \ldots)$ and $y = (1, -1, 0, 0, \ldots)$. Then $x + y = (2, 0, 0, \ldots)$ and $x - y = (0, 2, 0, \ldots)$. We have $\|x + y\|^2 + \|x - y\|^2 = 8$ and $\|x\| = (1+1)^{1/p} = 2^{1/p}$, so $2(\|x\|^2 + \|y\|^2) = 2(2^{2/p} + 2^{2/p}) = 2^{2+2/p} \neq 8$ for $p \neq 2$. So there's no way to define an inner product on ℓ^p that gives us the ℓ^p norm.

Example 8.24

$\mathcal{C}[a, b]$ is not an inner product space — for example, we can take $x(t) = 1$ and $y(t) = \frac{t-a}{b-a}$, and you can use these to show that the parallelogram equality (for the sup norm) is not satisfied.

§9 March 4, 2024

The midterm is on Wednesday; it covers content up to Kreyszig 3.4, including what we'll cover today and the work we did on Lebesgue measure, the Lebesgue integral, and L^p spaces (the emphasis will be on L^p spaces, but you can't really talk about L^p spaces without generally understanding the first two lectures). There will be 6 problems, and some will have subparts.

Today we'll talk more about the inner product, and state several more definitions.

§9.1 Completion of an inner product space

Proposition 9.1

Let $x_n \to x$ and $y_n \to y$. Then $\langle x_n, y_n\rangle \to \langle x, y\rangle$.

Proof. What $x_n \to x$ means is that $\|x_n - x\| \to 0$. In order to prove the desired statement, we have

$$\langle x_n, y_n\rangle - \langle x, y\rangle = (\langle x_n, y_n\rangle - \langle x_n, y\rangle) + (\langle x_n, y\rangle - \langle x, y\rangle),$$

so by the triangle inequality

$$|\langle x_n, y_n\rangle - \langle x, y\rangle| \leq |\langle x_n, y_n - y\rangle| + |\langle x_n - x, y\rangle|.$$

And now we can apply the Schwarz inequality — this is bounded by

$$\|x_n\| \|y_n - y\| + \|x_n - x\| \|y\|.$$

The terms $\|x_n\|$ and $\|y_n\|$ are bounded, while $\|x_n - x\|$ and $\|y_n - y\|$ go to 0; so this entire expression goes to 0. □

Definition 9.2. An *isomorphism of inner product spaces* X and $\widetilde{X}$ is a bijective linear map T preserving the inner product — i.e., for all $x, y \in X$, we have $\langle x, y\rangle = \langle Tx, Ty\rangle$.

Remark 9.3. Note that the two sides of this equation use different inner products.

Note that in particular, such an isomorphism is both an isomorphism of vector spaces and an isometry (as taking $x = y$ we get $\|x\|^2 = \|Tx\|^2$).

We can now complete an inner product space, in the following way.

Definition 9.4. Let X be an inner product space. Then there exists an isomorphism $A\colon X \to W$ where W is a dense subspace of a Hilbert space H, and H is unique up to isomorphism.

The proof simply builds on the proof for Banach spaces.

Proof. Recall that we proved earlier that we can complete any normed space X via an isometry $A\colon X \to W$ where W is a dense subspace of a *complete* normed space $\widetilde{X}$, and $\widetilde{X}$ is unique up to isometry. We'll extend this idea to capture the inner product in a natural way — given two sequences (x_n) and (y_n), we define

$$\langle x, y\rangle = \lim_{n\to\infty} \langle x_n, y_n\rangle.$$

(The fact that this works follows from the continuity of the inner product.) □

§9.2 Subspaces

Definition 9.5. A subspace Y of an inner product space X is a vector subspace of X with the same inner product, restricted to $Y \times Y$.

Note that a subspace need not be complete. But immediately from the earlier Banach space theory, we have the following theorem.

Theorem 9.6

Let Y be a subspace of a Hilbert space X. Then:

(1) Y is complete if and only if it is closed.

(2) If Y is finite-dimensional, then Y is complete.

(3) If H is separable, then Y is separable.

(We haven't talked about separability much, but we will later.)

Definition 9.7. A set is *separable* if it has a countable dense subset.

These all follow from the theory we did for Banach spaces, because the norm comes straight from the inner product.

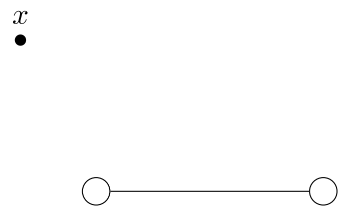
§9.3 Minimizing distance

Definition 9.8. Let X be a normed space, and M a subset of X. Then for $x \in X$, we define the *distance from x to M* as

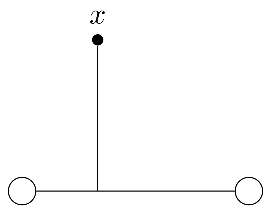
$$\delta = \delta(x) = \inf_{y\in M} \|x - y\|.$$

Question 9.9. Does there exist $y \in M$ achieving this infimum? And if so, is y unique?

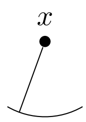
If we just take M to be an open interval in $\mathbb{R}^2$, then the answer might be no:



Meanwhile, if we move x over to the right, then the infimum is achieved (by the foot of the perpendicular), and it is unique.



But there are also cases where the infimum is achieved, but isn't unique — for example, if M is an arc of a circle centered at x, then all its points are at the same distance from x.



So we want to come to an understanding of *when* the infimum is achieved, and when it's *uniquely* achieved.

The natural result that comes from thinking about these examples is the following theorem.

Theorem 9.10

Let X be an inner product space, and let $M \subseteq X$ be convex, nonempty, and complete. Then for every $x \in X$, there exists $y \in M$ such that the infimum $\delta = \inf_{\widetilde{y} \in M} \|x - \widetilde{y}\|$ is achieved by the point y.

Remark 9.11. If X is finite-dimensional, then any closed set will be complete. But we're going to be thinking about inner product spaces that are infinite-dimensional, so being complete isn't the same thing as being closed.

Proof. For now, let's assume that $x = 0$ and $0 \notin M$ (to get an idea of how the proof works).

Claim 9.12 — There exists $y \in M$ such that $\delta = \|x - y\|$.

Proof. Since δ is defined as an infimum, we can define a sequence $(y_n) \subseteq M$ such that $\delta_n = \|x - y_n\| \to \delta$. Recall that we're taking $x = 0$, so this actually means $\|y_n\| \to \delta$. Our goal is to show that y_n converges to a point y; we'll do this by showing that the sequence is Cauchy. Then we can use the parallelogram equality on the sequence (y_n) — we have

$$\|y_n - y_m\|^2 + \|y_n + y_m\|^2 = 2(\|y_n\|^2 + \|y_m\|^2),$$

and rearranging this gives that

$$\|y_n - y_m\|^2 = -\|y_n + y_m\|^2 + 2(\|y_n\|^2 + \|y_m\|)^2.$$

We want to show the left-hand side goes to 0 as $m, n \to \infty$. To see this, first note that $y_n + y_m = 2 \cdot \frac{1}{2}(y_n + y_m)$, and by convexity we have $\frac{1}{2}(y_n + y_m) \in M$, which means that $\|\frac{1}{2}(y_n - y_m)\| \geq \delta$ (since $\inf_{z \in M} \|z\| = \delta$). (This is the key step.) This gives

$$\|y_n - y_m\|^2 \leq -4\delta^2 + 2(\delta_n^2 + \delta_m^2).$$

And now since the sequence δ_n converges to δ, as long as m and n are large enough the right hand side is arbitrarily small, and so we have that (y_n) is Cauchy.

So the sequence (y_n) is Cauchy, and since M is complete, this means (y_n) converges to a point $y \in M$. □

Claim 9.13 — The point y is unique.

Proof. Assume that there are two points $y_1, y_2 \in M$ with $\|x - y_1\| = \|x - y_2\| = \delta$ (i.e., $\|y_1\| = \|y_2\| = 0$). Then again using the parallelogram inequality, we have

$$\|y_1 - y_2\|^2 = -\|y_1 + y_2\|^2 + 2(\|y_1\|^2 + \|y_2\|^2) = -\|y_1 + y_2\|^2 + 4\delta^2.$$

We can then use the same convexity trick as before — since M is convex, we know that $\frac{1}{2}(y_1 + y_2) \in M$, and therefore $\|\frac{1}{2}(y_1 + y_2)\| \geq \delta$, and we get that

$$\|y_1 - y_2\|^2 \leq -4\delta^2 + 4\delta^2 = 0.$$

So we must have $\|y_1 - y_2\| = 0$, which means $y_1 = y_2$. □

For the general case (where $x \neq 0$), we can just translate everything by x (and if $x \in M$, then the statement is immediate). □

Remark 9.14. Why do we need convexity? Here we're using the fact that for $y_n, y_m \in M$ we have $\frac{1}{2}(y_n + y_m) \in M$ in order to get a *lower* bound on $\|y_n + y_m\|$.

Lemma 9.15

Let Y be a complete *subspace* of an inner product space X, and let $x \in X$. Then if y is the unique element of Y minimizing $\|x - y\|$ and $z = x - y$, then $z \perp Y$.

First note that any subspace is convex, so the earlier theorem applies.

Proof. We want to show that for all $y' \in Y$, we have $\langle z, y' \rangle = 0$. Assume not, so there exists $y_1 \in Y$ for which $\langle z, y_1 \rangle = \beta \neq 0$ (in particular, this means $y_1 \neq 0$). Because Y is a subspace, we have $y + \alpha y_1 \in Y$ for all α; we'll show that there exists some choice of α for which replacing y with this vector gives us a smaller distance. We have

$$\|x - (y + \alpha y_1)\|^2 = \|z - \alpha y_1\|^2 = \|z\|^2 - \overline{\alpha}\langle z, y_1 \rangle - \alpha \langle y_1, z - \alpha y_1 \rangle.$$

And now we can choose α such that $\langle y_1, z - \alpha y_1 \rangle = 0$ — to do this, we can write this as $\langle y_1, z \rangle - \overline{\alpha} \langle y_1, y_1 \rangle$, which means we want to take

$$\overline{\alpha} = \frac{\langle y_1, z \rangle}{\|y_1\|^2}.$$

Now when we plug this choice of α in, we get that

$$\|z - \alpha y_1\|^2 = \|z\|^2 - \frac{\langle y_1, z \rangle}{\|y_1\|^2} \langle z, y_1 \rangle = \|z\|^2 - \frac{|\langle y_1, z \rangle|^2}{\|y_1\|^2} < \|z\|^2$$

(since we assumed $\langle y_1, z \rangle \neq 0$). So we've obtained a vector $y' = y + \alpha y_1 \in Y$ with distance to x that's smaller than δ, which is a contradiction. □

§9.4 Orthogonal complements and direct sums

Definition 9.16. For a subset Y of an inner product space X, we define the *orthogonal complement* of Y as the set

$$Y^\perp = \{z \in X \mid \langle z, y \rangle = 0 \text{ for all } y \in Y\}.$$

Definition 9.17. We say a vector space X is the *direct sum* of two vector subspaces Y and Z if every element $x \in X$ can be written *uniquely* as $y + z$ for $y \in Y$ and $z \in Z$.

Theorem 9.18

Let H be a Hilbert space, and let $Y \subseteq H$ be a closed subspace. Then $H = Y \oplus Y^\perp$.

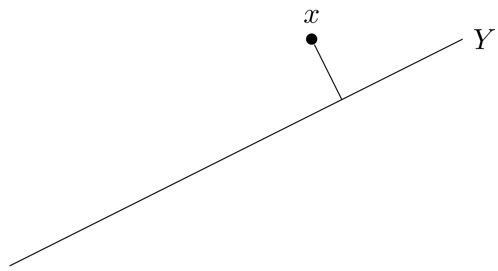
Proof. Since H is complete and Y is closed, then Y is complete; so given $x \in X$, we can define $y \in Y$ as the unique element of Y minimizing $\|x - y\|$, and z as $x - y$; then by the previous lemma we have $z \in Y^\perp$. So it's *possible* to decompose $x = y + z$ as a sum of elements in Y and $Y^\perp$.

Now we need to prove the uniqueness of this decomposition — suppose that we have two decompositions $x = y + z = y_1 + z_1$. Then this means $y - y_1 = z_1 - z$. But $y - y_1 \in Y$ and $z_1 - z \in Y^\perp$, and $Y \cap Y^\perp = \{0\}$ (since any element in both Y and $Y^\perp$ must satisfy $\langle y, y \rangle = 0$, and therefore $y = 0$). This implies $y = y_1$ and $z = z_1$, giving the uniqueness of our decomposition. □

Remark 9.19. The things we've seen here are things that we're familiar with in a geometric setting. In this course, we're asking, in what ways does this geometric intuition generalize to infinite-dimensional vector spaces? And it'll give us a way of understanding things like direct sum in infinite-dimensional vector spaces.

Definition 9.20. Given a closed subspace Y of a Hilbert space H, the *orthogonal projection* of a point $x \in H$ onto Y is defined as the unique point $y \in Y$ minimizing $\|x - y\|$ (i.e., satisfying $\|x - y\| = \delta(x)$). We define the *projection operator* $P \colon H \to Y$ as the corresponding map $x \mapsto y$.

This is a natural generalization of what we've seen in geometry — if we're in $\mathbb{R}^2$ and our subspace is a line, then the projection of a point x onto the line is the closest point on that line to x.



Fact 9.21 — Given any closed subspace $Y \subseteq X$, the projection operator P satisfies the following properties:

(1) P is idempotent, i.e., $P^2x = Px$.

(2) P maps H onto Y, and is the identity on Y.

(3) $Y^\perp = \mathcal{N}(P)$.

(These are all direct.)

Lemma 9.22
If $Y \subseteq H$ is closed, then $Y^{\perp\perp} = Y$.

Proof. One direction is immediate — if we have $x \in Y$ then $x \perp Y^\perp$, which means $x \in Y^{\perp\perp}$. So we automatically have that $Y \subseteq Y^{\perp\perp}$.

For the other direction (where we want to show that $Y^{\perp\perp} \subseteq Y$), we have to be more careful. Let $x \in Y^{\perp\perp}$. Then we can write $x = y + z$ with $y \in Y$ and $z \in Y^\perp$ (since H decomposes as a direct sum of Y and $Y^\perp$); our goal is to show that $z = 0$.

First, we know that $x \in Y^{\perp\perp}$ and $y \in Y \subseteq Y^{\perp\perp}$, so then $z = x - y \in Y^{\perp\perp}$ as well (since $Y^{\perp\perp}$ is a vector space). And now z is in both $Y^\perp$ and $Y^{\perp\perp}$, which means it must be 0, and therefore $x \in Y$. □

§9.5 Orthonormal sets

Definition 9.23. Let X be an inner product space. We say a set $M \subseteq X$ is *orthogonal* if its elements are pairwise orthogonal (i.e., for all distinct $x, y \in M$ we have $\langle x, y\rangle = 0$). We say M is *orthonormal* if it is orthogonal and $\|x\| = 1$ for all $x \in M$.

Fact 9.24 — An orthonormal set is linearly independent.

Example 9.25
In $\mathbb{R}^n$, the standard basis vectors e_i form an orthonormal set.

Example 9.26
In ℓ^2, the basis vectors e_i again form an orthonormal set.

Example 9.27
In $L^2[0, 2\pi]$, we can build an orthonormal set out of cosines and sines (we'll explore this more later).

Proposition 9.28 (Bessel's inequality)
Let (e_k) be an orthonormal sequence in an inner product space X. Then for all $x \in X$, we have

$$\sum_{k=1}^{\infty} |\langle x, e_k \rangle|^2 \le \|x\|^2 .$$

Definition 9.29. We refer to $\langle x, e_k \rangle$ as the *Fourier coefficients* of x with respect to (e_k).

Remark 9.30. Just as in linear algebra, we can use the Gram–Schmidt process to turn any sequence of vectors into an orthonormal one.

§10 March 11, 2024

Today we'll talk about totality and the adjoint operator.

§10.1 Bessel's inequality

Recall that at the end of last lecture, we talked about Bessel's inequality.

Theorem 10.1 (Bessel's inequality)
If we're given an orthonormal sequence $(e_k)_{k\in\mathbb{N}}$ in an inner product space X, then for all $x \in X$ we have

$$\sum_{k\in\mathbb{N}} |\langle x, e_k \rangle|^2 \le \|x\|^2 .$$

The reason that the limit on the left-hand side exists (and that this inequality is true) is that the partial sums $\sum_{k=1}^{n} |\langle x, e_k \rangle|^2$ form an increasing sequence (all terms are positive), and we can show that these partial sums satisfy $\sum_{k=1}^{n} |\langle x, e_k \rangle|^2 \le \|x\|^2$.

Remark 10.2. Note that this implies at most finitely many of the terms $|\langle x, e_k \rangle|$ are greater than $1/m$ for each $m \in \mathbb{N}$. (If there were infinitely many, then the series would be unbounded.) We're going to use this later on to talk about Fourier coefficients.

Definition 10.3. Given an orthonormal sequence $(e_k)_{k\in\mathbb{N}}$ in an inner product space X, let $(\alpha_k)_{k\in\mathbb{N}}$ be a sequence in $\mathbb{K}$, and let $s_n = \sum_{k=1}^{n} \alpha_k e_k$. If $\|s_n - s\| \to 0$ as $n \to \infty$, then we say the series $\sum_{k\in\mathbb{N}} \alpha_k e_k$ converges to s.

Remark 10.4. It's important to think carefully about where the convergence is taking place and what that means; here convergence means convergence in the norm.

Theorem 10.5

Given an orthonormal sequence $(e_k)_{k\in\mathbb{N}}$ in a Hilbert space H:

(1) $\sum_{k\in\mathbb{N}} \alpha_k e_k$ converges to some x if and only if $\sum |\alpha_k|^2$ is finite.

(2) If this sum converges to x, then for each k we have $\langle x, e_k\rangle = \alpha_k$.

(3) For *any* $x \in H$, we have $\sum_{k=1}^{\infty} \langle x, e_k\rangle e_k \to x$.

This should match our intuition — our intuition for everything with inner product spaces is based on $\mathbb{R}^n$. It's clear that (2) is true if we take a *finite* linear combination; and we're trying to extend it to *infinite* 'linear combinations' (which converge).

Definition 10.6. Given $x \in H$, we call $\alpha_k = \langle x, e_k\rangle$ the *Fourier coefficients* of x with respect to the orthonormal sequence (e_k).

Here's the proof of the theorem.

Proof. For the forwards direction of (1), suppose that $\|s_n - s\| \to 0$ (where $s_n = \sum_{k=1}^{n} \alpha_k e_k$). This means the sequence (s_n) is Cauchy. Then for any $m > n$, we have

$$\|s_n - s_m\| = \left\| \sum_{k=m+1}^{n} \alpha_k e_k \right\| = \sqrt{\sum_{k=m+1}^{n} |\alpha_k|^2}$$

(since we have a finite sum). Since (s_n) is Cauchy, for all $\varepsilon > 0$ there exists N such that for all $m > n > N$ we have $\sum_{k=n+1}^{m} |\alpha_k|^2 < \varepsilon$. This implies the sequence $\sum_{k=1}^{n} |\alpha_k|^2$ is Cauchy in $\mathbb{R}$; and $\mathbb{R}$ is complete, so $\sum_{k=1}^{\infty} |\alpha_k|$ converges (i.e., the sequence $\sum_{k=1}^{n} |\alpha_k|^2$ over n converges).

For the backwards direction of (1), we can make the same argument in reverse — if $\sum_{k=1}^{\infty} |\alpha_k|^2$ is convergent, then it's also Cauchy; this implies the sequence of partial sums is Cauchy in norm. And because H is a Hilbert space, this means the partial sums converge in H, as desired.

For (2), we want to show that if we have convergence in the norm, then the Fourier coefficients are what we'd expect. Note that we have $\|s_n - x\| \to 0$ and, looking at *finite* sums (replacing x with s_n), we have $\langle s_n, e_k\rangle = \langle \sum_{i=1}^{n} \alpha_i e_i, e_k\rangle = \alpha_k$ as long as $n \geq k$. By continuity of the inner product, we have $\langle s_n, e_k\rangle \to \langle x, e_k\rangle$ as $n \to \infty$; this gives (2).

Finally, for (3), again by Bessel's inequality we have

$$\sum_{k=1}^{\infty} |\langle x, e_k\rangle|^2 \leq \|x\|^2,$$

which in particular means the sum on the left-hand side is finite. And then by (1), this means $\sum_{k\in\mathbb{N}} \langle x, e_k\rangle e_k$ converges to *some* $y \in X$. □

Note that the fact that there are finitely many e_k for which $|\langle x, e_k\rangle| \geq 1/m$ for each m, combined with (ii) (which doesn't require completeness, and therefore is true for *any* inner product space X), gives the following statement:

Theorem 10.7

For any inner product space X and any $x \in X$, there are only countably many nonzero Fourier coefficients of x with respect to an orthonormal family $(e_k)_{k\in I}$.

Here your orthonormal family could be of any order (i.e., I doesn't have to be countable); but the point is that even if I is huge, only countably many of the Fourier coefficients are nonzero.

§10.2 Totality

Definition 10.8. Let X be a normed space, and let $M \subseteq X$. We say M is *total* if $\overline{\operatorname{Span} M} = X$.

Definition 10.9. For an inner product space X, a *total orthonormal set* M is a total set which is orthonormal.

Remark 10.10. A total orthonormal set is *not* a basis unless X is finite-dimensional. This makes sense because for a set to be a basis, we have to be able to express *every* element as a *finite* linear combination of elements of the basis. But if we have a countable total orthonormal set, we can't express all elements of the space as a *finite* linear combination.

We're going to state the results in this area and then move along to other things.

Theorem 10.11

Let X be an inner product space, and let $M \subseteq X$ be an orthonormal set.

(1) If M is total in X, then $M^\perp = \{0\}$.

(2) If X is a Hilbert space, then the converse holds as well — if $M^\perp = \{0\}$, then M is total in X.

So this gives an equivalent characterization of totality for Hilbert spaces — M is total if and only if the only element in X orthogonal to the entire set is 0.

Finally, there's another equivalent characterization of totality for Hilbert spaces:

Theorem 10.12

If X is a Hilbert space, then M is total in X if and only if for all $x \in X$ we have

$$\sum_{k \in \mathbb{N}} |\langle x, e_k \rangle|^2 = \|x\|^2 .$$

(This is called *Parseval's equality.*)

Remark 10.13. There exist inner product spaces that are not complete that do not contain a total orthonormal set. (The proof is outside the scope of this class.)

Theorem 10.14

Let H be a Hilbert space.

(1) If H is separable, then every orthonormal set is total. (?? I assume this should read countable.)

(2) If H contains a countable total orthonormal set, then H is separable.

What's going on is separability means you have a countable dense subset; and a set is total if the closure of its span is the entire set. So one would hope for this regarding the interaction between the geometry of orthogonality with the concept of separability.

§10.3 The Riesz representation theorem

Theorem 10.15 (Riesz representation theorem (v1))

Let H be a Hilbert space, and let $f \in H'$ be a bounded linear functional on H. Then there exists a unique $z \in H$ such that f can be written as $f(x) = \langle x, z\rangle$. Furthermore, we have $\|f\|_{B(H,\mathbb{K})} = \|z\|$.

(The norm of f is as an operator $H \to \mathbb{K}$; and the norm of z is in H.)

What this theorem is telling us is that if we take any functional on a Hilbert space, then we can write it as an inner product using an element of the Hilbert space.

Example 10.16

Let $H = \mathbb{R}^2$. What does it mean for a function to be an element of $(\mathbb{R}^2)'$? This means f is a bounded linear functional, so in particular f should be linear; and the linear functionals on $\mathbb{R}^2$ all look like $f\colon (x, y) \mapsto \alpha x + \beta y$ (i.e., f is a plane through the origin). And we can write this as $f(x, y) = \langle(\alpha, \beta), (x, y)\rangle$.

Intuitively, Riesz is a generalization of this idea to infinite-dimensional Hilbert spaces. (In general, most things with Hilbert spaces are about generalizing our intuition about geometry and orthogonality to infinite-dimensional spaces.)

Proof. We'll split the proof into three steps. First, we'll prove that there *exists* z such that $f(x) = \langle x, z\rangle$. Then we'll prove that z is unique. And finally, we'll prove that $\|f\|_{B(H,\mathbb{K})} = \|z\|$.

For (1), note that if $f = 0$ then $z = 0$, and the theorem holds. So we'll assume $f \neq 0$. Then $\mathcal{N}(f)$ is not the entire space H (by definition). And this means $\mathcal{N}(f)^\perp$ is not just $\{0\}$ — this is because we know $H = \mathcal{N}(f) \oplus \mathcal{N}(f)^\perp$. So this means we can choose some element $z_0 \in \mathcal{N}(f)^\perp$ with $\|z_0\| = 1$ (there must be some nonzero element, and we can scale it to have norm 1).

Now for $x \in H$, let

$$v_x = f(x) \cdot z_0 - f(z_0) \cdot x.$$

(Here $f(x)$ and $f(z_0)$ are scalars, and $z_0, x \in H$, so this defines some element of H.) Note that $f(v_x) = 0$ by linearity; and this means $v_x \in \mathcal{N}(f)$. In particular, this means

$$\langle v_x, z_0\rangle = 0$$

(because $z_0 \in \mathcal{N}(f)^\perp$). Now let's multiply out what this inner product is — we get

$$\langle v_x, z_0\rangle = \langle f(x)z_0 - f(z_0)x, z_0\rangle = f(x)\langle z_0, z_0\rangle - f(z_0)\langle x, z_0\rangle = f(x) - f(z_0)\langle x, z_0\rangle$$

(since $\|z_0\| = 1$). Solving for $f(x)$ gives that

$$f(x) = f(z_0)\langle x, z_0\rangle = \langle x, \overline{f(z_0)}z_0\rangle,$$

which gives a statement of the form we want (with $z = \overline{f(z_0)}z_0$).

Remark 10.17. What's going on geometrically is this argument — we take an element perpendicular to the null space, and use projection to figure out what the proper choice of z is.

So we've completed step 1 (showing that z exists); now we want to do step 2 (showing that z is unique). To do so, assume that z is not unique, so there are $z_1 \neq z_2$ with $f(x) = \langle x, z_1\rangle = \langle x, z_2\rangle$ for all x. Then in particular, we have

$$\langle x, z_1 - z_2\rangle = 0$$

for all x. Letting $x = z_1 - z_2$, we get that

$$\langle z_1 - z_2, z_1 - z_2 \rangle = \|z_1 - z_2\|^2 = 0,$$

and therefore $z_1 = z_2$. This proves uniqueness.

Finally, to prove that $\|f\|_{B(H,\mathbb{K})} = \|z\|$, note that since $f \neq 0$ we have $z \neq 0$. Then

$$\|z\|^2 = \langle z, z \rangle = f(z),$$

and since $f(z) \leq \|f\| \, \|z\|$ we get that $\|z\| \leq \|f\|$.

Meanwhile, for the reverse direction, we have

$$\|f\| = \sup_{\|x\|=1} |f(x)| = \sup_{\|x\|=1} |\langle x, z \rangle| \, .$$

But by the Schwarz inequality we have

$$|\langle x, z \rangle| \leq \|x\| \, \|z\|$$

for all x, so in particular this is at most $\|z\|$ for $\|x\| = 1$; so then we get $\|f\| \leq \|z\|$, as desired. □

This is a cool, natural geometric result that's very important in many fields of math.

We have a generalization of this result to *sesquilinear forms*; we need to talk about this generalization to discuss Hilbert adjoints. (We'll state this generalization but not prove it.)

Definition 10.18. Let X and Y be vector spaces over $\mathbb{K}$. We say a map $h\colon X \times Y \to \mathbb{K}$ is a *sesquilinear form* if h is linear in the first variable and conjugate-linear in the second.

Remark 10.19. If $\mathbb{K} = \mathbb{R}$ then sesquilinear is the same thing as bilinear.

Definition 10.20. If X and Y are normed spaces, then we say a sesquilinear form h is bounded if there exists $C > 0$ such that

$$|h(x, y)| \leq C \, \|x\|_X \, \|y\|_Y$$

for all $x \in X$ and $y \in Y$. In this case, we define

$$\|h\| = \sup_{\|x\|=\|y\|=1} |h(x, y)| \, .$$

(In other words, $\|h\|$ is the best constant C you can get; this is the generalization you'd expect from our discussion of bounded linear operators.)

We'll now state a generalization of Riesz's representation theorem to bounded sesquilinear forms.

Theorem 10.21

Let H_1 and H_2 be Hilbert spaces and let $h\colon H_1 \times H_2 \to \mathbb{K}$ be a bounded sesquilinear form. Then h has a representation

$$h(x, y) = \langle Sx, y \rangle$$

where $S\colon H_1 \to H_2$ is a bounded linear operator. Furthermore, S is unique, and we have $\|S\| = \|h\|$.

Note that this inner product is in H_2 (since S maps $H_1 \to H_2$). When we write $\|S\|$, we're using the operator norm on $B(H_1, H_2)$; and $\|h\|$ is in the sense defined above.

Proof sketch. First, note that $\overline{h(x,y)}$ is linear in y. Now we'll fix $x \in H_1$ and apply the first version of the Riesz representation theorem; this gives us that

$$\overline{h(x,y)} = \langle y, z_x \rangle$$

for some unique $z_x \in H_2$ (which depends on x). Immediately, we deduce that

$$h(x,y) = \langle z_x, y \rangle$$

(taking the conjugate of both sides).

Now consider the map $S\colon H_1 \to H_2$ mapping $x \mapsto z_x$; then we get that $h(x,y) = \langle Sx, y \rangle$.

We now need to check that S has the properties we want — checking that it's a bounded linear operator is straightforward (to prove linearity, for example if we take $x_1 + x_2$, we have

$$\overline{h(x_1 + x_2, \cdot\,)} = \overline{h(x_1, \cdot\,)} + \overline{h(x_2, \cdot\,)},$$

which ends up giving $z_{x_1+x_2} = z_{x_1} + z_{x_2}$). □

§10.4 Adjoints

Recall from linear algebra that if $A\colon \mathbb{C}^n \to \mathbb{C}^n$ is a linear transformation, then we can represent A as a $n \times n$ matrix. Then its *adjoint* A^* is defined as its conjugate transpose, and importantly, it satisfies

$$\langle Ax, y \rangle = \langle x, A^* y \rangle$$

for all $x, y \in \mathbb{C}^n$. This is the important equality we want to generalize to the abstract setting of Hilbert spaces.

Theorem 10.22

Let $T\colon H_1 \to H_2$ be a bounded linear operator, where H_1 and H_2 are Hilbert spaces. Then there exists a *unique* linear operator $T^*\colon H_2 \to H_1$ such that

$$\langle Tx, y \rangle = \langle x, T^* y \rangle$$

for all $x \in H_1$ and $y \in H_2$. Furthermore, we have $\|T\| = \|T^*\|$.

Proof. Let $h\colon H_2 \times H_1 \to \mathbb{K}$ be defined as

$$h(y, x) = \langle y, Tx \rangle.$$

Then h is a sesquilinear form — linearity in the first term is immediate, and conjugate linearity in the second follows from the linearity of T and conjugate linearity of $\langle y, - \rangle$ (explicitly, $h(y, x_1 + \alpha x_2) = \langle y, Tx_1 + \alpha T x_2 \rangle = \langle y, Tx_1 \rangle + \overline{\alpha} \langle y, Tx_2 \rangle$).

We also claim that h is bounded — to see this, by the Schwarz inequality we have

$$|h(x,y)| = |\langle y, Tx \rangle| \le \|y\| \, \|Tx\| \le \|y\| \, \|T\| \, \|x\| \, .$$

So this implies h is bounded, and $\|h\| \le \|T\|$.

In fact, we claim that $\|h\|$ is exactly equal to $\|T\|$. To see this, we can specifically plug in $y = Tx$; then we get

$$\|h\| \ge \sup_{\|x\| \ne 0, \|Tx\| \ne 0} \frac{\langle Tx, Tx \rangle}{\|x\| \, \|Tx\|} = \sup_{\|x\| \ne 0, \|Tx\| \ne 0} \frac{\|Tx\|}{\|x\|} = \|T\| \, .$$

These two inequalities together tell us that $\|h\| = \|T\|$.

And now we can apply the Riesz representation theorem for sesquilinear forms, and let T^* be the S that it gives. There's a small amount of algebra left to do — we conclude that there exists a unique $T^*: H_2 \to H_1$ such that $h(y, x) = \langle T^*y, x\rangle$, and taking conjugates gives that $\langle x, T^*y\rangle = \overline{h(y, x)} = \langle Tx, y\rangle$. So we've ended up with the adjoint that we promised, and it satisfies $\|T^*\| = \|h\| = \|T\|$. □

Remark 10.23. Theorem 3.9–4 in Kreyszig gives a list of properties for the adjoint; we'll talk about the particular ones we'll use in class next time.

§11 March 13, 2024

§11.1 Comments on the midterm

§11.1.1 Problem 1

On Problem 1, we defined $d(f, g) = \int_0^1 |f - g|^p \, dx$, where $p \in [0, 1]$. How do you prove the triangle inequality here? It's enough to show that

$$|f - g|^p \le |f - h|^p + |h - g|^p$$

(this is what we want to show). Marjie doesn't think this is obvious at all, from looking at it.

As another remark, suppose we're just given two nonnegative real numbers a and b; is $(a + b)^p \le a^p + b^p$? Our intuition is that it is, but this isn't obvious. Let's call this $(*)$. Without loss of generality we can assume $b \ge a$; then we can divide through by b^p and get

$$\left(1 + \frac{a}{b}\right)^p \le 1 + \left(\frac{a}{b}\right)^p.$$

And this is true if and only if $(1 + t)^p \le 1 + t^p$ for all $t \in (0, 1]$. And now if we raise this to the $1/p$th power, we want to show $1 + t \le (1 + t^p)^{1/p}$. And at this point, we should start to believe this — maybe we can think of p as $\frac{1}{2}$ (the simplest example), and in that case this says $(1 + t^{1/2})^2 \ge 1 + t$, which is true. So this is the first point at which Marjie believes we've almost proven the thing, but it's still not proved. You can do things with Taylor's theorem; there are also a few clever ways ('clever' means something that saves you time and is worth memorizing, but you often have to see it to know it).

The clever trick is to let $q = 1 - p$ (you can assume $p \in (0, 1)$). Then the left-hand side of $(*)$ becomes

$$(a + b)^{1-q} \le a(a + b)^{-q} + b(a + b)^{-q}.$$

But we have $(a + b)^q \ge a^q$, so $(a + b)^{-q} \le a^{-q}$; then we get that this is at most $a^{1-q} + b^{1-q} = a^p + b^p$.

Some of us remarked on concavity; there is a clever trick you can do to use the concavity of the function $x(t) = t^p$, but you can't just say 'concavity', you have to use it. In particular, the concavity isn't enough; if you use concavity of this function, you also have to use the fact that $x(0) \ge 0$.

Remark 11.1. This is a good note on how much detail to write in your proof.

§11.1.2 Problem 6

To prove problem 6, there are two theorems we proved in class, about compactness. One said that the unit ball in infinite dimensions is not compact; conversely, in finite dimensions the unit ball is compact. But you should be very careful about moving intuition from finite dimensions to infinite ones. A number of people tried to prove this using a generalization of orthogonality; but this problem is about normed spaces, not inner product spaces, so there's no concept of orthogonality here.

Marjie thinks of the exam as out of 75, even though there were 120 points. (Getting 75 points is reasonable for an advanced undergraduate course; getting 120 is not.)

§11.2 Zorn's lemma

§11.2.1 Partial orderings

Definition 11.2. A *partial order* on a set M is a binary relation $\preceq$ satisfying:

(i) Reflexivity — $a \preceq a$ for all $a \in M$.

(ii) Transitivity — if $a \preceq b$ and $b \preceq c$, then $a \preceq c$.

(iii) Anti-symmetry — if $a \preceq b$ and $b \preceq a$, then $a = b$.

Example 11.3

- $\leq$ is a partial ordering on $\mathbb{R}$ (this is actually a total ordering, and it makes a lot of interesting math problems simple in the one-dimensional case).
- Let X be a set, and $\mathcal{P}(X)$ its power set. Then containment of subsets gives an ordering on $\mathcal{P}(X)$ — where $A \preceq B$ if $A \subseteq B$.
- Consider the set of n-tuples of real numbers; then we can define the partial ordering $x \preceq y$ if $x_i \leq y_i$ for all $i \in [n]$.

Note that the word *partially* means that some elements may be incomparable — there can exist $a, b \in M$ such that neither $a \preceq b$ nor $b \preceq a$ holds.

Definition 11.4. If neither $a \preceq b$ nor $b \preceq a$ holds, we say a and b are *incomparable.*

Definition 11.5. A *totally ordered set* (or *chain*) is a set where every pair of elements is comparable.

Definition 11.6. An *upper bound* on a subset $W \subseteq M$ is an element $x \in M$ such that for all $y \in W$, we have $y \preceq x$.

(It's possible that x may not be in W.)

Notice that an upper bound may or may not exist.

Example 11.7

In our second example, given any chain $W \subseteq \mathcal{P}(X)$, the union of all elements of the chain is an upper bound on W. In fact, the entire set X is an upper bound on *any* set $W \subseteq \mathcal{P}(X)$.

Example 11.8

The set $\mathbb{R}$ has no upper bound.

Definition 11.9. We say an element $m \in M$ is *maximal* if for all $x \in M$, if $m \preceq x$ then $m = x$.

In other words, if we can compare any element of M to m, then it must be smaller than m. There might be some elements in M that you *can't* compare to m, but the ones that you can compare to m must all be smaller.

Remark 11.10. It's also true that M may or may not have maximal elements — for example, $\mathbb{R}$ doesn't have a maximal element, while X (the entire set) is a maximal element of W.

§11.2.2 Zorn's lemma and consequences

Note that Zorn's lemma is the analyst's version of the axiom of choice; so we'll take it as an axiom, and not prove it.

Theorem 11.11 (Zorn's lemma)

Let $M \neq \emptyset$ be a partially ordered set such that every chain has an upper bound. Then M has a maximal element.

Before we get to the Hahn–Banach theorem, we'll talk about some immediate consequences which came up earlier in the text.

Theorem 11.12

Every vector space $X \neq \{0\}$ has a Hamel basis.

Proof. To use Zorn's lemma, we need to come up with a binary relation where every chain has an upper bound; this is a natural thing to do here. Let M be the set of all linearly independent subsets of X. Note that M is nonempty (since $X \neq \{0\}$, there exists some $x \in X$ such that $x \neq 0$, and then we have $\{x\} \in M$).

Now define our partial ordering $\preceq$ by set inclusion — we say $A \preceq B$ if $A \subseteq B$.

Let $\mathcal{C}$ be any chain; we want to prove that $\mathcal{C}$ has an upper bound (so that we can use Zorn's lemma). To see this, we can simply note that $\bigcup_{S \in \mathcal{C}} S$ is an upper bound on $\mathcal{C}$.

So by Zorn's lemma, M has a maximal element; now we just need to prove that the maximal element generates the entire space. Call the maximal element B; we want to show that B is a Hamel basis for X.

Let's suppose not. Then we can take some $x \in X \setminus \operatorname{Span}(B)$, and let $B' = B \cup \{x\}$; then $B' \in M$ as well (because x wasn't in $\operatorname{Span}(B)$). Then $B \subsetneq B'$, but this is a contradiction (contradicting the maximality of B). □

The proof of the following theorem is very similar, so we'll just state it.

Theorem 11.13

In every Hilbert space, there exists a total orthonormal basis.

To prove this, we just need to choose an orthonormal set of vectors (with relation by containment again). Each chain has an upper bound, so there's a maximal element; and then this looks quite similar.

§11.3 The Hahn–Banach theorem

(This is the version in Kreyszig 4.3–1. We'll remark on the other version later.)

Theorem 11.14 (Hahn–Banach theorem)

Let X be a normed space over $\mathbb{K}$, and suppose Y is a (not necessarily closed) subspace of X. If $\varphi_0 \colon Y \to \mathbb{K}$ is a bounded linear functional, then there exists $\varphi \in X'$ (i.e., $\varphi \colon X \to \mathbb{K}$ is a bounded linear functional) extending φ_0 (i.e., $\varphi|_Y = \varphi_0$) and such that $\|\varphi\|_{\text{op}} = \|\varphi_0\|_{\text{op}}$.

(The last condition is really the strength of the theorem.)

We'll go through the proof of this, and then talk about some of the immediate applications (that imbue the concept of the dual space with value — it's not yet clear why the dual space is useful). And then there are some more important uses of the Hahn–Banach theorem as we move along — for example, extending the idea of the adjoint operator to normed spaces (instead of just Hilbert spaces).

But for now, let's look at the proof.

Proof. We're going to look at the proof for $\mathbb{R}$. Notice first that if $\|\varphi_0\|_{\text{op}} = 0$, then $\varphi_0 = 0$; and so we can just let $\varphi = 0$, and the theorem holds. So we can assume that $\|\varphi_0\|_{\text{op}} \neq 0$. In fact, we can reduce to the case $\|\varphi_0\|_{\text{op}} = 1$, just by dividing through by a constant (and multiplying φ through by the same constant in the end).

Let $z \in X \setminus Y$, and let $Y_1 = \{\alpha z - y \mid y \in Y, \alpha \in \mathbb{R}\}$ — equivalently $Y_1 = \text{Span}(Y \cup \{z\})$, but we've written it in this way to make the math in future steps a bit clearer.

The idea here is going to be that we extend φ_0 to φ_1, and then we're going to do this iteratively, one point at a time; and we'll use Zorn's lemma to put an ordering on the subspace-functional pairs (Y', φ') that we create (where φ' is a functional on Y), and apply it to come up with a maximal element; and we'll show that maximal element is actually the entire space.

In the first step, we can define φ_1, but then we need to make sure φ_1 has norm 1. So we define

$$\varphi_1(\alpha z - y) = \alpha c - \varphi_0(y),$$

where we'll choose c momentarily. First, note that $\varphi_1 \colon Y_1 \to \mathbb{R}$ is well-defined (because every element of Y_1 can be written uniquely as $\alpha z - y$). Notice also that φ_1 is an extension of φ_0, in that $\varphi_1|_Y = \varphi_0$ (if we're looking at an element of Y, then $\alpha = 0$, and this just says $\varphi_1(-y) = -\varphi_0(y)$). And the fact that φ_1 is linear is also immediate ($\alpha \mapsto \alpha c$ and φ_0 are both linear).

So the important thing now is that we want to choose c so that $\|\varphi_1\|_{\text{op}} = \|\varphi_0\|_{\text{op}}$. We have $\|\varphi_0\| = 1$, which means we want

$$\|\varphi_1\|_{\text{op}} = \sup_{\alpha z - y \neq 0} \frac{|\alpha c - \varphi_0(y)|}{\|\alpha z - y\|} = 1,$$

or equivalently that $|\alpha c - \varphi_0(y)| \leq \|\alpha z - y\|$ (for all $\alpha \in \mathbb{R}$ and $y \in Y$; note that z is fixed). Dividing through by α, this is true if and only if

$$|c - \varphi_0(y/\alpha)| \leq \|z - y/\alpha\|$$

for all $\alpha \in \mathbb{R}$ and $y \in Y$. Replacing y/α by y, this is true if and only if

$$|c - \varphi_0(y)| \leq \|z - y\|$$

for all $y \in Y$.

And now, solving for c, this is true if and only if

$$\varphi_0(y) - \|z - y\| \leq c \leq \varphi_0(y) + \|z - y\|$$

for all $y \in Y$. So in order to show that we can choose c, we want to show that

$$\bigcap_{y \in Y} [\varphi_0(y) - \|z - y\|, \varphi_0(y) + \|z - y\|]$$

is nonempty.

And we claim that this intersection is nonempty if and only if $\varphi_0(y) - \|z - y\| \leq \varphi_0(\nu) + \|z - \nu\|$ for all $y, \nu \in Y$. (This is because we've got an intersection of intervals; there's a detailed proof in the notes.)

So then it suffices to prove this inequality. But it's actually immediate, because

$$\varphi_0(y) - \varphi_0(\nu) = \varphi_0(y - \nu),$$

and since $\|\varphi_0\|_{\text{op}} = 1$ we have

$$|\varphi_0(y - \nu)| \leq \|\varphi_0\|_{\text{op}} \|y - \nu\| = \|y - \nu\|.$$

And then we can just use the triangle inequality to conclude that

$$|\varphi_0(y) - \varphi_0(\nu)| \leq \|y - \nu\| \leq \|z - y\| + \|z - \nu\|,$$

which exactly gives what we wanted.

So we've proven that we can choose c (we've proven this intersection is nonempty, and c can have any value in this intersection). So we have completed the proof that there exists c which works.

Now we've defined φ_1, and now we're prepared to use Zorn's lemma. Let $\mathcal{P}$ be the collection of all pairs (Y', φ') where Y' is a subspace of X and $\varphi' \colon Y' \to \mathbb{R}$ is a bounded linear operator with $\|\varphi'\|_{\text{op}} = 1$.

We say $(Y', \varphi') \preceq (Y'', \varphi'')$ if $Y' \subseteq Y''$ and $\varphi''|_{Y'} = \varphi'|_{Y'}$ — in other words, φ'' is an extension of φ'.

Now we need to show that every chain has an upper bound. To do so, let $\mathcal{C} = \{(Y_\beta, \varphi_\beta)\}_{\beta \in J}$ be a chain (where J is some index set), i.e., a totally ordered subset of $\mathcal{P}$. Let $N = \bigcup_{\beta \in J} Y_\beta$; because each Y_β is a subspace of X, then N is a subspace of X as well. And let $\widetilde{\varphi} \colon N \to \mathbb{R}$ be defined as follows — for any $y \in N$, we define

$$\widetilde{\varphi}(y) = \varphi_\beta(y)$$

for some $\beta \in J$ such that $y \in Y_\beta$. (Because of the way we defined the ordering and the fact that $\mathcal{C}$ is totally ordered, this is well-defined — if there are multiple sets containing y, they'll have the same values of $\varphi_\beta(y)$.)

Further, $\widetilde{\varphi}$ is linear (this is immediate from the fact that each φ_β is linear; and every time we choose a linear combination, the components of that linear combination must be contained in some subspace Y_β, so we get linearity.)

And we have $|\widetilde{\varphi}(y)| = |\varphi_\beta(y)| \leq \|y\|$, which implies that $\|\widetilde{\varphi}\| = 1$.

So we've shown that $(N, \widetilde{\varphi})$ is an element of $\mathcal{P}$; and it's clear that it's an upper bound for $\mathcal{C}$ (if we take any element (Y_β, φ_β) of the chain, then by construction this is an upper bound).

So we can apply Zorn's lemma; and by Zorn's lemma, $\mathcal{P}$ has a maximal element. Let's call that maximal element $(X_\infty, \varphi_\infty)$.

Claim 11.15 — We have $X_\infty = X$.

Proof. This is very similar to the earlier argument — it's immediate that $X_\infty \subseteq X$. Now assume there exists $x \in X \setminus X_\infty$. Then we can perform the same extension process we used to go from Y to Y_1, replacing $z = x$ and $Y = X_\infty$ — this leads to a contradiction of the maximality of $(X_\infty, \varphi_\infty)$. So we must have $X = X_\infty$. □

And this completes the proof of the Hahn–Banach theorem (where our final φ is φ_∞). □

§11.4 Some consequences

There are some immediate consequences of the Hahn–Banach theorem, whose proofs are very pretty.

Question 11.16. Let $X \neq \{0\}$ be a normed linear space. Can X' be trivial?

The answer is no, and the proof follows from the following corollary to the Hahn–Banach theorem.

Corollary 11.17

Let $X \neq \{0\}$ be a nontrivial normed linear space, and let $x_0 \in X$ be a nonzero element of X. Then there exists $\varphi \in X'$ such that $\varphi(x_0) = \|x_0\|$.

Proof. We define $M = \text{Span}\{x_0\} = \{\alpha x_0 \mid \alpha \in \mathbb{K}\}$. Then M is a subspace of X. And we can define $\varphi_0 \colon M \to \mathbb{K}$ as

$$\varphi_0(\alpha x_0) = \alpha \|x_0\| .$$

This is a bounded linear functional on M (this subspace). And so we can apply the Hahn–Banach theorem to say there is a bounded linear functional of the same norm on the entire space, extending this one. □

Here's our next corollary.

Corollary 11.18

Suppose that X is a nontrivial normed linear space, and let $x_1 \neq x_2$ be distinct elements of X. Then there exists $\varphi \in X'$ such that $\varphi(x_1) \neq \varphi(x_2)$.

Remark 11.19. This tells us that the dual of a nontrivial normed space X (i.e., a space $X \neq \{0\}$) 'separates the points of X'; people often refer to this as a result about the 'richness' of the dual space.

Proof. We apply the first corollary to $x_0 = x_1 - x_2$; then we get some $\varphi \in X'$ with $\varphi(x_1 - x_2) = \|x_1 - x_2\| \neq 0$. □

Corollary 11.20

Let X be a normed linear space and $x_0 \in X$. Then we have

$$\|x_0\| = \sup\{|\varphi(x_0)| \mid \varphi \in X',\ \|\varphi\| = 1\}.$$

Proof. We can assume $x_0 \neq 0$ (if it's 0, then this is immediate). Then if $\|\varphi\| = 1$, we have $|\varphi(x_0)| \leq \|\varphi\|_{\text{op}} \|x_0\| \leq \|x_0\|$. Conversely, we can show that equality is achieved by using Corollary 1 — there exists $\widehat{\varphi}$ with $\|\widehat{\varphi}\|_{\text{op}} = 1$ and $|\widehat{\varphi}(x_0)| = \|x_0\|$, which means the supremum is achieved. □

So the Hahn–Banach theorem tells us something about the power of the dual space. We're going to keep playing it; it's worth thinking a bit more about what the theorem and the corollaries say (they're easy from the theorem, but interesting results on their own, and it may take some thinking about them to see that).

On Monday, we'll do some more with this theorem.

§12 March 18, 2024

Today we'll see some bigger applications of the Hahn–Banach theorem. First, we'll restate the theorem and the corollary that we're going to use.

Theorem 12.1 (Hahn–Banach)

Let X be a normed linear space over a field $\mathbb{K}$, let Y be a (not necessarily closed) subspace of X, and let $\varphi_0 : Y \to \mathbb{K}$ be a bounded linear functional on Y. Then we can extend φ_0 to a bounded linear functional φ on X satisfying $\|\varphi\|_{\text{op}} = \|\varphi_0\|_{\text{op}}$ — in other words, there exists a bounded linear functional $\varphi : X \to \mathbb{K}$ satisfying $\varphi|_Y = \varphi_0$ and $\|\varphi\|_{\text{op}} = \|\varphi_0\|_{\text{op}}$.

Corollary 12.2

Let X be a nontrivial normed linear space, and let $x_0 \in X$. Then there exists $\varphi \in X'$ satisfying $\|\varphi\|_{\text{op}} = 1$ and $\varphi(x_0) = \|x_0\|$.

We're going to use this today to develop theory behind the adjoint operator.

§12.1 The adjoint operator

Previously, we've talked about the Hilbert adjoint. The adjoint we'll talk about today is different.

Definition 12.3. Let X and Y be normed linear spaces and let $T : X \to Y$ be a bounded linear operator. We define its *adjoint operator* $T^\times : Y' \to X'$ in the following way — let $g \in Y'$. Then $T^\times g$ is defined as the functional $f \in X'$ such that $f(x) = g(Tx)$ for all $x \in X$.

Note that Tx is an element of Y. The adjoint $T^\times$ has to be an operator $Y' \to X'$, so to say what it does, we have to say what it does to an element of Y'. So we take some $g \in Y'$, and we want to map it to some $f \in X'$. And in order to tell us what an element in X' does, we have to say what it does to elements of x — so $f(x)$ is defined as $g(Tx)$.

It's not immediately clear that this is well-defined (i.e., it's not immediately obvious that this map f is an element of X'). So let's prove this.

Claim 12.4 — For f defined as above, we have $f \in X'$.

Proof. First we need to check that f is linear (in terms of how it acts on elements of X) — we have

$$f(\alpha x + y) = g(T(\alpha x + y)) = g(\alpha Tx + Ty) = \alpha g(Tx) + g(Ty) = \alpha f(x) + f(y)$$

(by the linearity of T and $g \in Y'$). So f is linear.

The next thing we need to show is that f is bounded. We have $|f(x)| = |g(Tx)| \leq \|g\| \, \|Tx\|$ (since $g \in Y'$ is bounded); and since T is also a bounded operator, this is bounded by $\|g\| \, \|T\| \, \|x\|$. (Here we've used the boundedness of both the operators g and T.) This proves that the map $T^\times$ we've defined is genuinely a map from Y' to X'. □

Remark 12.5. We're using T^* to denote the Hilbert adjoint, and $T^\times$ to denote the adjoint here. When we defined the Hilbert adjoint, given $T : X \to Y$, we defined T^* such that $\langle Tx, y \rangle = \langle x, T^* y \rangle$ — so T^* is a map $Y \to X$. Meanwhile, the adjoint operator is actually an operator $Y' \to X'$ — so it maps linear functionals to linear functionals. So these aren't the same thing.

Theorem 12.6

The adjoint operator $T^\times$ is linear and bounded, with $\|T^\times\| = \|T\|$.

This says something different from what we just proved — earlier, we proved that $T^\times g$ is an element of the dual space X'. Here, saying that $T^\times$ is linear means that the way it acts on elements of the dual of Y is linear.

Proof. To prove linearity here, we need to show that for $\alpha \in \mathbb{K}$ and $g_1, g_2 \in Y'$, we have $T^\times(\alpha g_1 + g_2) = \alpha T^\times g_1 + T^\times g_2$ (since we want to show that $T^\times$ is a linear operator on Y').

To do so, $T^\times(\alpha g_1 + g_2)$ is the operator such that for each $x \in X$ we have

$$T^\times(\alpha g_1 + g_2)(x) = (\alpha g_1 + g_2)(Tx) = \alpha g_1(Tx) + g_2(Tx) = \alpha T^\times g_1(x) + T^\times g_2(x).$$

This means $T^\times(\alpha g_1 + g_2) = \alpha T^\times g_1 + T^\times g_2$, as desired.

The next step is to prove that $T^\times$ is bounded (in the norm on $B(Y', X')$ — the space of bounded linear operators $Y' \to X'$), and that $\|T^\times\|_{B(Y',X')} = \|T\|$. By definition, we have

$$\|T^\times\|_{B(Y',X')} = \sup_{\|g\|_{Y'}=1} \|T^\times g\|_{X'}$$

(note that here g is an element of Y', and $T^\times g$ is an element of X'). And $T^\times g$ is defined as gT (as an element of X'); and for $\|g\| = 1$ we have $\|gT\| \le \|g\| \, \|T\| = \|T\|$. So that's one direction.

The other direction is where we use the Hahn–Banach theorem — by the corollary above, for each $y \in Y$, there exists $g \in Y'$ such that $\|g\| = 1$ and $g(y) = \|y\|$; in particular, applying this to $y = Tx_0$ for $x_0 \in X$, there exists $g_0 \in Y'$ with $\|g_0\| = 1$ and $g_0(Tx_0) = \|Tx_0\|$. But now we have

$$\|Tx_0\| = g_0(Tx_0) \le \|T^\times g_0\| \, \|x_0\| \le \|T^\times\| \, \|g_0\| \, \|x_0\|,$$

which gives that $\|T\| \le \|T^\times\| \, \|g_0\|$, which means $\|T^\times\| \ge \|T\|$ (since $\|g_0\| = 1$).

So in combination, these give that $\|T^\times\| = \|T\|$. □

Remark 12.7. The important thing here is that $\|T^\times\| \le \|T\|$ isn't hard, while we used Hahn–Banach to show that $\|T\| \le \|T^\times\|$; this depends on the idea that there is some $g_0 \in Y'$ satisfying $g_0(Tx_0) = \|Tx_0\|$.

§12.2 Hilbert adjoint vs. adjoint operator

To re-emphasize this point, if T is a map $X \to Y$, then T^* is a map $Y \to X$, but $T^\times$ is a map $Y' \to X'$.

Next, we'll use the Riesz representation theorem to understand the relationship between the Hilbert adjoint T^* and the adjoint operator $T^\times$. First we'll consider $\mathbb{R}^n$ — there we know that the Hilbert adjoint is the conjugate transpose. In that context, what is the adjoint operator? Then we'll use the Riesz representation theorem to talk about general (not necessarily finite-dimensional) Hilbert spaces.

§12.2.1 Finite-dimensional spaces

First suppose that T is a linear operator $\mathbb{R}^n \to \mathbb{R}^n$; then we can represent T by a $n \times n$ matrix (depending on our choice of bases). Let $E = \{e_1, \ldots, e_n\}$ be a basis for $\mathbb{R}^n$, and let $T_E = (\tau_{ij})$ be the matrix representing

T with respect to this basis. Let $x, y \in \mathbb{R}^n$ be such that $y = T_E x$ (here x and y are column vectors, and T_E is a $n \times n$ matrix). This tells us

$$y_i = \sum_{j=1}^{n} \tau_{ij} x_j$$

(by the definition of matrix multiplication).

Let $f = \{f_1, \dots, f_n\}$ be the dual basis corresponding to E — recall that this means $f_i(e_j) = \delta_{ij} = 1_{i=j}$. We've seen that this is a basis for $(\mathbb{R}^n)'$ (we've seen that $(\mathbb{R}^n)^* = (\mathbb{R}^n)' = \mathbb{R}^n$ — this isn't true in general, but it is in finite dimensions).

Now if $g \in (\mathbb{R}^n)'$, there exists $\alpha_i \in \mathbb{R}$ (for $1 \le i \le n$) such that $g = \sum_{i=1}^n \alpha_i f_i$. Then we have

$$g(T_E x) = g((\sum_{j=1}^{n} \tau_{ij} x_j)_{i=1}^n) = \sum_{i=1}^{n} \alpha_i \sum_{j=1}^{n} \tau_{ij} x_j.$$

And now we can just change the order of the summation and rewrite this as

$$\sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \tau_{ij} x_j = \sum_{j=1}^{n} x_j \left(\sum_{i=1}^{n} \alpha_i \tau_{ij} \right).$$

Now let $\beta_j = \sum_{i=1}^n \alpha_i \tau_{ij}$; this tells us that

$$g(T_E x) = \sum_{j=1}^{n} \beta_j x_j.$$

And this tells us that $T^\times g$ is the above function (as a function of x). But we can write the above function as $f(x) = \sum_{j=1}^n \beta_j f_j(x)$, and so this gives us that

$$T^\times g = \sum_{j=1}^{n} \beta_j f_j.$$

So what we have learned here is that the matrix representing $T^\times$ is $T_E^\intercal = (\tau_{ji})$.

Remark 12.8. Presenting linear algebra on the board is unpleasant, but what we've proved is that the matrix for the adjoint operator is the transpose. Meanwhile, recall that the matrix for T^* is the *conjugate* transpose. In the real case, these are the same — so if we're working over $\mathbb{R}^n$, then the Hilbert adjoint is the same as the adjoint operator.

Notice that this is really importantly using the Riesz representation theorem — the Hilbert adjoint maps $Y \to X$ and the adjoint operator maps $Y' \to X'$, but here $Y' \cong Y$ by the Riesz representation theorem (in finite dimensions). So that's why we see these are the same.

§12.3 General Hilbert spaces

The natural next question is what the relationship between T^* and $T^\times$ is for a general Hilbert space. So we're going to talk about that.

Suppose we have a map $T\colon H_1 \to H_2$, where H_1 and H_2 are Hilbert spaces; we'll use the Riesz representation theorem to understand the relationship between T^* and $T^\times$. In this context, we have $T^\times\colon H_2' \to H_1'$, where $T^\times g$ is defined as the function f where $g(Tx) = f(x)$.

By the Riesz representation theorem, there exists $x_0 \in H_1$ and $y_0 \in H_2$ such that $f(x) = \langle x, x_0 \rangle$ (for $x \in H_1$) and $g(y) = \langle y, y_0 \rangle$ (for $y \in H_2$). Let $A_1\colon H_1' \to H_1$ be the map $f \mapsto x_0$, and $A_2\colon H_2' \to H_2$ be the map

$g \mapsto y_0$. From the Riesz representation theorem, we have $\|A_1 f\| = \|x_0\| = \|f\|$ and $\|A_2 g\| = \|y_0\| = \|g\|$. Also, by construction we have that A_1 and A_2 are conjugate-linear. (This means $A_1(\alpha f) = \overline{\alpha} A_1 f$.)

Let $\overline{T^*}: H_2 \to H_1$ be defined by the formula $\overline{T^*} = A_1 T^\times A_2^{-1}$. We claim that this really gives the formula for the Hilbert adjoint. To see this, applying $\overline{T^*}$ to some y_0, we get that

$$\overline{T}^*(y_0) = A_1 T^\times A_2^{-1} y_0.$$

By definition $A_2^{-1} y_0 = g$, we have $T^\times g = f$, and $A_1 f = x_0$. Further, notice that $\overline{T}^*$ is linear — this follows because A_1 and A_2^{-1} are both conjugate-linear, and the adjoint operator is linear.

And now we have

$$\langle Tx, y_0 \rangle = g(Tx) = f(x) = \langle x, x_0 \rangle$$

(the left-hand side is by our definition of g as $\langle -, y_0 \rangle$). But the left-hand side is $\langle x, T^* y_0 \rangle$ by the original definition of the adjoint operator. Meanwhile, we have $\langle x, x_0 \rangle = \langle x, \overline{T}^* \rangle$. So $\overline{T}^* = T^*$ — meaning this formula we gave really is a formula for the Hilbert adjoint.

§12.4 Reflexivity of normed spaces

So now we've defined the adjoint operator and related it to the Hilbert adjoint, and we can start to use this idea to study function spaces again; specifically, we're going to go back to reflexivity of normed spaces. We'll begin studying this now, and then we'll use the idea of reflexivity to differentiate between spaces.

> **Definition 12.9.** We say a vector space X is *algebraically reflexive* if the canonical embedding $C: X \hookrightarrow X^{**}$ (where X^{**} represents the second algebraic dual of X) is onto.

This is just an embedding of vector spaces; it doesn't say anything about normed spaces. Recall that this maps $x \mapsto g_x$ where g_x is the map on X^* defined as $g_x(f) = f(x)$ — corresponding to evaluation at x.

Now we want to look at dual spaces rather than algebraic duals.

> **Definition 12.10.** Let X be a normed linear space, and let X' be its dual space (i.e., the space of *bounded* linear functionals). We define the map $C: X \to X''$ as $x \mapsto g_x$ where $g_x(f) = f(x)$ for all $f \in X'$.

So again, we're defining a canonical embedding; this looks the same as before, but now the domain space is different (we're taking $f \in X'$ rather than $f \in X^*$).

> **Lemma 12.11**
>
> Let X be a normed linear space. Then the map $C: X \to X''$ is well-defined (i.e., $g_x \in X''$ for all $x \in X$), and $\|g_x\| = \|x\|$.

Proof. We have to prove that g_x is linear and bounded (and that it satisfies this bound). For linearity, we want to take $\alpha x + y$; then we get

$$g_{\alpha x + y}(f) = f(\alpha x + y) = \alpha f(x) + f(y) = \alpha g_x(f) + g_y(f)$$

because f is an element of the dual space; so linearity is true.

To prove the bound, here we're going to use the corollary to Hahn–Banach again. By definition, we have

$$\|g_x\| = \sup_{f \neq 0} \frac{|g_x(f)|}{\|f\|} = \frac{|f(x)|}{\|f\|}.$$

But this is equal to $\|x\|$, as a corollary of the Hahn–Banach theorem (we used the theorem to prove the existence of f achieving equality). □

Remark 12.12. The Hahn–Banach theorem may look a bit awkward at first, but we've seen that it's really important to knowing that the adjoint operator has the same norm as the original, and here that C preserves norms. So the math that goes into studying dual spaces isn't as interesting without this theorem.

We call $C: X \to X''$ the canonical mapping of X into X''; it is linear, injective, and norm-preserving.

Remark 12.13. To see that C is injective, suppose that $g_x = g_y$; then $\|g_x - g_y\| = 0$. But this is $\|g_{x-y}\| = \|x - y\|$, which means $\|x - y\| = 0$ and therefore $x - y = 0$.

This means C is an isomorphism onto its range — it's not necessarily an isomorphism onto all of X'', but it *is* an isomorphism onto its range $\mathcal{R}(C)$. So we say that C is the *canonical embedding* of X into X''.

Remark 12.14. We say X is embeddable in Y if it is isomorphic to a subspace of Y, using an isomorphism of normed spaces. (Here we're saying X is embeddable in X''.)

Definition 12.15. We say a normed space X is *reflexive* if $C: X \to X''$ is onto.

Next time, we're going to use a bit more of this theory to say something about some different function spaces, and then we're going to talk about the uniform boundedness principle.

§13 March 20, 2024

§13.1 Reflexivity

Last class, we defined the *canonical embedding* $C: X \to X''$ as $C(x) = g_x$, where $g_x f = f(x)$ for all $f \in X'$. We proved that C is a bounded linear map satisfying $\|g_x\| = \|x\|$ for all x. This proof relied on the Hahn–Banach theorem.

Definition 13.1. If $C: X \to X''$ is onto, then we say X is *reflexive*.

(As usual, X is a normed linear space.)

Recall that for normed spaces X and Y, if Y is complete, then $B(X, Y)$ (the space of bounded linear operators from X to Y) is also complete. In particular, because $\mathbb{K}$ is complete, this means $X' = B(X, \mathbb{K})$ is always complete.

And so because $X'' = (X')'$, this means X'' is always complete. So this means if X is reflexive, then X is complete (because $X = X''$ and X'' is always complete). This gives the following statement:

Theorem 13.2

If X is reflexive, then X is complete.

Theorem 13.3

Every finite-dimensional normed linear space is reflexive.

Proof. We discussed this in the last class — the proof is that in finite dimensions, we just have $X^* = X'$ (every linear functional is bounded), and we also saw earlier that $X = X^{**}$, so this means $X = X''$. □

So we understand the dual space in finite dimensions; a natural question is what the dual space looks like for Hilbert spaces.

Theorem 13.4

Every Hilbert space H is reflexive.

Proof. To show that a Hilbert space is reflexive, we want to show that the map $C\colon H \to H''$ is onto.

First, we define a map $A\colon H' \to H$ as follows — given any $f \in H'$, we know there exists some $x_0 \in H$ such that $f(x) = \langle x, x_0 \rangle$. (This is the Riesz representation theorem.) So we define $Af = x_0$ — so A sends a functional to the element it corresponds to via the Riesz representation theorem.

Next, we claim that H' is itself a Hilbert space. We'll do this by defining an inner product on H' — we define $\langle f_1, f_2 \rangle_{H'} = \langle Af_2, Af_1 \rangle$ (we use the subscripts to denote where the inner product is being taken, since we have different inner products on H and H'). This is a natural generalization of what happens in $\mathbb{R}^n$. It's easy to verify that this is an inner product.

Now to show that C is onto, start with some $g \in H''$; we want to show there exists $x \in H$ such that $C(x) = g$. Since $g \in H''$, by the Riesz representation theorem there exists a function $f_0 \in H'$ such that $g(f) = \langle f, f_0 \rangle_{H'}$. But applying the definition of the inner product in H', this is $\langle Af_0, Af \rangle_H$. Now let $x = Af_0 \in H$ (this is sort of what we'd expect to happen).

Consider any $f \in H'$; then there exists $z_0 \in H$ such that $f(x) = \langle x, z_0 \rangle$ (here $z_0 = Af$). Then we have $g(f) = \langle f, f_0 \rangle_{H'} = \langle x, z_0 \rangle_H$. But this is precisely $f(x_0)$.

So we've shown that $g(f) = f(x)$; this means $g = Cx$, and therefore C is onto. □

§13.2 Separability and dual spaces

Theorem 13.5

If the dual space X' of a normed linear space X is separable, then X is separable.

The converse is not true — the dual space of a separable space need not be separable. However, it is true if we also have reflexivity.

Corollary 13.6

If X is reflexive *and* separable, then X' is also separable.

Proof. By reflexivity we have $X \cong X''$, which means X'' is separable, and therefore by this theorem X' must also be separable. □

This gives us one way of proving that a space is *not* reflexive — if X is separable and X' is *not* separable, then X cannot be reflexive (because of the above corollary). As an example:

Claim 13.7 — The space ℓ^1 is separable.

Proof. To see that ℓ^1 is separable, a space is separable if it has a countable dense subset, and a countable dense subset for ℓ^1 is the space of eventually 0 sequences with rational coefficients — we say a sequence x is eventually zero if there exists N (depending on x) such that $x_n = 0$ for all $n \geq N$. □

On the other hand, the dual of ℓ^1 is ℓ^∞.

Claim 13.8 — The space ℓ^∞ is not separable.

Proof. Let $S \subseteq \ell^\infty$ be the set of $\{0,1\}$-sequences. For each $y \in S$, let $\widehat{y} \in [0,1]$ be the number whose binary representation corresponds to y — i.e.,

$$\widehat{y} = \frac{y_1}{2} + \frac{y_2}{4} + \frac{y_3}{8} + \cdots.$$

Then because $[0,1]$ is uncountable and every element in it has a binary representation, S is uncountable. But the distance between any two elements in S is 1.

This means ℓ^∞ can't be the closure of a countable set, so it's not separable. □

So this implies that ℓ^1 is *not* reflexive. (We proved this a different way earlier just by noting that the dual space of ℓ^∞ is not ℓ^1.)

Now we'll prove this theorem. For this, we need the following lemma.

Lemma 13.9

Let $Y \subsetneq X$ be a proper closed subspace of the normed linear space X, and let $x_0 \in X \setminus Y$ and $\delta = \inf_{y \in Y} \|x_0 - y\|$ be the distance from x_0 to Y. Then there exists a functional $\widetilde{f} \in X'$ such that $\|\widetilde{f}\| = 1$ and such that $\widetilde{f}(y) = 0$ for all $y \in Y$ and $\widetilde{f}(x_0) = \delta$.

We're going to use this lemma to prove the theorem. It should immediately bring to mind the Hahn–Banach theorem, because we want to prove there exists an element of Y' that's 0 on Y and is equal to the distance to Y at x_0. So naturally, we're going to first define our function on a subspace, and then use Hahn–Banach to extend it to the entire space.

Proof. Let $Z = \operatorname{Span}(Y \cup \{x_0\})$, so every element of Z can be written as $y + \alpha x_0$. Define $f \colon Z \to \mathbb{K}$ as

$$f(y + \alpha x_0) = \alpha\delta,$$

where δ is as above. It's immediate that $f(y) = 0$ for all $y \in Y$ (because then $\alpha = 0$), and that $f(x_0) = f(0 + 1 \cdot x_0) = \delta$. It's also clear that f is linear.

We now need to check that $\|f\| = 1$. To do so, we have

$$\|f\| = \sup_{z \neq 0} \frac{f(z)}{\|z\|} = \sup_{y + \alpha x_0 \neq 0} \frac{|\alpha\delta|}{\|y + \alpha x_0\|} = \sup \frac{\delta}{\|x_0 - (-1/\alpha)y\|}.$$

Now because $\delta = \inf_{z \in Y} \|x_0 - z\|$, we must have $\|x_0 - (-1/\alpha)y\| \geq \delta$; and so we get that $\|f\| \leq 1$.

To prove the other direction, let $(y_n)_{n \in \mathbb{N}} \subseteq Y$ be a sequence such that $\|x_0 - y_n\| \to \delta$ as $n \to \infty$. Then we have

$$\|f\| = \sup \frac{|f(z)|}{\|z\|} \geq \frac{|f(y_n - x_0)|}{\|y_n - x_0\|} == \frac{|\delta|}{\|y_n - x_0\|}$$

(note that $f(y_n - x_0) = -\delta$ for any $y_n \in Y$). And since $\|y_n - x_0\| \to \delta$, the right-hand side goes to 1, and so we get $\|f\| \geq 1$.

Together, this means $\|f\| = 1$.

And so we've now defined $f \colon Z \to \mathbb{K}$ with all the desired properties, and we can use Hahn–Banach to extend f to $\widetilde{f} \in X'$ satisfying $f|_Z = \widetilde{f}|_Z$ (which means we still have $\widetilde{f}(y) = 0$ for $y \in Y$ and $\widetilde{f}(x_0) = \delta$) and $\|\widetilde{f}\| = \|f\| = 1$. □

Now we're going to use this to prove our theorem — that if the dual space is separable, then the space itself is separable.

Proof of theorem. Define the set

$$U' = \{f \in X' \mid \|f\| = 1\} \subseteq X'.$$

Because X' is separable, so is U' — in particular, this means U' contains a countable dense subset; let $(f_n)_{n\in\mathbb{N}}$ be a countable dense subset of U'.

In particular, we have

$$\|f_n\| = \sup_{\|x\|=1} |f_n(x)| = 1$$

for all n (by the definition of U'). So there exists a sequence of points $(x_n)_{n\in\mathbb{N}}$ inside X such that $\|x_n\| = 1$ and $|f_n(x_n)| \to 1$ — in particular, we can ensure $|f_n(x_n)| \geq \frac{1}{2}$.

Now let $Y = \mathrm{Cl}(\mathrm{Span}\{x_1, x_2, \ldots\})$ be defined as the closure of the span of x_1, x_2, ...

Claim 13.10 — The set Y is separable.

Proof. We can consider the set of linear combinations of the x_i with *rational* coefficients; Y is a closure of this set as well. (And this set is countable.) □

Claim 13.11 — We have $Y = X$.

(This is the more important claim; once we've proven it, this completes the proof because Y is separable.)

Proof. Assume not. Then by the lemma from earlier, there exists $\widetilde{f} \in X'$ satisfying $\widetilde{f}|_Y = 0$ and with $\|\widetilde{f}\| = 1$. This means $\widetilde{f}(x_n) = 0$ for all n, because each x_n is in Y.

And now $\left\|\widetilde{f}\right\| = 1$; in particular, this means $\widetilde{f} \in U'$. And so because $(f_n)_{n\in\mathbb{N}}$ is a countable dense subset of U', there should be some subsequence of f_n converging to $\widetilde{f}$.

But for all n, we have

$$\frac{1}{2} \leq |f_n(x_n)| = |f_n(x) - \widetilde{f}(x_n)| = |(f_n - \widetilde{f})(x_n)| \leq \|f_n - \widetilde{f}\| \, \|x_n\| \, .$$

But we chose x_n such that $\|x_n\| = 1$, and we can choose a subsequence of f_n for which $\|f_n - \widetilde{f}\| \to 0$; this is a contradiction (because the right-hand side then goes to 0, while the left-hand side is $\frac{1}{2}$). □

□

§13.3 Uniform boundedness theorem

Definition 13.12. A set S in a metric space M is *nowhere dense* if the interior of its closure is empty.

Theorem 13.13 (Baire category theorem)
A complete metric space cannot be written as a countable union of nowhere dense sets.

Proof. Let (M,d) be a complete metric space, and assume for contradiction that there exists a countable sequence of sets $(M_i)_{i\in\mathbb{N}}$ such that $M = \bigcup_{i\in\mathbb{N}} M_i$ and $\text{int}(\overline{M_i}) = \emptyset$ for all i.

First, there exists an open ball $B_1 = B(p_1, \delta_1) \subseteq M \setminus \overline{M_1}$ — this is because $M \setminus \overline{M_1}$ is open and nonempty (of course the interior of M is M itself, which is nonempty).

Now we can define $B_2 = B(p_2, \delta_2) \subseteq B(\rho_1, \delta_1/2) \setminus \overline{M_2}$. (Note that B_2 is in both $\overline{M_1}^{\mathsf{c}}$ and $\overline{M_2}^{\mathsf{c}}$.)

If we keep on doing this, then we get a sequence of balls $B_k = B(p_k, \delta_k) \subseteq M \setminus (\bigcup_{i=1}^{k} \overline{M_i})$, such that $B_k \subseteq B_{k-1}(p_{k-1}, \delta_{k-1}/(k-1))$. This implies in particular that the sequence $(p_k)_{k\in\mathbb{N}}$ is Cauchy. And because M is complete, this means (p_k) converges to some point $p \in M$.

But then p can't be in $\overline{M_i}$ for any i; this is a contradiction, because we assumed $\bigcup_{i\in\mathbb{N}} \overline{M_i} = M$. □

Now we'll use this to prove the uniform boundedness principle.

Theorem 13.14 (Uniform boundedness principle)

Let X be a Banach space, Y a normed linear space, and $\mathcal{F} \subseteq B(X,Y)$ a family of bounded linear operators. If for all $x \in X$ we have $\sup\{\|Tx\| \mid T \in \mathcal{F}\}$ is finite, then $\sup\{\|T\| \mid T \in \mathcal{F}\}$ is finite.

Notice that here $\mathcal{F}$ can have any size. So this says that if at each point x there's a bound, then the entire family of *operators* is bounded. This is not immediately clear — in particular, this doesn't hold if X is not a Banach space.

Proof. Let $A_n = \{x \in X \mid \|Tx\| \le n \text{ for all } T \in \mathcal{F}\}$. Because $\sup\{\|Tx\| \mid T \in \mathcal{F}\}$ is finite for every x, we have $\bigcup_{n\in\mathbb{N}} A_n = X$.

We claim that A_n is closed. To see this, we'll show it's sequentially closed; let $(x_m)_{m\in\mathbb{N}}$ be a sequence in A_n converging to some point $x \in X$, so that we want to show $x \in A_n$ as well. But by continuity of the norm (and of T) we have $\|Tx_m\| \to \|Tx\|$ for all $T \in \mathcal{F}$, and in particular this means $\|Tx\| \le n$ (since $\|Tx_m\| \le n$ for all m). This is true for all $T \in \mathcal{F}$, so that implies $x \in A_n$.

Now because X is complete, we can apply the Baire category theorem — there must exist n such that $\text{int}(\overline{A_n}) = \text{int}(A_n)$ is nonempty. (We're applying the fact that we've written X as a countable union of sets, so they can't *all* be nowhere dense; and so there is some n_0 such that A_n is not nowhere dense, which is exactly the above condition.)

So there exists a closed ball $\overline{\mathbb{B}}(x_0, \varepsilon) \subseteq A$ contained in A_{n_0} (with $\varepsilon > 0$).

Now for any T, we know

$$\|T\| = \sup_{\|x\|=1} \|Tx\|.$$

But if $\|x\| = 1$, then for any $T \in \mathcal{F}$ we can write

$$|Tx| = \frac{1}{\varepsilon}\|T(\varepsilon x)\| = \frac{1}{\varepsilon}\|T(\varepsilon x + x_0 - x_0)\| = \frac{1}{\varepsilon}\|T(\varepsilon x + x_0) - T(x_0)\| \le \frac{1}{\varepsilon}(\|T(\varepsilon x + x_0)\| + \|Tx_0\|).$$

And now, because $\varepsilon x + x_0$ and x_0 are both in the ball $\overline{\mathbb{B}}(x_0, \varepsilon) \subseteq A_n$, this means both of these norms are at most n; and therefore we get $\|Tx\| \le 2n_0/\varepsilon$.

So we've shown that for any $\|x\| = 1$ we have $\|Tx_0\| \le 2n_0/\varepsilon$; this in particular implies $\|T\| \le 2n_0/\varepsilon$ for all $T \in \mathcal{F}$, and therefore $\sup \|T\|$ is bounded. □

§14 April 1, 2024

Today we'll talk about the open mapping theorem.

§14.1 Open mapping theorem

Theorem 14.1 (Open mapping theorem)

Let X and Y be Banach spaces, and let $T \in B(X, Y)$ (so T is a bounded linear operator $X \to Y$). If T is surjective (i.e., T maps X onto Y), then T is an open map (i.e., $T(G)$ is open in Y for each open set G in X).

Corollary 14.2

Let X and Y be Banach spaces, and let $T \in B(X, Y)$ be bijective. Then its set-theoretic inverse T^{-1} satisfies $T^{-1} \in B(Y, X)$.

(We say 'set-theoretic inverse' because we can define an inverse of any map — we define $T^{-1}(y) = \{x \in X \mid Tx = y\}$, and when T is bijective this is a single-valued map. You can use injectivity to show that it's linear, but it's hard to prove boundedness without the open mapping theorem.)

Proof of corollary. By the linearity of T we have

$$T(\alpha T^{-1}(x) + T^{-1}(y)) = \alpha T(T^{-1}(x)) + T(T^{-1}(y)) = \alpha x + y$$

for all $x, y \in Y$ and $\alpha \in \mathbb{K}$; and since T is injective, this means T^{-1} is linear (as stated earlier).

Next, we want to show that T^{-1} is bounded. Recall that T^{-1} is bounded if and only if it is continuous, which is true if and only if $(T^{-1})^{-1}(G)$ is open for every open $G \subseteq X$. But $(T^{-1})^{-1}(G)$ is just $T(G)$, and we know this statement is true because T is an open map (by the open mapping theorem, which applies since T is onto). $\square$

So injectivity gives us linearity, and surjectivity tells us that T is an open map, which means T^{-1} is continuous and therefore bounded.

Next, we'll state a theorem that we'll prove in order to prove the open mapping theorem.

Theorem 14.3

Let X and Y be Banach spaces, and suppose that $A \in B(X, Y)$ is onto. Then there exists $\delta > 0$ such that $\delta \cdot \mathbb{B}_Y \subseteq A(\mathbb{B}_X)$ (where $\mathbb{B}_X$ and $\mathbb{B}_Y$ are the balls of radius 1 centered at the points 0 in X and Y, respectively).

This immediately implies the open mapping theorem:

Proof of open mapping theorem. Suppose that the above theorem holds. Consider any open $G \subseteq X$. We want to show that $A(G)$ is open, which means we want to show that for every $x_0 \in X$, there is an open ball centered at $A(x_0)$ contained in $A(G)$.

To see this, let $G' = G - x_0$ and let $t > 0$ be such that $\mathbb{B}(x_0, t) \subseteq G$, which means $\mathbb{B}(0, t) \subseteq G'$, or in other words that $t \cdot \mathbb{B}_X \subseteq G'$ (this is just a remark on the notation — $\mathbb{B}_X$ denotes the unit ball $\mathbb{B}(0, 1)$ in X). Then $A(G')$ contains $A(t\mathbb{B}_x) = t \cdot A(\mathbb{B}_x)$, which contains $t\delta \cdot \mathbb{B}_Y$ (by the above theorem).

Then because A is linear, this means $A(G) = A(G' + x_0) = A(G') + A(x_0) \supseteq A(x_0) + t\delta \cdot \mathbb{B}_Y$, which proves there is an open ball centered at $A(x_0)$ contained in $A(G)$, proving the open mapping theorem. $\square$

Now it remains to prove this second theorem. We'll do this by iterating the following lemma.

Lemma 14.4

Let X and Y be Banach spaces, and let $A \in B(X, Y)$ be onto. Then there exists $d > 0$ such that the following holds: Given $\varepsilon > 0$ and $y \in Y$, there exists $x \in X$ such that $\|Ax - y\| < \varepsilon$ and $\|x\| < \|y\| / d$.

So we have x, and a bounded linear map A sending X onto Y;this lemma says that given any $x \in X$, we can find a point $x \in X$ with Ax close to y and with $\|x\|$ bounded in terms of $\|y\|$.

Proof. First, since A is onto, we can write $Y = \bigcup_{k \in \mathbb{N}} A(k\mathbb{B}_X)$. (This is because every $z \in Y$ can be written as Ax for some x, and $x \in k\mathbb{B}_X$ for some k.)

Now we use the Baire category theorem, which tells us that Y cannot be a countable union of nowhere dense sets (sets whose closures have empty interior). So by the Baire category theorem, there exists k_0 such that $\text{int}(\overline{A(k_0 \cdot \mathbb{B}_X)})$ is nonempty (since the theorem says it can't be true that for every k this interior is empty). This means there exists some open ball contained in this closure — so there exist $y_0 \in Y$ and $r > 0$ such that $\mathbb{B}(y_0, r) \subseteq \overline{A(k_0 \cdot \mathbb{B}_X)} \subseteq Y$.

Now let $(x'_n)_{n \in \mathbb{N}}$ be a sequence in $k_0 \cdot \mathbb{B}_X$ such that $Ax'_n \to y_0$. Now for any z with $\|z\| < r$, we can let $(x''_n)_{n \in \mathbb{N}}$ be another sequence in $k_0 \cdot \mathbb{B}_X$ such that $Ax''_n \to z + y_0$. (This is because $\mathbb{B}(y_0, r)$ is contained in our closure, so if $\|z\| < r$ then $z + y_0$ is also contained in our closure.) We're eventually going to take z to be a rescaling of y.

Define $x_n = x''_n - x'_n$ for all $n \in \mathbb{N}$, so that $Ax_n \to z$. Note that $\|x_n\| \leq \|x''_n\| + \|x_n\| \leq 2k_0$ for all $n \in \mathbb{N}$.

To recap, so far we've expressed Y as a countable union, which means one has to have a closure with nonempty interior, which means it contains the closure of an on open ball. And then for z smaller than the radius of that ball, we can find two sequences of points inside $k_0\mathbb{B}_X$, one with images converging to y_0 and the other to $z + y_0$ (since y_0 and $z + y_0$ are both inside our closure).

Now given any $y \in Y$, let

$$z = \frac{r}{2} \cdot \frac{y}{\|y\|}.$$

This means $\|z\| = r/2 < r$. And this implies there exists a sequence $(x_n)_{n \in \mathbb{N}}$ in $2k_0\mathbb{B}_X$ such that $Ax_n \to z$. We want to scale back to get y, so we define

$$w_n = \frac{2\|y\|}{r} \cdot x_n \in X$$

for each $n \in \mathbb{N}$; then we have

$$\|w_n\| = \frac{2\|y\|}{r} \|x_n\| \leq \frac{4k_0}{r} \cdot \|y\|.$$

And by the linearity of A we have

$$Aw_n = \frac{2\|y\|}{r} \cdot Ax_n \to \frac{2\|y\|}{r} \cdot z = y.$$

Now let $d = r/4k_0$; and we can let $x = w_n$ for any n large enough such that $\|Aw_n - y\| < \varepsilon$ (since $Aw_n \to y$), which completes the proof. □

Now we want to iterate the lemma to prove our second theorem.

Proof of second theorem. Recall that X and Y are Banach spaces and $A \in B(X, Y)$ is onto; we fix d from the above lemma, and fix $y \in d \cdot \mathbb{B}_Y$ (i.e., the ball of radius d). Applying the lemma with $\varepsilon = d/2$, we can produce $x_1 \in X$ satisfying $\|y - Ax_1\| < \varepsilon = d/2$ and $\|x_1\| < \|y\| / d < 1$.

Then applying the lemma again to the point $y - Ax_1$ (in place of y) with $\varepsilon = d/2^2$, we can produce $x_2 \in X$ satisfying $\|(y - Ax_1) - Ax_2\| < d/2^2$ and $\|x_2\| < \|y - Ax_1\| / d < 1/2$.

Continuing inductively, we can produce a seqeunce of points x_1, x_2, ... satisfying

$$\|y - Ax_1 - Ax_2 - \cdots - Ax_n\| < \frac{d}{2^n} \text{ and } \|x_n\| < \frac{1}{2^{n-1}}$$

for each n (in particular, this means $\|x_k\| < 1/2^{k-1}$ for each $1 \le k \le n$).

Now define $v_n = x_1 + x_2 + \cdots + x_n$. Then $(v_n)_{n\in\mathbb{N}}$ is a Cauchy sequence, because

$$\|v_n - v_m\| = \|x_{m+1} + \cdots + x_n\| \le \frac{1}{2^m} + \cdots + \frac{1}{2^{n-1}} \to 0.$$

Because X is complete, this means v_n converges to some point $x \in X$. And we have $\|x\| \le \sum_{k=1}^{\infty} 1/2^{k-1} = 2$, which implies $x \in 2 \cdot \mathbb{B}_X$.

And we have $\|y - Av_n\| < d/2^n$ for each n (by the linearity of A we have $Ax_1 + \cdots + Ax_n = Av_n$), which means $Av_n \to y$. But A is bounded and therefore continuous, so since $v_n \to x$ we have $Av_n \to Ax$; this implies $Ax = y$.

So we've proved that each point $y \in d\mathbb{B}_Y$ can be written as Ax for some $x \in 2\mathbb{B}_X$; this means $A(2\mathbb{B}_X) \supseteq d\mathbb{B}_Y$, and therefore $A(\mathbb{B}_X) \supseteq \frac{d}{2}\mathbb{B}_Y$; this proves the theorem with $\delta = d/2$. □

Note that here, we importantly used the Baire category theorem, and that X is a Banach space.

§14.2 The closed graph theorem

Next we'll see a short consequence of this theorem — that the graph of T is closed if and only if T is bounded (in the same setup above).

Definition 14.5. Let X and Y be normed linear spaces, and let $T\colon X \to Y$. We define the *graph* of T as $\operatorname{graph}(T) = \{(x, Tx) \mid x \in X\}$, viewed as a subspace of $X \times Y$.

Remark 14.6. The space $X \times Y$ is a vector space, with a natural norm — we can define $\|(x, y)\| = \|x\|_X + \|y\|_Y$ (for $x \in X$ and $y \in Y$). If X and Y are complete, then so is $X \times Y$.

Theorem 14.7 (Closed graph theorem)

Let X and Y be Banach spaces, and let $T\colon X \to Y$ be linear. Then T is bounded (i.e., $T \in B(X, Y)$) if and only if $\operatorname{graph}(T)$ is closed in $X \times Y$.

Proof. The forwards direction is immediate when we think about what it means for the graph to be closed, so we'll prove it first — suppose that T is bounded. To prove that $\operatorname{graph}(T)$ is closed, we'll use the concept of being sequentially closed — we want to show that if we have a sequence $(x_n, Tx_n) \to (x, y)$, then $y = Tx$.

Since $(x_n, Tx_n) \to (x, y)$, this means

$$\|x_n - x\|_X + \|Tx_n - y\|_Y \to 0,$$

which in particular means $\|x_n - x\|_X \to 0$, and therefore $\|T(x_n - x)\| = \|Tx_n - Tx\| \to 0$ (by the continuity of T — recall that if T is a bounded linear operator then it is continuous). But now we have both $Tx_n \to Tx$ and $Tx_n \to y$, which means $y = Tx$.

Now we'll do the backwards direction — suppose that $\operatorname{graph}(T)$ is closed in $X \times Y$. Let $P_1\colon \operatorname{graph}(T) \to X$ be the map $P_1(x, Tx) = x$, and $P_2\colon \operatorname{graph}(T) \to Y$ be the map $P_2(x, Tx) = Tx$. Both of these are bounded linear maps (with $\|P_1\|, \|P_2\| \leq 1$ — this is because $\|(x, Tx)\| = \|x\| + \|Tx\| \leq \|x\|, \|Tx\|$), which in particular means they are continuous.

And P_1 is also bijective, which means we can apply the open mapping theorem, or rather, its corollary — by this corollary we know that P_1^{-1} is a bounded linear map $X \to \text{graph}(T)$, i.e., $P_1^{-1} \in B(X, \text{graph}(T))$. Then notice that $T = P_2 P_1^{-1}$; this means T is a composition of bounded linear maps, and is therefore bounded. $\square$

Remark 14.8. In the forwards direction, we used very little — we didn't need that X and Y were Banach spaces, we just used that T is a bounded linear operator. So for *any* bounded linear operator, its graph is closed. But for the other direction we do need them to be Banach spaces.

Remark 14.9. This proof seems very simple; but that's because most of the work goes into the proof of the open mapping theorem.

§15 April 3, 2024

Today we'll talk about convergence, specifically weak* convergence.

§15.1 Weak convergence

Often in analysis or applied math, we have a family of functions or functionals or operators, and we watn to know if the family converges to some limit. Coming up with a candidate for the limit, or finding a 'weaker' limit (something that functions like a limit, even if it isn't actually a limit) is often enough. And weak convergence is a natural definition coming from our theory about the dual space.

Definition 15.1. A sequence $(x_n)_{n \in \mathbb{N}}$ in X is *strongly convergent* if it is convergent in norm — i.e., there exists $x \in X$ such that $\|x_n - x\| \to 0$ as $n \to \infty$.

This is what we've talked about in terms of convergence so far. We denote this by $x_n \to x$ (so if we just write $\to$, this means convergence in norm), or by $\lim_{n \to \infty} x_n = x$.

Definition 15.2. A sequence $(x_n)_{n \in \mathbb{N}}$ is *weakly convergent* if there exists $x \in X$ such that $f(x_n) \to f(x)$ for all $f \in X'$.

Here $f(x_n) \to f(x)$ refers to (strong) convergence in $\mathbb{R}$. We dentoe this as $x_n \xrightarrow{w} x$, and we say that x is the *weak limit* of x_n.

Lemma 15.3

Let $x_n \xrightarrow{w} x$ in a normed linear space X. Then we have the following:

(a) The weak limit x of (x_n) is unique.

(b) Every subsequence (x_{n_j}) of (x_n) also weakly converges to x.

(c) There exists a constant C such that $\|x_n\| \leq C$ for all n (i.e., $\|x_n\|$ is bounded).

These sound simple, but there's work that goes into them — we'll need lots of the big theorems we've proven. In particular, (a) is a consequence of Hahn–Banach and (c) of the uniform boundedness theorem. So these seem like obvious results — they're what you'd hope would be true — but they're not simple to prove.

Proof. For (a), suppose that $x_n \xrightarrow{w} y$ as well (so we want to show $x = y$). Then for every $f \in X'$, we have $f(x_n) \to f(y)$, which means $f(x) = f(y)$ (because $f(x_n)$ also converges to $f(x)$, and this is strong convergence in $\mathbb{R}$, so $(f(x_n))$ can only have one limit).

Now because f is linear, this means $f(x - y) = 0$ — and this is true for *all* $f \in X'$. But it was a corollary of Hahn–Banach that if $f(z) = 0$ for all $f \in X'$, then $z = 0$. So we have $x - y = 0$, which means $x = y$; so the (weak) limit is unique.

Remark 15.4. To remind us of this corollary to Hahn–Banach, the first corollary of Hahn–Banach we proved was that given a normed linear space X and some $x \in X$, we can find $f \in X'$ such that $\|f\| = 1$ and $f(x) = \|x\|$. And as a corollary to this, we can write

$$\|x\| = \sup_{f \neq 0} \frac{|f(x)|}{\|f\|}.$$

(This equality has two sides; one is immediate by definition, and the other comes from the first corollary.) And this immediately implies that if $f(x) = 0$ for all f, then $\|x\| = 0$, so $x = 0$.

Now we'll prove (b). For $f \in X'$, we have $f(x_n) \to f(x)$ in $\mathbb{R}$ (or more generally, $\mathbb{K}$); this means in particular that the subsequence $(f(x_{n_j}))$ must also converge to $f(x)$. And this is true for every $f \in X'$, which means $x_{n_j} \xrightarrow{w} x$. (The point is that we're using strong convergence in $\mathbb{R}$ to come back to weak convergence of the subsequence.)

Now we'll prove (c), which says that every weakly convergent sequence is bounded; this is again an important result in a lot of what follows. To prove this, for every $f \in X'$, we know that $(f(x_n))$ is convergent in the field $\mathbb{K}$, which means it is bounded — i.e., $|f(x_n)| \leq c_f$ for all n, for some constant c_f depending on f. And this is exactly the setup for the uniform boundedness theorem — let $g_n \in X''$ be defined as $g_n = Cx_n$, where C is the canonical embedding $C\colon X \mapsto X''$. This means $g_n(f) = f(x_n)$ for all $f \in X'$. Notice that for all $n \in \mathbb{N}$, $|g_n(f)| = |f(x_n)| \leq c_f$ is bounded by a constant (depending on f); equivalently, the family of bounded linear operators $\mathcal{F} = \{g_n \mid n \in \mathbb{N}\} \subseteq X''$ is bounded on every $f \in X'$ (note that $g_n \in X''$, which means it maps $X' \to \mathbb{K}$).

We're going to use the uniform boundedness theorem, which requires our space to be complete; note that X' is complete, because it's a dual space (and $\mathbb{K}$ is complete). So we've now set things up to apply the uniform boundedness theorem.

Now by the uniform boundedness theorem, this implies $\|g_n\|$ is bounded — i.e., there exists c such that $\|g_n\| \leq c$ for all $n \in \mathbb{N}$. And because $g_n = Cx_n$ and C is norm-preserving, we have $\|g_n\| = \|x_n\|$; this means $\|x_n\| \leq c$ for all $n \in \mathbb{N}$. And that is exactly what we wanted to prove. □

So these seem like very natural results, but they take the work we've put in (with the uniform boundedness theorem and Hahn–Banach) to prove.

Theorem 15.5

Let $(x_n)_{n \in \mathbb{N}}$ be a sequence in a normed linear space X.

(a) If $x_n \to x$, then $x_n \xrightarrow{w} x$. (In other words, strong convergence implies weak convergence.)

(b) The converse is false in general.

(c) If $\dim X < \infty$, then the converse is true — if $x_n \xrightarrow{w} x$ then $x_n \to x$.

The first statement is really important — if you're trying to prove some results and you can prove that a sequence strongly converges, and you can prove what its weak limit *is*, then you know they're the same. This is how a lot of theory will look — you prove that the strong limit *exists* and you prove what exactly the weak limit is, and that tells you it's actually the strong limit. This is a very important use of weak convergence.

Proof of (a). For (a), suppose that $x_n \to x$; this means $\|x_n - x\| \to 0$. And this means for all functionals $f \in X'$, we have $|f(x_n) - f(x)| = |f(x_n - x)| \le \|f\| \, \|x_n - x\| \to 0$ (because f is linear, and we know $\|x_n - x\| \to 0$ and $\|f\|$ is finite). So this proves $f(x_n) \to f(x)$. And this is true for all $f \in X'$, which means $x_n \xrightarrow{w} x$. □

Proof of (b). For (b), we'll see a counterexample — we'll come up with a space X and a sequence that weakly converges but doesn't strongly converge. Let H be a Hilbert space, and let $(e_n)_{n\in\mathbb{N}}$ be an orthonormal sequence in H. We claim that $e_n \xrightarrow{w} 0$. By the Riesz representation theorem, every $f \in H'$ can be written as $f(x) = \langle x, z\rangle$ for some $z \in H$; this in particular means $f(e_n) = \langle e_n, z\rangle$ for all n. And by Bessel's inequality, we have

$$\sum_{n=1}^{\infty} |\langle e_n, z\rangle|^2 \le \|z\|^2 .$$

In particular, this means the series on the left-hand side is convergent, so we must have $\langle e_n, z\rangle \to 0$ as $n \to \infty$. And that means $f(e_n) = \langle e_n, z\rangle \to 0$.

And since this is true for every $f \in H'$, this means $e_n \xrightarrow{w} 0$ (because of course $f(0) = 0$).

But (e_n) doesn't strongly converge to 0 — it's an orthonormal sequence, so $\|e_n\| = 1$ for all n, which means e_n can't strongly converge to 0. So we've built a sequence that has a weak limit of 0, but does not strongly converge to 0. □

Proof of (c). The next thing we want to do is prove (c) — let $\{e_1, \dots, e_k\}$ be a basis for X (here X is finite-dimensional, and we're letting $\dim X = k$). This means we can write

$$x_n = \alpha_1^n e_1 + \cdots + \alpha_k^n e_k$$

for each $n \in \mathbb{N}$, and we can similarly write

$$x = \alpha_1 e_1 + \cdots + \alpha_k e_k.$$

Because $x_n \xrightarrow{w} x$, for every $f \in X'$ we have $f(x_n) \to f(x)$. Now we can consider the dual basis $\{f_1, \dots, f_k\}$ of X', where

$$f_j(e_\ell) = \delta_{j\ell} = \begin{cases} 1 & \text{if } j = \ell \\ 0 & \text{otherwise.} \end{cases}$$

Then for each j we have $f_j(x_n) \to f_j(x)$, i.e., $\alpha_j^n \to \alpha_j$; and this is true for every $j \in [k]$, which means

$$\|x_n - x\| = \left\| \sum (\alpha_j^n - \alpha_j) e_j \right\| \le \sum |\alpha_j^n - \alpha_j| \, \|e_j\| \to 0,$$

which means $x_n \to x$. □

§15.2 Convergence of operators

Definition 15.6. Let X and Y be normed linear spaces, and let $(T_n)_{n\in\mathbb{N}}$ be a sequence in $B(X, Y)$. We say (T_n) is:

(i) *uniformly operator convergent* if there exists $T\colon X \to Y$ such that $\|T_n - T\|_{B(X,Y)} \to 0$.

(ii) *strongly operator convergent* if there exists $T\colon X \to Y$ such that $\|T_n x - Tx\|_Y \to 0$ for every $x \in X$ (equivalently, $T_n x \to Tx$ in Y for every $x \in X$).

(iii) *weakly operator convergent* if there exists $T\colon X \to Y$ such that $T_n x \xrightarrow{w} Tx$ for every $x \in X$ (i.e., $f(T_n x) \to f(Tx)$ for every $f \in Y'$).

Remark 15.7. Note that these definitions (at least (ii) and (iii)) don't require that $T \in B(X, Y)$. ((i) does immediately imply that $T \in B(X, Y)$, but (ii) and (ii) don't.)

§15.2.1 Implications between types of convergence

Note that (i) implies (ii), and (ii) implies (iii). To see that (i) implies (ii), we have $\|T_n x - Tx\| = \|(T_n - T)x\| \leq \|T_n - T\| \|x\|$. And (ii) implies (iii) because strong convergence implies weak convergence.

Next we'll see some counterexamples for the reverse implications.

Example 15.8
Let $T_n \colon \ell^2 \to \ell^2$ be defined by

$$T_n(x) = (\underbrace{0, \ldots, 0}_{n}, x_{n+1}, x_{n+2}, \ldots)$$

(where $x = (x_j)_{n \in \mathbb{N}}$). Then T_n is strongly but not uniformly operator convergent to 0.

It's clear that each T_n is bounded (we have $\|T_n x\| \leq \|x\|$, since replacing some coordinates with 0 can only make the norm smaller) and linear. We have $T_n x \to 0 = 0x$ for all $x \in \mathbb{N}$; but $\|T_n - 0\| = \|T_n\| = 1$ for all n. So this shows that strong operator convergence does not imply (uniform) operator convergence — (ii) does not imply (i).

Example 15.9
Let $T_n \colon \ell^2 \to \ell^2$ be defined by

$$T_n(x) = (\underbrace{0, \ldots, 0}, x_1, x_2, \ldots)$$

(so we're shifting x n places to the right). Then T_n is weakly operator convergent but not strongly operator convergent to 0.

Proof. It's clear that $\|T_n\| = 1$ and that T_n is linear, so $T_n \in B(\ell^2, \ell^2)$. Because ℓ^2 is a Hilbert space, we can write each $f \in (\ell^2)'$ as $f(x) = \langle x, z \rangle = \sum_j x_j \overline{z_j}$ for some $z \in \ell^2$. This means $f(T_n x) = \langle T_n x, z \rangle = \sum_{j=n+1}^{\infty} x_{j-n} \overline{z_j}$. By Cauchy–Schwarz, we have

$$|f(T_n x)|^2 = |\langle T_n x, z \rangle|^2 \leq \|T_n x\|^2 \|z\|^2 = \sum_{i=n+1}^{\infty} |x_{i-n}|^2 \sum_{i=n+1}^{\infty} |\overline{z_i}|^2 .$$

But the first term is $\|x\|_{\ell^2}^2$, while the second term converges to 0 as $n \to \infty$. This shows that $f(T_n x) \to 0$ as $n \to \infty$. And this is true for every $f \in X'$, which implies $T_n x \xrightarrow{w} 0$, so T_n is weakly operator convergent to 0.

But T_n doesn't strongly converge to 0 (or to anything) — for example, we can see that $\|T_n(1, 0, \ldots)\| = 1$ for all n.

So this shows that weak operator convergence doesn't imply strong operator convergence (i.e., (iii) does not imply (ii)). $\square$

§15.3 Functionals and weak* convergence

Note that for a sequence of linear functionals $(f_n)_{n \in \mathbb{N}}$ in $X' = B(X, \mathbb{K})$, because $\mathbb{K}$ is finite-dimensional, strong and weak operator (functional) convergence are the same.

Definition 15.10. Let $(f_n)_{n\in\mathbb{N}}$ be a sequence in X' (for a normed linear space X).

- We say $(f_n)_{n\in\mathbb{N}}$ is *strongly convergent* if there exists $f \in X'$ such that $\|f_n - f\| \to 0$; this is denoted $f_n \to f$.
- We say $(f_n)_{n\in\mathbb{N}}$ is weak* convergent if there exists $f \in X'$ such that $f_n(x) \to f(x)$ for all $x \in X$; this is denoted $f_n \xrightarrow{w^*} f$.

Strong convergence isn't a new concept — it's just the usual notion of strong convergence applied to X'. The new notion here is weak* convergence. It's kind of the dual idea to weak convergence — in weak convergence we have a sequence in X, and it weakly converges if $f(x_n) \to f(x)$ for all $f \in X'$. Here we have a sequence in the dual space, and we say $f_n(x) \to f(x)$ for every $x \in X$. So we're going back to the original space, instead of ahead to the dual space.

This actually ends up being a more valuable idea than weak convergence (for functionals).

Remark 15.11. Weak convergence (in X') implies weak* convergence — so weak* convergence is less strong than weak convergence. (This is because of the canonical embedding — you can prove it by considering $C\colon X \hookrightarrow X''$. If we have weak convergence of (f_n), then this means for every $g \in X''$ we have $g(f_n) \to g(f)$; and taking $g = C(x)$ for $x \in X$ gives weak* convergence.)

Lemma 15.12

Let $(T_n)_{n\in\mathbb{N}}$ be a sequence in $B(X, Y)$, where X is a Banach space and Y is a normed linear space. If T_n is strongly operator convergent to T (i.e., $T_n x \to Tx$ for each $x \in X$), then $T \in B(X, Y)$.

Proof. The idea is to use the uniform boundedness theorem. We have that $T_n x$ is bounded (over n) for each $x \in X$ (since it converges to some Tx), and X is complete (by our hypothesis). So the uniform boundedness theorem applies, and $(\|T_n\|)$ is bounded (over $n \in \mathbb{N}$). Now suppose that $\|T_n\| \le c$ for all n (where c is some constant).

We want to show that T is bounded. To see this, for every $x \in X$ we have

$$\|Tx\| = \|(T_n - T)x\| + \|T_n x\| \le \|(T_n - T)x\| + \|T_n\| \|x\|.$$

We know $\|(T_n - T)x\| \to 0$, while $\|T_n x\| \le c \|x\|$; this means T is bounded with $\|T\| \le c$. □

If X is not complete, then the above lemma does not hold. As a counterexample:

Example 15.13

Let $X \subseteq \ell^2$ be the set of sequences with finitely many nonzero terms. (It's easy to see that X is not complete.) Let $T_n \in B(X, X)$ be defined as

$$T_n(x) = (x_1, 2x_2, \ldots, nx_n, x_{n+1}, x_{n+2}, \ldots).$$

Then T_n is a bounded linear operator, but $\|T_n\| = n \to \infty$.

We see that T_n is strongly operator convergent to the operator T defined as $T(x) = (x_1, 2x_2, 3x_3, \ldots)$: but T is not bounded.

So we really need completeness here to know that our strongly operator convergent sequence really has a limit in $B(X, Y)$.

§16 April 8, 2024

Today we'll talk about a standalone topic, the Banach fixed point theorem.

§16.1 Fixed points

Definition 16.1 (Fixed point). Let X be a set and $T\colon X \to X$ a map. Then a *fixed point* of T is an element $x \in X$ such that $Tx = x$.

In math, it's very important to know when a function has a fixed point — this is because you can transfer tons of problems into questions about whether there exists a fixed point.

Example 16.2

A linear map $A\colon v \mapsto Av$ always has 0 as a fixed point; and the map $T = A - \lambda\,\mathrm{Id}$ has a nonzero fixed point if and only if λ is an eigenvalue of A.

Example 16.3

A translation $v \mapsto v + b$ (with $b \neq 0$) has no fixed points.

Example 16.4

More generally, an affine transformation $T\colon v \mapsto Av + b$ with $b \neq 0$ has a *unique* fixed point if $A - \mathrm{Id}$ is invertible (since $v = Av + b$ rearranges to $(A - \mathrm{Id})v = b$; so if we can invert $A - \mathrm{Id}$ then we're sure to have a unique fixed point).

This isn't the only case where the affine transformation has a fixed point — we'll later hopefully come back to affine transformations and give another condition under which they have fixed points.

Example 16.5

As a silly example, solving $f(x) = 0$ is equivalent to solving $f(x) + x = x$; so if we define $Tx = f(x) + x$, then a solution to $f(x) = 0$ is a fixed point of f.

Example 16.6

If we consider $T\colon \mathcal{C}^\infty(\mathbb{R}) \to \mathcal{C}^\infty(\mathbb{R})$ sending $f \mapsto f'$, the fixed points are ke^t for $k \in \mathbb{R}$.

So fixed points are pretty important, and lots of problems can be rephrased in terms of them.

§16.2 Contractions

We'll talk about a specific class of functions and spaces guaranteed to have a fixed point.

Definition 16.7 (Contractions). Let (X, d) be a metric space. Then a map $T\colon X \to X$ is called a *contraction* if there exists $\alpha < 1$ such that $d(Tx, Ty) \leq \alpha d(x, y)$ for all $x, y \in X$.

We call α the *contraction factor*. So there's a scalar strictly less than 1 such that when we apply our function T, we decrease distances by a factor of at least α.

In lots of problems, the idea is to find a metric on the space you're studying such that your function T becomes a contraction, so that we can apply the theorem we're about to see.

Remark 16.8. It's not enough to have $d(Tx, Ty) < d(x, y)$ for all $x \neq y$ — this condition is something that might seem like it's contracting distances, but we really need a *fixed* scalar (the contraction ratio shouldn't be allowed to approach 1).

Here's an easy lemma:

Lemma 16.9

Contractions are continuous.

Proof. Given $\varepsilon > 0$, we can simply take $\delta = \varepsilon/\alpha$; if $d(x, y) < \delta$, then $d(Tx, Ty) \leq \alpha d(x, y) < \varepsilon$. □

§16.3 The Banach fixed point theorem

Theorem 16.10 (Banach fixed point theorem)

Assume that $T\colon X \to X$ is a contraction and that X is complete. Then T has a unique fixed point.

Remark 16.11. We're going to see that the whole point of this theorem is that it gives you a very explicit way to get the fixed point; so ideally that should really be part of the theorem. (This isn't just an existence statement — it really tells you how to get the fixed point.) But this is how the theorem is stated in all books, so it's how we'll state it as well.

Proof. Let α be the contraction factor.

There's two things we have to prove — that there is a fixed point, and that it's unique. Uniqueness is easier, so we'll prove it first — assume that $x, y \in X$ are fixed points. Then

$$d(x, y) = d(Tx, Ty) \leq \alpha d(x, y)$$

(the first equality is because $Tx = x$ and $Ty = y$, and the second is because T is a contraction). This means $(1 - \alpha)d(x, y) \leq 0$. But $1 - \alpha > 0$ and $d(x, y) \geq 0$, so we must have $d(x, y) = 0$, which means $x = y$ (since d is a metric).

Existence is the fun part. We're going to give an explicit construction — pick an arbitrary $x_0 \in X$ and define the sequence $x_n = Tx_{n-1} = T^n x_0$ (for all $n \in \mathbb{N}$). We'll show that $(x_n)_{n=0}^{\infty}$ converges to a fixed point of T. There's two parts of this statement — that it converges and that the limit is a fixed point. But we just proved that contractions are continuous, so if we can prove that it converges — so that $\ell = \lim_{n\to\infty} x_n$ — then we also have $\ell = \lim_{n\to\infty} Tx_{n-1} = T(\lim_{n\to\infty} x_{n-1}) = T\ell$ (since T is continuous), so ℓ is a fixed point. So now we really just need to prove that $(x_n)_{n=0}^{\infty}$ converges. And since X is complete, we just need to show that $(x_n)_{n=0}^{\infty}$ is Cauchy.

To do so, first note that for all $n \geq 0$ we have

$$d(x_{n+1}, x_n) \leq d(Tx_n, Tx_{n-1}) \leq \alpha d(x_n, x_{n-1}) \leq \cdots \leq \alpha^n d(x_1, x_0).$$

So for any $n > m \geq N$, then

$$d(x_n, x_m) \leq d(x_n, x_{n-1}) + \cdots + d(x_{m+1}, x_m) \leq (\alpha^{n-1} + \cdots + \alpha^m) d(x_1, x_0).$$

And by the geometric series formula, this is

$$\alpha^m \cdot \frac{1-\alpha^{n-m}}{1-\alpha} d(x,y) \le \frac{\alpha^m}{1-\alpha} d(x,x_0) \le \frac{\alpha^N}{1-\alpha} d(x_1,x_0).$$

And this goes to 0 as $N \to \infty$; so given any $\varepsilon > 0$, we can find N such that this is smaller than ε (the only thing in this expression that isn't fixed is N), which means $d(x_n, x_m) < \varepsilon$ for all $n > m \ge N$. So $(x_n)_{n=0}^\infty$ is Cauchy, as desired. □

In fact, there's a way to estimate the error, which comes from the above bound.

Corollary 16.12 (Error bound)

In the same notation as the proof of the Banach fixed point theorem, we have the 'prior estimate'

$$d(x_n, \ell) \le \frac{\alpha^n}{1-\alpha} d(x_1, x_0).$$

The reason this is called a 'prior estimate' is that the right-hand side is very explicit — so if we know what kind of error we want to have, we can just find how large n should be to get that error, which tells you how many steps you need to approximate your fixed point up to a certain precision. And once you've computed all your points you can do a posterior estimate — we've found which n we need in order to get a good approximation, and then once we've computed x_n (which means we've already computed $x_0, \ldots, x_{n-1}$) we actually have the *posterior estimate*

$$d(x_n, \ell) \le \frac{\alpha}{1-\alpha} d(x_{n-1}, x_n).$$

(So we use the prior estimate before doing our computations, to find n; and then we use the posterior estimate to estimate our error after doing computations.)

Remark 16.13. This is again not ideal mathematical writing — since the corollary involves x_n, it should really be defined in the theorem statement.

§16.4 Some applications

§16.4.1 Finding zeros

One famous application is to Newton's method for solving $f(x) = 0$ (assume for simplicity that $f'(x) \ne 0$). We take the operator

$$Tx = x - \frac{f(x)}{f'(x)}$$

(so that a fixed point of T is a zero of f). If x_0 and f are sufficiently 'nice,' then this is a contraction (you need f' to not be too wild, so that you don't start diverging from a point close to 0 — if your derivative is too extreme you might just start going off to ∞ instead of converging).

Suppose we want to solve $x^3+x-1 = 0$. Then we have many choices. For example, we can take $Tx = 1-x^3$. But if you use this method, it's actually very bad. So another thing you could do is take

$$Tx = \frac{1}{1+x^2}.$$

This is better because it shrinks distances much more than the first choice (which may not even shrink distances). Even better, you can take

$$Tx = \frac{x^{1/2}}{(1+x^2)^{1/2}},$$

which converges even better. So you can see that finding which transformations to take that will give us the fastest approximation is also a big game — you want to have the best contraction possible.

So there's two things — finding a good T, and finding a good starting point x_0.

§16.4.2 Affine-linear maps

To return to one of our initial examples, suppose that $Tx = Ax + b$. We need to figure out two things — what metric we're using on $\mathbb{R}^n$, and when T is contracting. We'll use the metric

$$d(x, y) = \max_{i\in[n]} |x_i - y_i| = \|x - y\|_\infty .$$

Suppose we write $A = (a_{ij})_{i,j\in[n]}$.

Theorem 16.14

If $\sum_{j=1}^{n} |a_{ij}| < 1$ for all i, then T is a contraction, and there is a unique fixed point.

So if for *every* row the sum of absolute values of its entries is less than 1, then T is a contraction.

Remark 16.15. The story doesn't end here — there's ways of transforming A (e.g., Jacobi iterations).

§16.4.3 Differential equations

In an *initial value problem*, we're given a differential equation of the form $y' = f(t, y)$, together with an initial condition $y(t_0) = y_0$.

Question 16.16. When does a solution exist?

The function f could be anything (e.g., $t + y^2$); we want some conditions on f under which we know there's a solution for y.

Theorem 16.17 (Picard's existence and uniqueness theorem)

Assume that f is continuous on a rectangle

$$\mathcal{R} = \{(t, x) \mid |t - t_0| \le a,\ |x - y_0| \le b\},$$

and let c be such that $|f(t, x)| \le c$ for all $(t, x) \in \mathcal{R}$. Also suppose that f is k-Lipschitz on $\mathcal{R}$ with respect to the second argument — i.e.,

$$|f(t, x) - f(t, y)| < k\,|x - y|$$

for all t, x, and y. Then the initial value problem

$$\begin{cases} y' = f(t, y) \\ y(t_0) = y_0 \end{cases}$$

has a unique solution defined on $[t_0 - \beta, t_0 + \beta]$ for some $0 < \beta < a$.

So we've got a rectangle containing our initial condition; this automatically means (since this rectangle is compact) that f is bounded on $\mathcal{R}$, so we can always define some c (this c will play a role in our proof).

In the proof, we'll get an explicit definition of β.

Proof. There's a lot of things we need to do. First we need to come up with an operator T — we simply define Ty as

$$(Ty)(t) = y_0 + \int_{t_0}^{t} f(s, y(s))\, ds.$$

If $Ty = y$, then taking $t = t_0$ gives $y(t_0) = y_0 + 0 = y_0$, and more generally

$$y(t) = y_0 + \int_{t_0}^{t} f(s, y(s))\, ds,$$

which by the fundamental theorem of calculus means that

$$y'(t) = f(t, y(t)).$$

So if y is a fixed point of T, then it is a solution of the initial value problem.

We haven't yet said what space T takes place in and what the metric on this space is, so we'll take care of that next. Let $J = [t_0 - \beta, t_0 + \beta]$ (we'll define β later, such that $\beta < a$). Then T is defined on $\mathcal{C}(J)$ (the space of continuous functions on J). We need a metric on this space; we'll use the ∞-norm

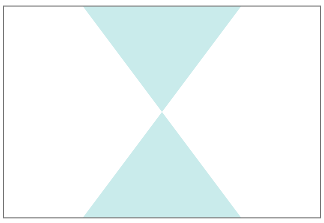
$$d(x, y) = \|x - y\|_\infty = \max_{t \in J} |x(t) - y(t)|$$

(here x and y are functions; in general the ∞-norm is defined by a sup, but since J is compact here we can just use a max).

The problem is that this space is too big — if we have a continuous function on J, its image is not necessarily going to remain in this rectangle $\mathcal{R}$. So we need to restrict it to stay within this rectangle — we define $\widetilde{\mathcal{C}} \subseteq \mathcal{C}(J)$ as

$$\widetilde{\mathcal{C}} = \{y \in \mathcal{C}(J) \mid d(y, y_0) \leq c\beta\}.$$

Here by y_0 we mean the constant function $t \mapsto y_0$, and we'll choose β such that $c\beta < b$.



The reason for this is that we want to bound $\int f(s, y(s))\, ds$, and this is at most c times the length of the interval we're integrating over. If c is very big then we need to take β small enough so that we stay within this rectangle (since we need T to never send things outside the rectangle).

If $y \in \widetilde{\mathcal{C}}$, then we have

$$d(Ty, y_0) = \max_{t \in J} \left| \int t_0{}^t f(s, y(s))\, ds \right|,$$

and since $|f(s, y(s))| \leq c$ this is smaller than $c\beta$; this ensures that Ty is also in $\widetilde{\mathcal{C}}$. (In fact, this computation just requires that $y \in \mathcal{C}$.)

So now we have a subspace $\widetilde{\mathcal{C}}$ that is preserved by our operator, which is pretty good. And $\widetilde{\mathcal{C}}$ is closed in $\mathcal{C}(J)$, so it is complete.

Now we have a space preserved under this map, so the last thing we need to show is that the map is a contraction. Suppose $x, y \in \widetilde{\mathcal{C}}$; then we have

$$d(Tx, Ty) = \max_{t \in J} |Tx - Ty| = \max_{t \in J} \left| \int_{t_0}^{t} f(s, x(s)) - f(s, y(s))\, ds \right|$$

(the y_0's cancel out). By the triangle inequality,

$$d(Tx, Ty) \le \max_{t \in J} \int_{t_0}^{t} |f(s, x(s)) - f(s, y(s))|\, ds.$$

And this is where we use the fact that f is Lipschitz — we have the same first argument, and we want to bound this difference in terms of the difference between x and y. And we can do this since f is Lipschitz, so

$$|f(s, x(s)) - f(s, y(s))| \le k\, |x(s), y(s)| \le k d(x, y).$$

So since we're integrating over an interval of length β, we get that

$$d(Tx, Ty) \le \beta k d(x, y).$$

And so for T to be a contraction, we just need $\beta k < 1$ — so we can define β such that

$$\beta < \min \left\{ a, \frac{b}{c}, \frac{1}{k} \right\}$$

(and as long as we choose β satisfying this, it'll work).

So after restricting our functions pretty heavily, we now have $\widetilde{\mathcal{C}}$ such that T is a contraction on $\widetilde{\mathcal{C}}$. And since $\widetilde{\mathcal{C}}$ is complete, this means there is a unique fixed point of T on $\widetilde{\mathcal{C}}$. □

Remark 16.18. The point of this was basically that you want your space to be preserved by T and you want T to be a contraction; all the conditions on β come from this.

Remark 16.19. We assumed f to be continuous and Lipschitz; both conditions are sufficient but may not be necessary, but it's still open to find necessary and sufficient conditions.

§16.5 A variant of the Banach fixed point theorem

Theorem 16.20 (Contraction of a ball)

Let X be a complete metric space and $T\colon X \to X$ a contraction on a closed ball $B = \{x \in X \mid d(x, x_0) \le r\}$. If $d(x_0, Tx_0) < (1 - \alpha)r$, then there is a unique fixed point of T on B, and $x_n = T^n x_0$ converges to this fixed point.

So here we have a map T which we don't know is a contraction everywhere, but it is a contraction on some ball. And this says that if T doesn't move the center of this ball by too much, then we still get a fixed point.

Proof. The idea is basically that in our earlier bound we had

$$d(x_0, x_n) \le \frac{1}{1 - \alpha} d(x_0, x_1) = \frac{1}{1 - \alpha} d(x_0, Tx_0),$$

so if we don't move the center by too much (so that $d(x_0, Tx_0)$ is small), then all x_n are in B; and B is closed, so their limit is also in B. □

§17 April 10, 2024

§17.1 Midterm announcements

The next midterm will cover everything since the last midterm — the corresponding Kreyszig sections are 3.6 (total orthonormal sets and sequences), 3.8–10 (the Riesz representation theorem, Hilbert adjoints, and adjoints), 4.1–4.9 (the big theorems), 4.12–4.13, 5.1 (the Banach fixed point theorem), and 5.3. Marjie aspires to make the second midterm a little more accessible and more appopriate in length; she is optimistic that this will be the case. All but one of the problems should be a direct application of the theory and definitions we've learned, and one problem brings together more tools from the entire course. But it can be hard to tell.

Right now the midterm is 6 problems; a couple have 2–3 parts.

Today we're going to talk about compact operators. This will *not* be covered on the midterm. (The material from here on will be emphasized on the final; we'll talk about this more closer to the final. There isn't a third midterm, so the final will be cumulative but with an emphasis on the last third of the material.) The nice thing about compact operators is that it refers to lots of the stuff that will be covered on the midterm, and will set us up for the next lecture on spectral theory — so it should be a pleasant review. We're following MacCluer here, because she does a nice job of setting up the connection between all the operator theory we've been doing and spectral theorems.

§17.2 Compact operators

Definition 17.1. Let X and Y be Banach spaces, and let $T\colon X \to Y$ be a map from X to Y. We say T is *compact* if whenever a sequence $(x_n)_{n\in\mathbb{N}} \subseteq X$ is bounded, we have that $(Tx_n)_{n\in\mathbb{N}} \subseteq Y$ has a convergent subsequence.

Remark 17.2. Recall that a metric space is *sequentially compact* if every sequence has a convergent subsequence. So this definition is equivalent to saying that T is compact if the image of every bounded set under T has compact closure — i.e., for every bounded $X' \subseteq X$, the closure of its image is compact in Y.

Let's first talk about some immediate implications of this definition; and then we'll see some examples.

Proposition 17.3

Let X and Y be Banach spaces, and $T\colon X \to Y$ a linear map. If T is compact, then T is bounded.

Proof. Suppose not — so we're supposing that T is not bounded. Then there exists a sequence $(v_n)_{n\in\mathbb{N}} \subseteq X$ with $\|v_n\| = 1$ such that $\|Tv_n\| \to \infty$. But $(v_n)_{n\in\mathbb{N}}$ is a bounded sequence, and if T is compact, then the image of every bounded sequence must have a convergent subsequence — so there exists a subsequence $(Tv_{n_k})_{k\in\mathbb{N}}$ that converges to some $y \in Y$, which means $\|Tv_{n_k}\| \to \|y\|$. But $\|y\|$ cannot be infinite, which is a contradiction. □

Remark 17.4. In class, we originally stated this for arbitrary maps (not necessarily linear). This proof requires T to be linear (in order to ensure $\|v_n\| = 1$), so for now we'll add 'linear' to the hypothesis; and if there's some error, Marjie will comment later.

§17.2.1 Some examples

Example 17.5

Let $S\colon \ell^2 \to \ell^2$ be the *forwards shift operator*

$$S\colon (x_1, x_2, \dots) = (0, x_1, x_2, \dots)$$

(where we shift all elements one step to the right, and put a 0 at the beginning). Then S is *not* compact — to see this, let $e_n = (0, 0, \dots, 1, 0, \dots)$ (with a 1 in the nth coordinate). Then $(e_n)_{n\in\mathbb{N}}$ is bounded, but $(Se_n)_{n\in\mathbb{N}} = (e_{n+1})_{n\in\mathbb{N}}$ does not have a convergent subsequence. So S is not compact.

For our next example, we'll need a few results we proved earlier.

Proposition 17.6

If X is a finite-dimensional normed linear space, then:

(1) Any two norms on X are equivalent.

(2) X is a Banach space.

Furthermore, for *any* normed linear space X, any finite-dimensional subspace of X is closed.

We proved (1) by showing that if we have a finite-dimensional space, then we can fix a basis and use that basis to determine a norm; we used this to show that all norms are equivalent. And for (2), the coefficients in that basis lie in the field $\mathbb{K}$, which is complete; we used this to show X is a Banach space.

Corollary 17.7

Let X be a finite-dimensional normed linear space, and $T\colon X \to Y$ a linear map. Then T is bounded.

Proof. As in the proofs of (1) and (2), let $\{e_1, \dots, e_n\}$ be a basis for X. Then any norm on X is equivalent to the ∞-norm defined with respect to this basis, i.e., the norm $\|x\|_\infty = \max_{i\in[n]} |\alpha_i|$ where $x = \sum \alpha_i e_i$. (So if we write any element of x with respect to this basis, we can define a norm looking at the largest coefficient in this basis — we used this in the earlier proofs, but now we're going to use it for this result.)

Then the idea is that for $x = \sum \alpha_i e_i$, we have

$$\|Tx\|_Y = \left\|\sum \alpha_i T e_i\right\| \le \|x\|_\infty \sum \|Te_i\|_Y .$$

And this implies T is bounded, with $\|T\| \le \sum \|Te_i\|$. □

We're going to use this in the next example.

Example 17.8

Let $T\colon \mathbb{C}^n \to \mathbb{C}^n$ be linear. Then T is compact.

Proof. Let $(x_n)_{n\in\mathbb{N}}$ be a bounded sequence in $\mathbb{C}^n$, with $\|x_n\| \le M$ for all $n \in \mathbb{N}$. By the above corollary, since T is linear, it is also bounded; this means $\|Tx_n\| \le \|T\| M$ for all $n \in \mathbb{N}$. So $(Tx_n)_{n\in\mathbb{N}} \subseteq \overline{\mathbb{B}}(0, \|T\| m)$ (we use $\overline{\mathbb{B}}$ to denote the closed ball with given center and radius). And this closed ball is compact (in $\mathbb{C}^n$), so our sequence (Tx_n) must have a convergent subsequence. This shows T is compact. □

The next example is finite rank operators.

Definition 17.9. Let X and Y be Banach spaces, and $T \in B(X, Y)$ a bounded linear map whose range $\mathcal{R}(T)$ is a finite-dimensional subspace of Y. Then we say T is a *finite rank operator*.

Lemma 17.10

Every finite rank operator is compact.

We previously considered the example of an operator $T\colon \mathbb{C}^n \to \mathbb{C}^n$; the exact same ideas in that proof tell you that every finite rank operator is compact. The point is that $\mathcal{R}(T)$ is finite-dimensional, so for any bounded $(x_n)_{n\in\mathbb{N}}$, its image will again be contained in the closure of a ball in finite-dimensional space, which is compact; and we can again use this to extract a convergent subsequence.

§17.3 Sequences of compact operators

Next, we'll get to one of our bigger theorems for today.

Theorem 17.11

Let X be a Banach space, and let $(T_n)_{n\in\mathbb{N}} \subseteq B(X, X)$ be a sequence of *compact* bounded linear operators $X \to X$ satisfying $\|T_n - T\| \to 0$ for some $T \in B(X, X)$. Then T is compact.

So this says that if we have a sequence of compact operators that converge in norm, then their limit is also a compact operator.

Proof. The proof is essentially a 'diagonal trick.' We want to show that T is compact — so let $(x_n)_{n\in\mathbb{N}}$ be a bounded sequence in X; we want to show that $(Tx_n)_{n\in\mathbb{N}}$ has a convergent subsequence. And because X is Banach, it suffices to show that $(Tx_n)_{n\in\mathbb{N}}$ has a *Cauchy* subsequence.

First, T_1 is a compact operator; this means the sequence (T_1x_n) has a convergent subsequence. (We'll pass to a list of subsequences, so we'll omit $n \in \mathbb{N}$.) In other words, this means there exists a subsequence $(x_{1n}) \subseteq (x_n)$ such that (T_1x_{1n}) is convergent.

And (x_{1n}) is bounded in X as well (since it's a subsequence of (x_n)), and T_2 is compact; this means there exists a subsequence $(x_{2n}) \subseteq (x_{1n})$ such that (T_2x_{2n}) is convergent. Note that (T_1x_{2n}) is also convergent, because (x_{2n}) is a subsequence of (x_{1n}).

Continuing inductively, we get a bunch of nested subsequences (x_{kn}) such that $(T_kx_{kn})_{n\in\mathbb{N}}$ is convergent as $n \to \infty$, and $(T_jx_{kn})_{n\in\mathbb{N}}$ is also convergent for all $j = 1, \ldots, k-1$.

So visually, we have a bunch of sequences (x_{kn}) (one in each row) where each row is a subsequence of the preceding row, and if we apply any of $T_1, \ldots, T_k$ to row k, we get a convergent sequence. And then we just take the diagonal — explicitly, we consider the sequence $(x_{nn})_{n\in\mathbb{N}}$.

$$\begin{array}{cccc} x_{11} & x_{12} & x_{13} & \cdots \\ x_{21} & x_{22} & x_{23} & \cdots \\ x_{31} & x_{32} & x_{33} & \cdots \end{array}$$

Then $(x_{nn})_{n\in\mathbb{N}}$ is a subsequence of our original sequence (x_n), and for each $k \in \mathbb{N}$ the sequence $(T_kx_{nn})_{n\in\mathbb{N}}$ is convergent (since $(x_{nn})_{n\geq k}$ is a subseqnece of $(x_{kn})_{n\in\mathbb{N}}$, which was chosen such that applying T_k to it produces a convergent sequence).

Now we're nearly done — because (x_n) was bounded, so is (x_{nn}); suppose that $\|x_{nn}\| \leq M$. Fix $\varepsilon > 0$; since $\|T_n - T\| \to 0$, there exists $k \in \mathbb{N}$ such that $\|T - T_k\| < \varepsilon/3M$. Then we know $(T_k x_{nn})_{n\in\mathbb{N}}$ is convergent, which means it is Cauchy; so there exists N such that for all $n, m \geq N$ we have $\|T_k x_{nn} - T_k x_{mm}\| < \varepsilon/3$. And now we want to show (Tx_{nn}) is Cauchy, so we'll divide up $\|Tx_{nn} - Tx_{mm}\|$ using the triangle inequality — we have

$$\|Tx_{nn} - Tx_{mm}\| \leq \|(T - T_k)x_{nn}\| + \|T_k(x_{nn} - x_{mm})\| + \|(T - T_k)x_{mm}\|.$$

The middle term is less than $\varepsilon/3$ (using the bound coming from $(T_k x_{nn})$ being Cauchy); and for the first and third terms, we have

$$\|(T - T_k)x_{nn}\| \leq \|T - T_k\| \, \|x_{nn}\| < \frac{\varepsilon}{3M} \cdot M = \frac{\varepsilon}{3},$$

and similarly for the last term. So we get that $\|Tx_{nn} - Tx_{mm}\| < \varepsilon$. This shows (Tx_{nn}) is Cauchy, which is what we wanted. □

§17.4 Approximations with finite rank operators

So we've proved that in a Banach space, a sequence of compact norms converging in norm actually converge to a compact operator. We also saw every finite rank operator is compact. Finite rank operators are in some sense 'finite-dimensional,' so a natural question is whether we can approximate compact operators with a sequence of finite rank operators. This would be a sort of natural converse to these two results.

We're going to go through a version of this for Hilbert spaces.

Example 17.12

Let H be a Hilbert space, and let $A \colon H \to H$ be a *diagonal* operator — an operator defined by $Ae_j = \alpha_j e_j$, where $(e_j)_{n\in\mathbb{N}}$ is a countable orthonormal basis for H. Assume that $\alpha_j \to 0$ (as $j \to \infty$). Then A is compact.

(For a countable orthonormal basis to exist, we want to assume H is separable.)

So A is a linear operator where applying A to any vector in our orthonormal basis just multiplies it by a scalar. This should remind us of linear algebra (the word 'diagonal' is a reference to the diagonal of a matrix).

In order to prove this, we're going to come up with a sequence of compact operators whose limit is A. We'll do this by using our two results from earlier — we're going to approximate A with finite rank operators, which are compact, and then apply the theorem to show that in fact A is compact.

Proof. Let A_n be the diagonal operator with diagonal $(\beta_j)_{j\in\mathbb{N}}$, where

$$\beta_j = \begin{cases} \alpha_j & j \leq n \\ 0 & j > n \end{cases}$$

(so A_n is essentially a truncation of A). Then A_n is a finite rank operator, and $A - A_n$ is a diagonal operator with diagonal $(\gamma_j)_{j\in\mathbb{N}}$, where

$$\gamma_j = \begin{cases} 0 & j \leq n \\ \alpha_j & j > n. \end{cases}$$

Because we have $\alpha_j \to 0$ as $n \to \infty$, we have

$$\|A - A_n\| = \sup_{j>n} |\alpha_j| \to 0$$

as $n \to \infty$. (Here we're using that (e_j) is orthonormal.) So we can apply the theorem to see that A is a compact operator. □

So we've shown that a *diagonal* operator whose diagonal converges to 0 is compact.

Theorem 17.13

Let H be a Hilbert space with a countable orthonormal basis, and let $T \in B(H, H)$ be compact. Then there exists a sequence $(T_n)_{n \in \mathbb{N}} \subseteq B(H, H)$ of finite rank operators such that $\|T_n - T\| \to 0$.

So this says every compact operator on a separable Hilbert space can be written as a limit of a sequence of finite rank operators.

Proof. Let $\{e_1, e_2, \ldots\}$ be a countable orthonormal basis for H. We'll use projection onto the first n vectors to set up our construction of T_n — let π_n be the projection map onto $\mathrm{Span}\{e_1, \ldots, e_n\}$, and let $T_n \in B(H, H)$ be defined as

$$T_n h = \pi_n T h = \sum_{i=1}^{n} \langle Th, e_i \rangle e_i.$$

Because $Th = \sum_{i \in \mathbb{N}} \langle Th, e_i \rangle e_i$, this means

$$\|T_n h - Th\|^2 = \sum_{i=n+1}^{\infty} |\langle k, e_i \rangle|^2,$$

which converges to 0 as $n \to \infty$ (because $\|Th\|$ is finite).

In the last class, we talked about different types of convergence of operators. This is not strong operator convergence — it's an analog of pointwise convergence, where we've shown that $T_n h$ converges in norm to Th for each h (in the norm on H, not in the norm on $B(H, H)$). So we want to show the stronger statement of *operator convergence*, that $\|T_n - T\| \to 0$.

To do this, let $\overline{\mathbb{B}}$ be the closed unit ball in H. Since T is a compact operator, by the equivalent definition of a compact operator (in the remark after the definition), we have that the closure of $T(\overline{\mathbb{B}})$ is compact in H. We're now going to look at a cover of this and use the fact that it has a finite subcover — given any $\varepsilon > 0$, the collection $(\mathbb{B}(Th, \varepsilon))_{h \in \overline{\mathbb{B}}}$ is an open cover of $\overline{T(\overline{\mathbb{B}})}$, so there exists a finite subcover — this means there exist $h_1, \ldots, h_m$ such that $\bigcup_{j=1}^{m} \mathbb{B}(Th_j, \varepsilon) \supseteq \overline{T(\overline{\mathbb{B}})}$.

And then because $\|T_n h - Th\| \to 0$ for each h, in particular this is true for each of $h_1, \ldots, h_m$; and since there's finitely many of them, we can choose $N_1, \ldots, N_m$ such that $\|T_n h_j - Th_j\| < \varepsilon$ for all $n \geq N_j$. Let $N^* = \max_{j \in [m]} N_j$; then $\|T_n h_j - Th_j\| < \varepsilon$ for all $j \in [m]$.

Now consider any $h \in \overline{\mathbb{B}}$. Because our balls cover this closure, we know there exists $h_j \in \overline{\mathbb{B}}$ such that $\|Th_j - Th\| < \varepsilon$. (This is because if $h \in \overline{\mathbb{B}}$ then $Th \in T(\overline{\mathbb{B}})$, so Th must lie in $\mathbb{B}(Th_j, \varepsilon)$ for some j; which in particular means $\|Th - Th_j\| < \varepsilon$.)

Then we have

$$\|T_n h - Th\| \leq \|(T_n - T)h_j\| + \|T_n(h_j - h)\| + \|T(h_j - h)\|.$$

we know $\|Th_j - Th\| < \varepsilon$, and $\|(T_n - T)h_j\| < \varepsilon$.

And we've shown this for every $h \in \overline{\mathbb{B}}$ (for every ε), which gives that $\|T_n - T\| \to 0$. □

Next time we'll review this final result, and we'll start on spectral theory (specifically, for compact self-adjoint operators on a Hilbert space).

§18 April 22, 2024

Today we'll start covering spectral theory.

§18.1 Compact operators

First we'll pick up where we left off last time, discussing compact operators. (This is from Chapter 4 of MacCluer and Chapter 8 of Kreyszig.)

First we'll review what we talked about last class.

Definition 18.1. Let X and Y be Banach spaces, and let $T\colon X \to Y$. We say T is *compact* if whenever $(x_n)_{n\in\mathbb{N}} \subseteq X$ is bounded, $(Tx_n)_{n\in\mathbb{N}} \subseteq Y$ has a convergent subsequence.

We then defined what it means for an operator to have *finite rank*, and used it to prove a bit of theory.

Definition 18.2. Let X and Y be Banach spaces, and let $T \in B(X,Y)$. We say T is of *finite rank* if its range $\mathcal{R}(T)$ is finite-dimensional (as a subspace of Y).

(This is the analog of rank from linear algebra.)

Then we proved the following three theorems, which we're going to use today to build spectral theory.

Theorem 18.3

If X and Y are Banach and $T\colon X \to Y$ is compact and linear, then $T \in B(X,Y)$.

Theorem 18.4

Every finite rank operator is compact.

The next theorem says that a limit of compact operators is compact.

Theorem 18.5

Let X be a Banach space, and $(T_n)_{n\in\mathbb{N}} \subseteq B(X,X)$ a sequence of compact operators with

$$\lim_{n\to\infty} \|T_n - T\| \to 0$$

for some $T \in B(X,X)$. Then T is compact.

(We proved this using a diagonal argument.) Our final result was a sort of density result, saying that every compact operator is a limit of finite rank operators (in a Hilbert space):

Theorem 18.6

Let H be a Hilbert space, and let $T \in B(H,H)$ be compact. Then there exists a sequence $(T_n)_{n\in\mathbb{N}}$ of finite rank operators with $\|T_n - T\| \to 0$.

We proved this when H was a Hilbert space with a countable orthonormal basis, but in fact this theorem holds even without that hypothesis.

§18.2 More with Hilbert adjoints

We're close to done setting things up. We're planning to prove our first spectral theorems for compact, self-adjoint operators on a Hilbert space. Here we're talking about the Hilbert adjoint. First let's recall the definition of the Hilbert adjoint:

Theorem 18.7

Let H be a Hilbert space, and $T \in B(H, H)$. Then there is a unique $T^* \in B(H, H)$ such that $\langle Tf, g \rangle = \langle f, T^* g \rangle$ for all $f, g \in H$.

Definition 18.8. We call T^* the *(Hilbert) adjoint* of T. If $T = T^*$, we say T is *self-adjoint.*

We'll need a few facts, which are pretty immediate (except for the last one):

Fact 18.9 — For all $T, S \in B(H, H)$, we have:

(i) $T^{**} = T$.

(ii) $(T + S)^* = T^* + S^*$.

(iii) $(TS)^* = S^* T^*$.

(iv) $\|T\| = \|T^*\|$, and $\|T^* T\| = \|T^2\|$.

Proposition 18.10

Let H be a Hilbert space and $T \in B(H, H)$. Then T is compact if and only if T^* is compact.

Proof. By the first property $T^{**} = T$, we only need to show one direction; we'll show the forwards direction. (If we show this direction, then the compactness of T^* implies that of $T^{**} = T$.)

Suppose that T is compact. By Theorem 18.5, there exists a sequence of finite rank operators $(T_n)_{n \in \mathbb{N}}$ satisfying $\|T_n - T\| \to 0$.

Claim 18.11 — For each $n \in \mathbb{N}$, T_n^* is a compact operator.

Proof. To see this, we'll look at the projection of H onto the finite-dimensional subspace $\mathcal{R}(T_n) \subseteq H$ — let $\pi_n \colon H \to H$ be the projection operator onto $\mathcal{R}(T_n)$, so immediately we have $\pi_n T_n = T_n$ (because the projection of something in $\mathcal{R}(T_n)$ onto $\mathcal{R}(T_n)$ is just itself). Then we can take the adjoint of both sides to get $(\pi_n T_n)^* = T_n^*$, which means $T_n^* \pi_n^* = T_n^*$. Further, because π_n is a projection, $\pi_n^* = \pi_n$ — this is because $\langle \pi_n x, y \rangle = \langle x, \pi_n y \rangle$ for all $x, y \in H$.

So now we have $T_n^* \pi_n = T_n^*$. This implies T_n^* is finite rank (since π_n is a projection onto a finite-dimensional subspace, so $T_n^* \pi_n$ also has finite-dimensional range). This implies T_n^* is compact. □

Finally, by (iv) we have $\|T_n^* - T^*\| = \|T_n - T\|_{\text{op}} \to 0$, which means by Theorem 18.5 that T is compact. □

§18.3 Spectral theory

In linear algebra, we've seen the *spectral theorem* — that if $n \times n$ matrix M is self-adjoint (i.e., $M = M^*$, where M^* is the conjugate tranpose of M), then:

- All eigenvalues of M are real.
- There exists a unitary matrix U (i.e., a matrix with $U^* = U^{-1}$) such that UMU^{-1} is a diagonal matrix.
- There exists an orthonormal basis of eigenvectors of M.

This is the first spectral theorem we learn (there are improvements when M is normal). We want to translate this to our situation of Hilbert spaces and Banach spaces. Generalizations of this result to Hilbert and Banach spaces are called *spectral theorems*.

Remark 18.12. This is Chapter 4 of MacCluer and Chapters 7–9 of Kreyszig; there's more applications and some theory from Chapter 10 and onwards.

For now, we'll assume $\mathbb{K} = \mathbb{C}$.

Definition 18.13. For a Banach space X and $T \in B(X, X)$, we say $\lambda \in \mathbb{C}$ is an *eigenvalue* of T if there exists $x \in X \setminus \{0\}$ such that $Tx = \lambda x$.

Definition 18.14. When λ is an eigenvalue of T, its *associated eigenspace* is $\mathcal{N}(T - \lambda \operatorname{Id})$, and the nonzero vectors in this eigenspace are the *eigenvectors* of T.

We're going to work with compact self-adjoint operators. Note that if $T \in B(H, H)$ is self-adjoint (i.e., $T = T^*$), then $\langle Tx, y\rangle = \langle x, Ty\rangle = \overline{\langle Ty, x\rangle}$. In particular, taking $x = y$ gives $\langle Tx, x\rangle = \overline{\langle Tx, x\rangle}$, which means $\langle Tx, x\rangle$ is real.

Lemma 18.15

Let H be a Hilbert space, and suppose that $T \in B(H, H)$ is self-adjoint. Then $\|T\| = \sup_{\|x\|=1} |\langle Tx, x\rangle|$.

This looks like some of the results we've had earlier in the course about the norm, but now we're talking about the norm of an *operator*, so this is different.

Remark 18.16. In linear algebra you might recognize this as the *Rayleigh quotient*, and in particular, taking the supremum is another way to compute the maximum eigenvalue — in linear algebra $\sup_{\|x\|=1} |\langle Tx, x\rangle|$ is the magnitude of the largest eigenvalue, which is in fact the norm of a matrix.

Proof. Let $M = \sup_{\|x\|=1} |\langle Tx, x\rangle|$. First, here are some useful observations:

- For every $h \in H$ with $h \neq 0$, we can scale h to $h' = h/\|h\|$, which has norm 1; then we have
$$|\langle Th, h\rangle| = \|h\|^2 |\langle Th', h'\rangle| \le M \|h\|^2 .$$
- We have $|\langle Th, h\rangle| \le \|Th\| \|h\| \le \|T\| \|h\|^2$, where the first inequality is by Cauchy–Schwarz; taking the supremum over $\|h\| = 1$ gives that $M \le \|T\|$.
- We have $\langle T(f + g), f + g\rangle - \langle T(f - g), f - g\rangle = 4 \operatorname{Re}\langle Tf, g\rangle$. This is a direct computation — the first term expands out to $\langle Tf, f\rangle + \langle Tf, g\rangle + \langle Tg, g\rangle + \langle Tg, f\rangle$. But $\langle Tg, f\rangle = \langle g, Tf\rangle = \overline{\langle Tf, g\rangle}$ (because T is self-adjoint), which means
$$\langle Tf, g\rangle + \langle Tg, f\rangle = 2 \operatorname{Re}\langle Tf, g\rangle.$$
And if we write the second term out too, the $\langle Tf, f\rangle$ and $\langle Tg, g\rangle$ terms cancel, and we get another term of $+2 \operatorname{Re}\langle Tf, g\rangle$.

Now since T is self-adjoint, we have that $\langle Tx, x\rangle$ and $\langle Ty, y\rangle$ are real for all $x, y \in H$, and

$$\langle Tx, x\rangle - \langle Ty, y\rangle \le |\langle Tx, x\rangle| + |\langle Ty, y\rangle| \le M(\|x\|^2 + \|y\|^2).$$

By the parallelogram equality, this is equal to

$$\frac{M}{2}(\|x + y\|^2 + \|x - y\|^2).$$

We've shown that $M \le \|T\|$, and we're trying to show that $\|T\| \le M$. In order to do this, let $v \in H$ satisfy $\|v\| = 1$ and $Tv \neq 0$, so that $\|Tv\| = s > 0$ for some s. Now take

$$x = v + \frac{Tv}{s} \text{ and } y = v - \frac{Tv}{s}.$$

Now we're going to plug this in; this tells us that

$$\langle Tx, x\rangle - \langle Ty, y\rangle \le \frac{M}{2}\left(\|2v\|^2 + \left\|\frac{2Tv}{s}\right\|^2\right) = 4M$$

(note that $\|Tv\| = s$ and $\|v\| = 1$). Then by the third observation, we have

$$\langle Tx, x\rangle - \langle Ty, y\rangle = 4\,\mathrm{Re}\left\langle Tv, \frac{1}{s}Tv\right\rangle = 4\,\|Tv\|.$$

This tells us $\|Tv\| \le M$. This is true for all v, so $\|T\| \le M$, which is what we wanted to prove. □

§18.4 Three spectral theorems

Theorem 18.17

If T is a compact self-adjoint operator in $B(H, H)$ (where H is a Hilbert space), then at least one of $\|T\|$ and $-\|T\|$ is an eigenvalue of T.

Proof. If $\|T\| = 0$ then $T = 0$, and there's nothing to prove; so assume $\|T\| > 0$.

By the above lemma, we have

$$\|T\| = \sup_{\|x\|=1} |\langle Tx, x\rangle|.$$

Let $(x_n)_{n\in\mathbb{N}} \subseteq H$ be a sequence approaching this supremum, so that $\|x_n\| = 1$ for all n and $|\langle Tx_n, x_n\rangle| \to \|T\|$. Now because T is self-adjoint, $\langle Tx_n, x_n\rangle$ is always real; this means we can find a subsequence of $(x_n)_{n\in\mathbb{N}}$ such that $\langle Tx_n, x_n\rangle$ converges to either $\|T\|$ or $-\|T\|$; we'll call this limit λ.

Then $\|Tx_n - \lambda x_n\|^2 = \|Tx_n\|^2 - 2\,\mathrm{Re}\langle Tx_n, \lambda x_n\rangle + |\lambda|^2\,\|x_n\|^2$ (just by multiplying things out). Because λ is real, and so is $\langle Tx_n, x_n\rangle$, the term in the middle is actually real; and $|\lambda|^2 = \|T\|^2$ and $\|x_n\| = 1$, so we get

$$\|Tx_n - \lambda x_n\|^2 = \|Tx_n\|^2 - 2\lambda\langle Tx_n, x_n\rangle + |\lambda|^2 \le 2\lambda^2 - 2\lambda\langle Tx_n, x_n\rangle$$

(note that $\|Tx_n\| \le \|T\|\,\|x_n\| = \|T\|$).

But applying this to our subsequence (x_{n_k}) with $\langle Tx_{n_k}, x_{n_k}\rangle \to \lambda$, the right-hand side converges to 0, and so we get $\|Tx_{n_k} - \lambda x_{n_k}\| \to 0$. This means $Tx_{n_k} - \lambda x_{n_k} \to 0$ (as a sequence of elements in H).

Now because T is compact and $\|x_{n_k}\| = 1$ (which means the sequence (x_{n_k}) is bounded), we know $(Tx_{n_k})_{k\in\mathbb{N}}$ has a convergent subsequence — $T_{x_{n_{k_\ell}}} \to y$. But we know $Tx_{n_k} - \lambda x_{n_k} \to 0$, so this tells us $\lambda x_{n_{k_\ell}} \to y$, which in particular means (since $\lambda \neq 0$) that $\lambda T_{x_{n_{k_\ell}}} \to Ty$. Meanwhile, by multiplying $T_{x_{n_{k_\ell}}} \to y$ by λ we get that $\lambda T_{x_{n_{k_\ell}}} = \lambda y$. Combining these gives $Ty = \lambda y$.

This is essentially what we wanted to show, but to complete the proof, we need to show that $y \neq 0$. But this is not hard — we have $\left\|Tx_{n_{k_\ell}}\right\| \ge \left\|\lambda x_{n_{k_\ell}}\right\| - \left\|Tx_{n_{k_\ell}} - \lambda x_{n_{k_\ell}}\right\|$ by the triangle inequality, and the right-hand side goes to $|\lambda|$, which implies that $|y| \ge |\lambda| \neq 0$. □

Theorem 18.18

Let H be a Hilbert space, and let $T \in B(H, H)$ be self-adjoint. Then every eigenvalue of T is real, and the eigenvectors for distinct eigenvalues are orthonormal.

Remark 18.19. This does not necessarily imply that T has any eigenvalues — there do exist self-adjoint operators that don't have any eigenvalues. (The first theorem we proved was only for *compact* operators — but we do know that those have at least one real eigenvalue, namely either $\|T\|$ or $-\|T\|$.)

Proof. Suppose λ is an eigenvalue of T, so that $Th = \lambda h$ for some $h \neq 0$. Then $\langle Th, h\rangle = \langle \lambda h, h\rangle = \lambda \|h\|^2$. But we also have $\langle Th, h\rangle = \langle h, Th\rangle = \langle h, \lambda h\rangle = \overline{\lambda} \|h\|^2$. Combining these gives $\lambda = \overline{\lambda}$, so $\lambda \in \mathbb{R}$.

For the second statement, we need to show that if we have distinct eigenvalues, then their eigenspaces are orthogonal. Suppose $\lambda \neq \mu$ are distinct eigenvalues of T, and let h and g be such that $Th = \lambda h$ and $Tg = \mu g$. Then we have $\langle Th, g\rangle = \langle g, Th\rangle$ because T is self-adjoint. But

$$\langle Th, g\rangle = \langle \lambda h, g\rangle = \lambda\langle h, g\rangle,$$

while similarly

$$\langle h, Tg\rangle = \langle h, \mu g\rangle = \mu\langle h, g\rangle$$

(note that we get μ instead of $\overline{\mu}$ when pulling it out because μ is real). This means

$$(\lambda - \mu)\langle h, g\rangle = 0,$$

and since $\lambda \neq \mu$ this means $\langle h, g\rangle = 0$. □

Theorem 18.20

Let H be a Hilbert space, and let $T \in B(H, H)$ be a compact self-adjoint operator. Then the set of eigenvalues of T is finite or countably infinite with $\lim_{i\to\infty} \lambda_i = 0$.

This is a super strong result. We'll start with the proof next time — it's very satisfying. But let's think about how this could work.

The idea is that we'll assume for contradiction that there are infinitely many λ_i that are bounded below by ε. This will contradict the compactness of T. (We define Λ_ε as the set of eigenvalues with $|\lambda| \geq \varepsilon$. Then we get $Ty_j = \lambda_j y_j$ for some nonzero y_j, and using the orthogonality of the y_j we can show that this gives a sequence without a convergent subsequence, contradicting compactness.)

§19 April 24, 2024

Today we'll continue spectral theory. Last class, the last thing we did was state the following theorem:

Theorem 19.1

Let H be a Hilbert space and let $T \in B(H, H)$ be compact and self-adjoint. Then the set of eigenvalues of T is finite or countably infinite with $\lambda_i \to 0$.

Proof. Recall that last time we proved that if T is self-adjoint, then all its eigenvalues λ_i are real (right now we're letting the index set be arbitrary). We also proved that eigenvectors for distinct eigenvalues are orthogonal.

Now if λ is an eigenvalue, then there exists $x \neq 0$ such that $Tx = \lambda x$; this means $|\lambda| \leq \|T\|$.

If the set of eigenvalues is finite, then there's nothing to prove. So now suppose that the set of eigenvalues is infinite; we'll let this set be $(\lambda_i)_{i \in I}$ (where I is some index set; we haven't yet proved it's countable). Let $\Lambda_\varepsilon = \{i \in I \mid |\lambda_i| > \varepsilon\}$. We claim that Λ_ε is finite for every $\varepsilon > 0$ — if we can show this, then we can put the eigenvalues in a countable sequence by looking at $\varepsilon = 1, \frac{1}{2}, \frac{1}{3}, \ldots$. Assume for contradiction that Λ_ε is infinite for some $\varepsilon > 0$. Then there exist distinct eigenvalues $(\lambda_j)_{j \in \mathbb{N}} \subseteq \mathbb{R}$ and eigenvectors $(y_j)_{j \in \mathbb{N}} \subseteq H$ such that $Ty_j = \lambda_j y_j$; and we can assume $\|y_j\| = 1$ for all j. As mentioned earlier, the y_j are orthogonal; and since they have norm 1, this means they're actually orthonormal. So then we have

$$\|Ty_j - Ty_k\|^2 = \|\lambda_j y_j - \lambda_k y_k\|^2 = |\lambda_j|^2 + |\lambda_k|^2 \geq 2\varepsilon^2$$

(by the orthonormality of the y_j). This means the sequence $(Ty_j)_{j \in \mathbb{N}}$ does not have a convergent subsequence; and this contradicts that T is compact (as $(y_j)_{j \in \mathbb{N}}$ is a bounded sequence).

So we've shown that Λ_ε is finite for all $\varepsilon > 0$. In particular $\Lambda_{1/n}$ is finite for each $n \in \mathbb{N}$, allowing us to enumerate all the eigenvalues (countably). □

Remark 19.2. Here we really used compactness in this proof, but the hypothesis that T is self-adjoint (giving us real eigenvalues) is not necessary — even if T is not self-adjoint, we'll still have a countable collection of eigenvalues converging to 0.

§19.1 Invariant and reducing subspaces

Lemma 19.3

Let H be a Hilbert space, and let $T \in B(H, H)$. Let $M \subseteq H$ be a closed subspace of H. If $TM = \{Tm \mid m \in M\}$ is contained in M, then $T^*(M^\perp) \subseteq M^\perp$. Conversely, if $T^*(M^\perp) \subseteq M^\perp$, then $TM \subseteq M$.

Proof. First, because $T^{**} = T$ and $M^{\perp\perp} = M$, we only need to prove the first direction. Suppose $TM \subseteq M$; then we want to show $T^*(M^\perp) \subseteq M^\perp$. Let $n \in M^\perp$; then we want to show that T^*n is also in $M^\perp$, meaning that $\langle T^*n, m\rangle = 0$ for all $m \in M$. In order to show this, we can simply move T^* to the other side — we have

$$\langle T^*n, m\rangle = \langle n, Tm\rangle.$$

But $n \in M^\perp$ and $Tm \in TM \subseteq M$, so this is 0; and this means $\langle T^*n, m\rangle = 0$. So $T^*n \in M^\perp$, which is what we wanted to prove. □

Corollary 19.4

Under the same hypotheses, if $T = T^*$ and $TM \subseteq M$, then $T(M^\perp) \subseteq M^\perp$.

We'll now write a few definitions.

Definition 19.5. We say a closed subspace $M \subseteq H$ is *invariant for T* if $TM \subseteq M$, and *reducing for T* if $TM \subseteq M$ and $TM^\perp \subseteq M^\perp$.

(These definitions are for H a Hilbert space and $T \in B(H, H)$, as in the lemma.)

In particular, the above corollary says that if T is self-adjoint and M is invariant for T, then M is also reducing for T.

Why do we say *reducing* here? If $TM \subseteq M$ and $TM^\perp \subseteq M^\perp$ and we want to understand T, then we can reduce the problem to understanding what T does on M and what it does on $M^\perp$ — so this corresponds to a reduction in the problem.

Example 19.6

Let $E \subseteq [0, 1]$ be measurable, and let $N_E = \{f \in L^2[0, 1] \mid f(x) = 0 \text{ for a.e.} x \in E\}$. Let $M_x : L^2[0, 1] \to L^2[0, 1]$ be the multiplication operator $M_x f = f \cdot x$ (where x is the independent variable in our L^2 functions). Then N_E is a reducing subspace for M_x.

Proof. First note that M_x is self-adjoint (assuming we're working with $\mathbb{R}$) — we have $\langle f, g\rangle = \int fg$, so $\langle fx, g\rangle = \int fxg = \langle f, gx\rangle$.

Now notice that if $f \in N_E$, then $M_x f = fx$ is also zero for almost every element of E, which means $M_x f \in N_E$. □

As mentioned earlier, the point is that reducing subspaces correspond to a reduction in the problem — in some cases (though maybe not this one), it might be easier to think of $L^2[0, 1]$ as $N_E \oplus N_E^\perp$; and then we can study what M_x does on N_E and on $N_E^\perp$.

Theorem 19.7

Let $T \neq 0$ be a compact self-adjoint operator in $B(H, H)$ (where H is a Hilbert space). Then there exists a finite or countably infinite orthonormal set (g_n) of eigenvectors of T with corresponding real eigenvalues (λ_n) such that

$$Tx = \sum_n \lambda_n \langle x, g_n\rangle \cdot g_n.$$

(And $\lambda_n \to 0$.)

(This is our strongest spectral theorem so far.)

Proof. Last class, we proved that the eigenvalues (λ_n) are real, and we proved that the one with largest magnitude, which we'll call λ_1, is either $\|T\|$ or $-\|T\|$. Let g_1 be the corresponding unit eigenvector, so that $Tg_1 = \lambda_1 g_1$. Let $M_1 = \text{Span}\{g_1\}$. Then $TM_1 \subseteq M_1$ (because every element of M_1 is of the form αg_1, and $T\alpha g_1 = \alpha T g_1 = \alpha\lambda_1 g_1$). This means M_1 is an invariant subspace for T; and since T is self-adjoint, this means M_1 is a *reducing* subspace for T.

Let $H_2 = M_1^\perp$ be the orthogonal complement of M_1 in H, and let $T_2 = T|_{H_2}$ be the restriction of T to H_2.

Claim 19.8 — This restriction T_2 is also compact and self-adjoint.

Proof. We'll prove self-adjointness first — let $x, y \in H_2$. Then

$$\langle T_2^* x, y\rangle = \langle x, T_2 y\rangle.$$

And because $y \in H_2$ and T_2 is just the restriction of T to H_2, this is just equal to

$$\langle x, Ty\rangle = \langle Tx, y\rangle = \langle T_2 x, y\rangle$$

(again using the definition of T_2 and the fact that T is reducing), proving that T_2 is self-adjoint.

For the second statement, suppose (x_n) is a bounded sequence in H_2; we need to show that $(T_2 x_n)$ has a convergent subsequence in H_2. This follows from the reducing property as well — we know $T_2 x_n = T x_n$ on H_2, and $(T x_n) \subseteq H_2$ (for $x_n \in H_2$); and we know it has a convergent subsequence, and that convergent subsequence must be in H_2. □

Now we have a compact self-adjoint operator on H_2. And we're going to repeat this — again we apply the theorem from last class to say that there exists λ_2 equal to either $\|T_2\|$ or $-\|T_2\|$ — and notice $\|T_2\| \leq \|T\|$, because T_2 is a restriction of T to a smaller set, and the norm is defined as a supremum, so $|\lambda_2| \leq |\lambda_1|$.

Let $g_2 \in H_2$ be the unit vector corresponding to λ_2, so $T_2 g_2 = \lambda_2 g_2$. Because g_2 is in H_2, it has to be orthogonal to g_1 (because $\operatorname{Span} g_1 = M_1$, and we defined $H_2 = M_1^\perp$). And this means g_2 is also an eigenvector for T with $T g_2 = \lambda_2 g_2$. Now we can define $M_2 = \operatorname{Span}\{g_2\} \subseteq H_2$, and this is a reducing subspace for T.

Continuing inductively, we produce an orthonormal sequence of eigenvectors $\{g_1, g_2, \ldots\}$ for T with corresponding eigenvalues $\lambda_1, \lambda_2, \ldots$, where this sequence is non-increasing in absolute value (i.e., $|\lambda_1| \geq |\lambda_2| \geq \cdots$) and where $|\lambda_j| = \|T_j\|$, where $T_1 = T$ and T_j is defined as the restriction of T to the set $H_j = \operatorname{Span}\{g_1, \ldots, g_{j-1}\}^\perp$.

Now one of two things can happen. Either this process terminates because at some point $T_j = 0$, or it continues indefinitely. We'll treat those two cases now.

Case 1 (The process terminates with $T_m = 0$ for some $m \in \mathbb{N}$). Let $x \in H$, and define $y = x - \sum_{j=1}^{m-1} \langle x, g_j\rangle g_j$. Then $y \in \operatorname{Span}\{g_1, \ldots, g_{m-1}\}^\perp$ (because we've subtracted off the components that are in the subspace spanned by these vectors), which means $y \in H_m$, and therefore $0 = T_m y = Ty = Tx - \sum \langle x, g_j\rangle T g_j = Tx - \sum \lambda_j \langle x, g_j\rangle g_j$. That tells us that Tx is equal to this sum.

Case 2 (The process goes on indefinitely). First, we'll prove on the problem set that each λ_i can appear at most finitely many times (i.e., each eigenvalue corresponds to a finite-dimensional eigenspace). In combination with the theorem from earlier, we have $\lambda_n \to 0$, and we want to show $Tx = \lim_{n\to\infty} \sum_{j=1}^{n-1} \lambda_j \langle x, g_j\rangle g_j$.

As before, let

$$y_n = x - \sum_{j=1}^{n-1} \langle x, g_j\rangle g_j,$$

As before, this implies $y_n \in H_n$. And we have $\|x\| \geq \|y_n\|$. Meanwhile

$$\left\| Tx - \sum_{j=1}^{n-1} \lambda_j \langle x, g_j\rangle g_j \right\| = \|T y_n\| = \|T_n y_n\| \leq \|T_n\| \, \|y_n\| \, .$$

But $\|T_n\| = |\lambda_n| \to 0$, while $\|y_n\| \leq \|x\|$ is bounded; so we're done. □

Remark 19.9. The idea was using reducing subspaces to decompose the problem, as in linear algebra.

Corollary 19.10

The sequence (λ_n) in the theorem is a complete list of the nonzero eigenvalues of T.

Proof. Suppose not. Then there exists μ which is nonzero and not equal to any of the λ_n such that $Tg = \mu g$, and we know $g \perp g_n$ for all n; in particular the statement of the theorem then says that $Tg = 0$ (since each term in the sum is 0). But $Tg = \mu g$, and since $\mu \neq 0$ this means $g = 0$, which is a contradiction. □

Corollary 19.11

If T is a compact self-adjoint operator on a *separable* Hilbert space H, then there exists an orthonormal *basis* $\{e_n\}$ of H consisting of eigenvectors of T and such that $Tx = \sum_n \lambda_n \langle x, e_n \rangle e_n$.

Proof. This is pretty immediate; all we need to do is append to the (g_n) a basis for $\mathcal{N}(T)$. (We use our orthonormal sequence from the theorem, and we add to it a countable basis for $\mathcal{N}(T)$, which is countable.) □

§19.2 An open question

There's a deep question in functional analysis:

Question 19.12. For X Banach (or Hilbert), does ever $T \in B(X, X)$ have a nontrivial closed invariant subspace?

Recall that an invariant subspace is one with $TM \subseteq M$. We have all the tools to understand this question now. We've proven that if T has an eigenvalue, then the answer is yes (because an eigenvalue immediately has an eigenspace, which is an invariant subspace). However, there exist operators T (we'll see one in the homework) such that T has no eigenvalues, but T still has invariant subspaces. So the converse is certainly not true.

If T is a compact operator on a Hilbert space (so X is Hilbert), then von Neumann and Arozajn proved that the answer is yes in 1950. This was generalized by Arozajn and another author for T a compact operator and X Banach — the answer is still yes. And since then there have been various results with conditions on T.

In 1987, Enflo constructed an operator on a Banach space with no invariant subspace; that's interesting because the other results are affirmative. (The construction of the Banach space is the challenging part.) Last year, he announced affirmative results for operators on Hilbert spaces (with no qualifications). (There was an update this month, and experts aren't yet sure if it's true.)

§20 April 29, 2024

§20.1 The spectrum

We're going to talk about the spectrum, so let's define it.

Definition 20.1. Let X be a Banach space (over $\mathbb{C}$), and $T \in B(X, X)$. The *spectrum* of T, denoted $\sigma(T)$, is the set of complex numbers $\lambda \in \mathbb{C}$ such that $T - \lambda I$ is not invertible.

(For this lecture — and the last one — we'll always be talking about Banach spaces over $\mathbb{C}$.)

(Note that $T - \lambda I$ could fail to be invertible either because it's not injective or because it's not surjective; we'll talk about these two cases later.)

There's a sense in which the spectrum is an invariant of an operator:

Proposition 20.2

Let X be Banach, and suppose that T, S, and S^{-1} are all bounded operators on X. Then $\sigma(T) = \sigma(S^{-1}TS)$.

So this is the sense in which we mean that the spectrum is an invariant of T.

Proof. Suppose that λ is *not* in $\sigma(T)$; we'll show that λ is also not in $\sigma(S^{-1}TS)$. For λ to be in the spectrum of T means that $T - \lambda I$ is invertible; let its inverse be V. Then we can rewrite

$$S^{-1}TS - \lambda I = S^{-1}(T - \lambda I)S,$$

and this has inverse $S^{-1}VS$; this implies $\lambda \notin \sigma(S^{-1}TS)$. So we've shown one direction of the containment. For the other direction, suppose that $\lambda \notin \sigma(S^{-1}TS)$, which means $S^{-1}TS - \lambda I = W - \lambda I$ is invertible; let its inverse be V. We'll essentially do the same thing as above — we have

$$SWS{-}1 - \lambda I = S(W - \lambda I)S^{-1}$$

(we've switched the order of S and S^{-1}, but the idea is the same) — because $W - \lambda I$ is invertible with inverse V, this means $S(W - \lambda I)S^{-1}$ has inverse SVS^{-1}. And this implies $\lambda \notin \sigma(SWS^{-1})$, and by definition $SWS^{-1} = T$ (because we set $W = S^{-1}TS$, so this is $SS^{-1}TSS^{-1} = T$); so $\lambda \notin \sigma(T)$.

So we've shown that the complements of $\sigma(T)$ and $\sigma(S^{-1}TS)$ are the same, which means these two spectra are also the same. □

The spectrum of T is where $T - \lambda I$ fails to be invertible. There are two different ways in which $T - \lambda I$ can fail to be invertible — $T - \lambda I$ is invertible if and only if it's bijective. One way this can fail is if $T - \lambda I$ is not one-to-one (or injective) — this means there exist $x \neq y$ in X such that $(T - \lambda I)x = (T - \lambda I)y$, or equivalently $(T - \lambda I)(x - y) = 0$. Letting $v = x - y$, this is the same as saying $Tv = \lambda v$, so in this case λ is an eigenvalue of T (with eigenvector v). If this happens, we say λ is in the *point spectrum* of T.

Definition 20.3. We define the *point spectrum* of T, denoted $\sigma_p(T)$, as the set of eigenvalues of T.

The second case is when $T - \lambda I$ is not onto (i.e., surjective). We'll split this into two cases. First we consider the case where $T - \lambda I$ is onto a dense subspace of X.

Definition 20.4. We define the *continuous spectrum* of T, denoted $\sigma_c(T)$, as the set of λ for which $T - \lambda I$ is not surjective, but its range is dense in X.

The final case is when $T - \lambda I$ is not surjective, and $\mathcal{R}(T - \lambda I) \subsetneq X$ is also not dense in X.

Definition 20.5. We define the *residual spectrum* of T, denoted $\sigma_r(T)$, as the set of λ for which the range of $T - \lambda I$ is not dense in X.

Practically, this ends up being the least tractable to understand. It makes sense that if you know what's going on with an operator on a dense subspace, you can still understand what's going on with the operator in general; but when the residual spectrum shows up, things get harder.

Remark 20.6. In finite dimensions, the spectrum is just the point spectrum — because in finite dimensions, T is bijective if and only if T is one-to-one, implying that $\sigma(T) = \sigma_p(T)$. So we can fully understand an operator just by looking at its eigenvalues (this is a linear algebra problem). So finite dimensions is sort of the solved case.

§20.2 An example

Example 20.7

Take $H = L^2([0,1])$, and define $M_x\colon H \to H$ as the multiplication operator $M_x(f)\colon x \mapsto x \cdot f(x)$.

On the problem set, we showed that M_x has no eigenvalues, meaning that $\sigma_p(M_x) = \emptyset$. But M_x does have lots of things in its spectrum:

Claim 20.8 — We have $[0,1] \subseteq \sigma(M_x)$.

(This means $\lambda \in \sigma(M_x)$ for all $\lambda \in [0,1]$. Note that these λ are real numbers, but we still view them as part of the complex plane.)

We'll just show this for $(0,1)$, but the argument extends to the endpoints too.

First, we'll recall a few facts from earlier in the course:

Fact 20.9 — Let $T \in B(H,H)$ be invertible. Then T is bounded below.

Proof. Note that $\|x\| = \|T^{-1}Tx\| \leq \|T^{-1}\|\,\|Tx\|$. This means we have

$$\|Tx\| \geq \frac{\|x\|}{\|T^{-1}\|}$$

for all $x \in H$. □

(The converse is also true, but this is the easy direction and all we need for this proof.)

Proof of claim. Let $\lambda \in (0,1)$. Then note that $M_x - \lambda I = M_{x-\lambda}$ (where this is the function defined as $M_{x-\lambda}f\colon x \mapsto xf(x) - \lambda f(x) = (x-\lambda)f(x)$). Choose n large enough such that $[\lambda - \frac{1}{n}, \lambda + \frac{1}{n}] \subseteq [0,1]$, and let $g_n \in L^2([0,1])$ be defined as

$$g_n(x) = \sqrt{\frac{n}{2}}\mathbf{1}_{E_n}$$

for $n \in \mathbb{N}$, where E_n is the interval $[\lambda - \frac{1}{n}, \lambda + \frac{1}{n}]$. Then

$$\|g_n\|^2 = \int_{E_n}\left(\sqrt{\frac{n}{2}}\right)^2 = 1.$$

Furthermore, we have

$$\|(M_x - \lambda I)g_n\|^2 = \int_{E_n}(x-\lambda)^2\frac{n}{2}\,dx$$

(the support of g_n is contained in E_n, so we'll always just be integrating over E_n). But $|x - \lambda| \leq \frac{1}{n}$ for all $x \in E_n$, so when we integrate, we get that this is bounded by $1/n^2$.

But this converges to 0. So we've taken $M_x - \lambda I$ and applied it to a sequence of norm 1, and shown that the norms of what we got converges to 0. This means $M_x - \lambda I$ is *not* bounded below, so it can't be invertible; this means $\lambda \in \sigma(M_x)$. □

(You can generalize this to the endpoints too, though we won't do this.)

Claim 20.10 — If $\lambda \in \mathbb{C} \setminus [0,1]$, then $M_x - \lambda I$ is invertible.

So this means $\sigma(M_x)$ is *precisely* $[0,1]$ — we've already proven that $[0,1]$ is part of the spectrum, and now we're going to prove that it's actually the entire spectrum.

Proof. Let $\lambda \notin [0,1]$. Then $\frac{1}{x-\lambda}$ is bounded on $[0,1]$, and $M_{1/(x-\lambda)}$ (multiplication by the function $\frac{1}{x-\lambda}$) is the bounded inverse of $M_x - \lambda I$. □

Remark 20.11. If we have a bounded linear map and we know it's invertible, then we know its inverse is bounded by the open mapping theorem.

So now we have one example about thinking about the spectrum when it's not just the point spectrum. Now we'll shift to talking about the Fredholm alternative. First we'll recall some facts about quotient spaces (since we haven't talked about them too much).

§20.3 Quotient spaces

Definition 20.12. For X a Banach space and $M \subseteq X$ a closed subspace, the *quotient space* X/M (pronounced X *mod* M) is the set of equivalence classes where $x \sim y$ if and only if $x - y \in M$ (equivalently, the set of cosets of M).

We can turn X/M into a normed space.

Definition 20.13. For $x \in X$, we define

$$\|x\|_{X/M} = \inf\{\|x+m\| \mid m \in M\}.$$

(We may also write $x+M$ to denote the element of X/M corresponding to x, so we'd write this as $\|x+M\|$.)

Given $T\colon X \to Y$, we can define a map $A\colon X/\ker T \to Y$ by

$$A(x + \ker T) = Tx.$$

It's easy to show that A is well-defined (it's independent of our choice of the element of $\ker T$ to use as our representative), one-to-one, and is bounded (with the same bound as T). If T is surjective, then $X/\ker T$ is isomorphic to $\mathcal{R}(T)$ (here we just mean an isomorphism, not necessarily an isometry) — this is because A is a bijection, and by the inverse mapping theorem (a consequence of the open mapping theorem), we have that A is invertible (i.e., $A^{-1} \in B(Y, X/\ker T)$).

Further, this means that if $T \in X'$, then $X/\ker T$ is isomorphic to the field $\mathbb{F}$ (again, this isomorphism is not necessarily an isometry).

We're going to use these facts in what follows.

§20.4 Fredholm alternative

Our first theorem to get to the Fredholm alternative is the following.

Theorem 20.14

Let X be Banach and let $T \in B(X,X)$ be compact. Then for $\lambda \neq 0$, the range of $T - \lambda I$ is closed.

Proof. We can rewrite $T - \lambda I = \lambda(\frac{1}{\lambda}T - I)$, and if T is compact then $\frac{1}{\lambda}T$ is compact (and if $\frac{1}{\lambda}T - I$ is closed, so is λ times this). So without loss of generality, we can assume $\lambda = 1$ — then we want to show that if T is compact, then $T - I$ has closed range. (This is not really important for the proof, but it means we don't have to write λ anymore.)

Assume for contradiction that the range of $T - I$ is not closed. Then the map $S\colon X/\ker T \to X$ defined as we've set up earlier, as

$$S(x + \ker T) = (T - I)x,$$

is well-defined, one-to-one, bounded, and because $\mathcal{R}(S) = \mathcal{R}(T - I)$, by our assumption $\mathcal{R}(S)$ is not closed.

Claim 20.15 — S is not bounded below.

Proof. We already know $S \in B(X, X)$; assume for contradiction that S is bounded below.

We want to use the fact that T is compact. There's something nice about compact operators having closed range — a compact operator means that any bounded sequence in the pre-image has a convergent subsequence in the image. So we want to use that idea to show in this case that S cannot be bounded below.

So given a sequence $(x_n)_{n\in\mathbb{N}}$, we define $(Sx_n) = (y_n)_{n\in\mathbb{N}}$ converging to some y. Then $(y_n)_{n\in\mathbb{N}}$ is Cauchy, which means

$$\|y_n - y_m\| = \|Sx_n - Sx_m\|.$$

Since S is bounded below, this is bounded below by $\delta \|x_n - x_m\|$, which means

$$\|x_n - x_m\| \le \delta \|y_n - y_m\|.$$

And since (y_n) is Cauchy, this means (x_n) is too, so $x_n \to x$ because X is Banach. This implies that $Sx = y$, because $y_n \to y$ and $x_n \to x$, so $Sx_n \to Sx$, which is y. This implies S is closed, which is a contradiction because we assumed S is not closed. □

Now because S is not bounded below, there exists a sequence of points $(x_n)_{n\in\mathbb{N}}$ in $X/\ker(T - I)$ such that $\|x_n\|_{X/\ker T} = 1$ (these really are $x_n + \ker(T - I)$), and such that $S(x + \ker(T - I)) = (T - I)x_n \to 0$. (Here we're just stating what it means for S to not be bounded below.)

What does it mean for (x_n) to have horm 1? The norm on the quotient space is an infimum, which means e.g. that for all n, there exists $y_n \in \ker(T - I)$ such that $\|x_n + y_n\|_X \le 2$. And $x_n + y_n \sim x_n$ (meaning they're in the same coset).

Because T is compact, we know (Tx_n) has a convergent subsequence, so Tx_{n_k} converges to some $y \in X$. And because $(T - I)x_n \to 0$, we must have $x_{n_k} \to y$. We now apply T to the last statement, and by continuity we get $Tx_{n_k} \to Ty$. This implies $Ty = y$ (because Tx_{n_k} converges to both y and Ty). This means $y \in \ker(T - I)$.

But this is a contradiction. Why? We have that (x_{n_k}) converges in the quotient space $X/\ker(T - I)$ to y. But in this quotient space, x_{n_k} all have norm 1; and meanwhile $\|y\|$ in this space is 0. So we can't have that these are converging in the quotient space yet have different norm.

So we must have that $\mathcal{R}(T - I)$ is closed. □

Remark 20.16. We should have defined $S\colon X/\ker(T - I) \to X$.

Remark 20.17. The first claim is true for any operator that's not closed — if you're not closed, then you're not bounded below.

Remark 20.18. Note that the sequence $(x_n + y_n)$ is bounded in X, so we're picking out a bounded sequence; and that bounded sequence is what gives us a contradiction because of the compactness of T.

Theorem 20.19

Let H be a Hilbert space, and let $T\colon H \to H$ be compact, and let $M_j = \mathcal{R}(T-I)^j$ (for any $j \in \mathbb{N}$). Then there exists j_0 such that $M_{j_0} = M_{j_0+1}$.

We won't make this argument right now, but it's believable — when $j = 1$ we just proved that $\mathcal{R}(T-I)$ is closed, and this implies immediately that $\mathcal{R}(T-I)^j$ is closed for all j. So you can show that M_j is closed (as in the previous theorem — it's just an iteration of the previous result) for all $j \in \mathbb{N}$. And then if we want to show that eventually $M_{j_0} = M_{j_0+1}$ for a compact operator, the argument looks like — suppose not, meaning that $M_{j+1} \subsetneq M_j$ for all j. Then we can build a bounded sequence where you build sequences that are in different M_j's — each element you choose is going to be in M_j but not M_{j+1}. It's immediate to come up with a sequence $x_j \in M_j \setminus M_{j+1}$ which is bounded but does not have a convergent subsequence; this will imply that T is not compact.

So we build a bounded sequence of points $x_j \in M_j \setminus M_{j+1}$ such that Tx_j does not have a convergent subsequence, but x_j is bounded; this contradicts compactness of T.

This result says something interesting — at some point $T - I$ stops changing your space. This actually tells us something very interesting, the Fredholm alternative.

Theorem 20.20 (Fredholm alterntive)

Let T be a compact operator on a Hilbert space H, and let $\lambda \in \mathbb{C} \setminus \{0\}$. Then:

(a) If $T - \lambda I$ is injective, then $T - \lambda I$ is invertible.

(b) If $T - \lambda I$ is surjective, then $T - \lambda I$ is invertible.

Again, in finite dimensions this isn't interesting; what's interesting is infinite dimensions where it's not immediate.

Remark 20.21. There's a nice interpretation of this setup — if you think of $(T - \lambda I)x = y$ as an equation where you're given y and you want to solve for x, then (a) says that if a solution always is unique then it always exists, and (b) says that if a solution always exists then the solutions are always unique.

§21 May 1, 2024

§21.1 Review

Theorem 21.1

Let H be a Hilbert space, and let $T\colon H \to H$ be compact. For each $j \in \mathbb{N}$, define $M_j = \mathcal{R}(T-I)^j$. Then there exists $j_0 \in \mathbb{N}$ such that $M_{j_0} = M_{j_0+1}$.

We didn't prove this last time, but it's worth going through.

The big-picture idea is that T is a compact operator; and we showed that then T can be approximated by finite-rank operators. So we have to have this sort of stability result.

Proof. Last class, we saw a theorem that if $T \in B(H,H)$ is compact, then $\mathcal{R}(T-I)$ is closed. In particular, this means M_1 is closed.

Now we can expand

$$(I-T)^j = I - jT + \frac{j(j-1)}{2}T^2 - \cdots + (-1)^j T^j.$$

(This si just the binomial expansion theorem.) Letting this be $I - A$, because T is compact we have that A is also compact. And then we can apply the theorem from last class to also say that $M_j = \mathcal{R}(I - A)$ is also closed.

Immediately we have $M_{j+1} \subseteq M_j$; assume for contradiction that this containment is strict (for all $j \in \mathbb{N}$), so that $M_{j+1} \subsetneq M_j$. Then $\dim(M_j/M_{j+1}) \geq 1$, which means there exists $x_j \in M_j$ such that $\|x_j + M_{j+1}\|_{M_j/M_{j+1}} = 1$. And like last time, if we have a bounded sequence in the quotient space norm, then in fact we can assume it's bounded in the original norm by e.g., 2 — so we can assume $\|x_j\|_H \leq 2$ for all $j \in \mathbb{N}$ (we choose an appropriate element of M_j to add to x_j).

Now we have a bounded sequence, so by the compactness of T its image should have a convergent subsequence; we're going to show that this is not the case.

Specifically, we claim that $\|Tx_j - Tx_k\| \geq 1$ for all $j \neq k$; if we prove this, it'll contradict the compactness of T.

Suppose that $j < j+1 \leq k < k+1$. Then $x_k \in M_k \subseteq M_{j+1}$, and $(T-I)x_j \in M_{j+1}$ as well, and finally $(T-I)x_k \in M_{k+1} \subseteq M_{j+1}$. Now let

$$y = (T-I)x_j - (T-I)x_k + x_k.$$

Each of the terms on the right is an element of M_{j+1}, so $y \in M_{j+1}$ as well. But now we have

$$\|x_j + y\| \geq \inf_{m \in M_{j+1}} \|x_j + m\| = \|x_j\|_{M_j/M_{j+1}} = 1,$$

because $y \in M_{j+1}$ and we defined x_j such that $\|x_j\|_{M_j/M_{j+1}} = 1$. But this is equal to $\|Tx_j - Tx_k\|$, so we've shown that $\|Tx_j - Tx_k\| \geq 1$, as desired. □

Again, the idea is that we looked at the expansion, which was also of the form I minus a compact operator and therefore has closed range. And then using the geometric structure of the Hilbert space, we picked out a sequence that is bounded but does not have a convergent subsequence, contradicting the compactness of T.

§21.2 Fredholm alternative

Theorem 21.2

Suppose that T is a compact operator on a Hilbert space H (over $\mathbb{C}$), and let $\lambda \neq 0$. If $T - \lambda I$ is not invertible, then λ is an eigenvalue of T.

This is equivalent to saying that $\sigma(T) = \sigma_p(T)$ — the only complex numbers in the spectrum are eigenvalues.

In the big picture, since compact operators are limits of finite-rank ones, we'd expect them to share properties. For *finite-rank* operators, $T - \lambda I$ fails to be invertible if and only if it fails to be injective; so those have this property. And now we're saying compact operators, which are the limit of finite-rank ones, also have this property.

Proof. Without loss of generality we can assume $\lambda = 1$. Assume that $T - I$ is not invertible, and suppose for contradiction that 1 is not an eigenvalue for T — so $1 \in \sigma(T) \setminus \sigma_p(T)$. This means $\ker(T - I) = \{0\}$ (in other words, we're just saying that $T - I$ is injective); but $T - I$ is *not* onto. In particular, this means $(T-I)H \subsetneq H$ (the notation on the left-hand side means the range of $T - I$).

And then $(T-I)^2H \subsetneq (T-I)H$ for the same reason (here you use the fact that $T-I$ is invertible), and so on. And this implies $(T-I)^{j+1}H \subsetneq (T-I)^jH$ for all j. But now this immediately contradicts the last theorem, which says that at some point this must stabilize. □

Theorem 21.3 (Fredholm alternative)

Let T be a compact operator on the Hilbert space H (over $\mathbb{C}$), and let $\lambda \in \mathbb{C} \setminus \{0\}$.

(a) If $T-\lambda I$ is injective, then it is invertible.

(b) If $T-\lambda I$ is onto, then it is invertible.

Proof. First, (a) is just the previous theorem. And we can actually get (b) by doing a little trick with the adjoint — suppose $(T-\lambda I)$ maps H onto H (i.e., it is surjective). Then the Hilbert adjoint $(T-\lambda I)^*$ is $T^* - \overline{\lambda}I$, and it must be injective. To see this, on a Hilbert space H, if we have $S \in B(H,H)$, we claim that $\ker(S) = \mathcal{R}(S^*)^\perp$. To see this, let $x \in H$. We have $x \in \ker(S)$ if and only if $\langle Sx, y\rangle = 0$ for all $y \in H$; this is equivalent to $\langle x, S^*y\rangle = 0$ for all $y \in H$, which is exactly the same as saying $x \perp \mathcal{R}(S^*)$. So this tells us $\ker(S) = \mathcal{R}(S^*)^\perp$, which means that if S^* is onto then S is injective.

So now $(T-\lambda I)^*$ is injective, and we can just apply (a) to $(T-\lambda I)^*$ — this gives that $(T-\lambda I)^*$ is invertible. Let W be its inverse. Then $\langle (T-\lambda I)x, Wy\rangle = \langle x, (T-\lambda I)^*Wy\rangle = \langle x, y\rangle$. But this is also $\langle W^*(T-\lambda I)x, y\rangle$, which implies that W^* is the inverse of $T-\lambda I$. So $T-\lambda I$ is invertible, since we've figured out what its inverse is. □

§21.3 The resolvent

(This part is from Sections 7.2 and 7.3 of Kreyszig.)

So far, we've been working with the spectum of compact operators because this works very nicely. But now we're going to talk about the spectrum in more general cases.

Definition 21.4. Let X be a nontrivial normed linear space (not necessarily Banach) over $\mathbb{C}$, and let $T\colon \mathcal{D}(T) \to X$ (where $\mathcal{D}(T) \subseteq X$) be a linear operator. Suppose $T-\lambda I$ is invertible; then we denote its inverse by $R_\lambda(T) = (T-\lambda I)^{-1}$ and call it the *resolvent operator* of T.

Remark 21.5. Last time, we discussed thinking of $(T-\lambda I)x = y$ as an equation we want to solve, where we're given y and want to solve for x; the reason we call this the resolvent is because when $T-\lambda I$ is invertible, the inverse is exactly how you solve for x — we apply the inverse to both sides, and we end up wtih $x = R_\lambda(T)y$.

Definition 21.6. We say $\lambda \in \mathbb{C}$ is a *regular value* of T if $T-\lambda I$ is invertible, and we define the *resolvent set* of T, denoted by $\rho(T)$, as the set of regular values of T.

In particular, this means $\mathbb{C} = \rho(T) \cup \sigma(T)$.

Our first result is about inverses.

Theorem 21.7

Let $T \in B(X,X)$ where X is Banach. If $\|T\| < 1$, then $(I-T)$ is invertible with inverse $\sum_{j=0}^\infty T^j = I + T + T^2 + \cdots$.

This is what we'd hope for out of an inverse, but it does require the completeness of X. Note that once we have that $I - T$ is invertible, we know that $(I-T)^{-1} \in B(X,X)$ by the open mapping theorem.

Proof. First note that $\|T^j\| = \|T \circ \cdots \circ T\| \leq \|T\|^j$. So because of this, we have that for $\|T\| < 1$, we have

$$\left\| \sum_{j=)}^{\infty} T^j \right\| \leq \sum_{j=0}^{\infty} \|T\|^j .$$

But this is a geometric series, and since $\|T\| < 1$ it's convergent. Since X is complete, so is $B(X,X)$ (recall that we proved earlier in the course that $B(X,Y)$ is complete whenever Y is complete). And this means the absolute convergence of the above sum implies that the operators $\sum_{j=0}^{\infty} T^j$ converge to some operator S. (This is only true because $B(X,X)$ is complete, and the above absolute convergence implies that this sequence is Cauchy.)

Now we need to see that $S = (I-T)^{-1}$. We'll do this by computing the products $(I-T)S$ and $(T-I)S$. We want to show that $(I-T)S = I$; but because S is this limit, we're going to actually show that $(I-T)\sum_{j=0}^{n} T^j \to I$. To see this, we can just write this out, and we'll get a telescoping sum $\sum_{j=0}^{n} T^j - \sum_{j=0}^{n} T^{j+1} = I - T^{n+1}$. And now the point is that $T^{n+1} \to 0$, because $\|T^{n+1}\| \leq \|T\|^{n+1} \to 0$ (as $n \to \infty$). So we have $(I-T)S = I$; you can also show that $S(I-T) = I$ (in the same way), implying that $S = (I-T)^{-1}$. □

Theorem 21.8

Let $T \in B(X,X)$ for X a Banach space over $\mathbb{C}$. Then $\rho(T)$ is open (equivalently, $\sigma(T) = \mathbb{C} \setminus \rho(T)$ is closed).

Proof. If $\rho(T)$ is empty, then there's nothing to show. (It's actually not possible for this to happen, which we will prove shortly, but that doesn't matter here.) Assume $\rho(T)$ is nonempty, and let $\lambda_0 \in \rho(T)$. Then for any $\lambda \in \mathbb{C}$, we can write

$$T - \lambda I = (T - \lambda_0 I) - (\lambda - \lambda_0)I = (T - \lambda_0)(I - (\lambda - \lambda_0)(T - \lambda_0 I)^{-1}).$$

Our textbook calls $T - \lambda I = T_\lambda$, so in this shorthand this equation would read

$$T_\lambda = T_{\lambda_0}(I - (\lambda - \lambda_0)R_{\lambda_0}(T)).$$

Now for this problem, we'll call $I - (\lambda - \lambda_0)R_{\lambda_0}(T) = V$; then this tells us $T_\lambda = T_{\lambda_0}V$. Since $\lambda_0 \in \rho(T)$ and $T \in B(X,X)$, we know $R_{\lambda_0}(T)^{-1} = T_{\lambda_0} \in B(X,X)$.

We defined $V = I - (\lambda - \lambda_0)(R_{\lambda_0}(T))$; and as long as $\|(\lambda - \lambda_0)R_{\lambda_0}(T)\| < 1$, we get that V is invertible with inverse $V^{-1} = \sum_{j=0}^{\infty} (\lambda - \lambda_0)^j R_{\lambda_0}(T)^j$. In particular, V is invertible (with $V^{-1} \in B(X,X)$) as long as $\|(\lambda - \lambda_0)R_{\lambda_0}(T)\| < 1$. And we can just pull out the scalar — this is true as long as

$$|\lambda - \lambda_0| < \frac{1}{\|R_{\lambda_0}(T)\|}.$$

(We might write R_{λ_0} instead of $R_{\lambda_0}(T)$ because this notation is confusing.)

In particular, V^{-1} exists and is bounded as long as λ is close to λ_0; and to finish, we had

$$T_\lambda = T_{\lambda_0}V.$$

And T_{λ_0} is invertible with inverse $R_{\lambda_0}(T) \in B(X,X)$. So we get that if

$$|\lambda - \lambda_0| < \frac{1}{\|R_{\lambda_0}(T)\|},$$

then T_λ is also invertible with inverse $(T_{\lambda_0}V)^{-1} = V^{-1}T_{\lambda_0}^{-1} = V^{-1}R_{\lambda_0}$.

(In fact, this even gives a formula for R_λ — it's $V^{-1}R_{\lambda_0}$, and we saw an explicit formula for V^{-1}.)

So what we've shown is that the radius-$\frac{1}{\|R_{\lambda_0}\|}$ ball centered at λ_0 is contained in $\rho(T)$. This shows $\rho(T)$ is open, so we're done. □

In the big picture, the way this proof works is we started with some λ_0 in the resolvent; this means in particular that $\|R_{\lambda_0}\|$ is some finite number. And we showed that if we choose λ sufficiently close to λ (depending on the above norm), then that is also in the resolvent, using the above theorem.

So the resolvent is open, which means the spectrum is closed.

Corollary 21.9

For every $\lambda_0 \in \rho(T)$, $R_\lambda(T)$ can be represented as

$$R_\lambda = \sum_{j=0}^{\infty} (\lambda - \lambda_0)^j R_{\lambda_0}^{j+1}$$

as long as $|\lambda - \lambda_0| < \frac{1}{\|R_{\lambda_0}\|}$.

Of course, this condtiion on λ defines an open disk in $\mathbb{C}$.

This is kind of a continuity result about what happens near a resolvent value and how you can write the resolvent operator using the resolvent operator nearby.

The next result is a nice consequence of this. We had a theorem about the eigenvalues of an operator — we had a theorem that says either $\|T\|$ or $-\|T\|$ is always an eigenvalue, under certain conditions (specifically, compact self-adjoint operators on a Hilbert space). Today we'll look at a more general result.

Theorem 21.10

Let $T \in B(X, X)$ where X is a Banach space over $\mathbb{C}$. Then $\sigma(T)$ is contained in the disk $|\lambda| \le \|T\|$.

In the case where the operator was compact and self-adjoint we showed the largest eigenvalue had absolute value *exactly* the norm of the operator. But that's just the point spectrum; now we're looking at the entire spectrum, and also the more general case that X is Banach.

This also immediately implies that $\rho(T) = \mathbb{C} \setminus \sigma(T)$ contains the complement of this disk, so in particular it's nonempty.

Proof. Let $\lambda \in \mathbb{C} \setminus \{0\}$ and let $\kappa = 1/\lambda$; and let $R_\lambda = (T - \lambda I)^{-1} = -\frac{1}{\lambda}(I - \kappa T)^{-1}$ (we're just moving things around to get things into a structure where we can use the earlier theorem on inverses). So if $|\kappa| \, \|T\| < 1$, then we can write this as

$$-\frac{1}{\lambda}\sum_{j=0}^{\infty}(kT)^j = -\frac{1}{\lambda}\sum_{j=0}^{\infty}(\frac{1}{\lambda}T)^j.$$

This is true if $\left\|\frac{1}{\lambda}T\right\| < 1$, equivalently $|\lambda| > \|T\|$. So we've shown that whenever $|\lambda| > \|T\|$. This tells us that if $|\lambda| > \|T\|$ then $\lambda \in \rho(T)$, which means the spectrum must be contained in the radius-$\|T\|$ ball. □

Definition 21.11. The *spectral radius* of T (for $T \in B(X,X)$ where X is Banach), denoted $r_\sigma(T)$, is defined as

$$r_\sigma(T) = \sup_{\lambda \in \sigma(T)} |\lambda| .$$

Immediately, what we just showed means $r_\sigma(T) \leq \|T\|$. Next time we're hopefully going to prove that in fact, $r_\sigma(T) = \lim_{n\to\infty} \|T^n\|^{1/n}$. This should look kind of like a result from real analysis.

§22 May 6, 2024

§22.1 Equivalence of two definitions

We've seen two different definitions from the two textbooks, but they're actually equivalent; we'll first point out why.

We started with the definition in MacCluer, where X is Banach and $T \in B(X,X)$. We said $\lambda \in \rho(T)$ if $(T - \lambda I)$ is invertible. And then we said there are two ways $T - \lambda I$ could fail to be invertible:

(1) It's not injective — in this case, we say λ is an element of the *point spectrum* $\sigma_p(T)$ (i.e., it's an eigenvalue).

(2) It's not onto, in which case there are two things that could happen:
- If it's onto a dense subspace of X, we say λ is in the *continuous spectrum* $\sigma_c(T)$.
- If not, we say it's in the *residual spectrum.*

Here we started with the assumption that X is Banach and T (and therefore $T - \lambda I$) is bounded. Note that here we view $T - \lambda I$ as a map $X \to X$.

In the Kreyszig textbook, there's no assumption on X (i.e., it's just a normed linear space, not necessarily Banach), and T is just a linear operator. Kreyszig is taking a more general definition of what it means to be an element of the resolvent; he says $\lambda \in \rho(T)$ if the following three conditions hold:

(1) $T - \lambda I$ is injective (equivalently, it's invertible as a map from X to its range $\mathcal{R}(T - \lambda I)$ — Kreyszig uses invertible in this sense).

(2) $(T - \lambda I)^{-1}$ is bounded.

(3) $T - \lambda I$ is onto a dense subset of X.

If X is Banach and T is bounded, then this gives the same definition as in MacCluer. Here's why the Kreyszig definition gives us the MacCluer theorem — if X is Banach and $T \in B(X,X)$, then if $T - \lambda I$ is injective *and* surjective, then $(T - \lambda I)^{-1}$ is bounded by the corollary to the open mapping theorem for inverse maps. And the third condition of Kreyszig says that $\overline{\mathcal{R}(T - \lambda I)} = X$; if this is true but $\mathcal{R}(T - \lambda I)$ itself is not X, then we can prove that $(T - \lambda I)$ is not bounded below (using the fact that T is bounded and X is Banach), which implies $(T - \lambda I)^{-1}$ is not bounded, contradicting (2).

So this implies $\mathcal{R}(T - \lambda I) = X$, so $T - \lambda I$ is onto; and that means (together with the fact it's injective) that $(T - \lambda I)^{-1}$ is bounded.

So in the MacCluer textbook we started with a Banach space and bounded linear operator; in the Kreyszig textbook we just start with a normed linear space and a general linear map, and we're trying to generalize things. And the Kreyszig definition, plus the assumption that X is Banach and T is bounded, gives the MacCluer definition.

§22.2 Properties of resolvent operators

We've seen a few small theorems from Chapter 9 about compact self-adjoint operators that we'll get to next class. But we're also building to a result about the spectral result for a general operator; today we're going to show a couple of technical components of that argument and recall some complex analysis (we're not assumed to know complex analysis, but we'll need some in order to prove this result).

Theorem 22.1

Let X be a complex Banach space, $T \in B(X, X)$, and $\lambda, \mu \in \rho(T)$. Then:

(a) $R_\lambda := (T - \lambda I)^{-1}$ satisfies

$$R_\mu - R_\lambda = (\mu - \lambda) R_\mu R_\lambda.$$

(b) R_λ commutes with any $S \in B(X, X)$ that commutes with T.

(c) R_λ commutes with R_μ (i.e., $R_\lambda R_\mu = R_\mu R_\lambda$).

We use R_λ to denote the resolvent operator of T with respect to λ. We also use T_λ to denote $T - \lambda I$.

Proof. These are pretty direct. For (a), because X is Banach and $T \in B(X, X)$ and $\lambda \in \rho(T)$, we know that $\mathcal{R}(T_\lambda) = X$ (this just says that T_λ must be surjective — this is immediate from the MacCluer definition, and we just saw how it follows from the Kreyszig one). And $R_\lambda T_\lambda = I = T_\lambda R_\lambda$ (by the definition of an inverse). Similarly, we have $R_\mu T_\mu = I = T_\mu R_\mu$.

Together, these mean we can write

$$R_\mu - R_\lambda = R\mu(T_\lambda R_\lambda) - (R_\mu T_\mu) R_\lambda$$

(we just threw in two extra factors of the identity). And we can now write this as

$$R_\mu (T_\lambda - T_\mu) R_\lambda.$$

But $T_\lambda - T_\mu = (T - \lambda I) - (T - \mu I) = (\mu - \lambda) I$ (the T's cancel), and I is just the identity while $\mu - \lambda$ are scalars, so we can pull them out (using the linearity of R_μ and R_λ) to get that this becomes $(\mu - \lambda) R_\mu R_\lambda$.

For (b), suppose that $S \in B(X, X)$ commutes with T, meaning that $ST = TS$. Then this means $ST_\lambda = S(T - \lambda I) = (T - \lambda I)S = T_\lambda S$ (since S commutes with both T and I). And then because $I = T_\lambda R_\lambda = R_\lambda T_\lambda$, we get that

$$R_\lambda S = R_\lambda S(T_\lambda R_\lambda) = R_\lambda T_\lambda S R_\lambda = S R_\lambda$$

(the first equality corresponds to just throwing in the identity, and the second is because T_λ and S commute).

Finally for (c), by (b) we have that R_μ commutes with T (since T commutes with itself), which then implies R_λ commutes with R_μ. □

§22.3 Polynomials and the spectrum

These are the technical results we need for our next theorem. Now we're going to recall a theorem from linear algebra.

Recall that in linear algebra, if λ is an eigenvalue of a matrix A, then λ^n is an eigenvalue of A^n. For example, if $Ax = \lambda x$, then $A^2 x = A(Ax) = A(\lambda x) = \lambda Ax = \lambda^2 x$; this immediately shows you λ^2 is an eigenvalue of A^2, and you can inductively continue to see that λ^n is an eigenvalue of A^n.

Furthermore, if $p(\lambda)$ is a polynomial in λ, i.e., $p(\lambda) = \alpha_n \lambda^n + \alpha_{n-1}\lambda^{n-1} + \cdots + \alpha_0$, then if λ is an eigenvalue of A, then $p(\lambda)$ is an eigenvalue of $p(A)$, where $p(A)$ is the polynomial of the matrix A — i.e., $\alpha_n A^n + \cdots + \alpha_0 I$.

This may seem a bit obscure, but this generalizes really beautifully to bounded linear operators as well; this is an important result that will help us deduce the mentioned theorem about the spectral radius.

Theorem 22.2

Let X be Banach and let $T \in B(X, X)$, and let $p(\lambda) = \alpha_n \lambda^n + \cdots + \alpha_0$ be a polynomial (where $\alpha_n \neq 0$). Then $\sigma(p(T)) = p(\sigma(T))$.

On the left-hand side $p(T)$ is an operator (namely, $\alpha_n T^n + \cdots + \alpha_0 I$), so $\sigma(p(T))$ is a subset of $\mathbb{C}$; on the right-hand side $\sigma(T)$ is a subset of $\mathbb{C}$, and then we're applying the polynomial p, so that produces another subset of $\mathbb{C}$.

Proof. We'll assume $\sigma(T) \neq \emptyset$; we'll prove this soon (without relying on this theorem, so it won't be circular).

If $n = 0$, this result is trivial — then $p(T)$ is just $\alpha_0 I$, so the left-hand side is $\sigma(\alpha_0 I) = \{\alpha_0\}$. Meanwhile on the right-hand side, applying p to any subset of $\mathbb{C}$ just gives α_0, so $p(\sigma(T)) = \{\alpha_0\}$ as well.

So we now assume $n \neq 0$. We'll first show that $\sigma(p(T)) \subseteq p(\sigma(T))$. To do so, we define $S = p(T)$, and $S_\mu = p(T) - \mu I$ (for any $\mu \in \mathbb{C}$). If S_μ^{-1} exists (i.e., S_μ is bijective), then $\mu \in \rho(p(T))$ and $S_\mu^{-1} = R_\mu(p(T))$ (this is essentially by definition).

Now fix $\mu \in \mathbb{C}$, and let $s_\mu(\lambda) = p(\lambda) - \mu$ be the corresponding polynomial in λ. This polynomial factors over $\mathbb{C}$ (polynomials always factor over $\mathbb{C}$), so we can write $s_\mu(\lambda) = \alpha_n(\lambda - \gamma_1) \cdots (\lambda - \gamma_n)$ for some $\gamma_1, \ldots, \gamma_n \in \mathbb{C}$. This then means that we can also factor

$$S_\mu = \alpha_n (T - \gamma_1 I)(T - \gamma_2 I) \cdots (T - \gamma_n I).$$

For each $j \in [n]$, if $\gamma_j \in \rho(T)$ then $T - \gamma_j I$ is invertible, and its inverse is in $B(X, X)$. This means if $\gamma_j \in \rho(T)$ for all $j \in [n]$, then we can write

$$S_\mu^{-1} = \frac{1}{\alpha_n}(T - \gamma_n I)^{-1} \cdots (T - \gamma_1 I)^{-1}.$$

This immediately implies S_μ^{-1} exists. Equivalently, if $\mu \notin \rho(p(T))$ then $\mu \in \sigma(p(T))$, which means there exists $j \in [n]$ such that $T - \gamma_j I$ is not invertible, meaning that $\gamma_j \in \sigma(T)$. But if we plug γ_j into our polynomial, then we get $s_\mu(\gamma_j) = p(\gamma_j) - \mu = 0$, so $p(\gamma_j) = \mu$. This tells us $\mu \in p(\gamma_j) \subseteq p(\sigma(T))$.

So we've shown one direction of the containment; now we'll show the other direction, that $p(\sigma(T)) \subseteq \sigma(p(T))$. To do so, let $\kappa \in p(\sigma(T))$, so there is some $\beta \in \sigma(T)$ with $\kappa = p(\beta)$.

Since $\beta \in \sigma(T)$, a couple of things can happen.

Case 1 ($T - \beta I$ is not injective). This case means β is an eigenvalue of T, or equivalently $T - \beta I$ has no inverse map $\mathcal{R}(T - \beta I) \to X$. Now if we define the polynomial $s_\kappa(\lambda) = p(\lambda) - \kappa$, we know that β is a root of $s_\kappa(\lambda)$. Then as we did a moment ago, we can factor $s_\kappa(\lambda)$ as $(\lambda - \beta)g(\lambda)$, where $g(\lambda)$ is a degree-$(n-1)$ polynomial. And similarly to before, we can define the polynomial *operator* $S_\kappa = p(T) - \kappa I = (T - \beta I)g(T)$. Note that $g(T)$ is again a degree-$(n-1)$ polynomial, so we actually know $T - \beta I$ commutes with every element in that factor expansion; so we could also write this as $g(T)(T - \beta I)$.

Now we claim that S_κ does not have an inverse. To see this, suppose that it does; then we're going to show this implies $T - \beta I$ also has an inverse, which we assumed is not true. We can write

$$I = S_\kappa S_\kappa^{-1} = (T - \beta I)g(T)S_\kappa^{-1}.$$

But $S_\kappa^{-1} g(T)(T - \beta I)$ is also the identity. And $g(T)S_\kappa^{-1} = S_\kappa^{-1} g(T)$ by the same remark about commuting; this implies $T - \beta I$ is invertible, which is a contradiction to our hypothesis.

And so S_κ^{-1} does not exist, which implies $\kappa \in \sigma(p(T))$.

So now we've proven this containment in the case $T - \beta I$ is not injective; we now just need to prove it in the case that $T - \beta I$ is not onto.

Case 2 ($T - \beta I$ is injective but not onto). Then if we look back at the definition of S_κ as $S_\kappa = (T - \beta I)g(T)$, since $T - \beta I$ is not onto, neither is S_κ; this means $\kappa \in \sigma(p(T))$. □

Remark 22.3. We're liberally using the first technical result when e.g. talking about S_κ^{-1} and $g(T)$ commuting.

§22.4 Complex analysis

We'll now talk a little about complex analysis. (We'll define enough things to be able to get to the theorem about spectral radius.)

Definition 22.4. A *domain* is an open, connected subset of $\mathbb{C}$.

(In $\mathbb{C}$, *connected* is the same as *path-connected* — so you can think of a set as connected if there's a continuous map from $[0, 1]$ connecting every pair of points.)

Definition 22.5. A map $h: \mathbb{C} \to \mathbb{C}$ is *holomorphic* (or *analytic*) on a domain $G \subseteq \mathbb{C}$ if h is differentiable on G, meaning that for every $\lambda \in G$, the limit

$$\lim_{\Delta\lambda \to 0} \frac{h(\lambda + \Delta\lambda) - h(\lambda)}{\Delta\lambda}$$

exists (we call this limit the derivative $h'(\lambda)$).

The important theorem from complex analysis is that this is equivalent to having a power series.

Definition 22.6. We say h is holomorphic at $\lambda_0 \in \mathbb{C}$ if h is holomorphic in a neighborhood of λ_0 (i.e., an open set containing λ_0).

Theorem 22.7

A function h is holomorphic on G if and only if for every $\lambda_0 \in G$, h has a power series representation at λ_0 with nonzero radius of convergence, i.e., for each λ_0, there exists some radius $r(\lambda_0) > 0$ such that we can write

$$h(\lambda) = \sum_{j=0}^{\infty} c_j (\lambda - \lambda_0)^j$$

for all λ such that $|\lambda - \lambda_0| < r(\lambda_0)$.

This is actually an incredible theorem — in calculus you introduce power series representations and there are conditions under which they converge. And here this says that a function is differentiable *if and only if* it can be represented as a power series in this way. This is very odd — we start with one derivative and get something involving infinitely many derivatives. (We won't prove this.)

Now we need to introduce the concepts we're going to use it with in operator theory.

Definition 22.8. Let $\Lambda \subseteq \mathbb{C}$. Then we call a function $S: \Lambda \to B(X, X)$ (we'll write $\lambda \mapsto S_\lambda$) an *operator function*.

(This is consistent with the notation we were using earlier — e.g., we used T_λ to denote $T - \lambda I$, which was a specific case of this.)

Definition 22.9. Let Λ be open and X a complex Banach space. Then we say an operator function S is *locally holomorphic on* Λ if for every $x \in X$ and for all $f \in X'$, the function $h(\lambda) = f(S_\lambda x)$ is holomorphic at every $\lambda \in \Lambda$.

We said S_λ is a bounded operator $X \to X$; then when we apply it to $x \in X$ we end up with an element of X. And then f is an element of the dual space, so the right-hand side gives us an element of $\mathbb{C}$.

> **Remark 22.10.** It's worth noting that there's a couple of ways we could say S_λ is 'differentiable,' which is what we're trying to do in some sense. We're doing this by saying that the corresponding complex-valued function $h: \mathbb{C} \to \mathbb{C}$ is differentiable. It's not immediately clear that we need complex analysis and the dual space, but as we continue we'll see that this is the right definition.

If h is holomorphic, this implies h' exists; that's the same as saying

$$\lim_{\Delta\lambda \to 0} \left| h'(\lambda) - \frac{h(\lambda + \Delta\lambda) - h(\lambda)}{\Delta\lambda} \right| = 0.$$

And our function h is supposed to be holomorphic for all $x \in X$ and $f \in X'$ (h depends on x and f, and this has to be true for each h). This will imply the other forms of convergence that you might expect — you get that

$$\left\| S'_\lambda - \frac{S_{\lambda+\Delta\lambda} - S_\lambda}{\Delta\lambda} \right\|_{B(X,X)} \to 0$$

as well (which also implies the other types of convergence).

Next class is the last new material before the final; we'll cover the theorem about the spectral radius, and then a few components of Chapter 9 building on things we've done in MacCluer. Next Monday we'll do a couple presentations, and then we'll try to look at the big picture of the material we've covered, talk a bit about the final, and answer any questions.

There's a pset due at the end of this week.

§23 May 8, 2024

PS7 is due Thursday; the last day we can have coursework due is Friday, but the graders won't look at it until Monday, so if we upload it by Monday it's fine.

The last content for the final will be today. On Monday there will be 1–2 presentations from us on some related interesting math; we'll go over the balance of the final (the content will be focused on the last third of the course — since the last midterm — but it will be comprehensive). And Marjie will talk about next courses at MIT — how functional analysis fits into other things in math. Her final office hours will be on the Tuesday before the final (May 21), rather than that Monday.

Today we'll continue where we left off, using some complex analysis to say something about the spectral radius of a bounded operator on a Banach space.

§23.1 Complex analysis

Recall that for $\Lambda \subseteq \mathbb{C}$ open and X a complex Banach space, we considered *operator functions* $S: \Lambda \to B(X, X)$ (so S takes in an element of Λ and sends it to a bounded linear operator — we denote it by $S_\lambda: \lambda \mapsto S_\lambda$).

> **Definition 23.1.** We say S is locally holomorphic if for all $x \in X$ and $f \in X'$, the function $h: \mathbb{C} \to \mathbb{C}$ defined by $h(\lambda) = f(S_\lambda x)$ is holomorphic.

(This means h is defined and differentiable on Λ.)

And we said that being holomorphic is equivalent to having a power series with nonzero radius of convergence — so for all $\lambda_0 \in \Lambda$ we can represent

$$h(\lambda) = \sum_{j=0}^{\infty} c_j (\lambda - \lambda_0)^j,$$

where this power series has a nonzero radius of convergence.

We're going to use this in a moment. There may also be other complex analysis tools that we'll comment on as we use them, but this is the setup we need for now.

§23.2 Holomorphicity of R_λ

We'll also recall one theorem from a couple of days ago.

Theorem 23.2

Let X be a complex Banach space and $T \in B(X, X)$, and let $\lambda_0 \in \rho(T)$. Then we can write $R_\lambda = (T - \lambda I)^{-1}$ as

$$R_\lambda = \sum_{j=0}^{\infty} (\lambda - \lambda_0)^j R_{\lambda_0}^{j+1}$$

for all λ with

$$|\lambda - \lambda_0| < \frac{1}{\|R_{\lambda_0}\|},$$

where this series is absolutely convergent.

We'll put these together to say something about R_λ.

Let X be a complex Banach space, T a bounded operator on X, $x \in X$, and $f \in X'$. Now we'll define h as before, but using R_λ instead of a general operator function — so we define $h(\lambda) = f(R_\lambda x)$, where $R_\lambda = (T - \lambda I)^{-1}$.

Claim 23.3 — The function h is holomorphic on $\rho(T)$.

Proof. Consider any $\lambda_0 \in \rho(T)$. Then Theorem 23.2 implies that $h(\lambda)$ can be represented as a power series $h(\lambda) = \sum_{j=0}^{\infty} (\lambda - \lambda_0)^j f(R_{\lambda_0}^{j+1} x)$ (note that f is linear, so we can move it inside the sum and pull out $(\lambda - \lambda_0)^j$). In particular, note that $f(R_{\lambda_0}^{j+1} x)$ is just some complex number (not depending on λ) for every j, and Theorem 23.2 tells us that this series is absolutely convergent if

$$|\lambda - \lambda_0| < \frac{1}{\|R_{\lambda_0}\|}.$$

This implies that h is holomorphic at λ_0 (for every $\lambda_0 \in \rho(T)$); so in particular, it's holomorphic on $\rho(T)$. □

So we've proved the following theorem:

Theorem 23.4

Let X be a complex Banach space and $T \in B(X, X)$. Then the resolvent operator $R_\lambda = R_\lambda(T)$ is locally holomorphic on $\rho(T)$.

§23.3 The spectrum is nonempty

Now we're finally going to show that the spectrum is nonempty for a bounded operator on a Banach space, using this idea and a little bit of complex analysis. Once we show the spectrum is nonempty, we can show other things — that the resolvent is in fact the *largest* set on which R_λ is locally holomorphic, and as promised, that the spectral radius of T follows the formula we discussed. That relies on this next result, which we mentioned last class.

Recall that for $\lambda \neq 0$, we can represent $R_\lambda = (T - \lambda I)^{-1}$ as

$$(-\lambda(I - \frac{1}{\lambda T}))^{-1} = -\frac{1}{\lambda}\left(I - \frac{1}{\lambda}T\right)^{-1}.$$

And if $|\lambda| > \|T\|$, then we can rewrite this using the inverse formula we proved — this is

$$-\frac{1}{\lambda}\sum_{j=0}^{\infty}\left(\frac{1}{\lambda}T\right)^{j}$$

(this is absolutely convergent if $\frac{1}{|\lambda|}\|T\| < 1$). Multiplying out, this says that $|\lambda| > \|T\|$.

In an earlier class, we used this to show that $\sigma(T) \subseteq \{|\lambda| \leq \|T\|\}$. But now we're going to use this formula to prove that the spectrum is nonempty.

Theorem 23.5

If $X \neq \{0\}$ is a nontrivial complex Banach space and $T \in B(X, X)$, then $\sigma(T)$ is nonempty.

Proof. If $T = 0$, then $\sigma(T) = \{0\}$ is nonempty; now suppose not, and assume for contradiction that $\sigma(T) = \emptyset$. This is equivalent to saying that $\rho(T) = \mathbb{C}$. Let $x \in X$ and $f \in X'$, and let $h\colon \mathbb{C} \to \mathbb{C}$ be defined as $h(\lambda) = F(R_\lambda x)$. Then we just proved in the last theorem that h is locally holomorphic on $\rho(T)$; here we're assuming $\rho(T) = \mathbb{C}$, so h is holomorphic on $\mathbb{C}$.

If we've taken complex analysis, you use the word *entire* to describe h. Now we're going to use a complex analysis trick — from our formula we can show h is bounded, and the only bounded entire functions are constant functions (this is Liouville's theorem).

Since a holomorphic function is differentiable, this means it's in particular continuous. And if h is continuous in $\mathbb{C}$, this means it's bounded on any compact subset — in particular, it's bounded on the set $\{\lambda \in \mathbb{C} \mid |\lambda| \leq 2\|T\|\}$. Further, using the above formula for R_λ as

$$R_\lambda = -\frac{1}{\lambda}\sum_{j=0}^{\infty}\left(\frac{1}{\lambda}T\right)^{j},$$

we see that this series converges absolutely for $|\lambda| > \|T\|$. So in particular, we get

$$\|R_\lambda\| \leq \frac{1}{|\lambda|}\sum_{j=0}^{\infty}\left\|\left(\frac{1}{\lambda}T\right)^{k}\right\| \leq \frac{1}{|\lambda|}\sum_{j=0}^{\infty}\frac{1}{|\lambda|^{j}}\|T\|^{j}.$$

In particular, if $|\lambda| \geq 2\|T\|$, then we get $\|R_\lambda\| \leq \frac{1}{|\lambda|}\sum_{j=0}^{\infty}(\frac{1}{2})^{j} = 2/|\lambda| \leq 1/\|T\|$.

And now we've shown R_λ is bounded (as a function of λ), and this will imply that h is bounded — we defined $h(\lambda) = f(R_\lambda x)$, so

$$|h(\lambda)| = |f(R_\lambda x)| \leq \|f\|\|R_\lambda x\| \leq \|f\|\|R_\lambda\|\|x\|.$$

We just proved that $\|R_\lambda\| \le 1/\|T\|$ for $|\lambda| \ge 2\|T\|$, and plugging this in gives that $|h(\lambda)| \le \|f\|\|x\|/\|T\|$ (this is a finite number not depending on λ). So this implies h is bounded for $|\lambda| \ge 2\|T\|$, and the continuity of h means it's bounded for $|\lambda| \le 2\|T\|$, so h is bounded on $\mathbb{C}$.

And then by Liouville's theorem from complex analysis, this means h must be a constant function — this means $h(\lambda) = f(R_\lambda x)$ is equal to a constant. But $R_\lambda x = (T - \lambda I)^{-1}x$; this means $(T - \lambda I)$ is independent of λ (since this is true for all f and x). This is a contradiction, which means $\sigma(T)$ is nonempty. □

Remark 23.6. The point is that we used the resolvent operator to show that if $\sigma(T)$ was empty, then the resolvent operator would give us an entire function which is bounded, and therefore must be constant; and this is a contradiction.

As an immediate consequence of the last two theorems:

Theorem 23.7
If $T \in B(X, X)$ for a complex Banach space X and $\lambda \in \rho(T)$, then

$$\|R_\lambda\| \ge \frac{1}{\delta(\lambda)},$$

where $\delta_\lambda = \inf_{s\in\sigma(T)} |\lambda - s|$ is the distance from λ to the spectrum $\sigma(T)$.

This in particular implies $\|R_\lambda\| \to \infty$ as $\delta(\lambda) \to 0$.

Proof. For all $\lambda_0 \in \rho(T)$, we've seen that

$$\left\{\lambda \in \mathbb{C} \mid |\lambda| \le \frac{1}{\|R_{\lambda_0}\|} \subseteq \rho(T)\right\}.$$

Because $\sigma(T)$ is nonempty, for $\lambda_0 \in \rho(T)$ we know that $\delta(\lambda_0) = \inf_{\mu\in\sigma(T)} |\lambda_0 - \mu|$ is well-defined and finite; and $|\lambda_0 - \mu| \ge 1/\|R_{\lambda_0}\|$ for all $\mu \in \sigma(T)$, so this gives $\delta(\lambda_0) \ge 1/\|R_{\lambda_0}\|$, which rearranges to what we want. □

We've showed that the resolvent operator is locally holomorphic on the resolvent, and in fact that's the largest set on which it's holomorphic — because outside that set we see $\|R_\lambda\| \to \infty$, so it can't be holomorphic (more precisely, we can't define it in such a way that it's holomorphic — right now, it's only defined on $\rho(T)$).

§23.4 The spectral radius

Theorem 23.8
If $T \in B(X, X)$ for a complex Banach space X, then

$$r_\sigma(T) = \lim_{n\to\infty} \|T^n\|^{1/n}.$$

Recall that $r_\sigma(T) = \sup_{\lambda\in\sigma(T)} |\lambda|$.

Proof. Last class, we proved that $\sigma(p(T)) = p(\sigma(T))$ for any polynomial p. So immediately, we know $\sigma(T^n) = \sigma(T)^n$. This in particular means $r_\sigma(T^n) = (r_\sigma(T))^n$.

We always have $r_\sigma(S) \le \|S\|$ (we proved this earlier), so applying this to T^n, we get that $r_\sigma(T^n) \le \|T^n\|$. Putting these together, we get that $r_\sigma(T) = (r_\sigma(T^n))^{1/n} \le \|T^n\|^{1/n}$ for all n. So that's the first side of the inequality — we've shown that $r_\sigma(T) \le \liminf_{n\to\infty}\|T^n\|^{1/n}$, which of course also means

$$r_\sigma(T) \le \liminf_{n\to\infty}\|T^n\|^{1/n} \le \limsup_{n\to\infty}\|T^n\|^{1/n}.$$

We're next going to show $\limsup_{n\to\infty}\|T^n\|^{1/n} \le r_\sigma(T)$; this implies equality must hold throughout (in particular, this means that $\lim_{n\to\infty}\|T^n\|^{1/n}$ exists, and gives us the theorem).

There's a couple more complex analysis components we need for this argument.

Theorem 23.9 (Hadamard's formula)

The power series $\sum_{n=0}^\infty c_n\kappa^n$ (over $\mathbb{C}$) converges absolutely for $|\kappa| < r$, where

$$r = \frac{1}{\limsup_{n\to\infty}|c_n|^{1/n}}.$$

Now we're going to use our formula for R_λ again — to set ourselves up to apply this, we'll let $\kappa = \frac{1}{\lambda}$, so

$$R_\lambda = -\frac{1}{\lambda}\sum_{j=0}^\infty\left(\frac{1}{\lambda}T\right)^j = -\kappa\sum_{j=0}^\infty(\kappa T)^j.$$

(In complex analysis, if we think about the set of complex numbers whose norm is bounded, this gives us a disk; and if we think about $1/\|\cdot\|$ being bounded, this gives you the complement of a disk. So there's an important game in complex analysis proofs of going between the two; this seems kind of silly, but it makes sense in context. And in particular it makes sense here because we've shown an upper bound on $\sigma(T)$.)

And just looking at the sum component, we have

$$\left\|\sum_{j=0}^\infty(\kappa T)^j\right\| \le \sum_{j=0}^\infty|\kappa|^j\,\|T^j\|.$$

So applying the Hadamard formula directly, we see that this sum $\sum_{j=0}^\infty(\kappa T)^j$ converges absolutely if

$$|\lambda| = \frac{1}{|\kappa|} > \frac{1}{r} = \limsup_{n\to\infty}\|T^n\|^{1/n}.$$

But from the previous theorem, R_λ is locally holomorphic on $\rho(T) \subseteq \mathbb{C}$. So we'll let $M = \{\frac{1}{\lambda} \mid \lambda \in \rho(T)\}$. Then thinking about the map $\kappa \mapsto R_\lambda := -\kappa\sum_{j=0}^\infty\kappa^nT^n$, we have that this is absolutely convergent on M (that's just a restatement of this result).

And its radius of convergence is precisely the largest disk contained in M. (This is also a complex analysis result; there are references in the text.)

And we know that this is $\{\kappa \mid \kappa < 1/\limsup_{n\to\infty}\|T^n\|^{1/n}\}$. Together, these give us that the map $\lambda \mapsto R_\lambda$ (in the inverted complex plane) satisfies $\limsup_{n\to\infty}\|T^n\|^{1/n} \le r_\sigma(T)$. This completes the proof. □

Remark 23.10. If you don't have a complex analysis background, it might seem we did some parlor tricks; this is maybe a dissatisfying final proof. But if we have some complex analysis, we can probably vouch for this being a reasonable argument. Maybe this is an argument to take a complex analysis course; it shows up in lots of fields of math and it's very neat and orderly, and everything works out nicely — because for example, a function just being differentiable means it has a power series representation everywhere, and the only way it can be bounded is if it's constant. This is extremely restrictive; and so it shows up in low-dimensional topology and real analysis and lots of things.

§23.5 More on self-adjoint operators in Hilbert spaces

There are some problems on the pset from Chapter 9; mostly these should look like applications of the stuff from MacCluer. But it's worth remarking on a few of the new components that show up — they're in the Kreyszig text too, but that's the next thing we'll do. So we'll remark on a couple of items from Chapter 9 of Kreyszig.

Now we're back in the context of Hilbert spaces and self-adjoint operators.

Theorem 23.11 (Theorem 9.1–2)
Let H be a Hilbert space, and let $T \in B(H, H)$ be self-adjoint (i.e., $T = T^*$). Then $\lambda \in \rho(T)$ if and only if there exists $c > 0$ such that for all $x \in H$ we have $\|T_\lambda x\| \geq c\,\|x\|$.

What we're really saying is that T_λ is bounded below, where $T_\lambda = T - \lambda I$.

We're going to use this in the pset.

Proof. One side of the proof is sort of easy — for the forwards direction, if $\lambda \in \rho(T)$ then $R_\lambda = T_\lambda^{-1}$ exists and is bounded (by the open mapping theorem). Then for any $x \in H$ we have

$$\|x\| = \|R_\lambda T_\lambda x\| \leq \|R_\lambda\|\,\|T_\lambda x\|,$$

which gets that

$$\|T_\lambda x\| \geq \frac{1}{\|R_\lambda\|}\,\|x\|.$$

The proof of the other direction is in the textbook; it uses some more careful theory from Hilbert spaces. But you can just apply this and use it in the pset. (This direction should give us some intuition as to why the theorem is sensible; the more powerful direction is the other direction, that if T_λ is bounded below then λ must be in the resolvent.) □

Theorem 23.12 (Theorem 9.2–1)
Let H be a complex Hilbert space, and $T \in B(H, H)$ a bounded self-adjoint operator. Then $\sigma(T) \subseteq [m, M] \subseteq \mathbb{R}$, where $m = \inf_{\|x\|=1}\langle Tx, x\rangle$ and $M = \sup_{\|x\|=1}\langle Tx, x\rangle$.

We've proved most of the things here — we proved it's contained in $\mathbb{R}$, and we proved this upper bound. (Or maybe we didn't — our proof used compactness.) The problem set also asks us to prove this lower bound. Back when we proved this, we mentioned a connection to linear algebra and the Rayleigh quotient; this extends that.

As a follow-up:

Definition 23.13. If $m \geq 0$, then we say T is a *positive* operator.

This is because for every $x \in X$, we have $\langle Tx, x\rangle \geq 0$. In fact, this is a nice way to create a (partial) ordering on the set of bounded self-adjoint operators:

Definition 23.14. We say $S \succeq T$ if $S - T$ is positive.

Not all operators are comparable, but this gives you a partial ordering. This should be familiar — in a linear algebra perspective, this is the analog of being positive semidefinite (and we can look at a difference of matrices and talk about that difference being positive semidefinite or positive definite).

§24 May 13, 2024

First we'll get questions about the exam out of the way; then a student has a nice presentation about something related to Hilbert spaces; then Marjie will talk a bit about related courses available in the fall; and if we want, she can chat a little about related research.

§24.1 The final

Marjie has written the final, and the graders have edited it.

The final will have 4 questions, each of which has a lot of parts (about 6). As promised, the emphasis will be on the material after the second midterm (really strongly) — all the questions will be about this material, but if you don't know earlier material, then you probably won't really be able to navigate them.

We'll remark briefly on a word cloud of the earlier content that seems relevant, that we should understand:

- Banach spaces — in particular, what it means to be complete.
- Hilbert spaces — as complete inner product spaces.
- What it means to have an orthonormal basis, orthogonal projections, and the Hilbert adjoint.
- All the operator properties we've discussed — the meaning of the descriptors bounded, compact, closed, self-adjoint.
- There might also be brief applications of the big theorems: uniform boundedness, Hahn–Banach, the open mapping theorem and its corollary the bounded inverse theorem, and the Baire category theorem and Zorn's lemma (which these relied on).
- Some more basic stuff, like what it means to be an open or closed subspace, what it means to be one-to-one, onto, have an inverse, be dense.
- As examples, L^p spaces and ℓ^p spaces.

This is the content from earlier in the course that shows up in the material since the second midterm; these are things we need to be able to use to do stuff since the second midterm. But the focus really will be on spectral theory and invariant subspaces and content from the psets and from class. (The point is that to do problems with spectral theory there's so much content from earlier in the course that you need anyways — nearly all these words will show up in some way or be referred to in the final. You can imagine that this is the foundation of the pyramid, but the final will actually be on spectral theory.)

The psets are good review, as are reading sections of the book or proofs we did in class.

Remark 24.1. Should we be comfortable with the complex analysis results used in the class? No. (Marjie did not think that would be fair.)

Lots of spectral theory relies on the fact that the field is complex — for example, we proved a compact self-adjoint operator has real eigenvalues or a real spectrum or something like this, and you have to understand how you'd prove something like this (you assume there are imaginary values, and understand complex numbers enough to show that the imaginary part of a complex number is zero) — but you don't need complex analysis theory from the last section.

(It is longer than the midterms.)

Emphasis is really on the content after the midterms, but you have to understand these terms or else it'll be a challenge.

Remark 24.2 (Grading)**.** How will grading work? Our homework is a strict percentage — your top 6 homeworks are taken, and each is 5%. For the midterms, Marjie's plan, which she claims only benefits people (it's a ceiling plan rather than a floor plan) is that they gave A, B, and C for the first midterm; these will be taken as 95, 85, and 75. And then if at the end you're at any boundary, she'll look at your actual fraction. (If you got a higher score, there will probably be some 100s to keep this fair. But if you are below 95, 85, or 75 you will just be picked up to that number.)

30% of your grade is psets, with the top 6 being 5% each. Then the two midterms are each 20%, and the final is 30%. So if a person got a B on midterm 1 then they'll get a $0.85 \cdot 20$; and maybe if you have 29 from the psets, and if the person got an A (but not a 100% type of A) then maybe this is $0.95 \cdot 20$, and then you end up with some total, and there's some number you need to get on the final to end up with over a 90%, which is an A.

(If you got over a 95%, then you get your actual percentage, because otherwise this is unfair. If you end up on the boundary of a grade, then Marjie will look at your actual score on the exam. She's not trying to be tricky.)

She'll also attempt to correctly transfer this from Gradescope to Canvas and update it in Canvas, so that we can see our grade before the final (so we know what we need to get). Her intent is for the course grade cutoffs to be 90% for an A, 80% for a B, 70% for a C, and 60% for a D. There may be some $-$'s and $+$'s if you're on the boundary.

The only mystery is whether the final will be graded directly; probably it will be graded in the way we've been grading so far.

§24.2 Classes in the fall

The relevant follow-ons to this class are:

- **18.101** (analysis on manifolds). (The name is 'and' instead of 'on.') This will be taking the ideas of measure theory and moving them to manifolds. It also looks like an introduction to differential topology. In real analysis, we're always thinking about $\mathbb{R}^n$, but you can put this on a surface and then ask the same questions — you started to do this in Calculus III, and this is an extension of those ideas, e.g., 'can I define a derivative on the sphere?' This is a nice course; whether your interests are more topological or more analytic, it has applications in both.
- **18.112** (complex analysis). We talked a little about complex analysis. This is sort of prerequisite knowledge moving on (e.g., what it means to be holomorphic) — it becomes a computational tool in basically every field of math, and it's a nice, neat, orderly, self-contained course.
- **18.152** (intro to PDEs). This will look at diffusional, elliptic, and hyperbolic PDEs. In the broad field of analysis, if you continue to study it, one major direction is differential equations; and the techniques we use vary vastly, and this will be an introduction to the techniques used. When Marjie studied this in graduate school, she was at a place where lots of people did elliptic PDEs, so they sort of focused in that area; but here it will be a little different. This will be an introduction to techniques.
- When Marjie took functional analysis in grad school, there was a follow-on course. Here the analog is **18.155** (differential analysis). This is more measure theory and L^p spaces and distributions and Fourier space and the Fourier transform, and a little bit of PDEs. To Marjie, this gives you access to today's PDE work almost more than the intro to PDE course, but any of these are nice.

  Similar to 18.152, this is a bit more axiomatic and structured — in differential equations you're often trying to figure out whether there exists a solution and whether it's unique. You want to study sequences of functions — we did that here, looking at function spaces. If you want to show a solution exists, there's often a weakening of the concept — here we literally talked about weak convergence of solutions, and that's a component of how you solve differential equations — you look for solutions in a weaker setting. One of those is distributions; another is the concept in measure theory of how we define the integral.

Later Marjie will chat about her own work; this is a natural bridge into some more analysis.

These are the four courses Marjie can most highly recommend; there's also geometry of manifolds (though 18.101 comes before that) and differential geometry (18.950).

If you haven't taken topology or algebra, you should probably also consider that.

Remark 24.3. Does the material we learned here appear in e.g. 18.101? The nice notion of the dual space is an important component of understanding differential forms and stuff like that; as you continue in math things might not be so serial anymore (meaning there's not really a complete ordering). What we've learned here should be beneficial in these courses, but it's not clear that any of those courses should strictly require this one. But at some point talking about functionals becomes expected in any part of analysis (you're kind of expected to understand function spaces). And some of the big theorems seem a little unwieldy right now, but they're sort of expected knowledge and often lead to generalizations in different fields. (None of the classes should be taught with this class as prerequisite knowledge, but we'll probably find that it's beneficial.)

§24.3 Reproducing kernel Hilbert spaces

This is a talk by one of the students in the class.

In this talk, we'll use X to denote some set of elements, and we'll use $\mathbb{F}$ to denote our field.

Definition 24.4. We let $\mathcal{F}(X, \mathbb{F})$ denote the set of all functions $X \to \mathbb{F}$. For a subset $\mathcal{S} \subseteq \mathcal{F}$, we define the *linear evaluation functional* for a fixed $x \in X$ as $E_x(f) = f(x)$ (where we're just evaluating $f \in \mathcal{S}$ at the given input).

A more useful definition (which is new):

Definition 24.5. A *reproducing kernel Hilbert space* (RKHS) on X is a subset $\mathcal{H} \subseteq \mathcal{F}(X, \mathbb{F})$ satisfying the properties that:

(1) $\mathcal{H}$ is a vector subspace of $\mathcal{F}(X, \mathbb{F})$.

(2) $\mathcal{H}$ has an inner product.

(3) $\mathcal{H}$ is complete (with respect to the norm) and separable — so in other words, $\mathcal{H}$ is a separable Hilbert space.

(4) For every $x \in X$, the linear evaluation functional $E_x \colon \mathcal{H} \to \mathbb{F}$ is bounded.

The *kernel* part of this definition comes from its representation from the Riesz representation theorem — by the Riesz representation theorem, we can represent the evaluation function E_x as $E_x(f) = \langle f, k_x \rangle$ (where k_x is our 'kernel').

We'll need one more definition to flesh out this representation a bit more.

Definition 24.6. We call k_x the *reproducing kernel* for x.

Since we have $k_x \in \mathcal{H}$ (and $\mathcal{H}$ is a function space), we can define $k_x(y) = E_y(k_x) = \langle k_x, k_y \rangle$ (where k_y is the element associated to E_y). So we see for a *fixed* y that $\langle f, k_y \rangle = f(y)$ by our definition of how the evaluation functional goes; this allows us to define the *reproducing kernel* of $\mathcal{H}$ as a function.

Definition 24.7. We define $K\colon X \times X \to \mathbb{R}$ (or $\mathbb{C}$) as $K(x, y) = k_y(x) = \langle k_x, k_y \rangle$.

So these are our definitions; now we're going to give the theorem. This theorem gives some intuition for why they're called 'reproducing.'

Theorem 24.8

If there exists a reproducing kernel K on a Hilbbert space $\mathcal{H}$ of functionals on X, then it is unique.

Furthermore, if there exists another Hilbert space $\mathcal{I}$ of functionals on X where the same K is also a reproducing kernel on $\mathcal{I}$, then then $\mathcal{I}$ and $\mathcal{H}$ are isomorphic.

Maybe the first statement is to be expected, but the thing that's really amazing is the second — the reproducing kernel uniquely defines its own Hilbert space. (You define one kernel and suddenly have an absurd amount of information about your Hilbert space.)

Proof. We'll first show uniqueness (which is pretty easy). As usual, suppowe we have two reproducing kernels K and K'; then we have

$$\|K - K'\|^2 = \langle K - K', K - K' \rangle = \langle K - K', K \rangle - \langle K - K', K' \rangle.$$

And substituting in our definition, we get

$$\langle k_y(x) - k'_y(x), k_y(x) \rangle - \langle k_y(x) - k'_y(x), k'_y(x) \rangle.$$

Now if you look at this definition, we see that all of these inner products actually just collapse under our definition, and we end up with

$$k_y(y) - k'_y(y) - (k_y(y) - k'_y(y)) = 0$$

(this is not immediate, but if you stare at the definitions for 30 seconds you'll probably get it).

Now we'll discuss the second part. (The proof is a bit convoluted.) Let's suppose that $\mathcal{H}$ and $\mathcal{I}$ have the same reproducing kernel K; we want to show that then $\mathcal{H} \cong \mathcal{I}$. To do this, we'll show that they're equal to each other and define the same norm.

Consider $\mathcal{H}$ and the functions $k_y(x) \in \mathcal{H}$, which are well-defined for all $y \in X$. So we then consider the linear space $\mathcal{H}_k$, which we define as the space where we take all these kernel functions as our basis — so we take our basis to be the functions $\{k_y(x)\}$ over all $x, y \in X$. We're going to argue that this space must be dense in both $\mathcal{I}$ and $\mathcal{H}$, and then we're going to use this to show that $\mathcal{H}$ and $\mathcal{I}$ are equivalent as sets and their norms are the same.

To see this, take any $f \in \mathcal{H}$ such that $f \perp \mathcal{H}_k$. By definition this means $\langle f, k_y \rangle = 0$ for every $y \in X$. But we see that this is the evaluation functional — because that's how we defined k_y — so this is basically just saying that $f(y) = 0$ for all $y \in X$. So f has to be the zero element. And then we have that $\mathcal{H}_k$ is dense.

What we're going to do next is try to write the norm on $\mathcal{H}$ in terms of K, and we're going to see that we can do everything with $\mathcal{I}$ as well, and soon things are going to start popping out.

We want to describe $\|\bullet\|_{\mathcal{H}}$ for $\mathcal{H}_k$ in terms of K. So consider some $f \in \mathcal{H}_k$. By density (since we have this basis), we have that

$$\|f\|_{\mathcal{H}}^2 = \sum \lambda_i \lambda_j \langle k_{x_i}, k_{x_j} \rangle = \sum_{i=1}^{n} \sum_{j=1}^{n} \lambda_i \lambda_j K(x_i, x_j)$$

(by how we defined K). And using the exact same argument, since we have the same kernel K on $\mathcal{I}$, then we still have all these functionals defined on it from our kernel, and so we also have that $\mathcal{H}_k$ is dense in $\mathcal{I}$ (by the same argument). So we also have that for $f \in \mathcal{H}_k$ that $\|f\|_{\mathcal{I}} = \sum \lambda_i \lambda_j K(x_i, x_j)$.

This doesn't yet say that anything is the same. But now we're going to argue for some arbitrary element in $\mathcal{H}$ and show that it's in $\mathcal{I}$. Suppose we have some $g \in \mathcal{H}$ (looking at the entire space now, instead of $\mathcal{H}_k$). Since $\mathcal{H}_k$ is dense, we can take some sequence of functions $(g_n) \subseteq \mathcal{H}_k$ such that $\|g_n - g\|_{\mathcal{H}} \to 0$. And we can see that then g_n is going to be Cauchy in both norms (with respect to $\mathcal{I}$ and $\mathcal{H}$). So because we've got a convergent sequence in $\mathcal{H}_k$ and it converges to some g, it's going to be Cauchy in both norms (because the norms are the same on $\mathcal{H}_k$ as a subset of $\mathcal{H}$ and of $\mathcal{I}$); so then g_n also has to converge to some $g_{\mathcal{I}}$ in $\mathcal{I}$. And reproducing kernels are bounded, so by the boundedness of reproducing kernels, then we get that g_n converges pointwise to both g and $g_{\mathcal{I}}$ (meaning $g_n(y) \to g(y), g_{\mathcal{I}}(y)$), which means we must actually have $g = g_{\mathcal{I}}$. So g is also in $\mathcal{I}$. This basically says that $\mathcal{H} \subseteq \mathcal{I}$, and this argument is symmetric, so we can show the reverse in the same way; so we get that $\mathcal{H} = \mathcal{I}$ (as sets).

The only thing left is to check that the norms are equivalent, but we can do this quickly. For any $f \in \mathcal{H} = \mathcal{I}$, we have $\|f\|_{\mathcal{H}} = \lim_{n\to\infty} \|f_n\|_{\mathcal{H}} = \lim_{n\to\infty} \|f_n\|_{\mathcal{I}} = \|f\|_{\mathcal{I}}$ (where f_n is a sequence in $\mathcal{I}$ — this is because the norms are equivalent in $\mathcal{H}_k$). □

§24.4 Marjie's research

Marjie often introduces what she does as the theoretical side of interpolation of data. As a nice metaphor, if someone collects some data, one thing you often do is find a linear best fit. This means you're minimizing the sum of squares of distances to some line.

But most things aren't linear, at least not on a large scale; then interpolation of data becomes more complicated. There's a celebrated theorem in the field, a partial converse to Taylor's theorem. (There's actually a couple of versions.) Taylor's theorem, broadly speaking, says something like if $f\colon \mathbb{R}^n \to \mathbb{R}$ is m times differentiable, then there exists a degree-m polynomial that approximates f at x_0 for all $x_0 \in \mathbb{R}^n$. Now in fact there's a very precise result about what it means to approximate f at a point x_0 — you get a bound between differences. And the partial converse asks, if you give me a bunch of polynomials, how do I know if there's a function interpolating the polynomials?

In one dimension, there's always a function interpolating two points, but how do we know the smallest possible derivative? You can use the mean value theorem — you know that at some point there's a derivative equal to the difference quotient.

Theorem 24.9 (Whitney's theorem)

Let $E \subseteq \mathbb{R}^n$ be a closed set (or more generally, any measurable set), and suppose we're given a set of degree-m polynomials $\{p_x\}_{x\in E}$ which are 'comparable' (in a specific quantitative sense related to MVT), then there exists a m-times differentiable function interpolating the polynomials.

When we say *interpolating*, this usually literally means that the function goes through the points. But since we have polynomials, this means you also want to say that the higher-order derivatives match — so interpolation goes to higher degrees of differentiability and smoothness.

So Taylor's theorem says if we're m times differentiable then we have nice polynomials; Whitney's theorem says if you give me nice comparable functions then we can interpolate them.

This is very nice, and Whitney proved this in 1934 using very pretty machinery that we continue to use in more general settings. The ability to come up with this bound is totally independent of the function values — the construction is really all about the set E (how good is the mean value estimate or how much you lose depends on E). From an interpolation perspective, this is saying how you collect your data gives your accuracy.

The way this works is you have a bunch of points in E, and to create an interpolator extension you decompose $\overline{E}$ into cubes where the distance from the cube to E is approximately teh side length of the cube. It turns out there's a nice reason this works, and it has to do with the mean value theorem. So you're decomposing $\overline{E}$ into cubes satisfying $\operatorname{diam}(Q)$ is comparable to $\operatorname{dist}(x, E)$ for all $x \in Q$. And basically the estimates exactly play well with this decomposition.

Then the follow-on is, this theorem was proved in 1934 and it's quite nice. But it turns out there's a harder question regarding the theoretical interpolation of data. Here we're given a degree-m polynomial. But what if we're instead given a function?

Question 24.10. Suppose $E \subseteq \mathbb{R}^n$ is closed and we're given $f\colon E \to \mathbb{R}$. Is there $F \in \mathcal{C}^m(\mathbb{R}^n)$ interpolating (extending) f?

We didn't talk about the $\mathcal{C}^m$ function spaces a ton, but we talked a bit about $\mathcal{C}^1$ — continuous functions with continuous first derivative — and this is a generalization.

It turns out this is much harder. The other thing is that $\mathcal{C}^m(\mathbb{R}^n)$ is m-times differentiable functions, but we can actually play with what function space we're using, and results vary widely. Another relevant function space that comes up in differential equations (which comes up in 18.155 and maybe 152) is Sobolev spaces. So you can replace $\mathcal{C}^m$ with other function spaces that embed in continuous functions.

Remark 24.11. You need continuity for interpolation to make sense — for example, we defined the L^p spaces up to a set of measure zero, which means that if you give me a function in a L^p space, you actually can't tell me what the function value is at any point, because it can change up to a set of measure zero — so it doesn't make sense to talk about interpolation with L^p functions, because up to a measure zero set the values could just be wrong. So when we talk about interpolation, we only talk about continuous functions so that you really have information about the function values.

But the point is that you can change the function space and ask this question again — Marjie did some work with Sobolev spaces and Besov spaces and stuff (they have to embed in continuous functions, usually $\mathcal{C}^0(\mathbb{R}^n)$).

The first big results were from Charlie Fefferman 2004–08. The very interesting realization he had was that like in the Whitney case where your extension depends entirely on the structure of E, here whether or not there is an extension depends entirely on the structure of E. The little selling point here is, why does this make sense? Let's think about a set E that looks — if E looks kind of like a curve, and we want to know if there's an extension in a particular function space of this curve, then if we collect data along this curve we have a ton of information about derivatives in one direction, but no information about derivatives in the other direction — so if we want to know if there's a $\mathcal{C}^m$ extension, we don't have any information in this direction. He realized there's a sort of inductive structure — if you can build an extension on this lower-dimensional set E, then you can extend it in the orthogonal direction very easily. But if you have more components of E up above, how do you make sure that the extensions match up together? So there's this pretty inductive structure that looks at E and $\overline{E}$ and identifies these lower-dimensional surfaces where the problem is easier to solve.

In Marjie's opinion, the structure of these proofs is also really pretty — just the machinery used where you decompose space is also used in harmonic analyis, and these are sort of 'stopping time rules,' so they're mathematically pretty. You have some set E, which can look like anything (think of this as where we're collecting data); and to find out information about what this set looks like, we create a stopping time rule that looks like I'm going to cut this cube up (in a *dyadic decomposition*), subdividing until one of my cubes only has a pretty 'flat' set in it So we cut in half and nobody is in a flat set in our four cubes; then we have to subdivide again. And now we have one cube where we've just got one little flat piece so we're done with that; but we have to cut the rest again. And so on. The one cube with a solid pice is going to keep on getting subdivided forever, so the solid piece is never going to show up in our subdivided cube. But maybe some of the other pieces are getting flatter — we cut a few times and eventually it becomes flat enough to stop cutting.

So you come up with this stopping time rule that lets you understand characteristics of E where you can understand how to extend a function.

She also does stuff with convex functions; the interpolation theory of convex functions is interesting because convexity is a 'nonlinear constraint' and that measn lots of arguments don't work the same way (you can't just patch together convex functions to get a convex function). So this thing where you zoom in decomposing space, solve the problem locally, and patch back together — this is really beautiful but doesn't work with convexity because you can't patch together solutions.

But it is still true morally that you can zoom in to gain information; but you need a different way to patch things together.