202B: Functional Analysis

Nir Elber

Spring 2024

Contents

How strange to actually have to see the path of your journey in order to make it.

—Neal Shusterman, [Shu16]

THEME 1
PRODUCT MEASURES

1.1 January 17

Let's just get started.

1.1.1 Course Notes

Here are some course notes.

- The professor for this course is Michael Christ.
- There is a bCourses, which I don't have access to.
- There will be an exam in the evening in February.
- Problem sets will be due on Fridays.
- We will assume analysis on the level of Math 202A; see something like [Elb22].
- The text for the course is [Fol99].

1.1.2 Measures

Our first topic is to integrate on product spaces. Roughly speaking, we might have some measure spaces $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ with some way to measure on them, and then we will want to measure $X \times Y$. Let's quickly recall what a measure is; we won't bother to recall the definition of a σ-algebra, but we will refer to [Elb22, Definition 5.25]. This requires the definition of a σ-algebra.

Definition 1.1 (σ-algebra)**.** Fix a set X. Then a collection $\mathcal{M} \subseteq \mathcal{P}(X)$ is a σ-*algebra* if and only if the following conditions are satisfied.

- $\varnothing \in \mathcal{M}$.
- $\mathcal{M}$ is closed under countable unions.
- $\mathcal{M}$ is closed under complements.

In the sequel, we will also want to produce σ-algebras.

Definition 1.2. Fix a set X. Given a collection $\mathcal{S} \subseteq \mathcal{P}(X)$, we will say that the smallest σ-algebra generated by $\mathcal{S}$ is the σ-algebra *generated by* $\mathcal{S}$.

It is lemma that a smallest (i.e., contained in all other such σ-algebras) such σ-algebra exists and is unique. Let's see this.

Lemma 1.3. Fix a set X and collection $\mathcal{S} \subseteq \mathcal{P}(X)$. Then there is a σ-algebra $\mathcal{M}$ containing $\mathcal{S}$ such that $\mathcal{M} \subseteq \mathcal{M}'$ for any σ-algebra $\mathcal{M}'$ containing $\mathcal{S}$. This $\mathcal{M}$ is also unique.

Proof. There is certainly some σ-algebra on X containing $\mathcal{S}$, namely $\mathcal{P}(X)$. So there is a nonempty collection $\underline{\mathcal{M}}$ of all σ-algebras containing $\mathcal{S}$, and then we define

$$\mathcal{M} := \bigcap_{\mathcal{M}' \in \underline{\mathcal{M}}} \mathcal{M}'.$$

Certainly $\mathcal{M}$ contains $\mathcal{S}$, and one can check directly that $\mathcal{M}$ is a σ-algebra. (See [Elb22, Lemma 5.28] for details.) And by construction, we see that $\mathcal{M} \subseteq \mathcal{M}'$ for any σ-algebra $\mathcal{M}'$ containing $\mathcal{S}$. Lastly, we note that $\mathcal{M}$ is unique because any two such σ-algebras $\mathcal{M}_1$ and $\mathcal{M}_2$ will be contained in each other and hence equal. ■

Anyway, here is our definition of a measure.

Definition 1.4 (measure)**.** Fix a σ-algebra $\mathcal{M}$ on a set X. Then a *measure* μ is a countably additive nonnegative function $\mu\colon \mathcal{M} \to \mathbb{R}_{\geq 0} \cup \{\infty\}$, and we require that $\mu(\varnothing) = 0$. Here, being countably additive means that

$$\mu\left(\bigsqcup_{i=1}^{\infty} E_i\right) = \sum_{i=0}^{\infty} \mu(E_i),$$

where the sum is allowed to be in ∞ (namely, diverge to infinity). We call the triple $(X, \mathcal{M}, \mu)$ a *measure space*.

Remark 1.5. If we have $\mu(\varnothing) > 0$, then the countably additive condition implies that $\mu(\varnothing) = \infty$ and then $\mu(E) = \infty$ for all $E \in \mathcal{M}$. This is in fact countably additive, but we would like to exclude it.

We will want to make our measures somewhat small.

Definition 1.6 (σ-finite)**.** Fix a measure space $(X, \mathcal{M}, \mu)$. Then μ is σ*-finite* if and only if X is a countable union of sets in $\mathcal{M}$ of finite measure.

This smallness condition is quite tame, and in practice all measures are σ-finite.

1.1.3 The Extension Theorem

We would like to discuss how to build measures from objects easier to construct. The following generalization of Definition 1.1 will be useful.

Definition 1.7 (algebra)**.** Fix a set X. Then a collection $\mathcal{A} \subseteq \mathcal{P}(X)$ is an *algebra* if and only if the following conditions are satisfied.

- $\varnothing \in \mathcal{A}$.
- $\mathcal{A}$ is closed under finite unions.
- $\mathcal{A}$ is closed under complements.

Example 1.8. Fix an uncountable set X, and let $\mathcal{A}$ denote the collection of finite and cofinite sets. Then $\mathcal{A}$ is an algebra (the finite union of finite sets is finite, and the finite union of cofinite sets is cofinite), but it need not be a σ-algebra because the countable union of finite sets need not be finite nor cofinite.

Example 1.9. Fix $X := \mathbb{R}$, and let $\mathcal{A}$ denote the collection of finite unions of open or closed intervals. Then $\mathcal{A}$ is an algebra but not a σ-algebra.

Additionally, the following generalization of Definition 1.4 will be useful.

Definition 1.10 (premeasure)**.** Fix an algebra $\mathcal{A}$ on a set X. Then a *premeasure* is a function $\rho\colon \mathcal{A} \to \mathbb{R}_{\geq 0} \cup \{\infty\}$ which is satisfies the following.

- $\rho(\varnothing) = 0$.
- Finitely additive: we have $\rho(A \sqcup B) = \rho(A) + \rho(B)$ for $A, B \in \mathcal{A}$.
- Countably additive: suppose $\{A_i\}_{i\in\mathbb{N}} \subseteq \mathcal{A}$ is pairwise disjoint, and $\bigsqcup_{i=1}^\infty A_i \in \mathcal{A}$. Then

$$\rho\left(\bigsqcup_{i=1}^\infty A_i\right) = \sum_{i=1}^\infty \rho(A_i).$$

And now here is our theorem.

Theorem 1.11 (Extension)**.** Fix a set X and a premeasure ρ on an algebra $\mathcal{A}$ over X. Then there exists a measure μ on the σ-algebra $\mathcal{M}$ generated by $\mathcal{A}$ such that $\mu|_{\mathcal{A}} = \rho$. Additionally, if ρ is σ-finite, then μ is unique on $\mathcal{M}$.

Here, σ-finiteness for ρ takes the same definition as Definition 1.6.

Proof of Theorem 1.11. For existence, combine [Elb22, Lemma 6.16 and Theorems 6.21, 6.24]. Further, uniqueness is [Elb22, Theorem 6.35]. It will be helpful to say a few words about the construction. Essentially, one builds an "outer measure" ρ^* on $\mathcal{P}(X)$ by

$$\rho^*(E) := \inf\left\{\sum_{n=0}^\infty \rho(A_n) : \{A_n\}_{n\in\mathbb{N}} \subseteq \mathcal{A} \text{ and } E \subseteq \bigcup_{n=0}^\infty A_n\right\}.$$

Then one restricts ρ^* to a smaller σ-algebra over which it becomes a bona fide measure. ■

1.1.4 Towards Product Measures

For our product measures, we take the following outline. Fix measure spaces $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$.

1. We will construct a special σ-algebra $\mathcal{M} \otimes \mathcal{N}$ on $X \times Y$. Then we will construct a measure $\mu \times \nu$ on $\mathcal{M} \otimes \mathcal{N}$.
2. Once the construction is in place, we will find a way to compare "double integrals" with "single integrals." Morally, one wants equalities comparing

$$\iint_{X\times Y} f\, d(\mu \times \nu) \qquad \text{and} \qquad \int_X \left(\int_Y f(x,y)\, d\nu(y)\right) d\mu(x).$$

   The moral of the story is that we will be able to compare our product measure with the measures on X and Y which we already understand.
3. Lastly, we will specialize to Euclidean space $\mathbb{R}^d$.

Let's go ahead and begin.

Definition 1.12 (measurable rectangle)**.** Fix measure spaces $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$. A *measurable rectangle* $E \subseteq X \times Y$ is a set of the form $A \times B$ where $A \in \mathcal{M}$ and $B \in \mathcal{N}$.

Example 1.13. The product of the circles $S^1 \subseteq \mathbb{R}^2$ and $S^1 \subseteq \mathbb{R}^2$ is the torus $S^1 \times S^1$ in $\mathbb{R}^4$ (identified with $\mathbb{R}^2 \times \mathbb{R}^2$).

Definition 1.14 (product algebra)**.** Fix measure spaces $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$. Then we define the *product algebra* $\mathcal{A}(X, Y)$ as the collection of all finite disjoint unions of measure rectangles.

Remark 1.15. The reason that we have taken finite disjoint unions of rectangles is because we know how to measure measurable rectangles, and we know how to sum their measures as disjoint unions.

It's not totally clear that we have actually defined an algebra. We'll show this next class.

1.2 January 19

Here we go.

1.2.1 The Product Algebra

We quickly pick up the following lemma.

Lemma 1.16. Fix finitely many subsets $A_1, \ldots, A_n \subseteq X$, and suppose that these subsets live in an algebra $\mathcal{A}$ on X. Then there exists a finite partition $\{C_\alpha\}_{\alpha \in \kappa}$ of X of sets in the algebra such that

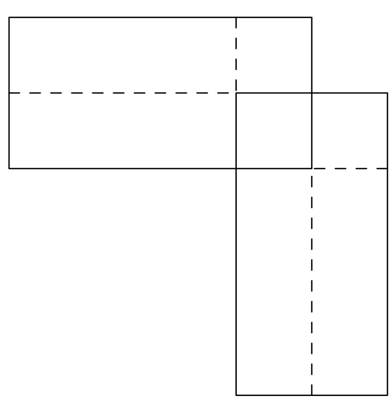
$$A_i = \bigsqcup_{\substack{\alpha \in \kappa \\ C_\alpha \subseteq A_i}} A_i.$$

Proof. We basically build a Venn diagram. Choose index set I to be $\{0, 1\}^n$, and define C_α for $\alpha \in I$ to be the set of $x \in X$ such that $x \in A_i$ if and only if $\alpha_i = 1$. Note that we can write C_α as

$$C_\alpha := \bigcup_{\alpha_i = 1} A_i \bigg\backslash \bigcup_{\alpha_i = 0} A_i.$$

Now, these $C_\bullet$s of course provide a partition satisfying the needed condition by its construction. ■

Anyway, let's return to showing that we have a product algebra. For example, it turns out that the union of two measure rectangles is again a measurable rectangle. Here's the image.



And here is our statement.

Lemma 1.17. Fix measure spaces $(X,\mathcal{M},\mu)$ and $(Y,\mathcal{N},\nu)$. Then $\mathcal{A}(X,Y)$ is actually an algebra.

Proof. Here are our checks.

- Note $\varnothing\times\varnothing=\varnothing$, so $\varnothing\in\mathcal{A}(X,Y)$.
- Finite union of rectangles: suppose that we have measurable rectangles $\{A_i\times B_i\}_{i=1}^n$. Then we show that the union is in $\mathcal{A}(X,Y)$. Now, the $A_\bullet$s produce some partition $\{C_\alpha\}_{\alpha\in I}\subseteq\mathcal{M}$ of X via Lemma 1.16, and the $B_\bullet$s produce some partition $\{D_\beta\}_{\beta\in J}\subseteq\mathcal{N}$ of Y via Lemma 1.16 again. Now
$$A_i\times B_i=\bigsqcup_{\substack{C_\alpha\subseteq A_i\\ D_\beta\subseteq B_i}}C_\alpha\times D_\beta,$$
so
$$\bigcup_{i=1}^n A_i\times B_i=\bigcup_{i=1}^n\bigsqcup_{\substack{C_\alpha\subseteq A_i\\ D_\beta\subseteq B_i}}C_\alpha\times D_\beta,$$
so our union is a union of measurable rectangles of the form $C_\alpha\times D_\beta$. But these measurable rectangles are all pairwise disjoint because the $C_\bullet$s and $D_\bullet$s are all pairwise disjoint, so the above union is in $\mathcal{A}$.
- Finite union: given $E_1,\ldots,E_n\in\mathcal{A}$, we need to show the union is in $\mathcal{A}$. Well, write
$$E_i=\bigsqcup_{j=1}^{n_i}A_{ij}\times B_{ij}$$
for some $A_\bullet\in\mathcal{M}$ and $B_\bullet\in\mathcal{N}$. Then
$$\bigcup_{i=1}^n E_i=\bigcup_{i=1}^n\bigcup_{j=1}^{n_i}A_{ij}\times B_{ij}$$
is a union of measurable rectangles and hence lives in $\mathcal{A}$ by the above check.
- Complement: given $E\in\mathcal{A}$, write
$$E=\bigcup_{i=1}^n A_i\times B_i$$
for measurable rectangles $A_\bullet\times B_\bullet$. As before, the $A_\bullet$s produce some partition $\{C_\alpha\}_{\alpha\in I}\subseteq\mathcal{M}$ of X via Lemma 1.16, and the $B_\bullet$s produce some partition $\{D_\beta\}_{\beta\in J}\subseteq\mathcal{N}$ of Y via Lemma 1.16 again. This allows us to write
$$E=\bigsqcup_{i=1}^n\bigsqcup_{\substack{C_\alpha\subseteq A_i\\ D_\beta\subseteq B_i}}C_\alpha\times D_\beta,$$
and then the complement $(X\times Y)\setminus E$ will be the union of the measurable rectangles $C_\alpha\times D_\beta$ not in the above union. But these are still disjoint measurable rectangles, so the union remains in $\mathcal{A}$. ■

1.2.2 The Product Measure

Let's now define our product premeasure. Given the measure spaces $(X,\mathcal{M},\mu)$ and $(Y,\mathcal{N},\nu)$, we would like to define
$$\rho\left(\bigsqcup_{i=1}^n A_i\times B_i\right)=\sum_{i=1}^n\mu(A_i)\nu(B_i),$$
but it is not obvious that this is well-defined. Instead of doing this, we will choose the following definition.

Definition 1.18 (product premeasure)**.** Fix measure spaces $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$. Given $E \in \mathcal{A}(X, Y)$, we define the *product premeasure* $\rho(E)$ as

$$\rho(E) := \int_X \nu(E_x)\, d\mu(x),$$

where $E_x := \{y \in Y : (x, y) \in E\}$.

Remark 1.19. One should perhaps check that E_x is always in $\mathcal{N}$ and hence measurable. But for this we simply write $E = \bigsqcup_{i=1}^n (A_i \times B_i)$ for measurable rectangles $A_i \times B_i$ and note that

$$E_x = \{y \in Y : (x, y) \in A_i \times B_i \text{ for some } i\} = \bigcup_{\substack{i=1 \\ x \in A_i}}^n B_i,$$

which is a finite union of measurable sets and hence in $\mathcal{N}$. In fact,

Remark 1.20. One should perhaps check that $x \mapsto \nu(E_x)$ is integrable. Continuing from the above, we can see that these B_i must be disjoint if $x \in A_i$ for each of these i, so actually

$$\nu(E_x) = \sum_{\substack{i=1 \\ x \in A_i}}^n \nu(B_i) = \sum_{i=1}^n 1_{A_i}(x)\nu(B_i),$$

which is a linear combination of indicators of μ-measurable sets, so this is a μ-integrable function. Notably, we see that the measure of a measurable rectangle $A \times B$ is in fact $\mu(A)\nu(B)$.

Remark 1.21. It is notable that we can write

$$\rho(E) = \int_X \nu(E_x)\, d\mu(x) = \int_X \int_Y 1_E(x, y)\, d\nu(y)\, d\mu(x),$$

where the equality follows because the measure $\nu(E_x)$ is simply integrating Y over the indicator of $1_E(x, y)$.

We now check that we have a premeasure.

Proposition 1.22. Fix measure spaces $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$. Then the defined product premeasure ρ on $\mathcal{A}(X, Y)$ is in fact a premeasure.

Proof. Here are our checks.

- Note $\rho(\varnothing) = 0$ because $\varnothing_x = \varnothing$ always.
- Finitely additive: fix disjoint $E_1, E_2 \in \mathcal{A}(X, Y)$, and we want to compute $\rho(E_1 \sqcup E_2)$. Well, we use Remark 1.21 to note

$$\begin{aligned} \rho(E_1 \sqcup E_2) &= \int_X \nu((E_1 \sqcup E_2)_x)\, d\mu(x) \\ &= \int_X \int_Y 1_{E_1 \sqcup E_2}(x, y)\, d\nu(y)\, d\mu(x) \\ &= \int_X \int_Y (1_{E_1}(x, y) + 1_{E_2}(x, y))\, d\nu(y)\, d\mu(x) \end{aligned}$$

Now, by linearity of integration, this is

$$\begin{aligned}\rho(E_1 \sqcup E_2) &= \int_X \int_Y 1_{E_1}(x,y)\,d\nu(y)\,d\,\mu(x) + \int_X \int_Y 1_{E_2}(x,y)\,d\nu(y)\,d\mu(x)\\ &= \rho(E_1) + \rho(E_2),\end{aligned}$$

as desired.

- Countably additive: we use the Monotone convergence theorem. Fix some disjoint subsets $\{E_i\}_{i=1}^\infty \subseteq \mathcal{A}(X,Y)$ such that $E := \bigcup_{i=1}^\infty E_i$ is in $\mathcal{A}(X,Y)$. Proceeding as in the previous check, we see that

$$\begin{aligned}\rho(E) &= \int_X \int_Y 1_E(x,y)\,d\nu(y)\,d\mu(x)\\ &= \int_X \int_Y 1_E(x,y)\,d\nu(y)\,d\mu(x)\\ &= \int_X \int_Y \left(\sum_{i=1}^\infty 1_{E_i}(x,y)\right) d\nu(y)\,d\mu(x).\end{aligned}$$

Now, the functions 1_{E_i} and 1_E are all integrable (for suitably fixed coordinates), so applying the Monotone convergence theorem [Elb22, Theorem 9.18] tells us that

$$\rho(E) = \sum_{i=1}^\infty \int_X \int_Y 1_{E_i}(x,y)\,d\nu(y)\,d\mu(x) = \sum_{i=1}^\infty \rho(E_i),$$

as desired. ■

We can now produce our product measure.

Definition 1.23 (product measure)**.** Fix measure spaces $(X,\mathcal{M},\mu)$ and $(Y,\mathcal{N},\nu)$. Define the *product σ-algebra* $\mathcal{M}\otimes\mathcal{N}$ to be the σ-algebra generated by $\mathcal{A}(X,Y) \subseteq \mathcal{P}(X\times Y)$. Then the product premeasure ρ on $\mathcal{A}(X,Y)$ extends by Theorem 1.11 to a measure $\mu\times\nu$ on $\mathcal{M}\otimes\mathcal{N}$.

Remark 1.24. By Theorem 1.11, if μ and ν are both σ-finite, then one can see that ρ is still σ-finite by some covering with measurable rectangles, so $\mu\times\nu$ becomes the unique measure on $\mathcal{M}\otimes\mathcal{N}$ extending ρ.

1.2.3 Tonelli's Theorem

The construction of our product premeasure in Definition 1.18 has a "handedness" in that we integrate with respect to Y and then with respect to X. This is somewhat upsetting, so we work to remedy this.

Theorem 1.25 (Tonelli)**.** Fix σ-finite measure spaces $(X,\mathcal{M},\mu)$ and $(Y,\mathcal{N},\nu)$. Fix a measurable function $f\colon X\times Y\to[0,\infty]$. Then the following hold.

(a) The function $y\mapsto f(x,y)$ is $\mathcal{N}$-measurable.

(b) The function $x\mapsto \int_Y f(x,y)\,d\nu(y)$ is $\mathcal{M}$-measurable.

(c) We have

$$\int_{X\times Y} f\,d(\mu\times\nu) = \int_X \int_Y f(x,y)\,d\nu(y)\,d\mu(x).$$

Remark 1.26. Note that, once measurable, we can integrate a nonnegative function if we allow for infinite values. For example, see something like [Elb22, Proposition 9.22]

Reductions of Theorem 1.25. We begin with two reductions.

- We reduce to the case where f is the indicator of a function 1_E. Indeed, having the result for indicators shows the conclusions for any linear combination of these, so we get the result for simple measurable functions, and then we can get the general case by taking monotone limits via the Monotone convergence theorem [Elb22, Theorem 9.18].

  (Namely, (a) is direct by taking limits, (b) follows by the Monotone convergence theorem to move out the limit out of the integral and then taking limits to get measurable, and (c) is achieved directly by the Monotone convergence theorem repeatedly.)

- We reduce to the case where X and Y are spaces of finite measure. Indeed, by the σ-finiteness of X and Y, we can partition each into countable disjoint union of sets of finite measure, and then by taking rectangles, we see that $X \times Y$ is a countable union of disjoint sets of finite measure. So achieving the result on these disjoint sets of finite measure, we can check the conclusions by summing over all the disjoint spaces, again concluding via the Monotone convergence theorem [Elb22, Theorem 9.18]. Namely, one can do an identical argument to the parenthetical remark of the previous reduction.

Before doing anything, we note that the σ-algebra $\mathcal{M} \otimes \mathcal{N}$ is not obviously generated at finite steps from $\mathcal{A}(X, Y)$; in fact, there is no countable constructive procedure to do this. So we are not going to proceed by trying to build up to $\mathcal{M} \otimes \mathcal{N}$; instead we will have to do something difficult. ■

1.3 January 22

Here we go.

1.3.1 Proof of Tonelli's Theorem

Last class we reduced the proof of Theorem 1.25 to having $f = 1_E$ for some measurable set E and having X and Y be finite measure spaces. Today we will complete the proof. We proceed by a sequence of lemmas.

Lemma 1.27. Fix measure spaces $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$. If $E \in \mathcal{M} \otimes \mathcal{N}$ and $x \in X$, then the slice

$$E_x \coloneqq \{y \in Y : (x, y) \in E\}$$

is in $\mathcal{N}$.

Proof. The problem is that we know very little about $\mathcal{M} \otimes \mathcal{N}$, so we will have to do something indirect. Continue with x fixed, but we let E vary to define

$$\mathcal{D}_x \coloneqq \{E \subseteq X \times Y : E_x \in \mathcal{N}\}.$$

Note $\mathcal{D}_x$ is a σ-algebra, as we now check.

- Note $\varnothing \subseteq X \times Y$ has $\varnothing = \varnothing_x$ in $\mathcal{N}$. So $\varnothing \in \mathcal{D}_x$.

- Complement: if $E \in \mathcal{D}_x$, then $((X \times Y) \setminus E)_x = Y \setminus E_x$ as this set contains exactly the $y \in Y$ such that $(x, y) \notin E$. Thus, $((X \times Y) \setminus E)_x \in \mathcal{N}$, so $(X \times Y) \setminus E \in \mathcal{N}$.

- Countable unions: fix a countable collection $\{E_i\}_{i=1}^\infty \subseteq \mathcal{D}_x$. Then
$$\left(\bigcup_{i=1}^\infty E_i\right)_x = \bigcup_{i=1}^\infty (E_i)_x$$
because some y lives in this set if and only if (x,y) belongs to one of the E_i. The right-hand side lives in $\mathcal{N}$ by assumption, so we see $\bigcup_{i=1}^\infty E_i \in \mathcal{D}_x$.

Furthermore, we note that $\mathcal{D}_x$ contains $\mathcal{A}(X,Y)$. Indeed, it suffices to check that $\mathcal{D}_x$ contains measurable rectangles because $\mathcal{A}(X,Y)$ contains disjoint unions of these. Well, for a measurable rectangle $A \times B$, we see
$$(A\times B)_x = \begin{cases} B & \text{if } x \in A, \\ \varnothing & \text{if } x \notin A, \end{cases} \tag{1.1}$$
always lives in $\mathcal{N}$, so $A \times B \in \mathcal{D}_x$. In total, it follows that $\mathcal{D}_x$ must contain the smallest σ-algebra containing $\mathcal{A}(X,Y)$, so $\mathcal{M}\otimes\mathcal{N} \subseteq \mathcal{D}_x$. This is what we wanted. ■

Remark 1.28. The above proof exemplifies how we will access $\mathcal{M}\otimes\mathcal{N}$: we will construct some σ-algebra characterizing the desirable properties, and then we will show that it contains $\mathcal{A}(X,Y)$ in order to contain $\mathcal{M}\otimes\mathcal{N}$.

We now prove (a) and (b) of Theorem 1.25.

Lemma 1.29. Fix measure spaces $(X,\mathcal{M},\mu)$ and $(Y,\mathcal{N},\nu)$, with Y finite. If $E \in \mathcal{M}\otimes\mathcal{N}$, then the function $f_E\colon x \mapsto \nu(E_x)$ is $\mathcal{M}$-measurable.

Proof. We consider
$$\mathcal{D} := \{E \subseteq X \times Y : f_E \text{ is } \mathcal{M}\text{-measurable and } E_x \text{ is always measurable}\}.$$
We would like to show that $\mathcal{D}$ contains $\mathcal{M}\otimes\mathcal{N}$. Let's show some properties of $\mathcal{D}$. We won't succeed at showing that $\mathcal{D}$ is actually a σ-algebra, but we will get close enough. Of course, $\mathcal{D}$ contains $\varnothing$ because $f_\varnothing$ is just the zero function. Additionally, taking complements uses finiteness of the measure spaces: if $\nu(Y) < \infty$, we can write
$$f_{(X\times Y)\setminus E}(x) = \nu(Y) - \nu(E_x) = \nu(Y) - f_E(x),$$
so we are done because the right-hand side is a measurable function in x. (Indeed, constant functions and sums of measurable functions are all measurable.)

Remark 1.30. There is an issue with taking unions: given $E, F \in \mathcal{D}$, we want to look at $f_{E\cup F}$, but there is no obvious way to access this only in terms of f_E and f_F because there may be some intersection.

In light of Remark 1.30, we need a trick. Do note that we can show that $\mathcal{D}$ is closed under disjoint unions because then $f_{E\sqcup F} = f_E + f_F$, so $f_{E\sqcup F}$ being measurable is recovered from summing f_E and f_F. Thus, because $\mathcal{D}$ contains measurable rectangles (note that the measure of the output of the function (1.1) is measurable as it's basically an indicator), so $\mathcal{A}(X,Y) \subseteq \mathcal{D}$.

For our trick, we proceed in steps.

1. We begin by showing that $\mathcal{D}$ is closed under countable ascending unions: given an ascending sequence of sets $\{E_i\}_{i=1}^\infty \subseteq \mathcal{D}$, then we set $E := \bigcup_{i=1}^\infty E_i$ and see
$$\lim_{n\to\infty} \nu((E_n)_x) = \nu(E_x)$$
because $\left(\bigcup_{i=1}^\infty E_i\right)_x = \bigcup_{i=1}^\infty (E_i)_x$ tells us that the $(E_n)_x$ are measurable sets ascending to E_x, so we get the above limit via [Elb22, Proposition 6.36]. Thus, f_E is the pointwise limit of the f_{E_n}s, so f_E is $\mathcal{M}$-measurable.

2. Additionally, $\mathcal{D}$ is closed under countable descending intersections: the same argument of the previous point works, exchanging the word "ascending" with "descending," exchanging unions with intersections, and exchanging the citation with [Elb22, Corollary 6.37]. Note that our sets are of finite measure because Y is finite!

To proceed with the proof, we pick up the following definition.

Definition 1.31 (monotone class)**.** Fix a set Ω. Then a *monotone class* is a collection $\mathcal{C} \subseteq \mathcal{P}(\Omega)$ which contains $\varnothing$ and is closed under countable ascending unions and countable descending intersections.

In particular, we have shown that $\mathcal{D}$ is a monotone class. We will want the following fact about monotone classes.

Lemma 1.32. Fix a set Ω, and let $\mathcal{A}$ be an algebra on Ω. Then the smallest monotone class $\mathcal{C}$ containing $\mathcal{A}$ is a σ-algebra.

Proof. Note that the notion of a "smallest monotone class" makes sense because the intersection of monotone classes is another monotone class, so we can take $\mathcal{C}$ to be the intersection of all monotone classes containing $\mathcal{A}$. Anyway, here are our checks.

1. Fix $\mathcal{D}$ to be the collection of subsets of Ω whose complement is in $\mathcal{C}$. We claim that $\mathcal{D}$ is a monotone class; this will imply that $\mathcal{D}$ contains $\mathcal{C}$ (because $\mathcal{D}$ of course contains $\mathcal{A}$, which is closed under complements), meaning that $\mathcal{C}$ is closed under complements. For countable ascending unions of $E_1 \subseteq E_2 \subseteq \cdots$, we note that the union E has
$$\Omega \setminus E = \bigcap_{i=1}^{\infty} \Omega \setminus E_i,$$
which is in $\mathcal{C}$, so $E \in \mathcal{D}$. Replacing unions with intersections shows that $\mathcal{D}$ is closed under

2. If $A \in \mathcal{A}$ and $B \in \mathcal{C}$, then we claim $A \cup B \in \mathcal{C}$. Well, fix A, and we set
$$\mathcal{D}_A \coloneqq \{E \subseteq \Omega : A \cup E \in \mathcal{C}\}.$$
We claim that $\mathcal{D}_A$ is a monotone class, and it contains $\mathcal{A}$ (which is closed under unions), so $\mathcal{D}_A$ will contain $\mathcal{C}$, proving the claim. For ascending unions $E_1 \subseteq E_2 \subseteq \cdots$, we note
$$\left(\bigcup_{i=1}^{\infty} E_i\right) \cap A = \bigcup_{i=1}^{\infty} (E_i \cap A),$$
so the union is still in $\mathcal{D}_A$. Replacing the big $\bigcup$ with a big $\bigcap$ and working with a descending intersection shows that $\mathcal{D}_A$ is a monotone class, as needed.

3. If $A \in \mathcal{C}$ and $B \in \mathcal{C}$, then we claim $A \cup B \in \mathcal{C}$. Once again, we fix A and set
$$\mathcal{D}_A\{E \subseteq \Omega : A \cup E \in \mathcal{C}\}.$$
The previous check tells us that $\mathcal{D}_A$ contains $\mathcal{A}$. The same proof as the previous check tells us that $\mathcal{D}_A$ is a monotone class, so we once again are allowed to conclude that $\mathcal{D}_A$ contains $\mathcal{C}$, so the claim follows.

4. Lastly, we show $\mathcal{C}$ is closed under countable unions. Well, given a countable collection $\{E_i\}_{i=1}^{\infty} \subseteq \mathcal{C}$, we set
$$F_j \coloneqq \bigcup_{i \le j} E_i,$$
which is in $\mathcal{C}$ by the previous check. Then the union of the $E_\bullet$s is the union of the $F_\bullet$s, but $\mathcal{C}$ is a monotone class, so it contains the union of the $F_\bullet$s (which are ascending), so we are done. ■

Now, we see that Lemma 1.32 finishes the proof: $\mathcal{D}$ must contain the smallest monotone class containing $\mathcal{A}(X,Y)$, which is a σ-algebra by Lemma 1.32, so $\mathcal{D}$ contains the smallest σ-algebra containing $\mathcal{A}(X,Y)$, so $\mathcal{D}$ contains $\mathcal{M}\otimes\mathcal{N}$, as needed. ■

We now complete the proof of Theorem 1.25; the following is the statement of (c) for one of the equalities where $f=1_E$.

Lemma 1.33. Fix σ-finite measure spaces $(X,\mathcal{M},\mu)$ and $(Y,\mathcal{N},\nu)$. Then

$$(\mu\times\nu)(E)=\int_X \nu(E_x)\,d\mu(x).$$

Proof. Proceed as in Lemma 1.29. Explicitly, set

$$\mathcal{D}:=\left\{E\subseteq X\times Y : x\mapsto\nu(E_x)\text{ is measurable and }(\mu\times\nu)(E)=\int_X\nu(E_x)\,d\mu(x)\right\}.$$

The construction of the measure $\mu\times\nu$ implies this equality when E is a measurable rectangle or even when E is a disjoint union of measurable rectangles, so $\mathcal{D}$ contains $\mathcal{A}(X,Y)$. A direct computation shows that $\mathcal{D}$ is closed under complements, and the Dominated convergence theorem [Elb22, Theorem 9.14] shows that $\mathcal{D}$ is closed under ascending unions and descending intersections. So $\mathcal{D}$ is a monotone class containing the algebra $\mathcal{A}(X,Y)$, which implies that $\mathcal{D}$ contains the smallest monotone class containing $\mathcal{A}(X,Y)$, which is a σ-algebra by Lemma 1.32, so $\mathcal{D}$ contains the smallest σ-algebra containing $\mathcal{A}$, so $\mathcal{D}$ contains $\mathcal{M}\otimes\mathcal{N}$. ■

1.4 January 24

Let's begin.

1.4.1 Addenda to Tonelli's Theorem

Last class we completed the proof of Theorem 1.25. We take a moment to note that there is a "mirror" of Tonelli's theorem as follows.

Theorem 1.34 (Tonelli). Fix σ-finite measure spaces $(X,\mathcal{M},\mu)$ and $(Y,\mathcal{N},\nu)$. Fix a measurable function $f\colon X\times Y\to[0,\infty]$. Then the following hold.

(a) The function $x\mapsto f(x,y)$ is $\mathcal{M}$-measurable.

(b) The function $y\mapsto\int_Y f(x,y)\,d\nu(y)$ is $\mathcal{N}$-measurable.

(c) We have

$$\int_{X\times Y} f\,d(\mu\times\nu)=\int_Y\int_X f(x,y)\,d\mu(x)\,d\nu(y).$$

We will not write this proof because one can simply interchange X and Y in the provided proof of Theorem 1.25. Perhaps one will complain that the definition of the product premeasure Definition 1.18 appears asymmetric, but in fact it does not. Indeed, Remark 1.20 explains that the measure of a measurable rectangle is symmetric, which then explains how to measure anything in $\mathcal{A}(X,Y)=\mathcal{A}(Y,X)$ symmetrically, and then the Extension Theorem 1.11 tells us that this uniquely measures anything in $\mathcal{M}\otimes\mathcal{N}$ symmetrically.

Corollary 1.35. Fix σ-finite measure spaces $(X,\mathcal{M},\mu)$ and $(Y,\mathcal{N},\nu)$. If $f\colon X\times Y\to[0,\infty]$ is $(\mathcal{M}\otimes\mathcal{N})$-measurable, then

$$\int_X\int_Y f(x,y)\,d\nu(y)\,d\mu(x)=\int_Y\int_X f(x,y)\,d\mu(x)\,d\nu(y).$$

Proof. Combine Theorems 1.25 and 1.34. ■

Perhaps one might worry about spaces which are not σ-finite. Here are some examples.

Example 1.36. Fix an uncountable set X. Then one can define a measure μ on $\mathcal{M} := \mathcal{P}(X)$ by $\mu(E) := \#E$. This is not σ-finite because subsets of X has finite measure if and only if it is finite, and X cannot be covered by countably many finite sets.

Example 1.37. Fix an uncountable set X, and let $\mathcal{M}$ be the collection of countable and cocountable subsets. Then the function $\mu\colon \mathcal{M} \to [0,\infty]$ defined by

$$\mu(E) := \begin{cases} 0 & \text{if } E \text{ is countable,} \\ \infty & \text{if } E \text{ is cocountable} \end{cases}$$

is a measure. Now, X fails to be σ-finite because the sets of finite measure are exactly the countable ones, and X cannot be covered by countably many countable subsets.

1.4.2 Fubini's Theorem

We are now ready to state Fubini's theorem. This requires the following definition.

Definition 1.38. Fix a measure space $(X,\mathcal{M},\mu)$. Then we define $L^1(\mu)$ consists of the measurable functions $f\colon X \to \mathbb{C}$ (defined almost everywhere) such that

$$\int_X |f| \, d\mu < \infty.$$

Remark 1.39. If $f \in L^1(\mu)$, then one sees that $\int_X f\,d\mu$ makes sense. Namely, one has

$$\int_X f\,d\mu = \int_X \operatorname{Re} f\,d\mu + i\int_X \operatorname{Im} f\,d\mu,$$

and the integrals $\int_X \operatorname{Re} f\,d\mu$ and $\int_X \operatorname{Im} f\,d\mu$ are both bounded by $f \in L^1(\mu)$. Something like [Elb22, Proposition 9.22] assures us that this makes sense (upon taking differences).

Theorem 1.40 (Fubini)**.** Fix σ-finite measure spaces $(X,\mathcal{M},\mu)$ and $(Y,\mathcal{N},\nu)$. Fix a measurable function $f\colon X\times Y \to \mathbb{C}$ such that $\int_{X\times Y} |f| \, d(\mu\times\nu) < \infty$. Then the following hold.

(a) For μ-almost every $x \in X$, the function $f_x\colon y \mapsto f(x,y)$ is defined and $\mathcal{N}$-measurable and in $L^1(\nu)$.

(b) The function $x \mapsto \int_Y f(x,y)\,d\nu(y)$ is defined almost everywhere and $\mathcal{M}$-measurable and in $L^1(\mu)$.

(c) We have
$$\int_{X\times Y} f\,d(\mu\times\nu) = \int_X \int_Y f(x,y)\,d\nu(y)\,d\mu(x).$$

Proof. Note that we have the result for nonnegative functions by Theorem 1.25. The idea is to reduce to this case. Here we go.

- By writing $f = u + iv$ for $u := \operatorname{Re} f$ and $v := \operatorname{Im} f$, we may assume that f is real-valued. Explicitly, note the set of functions f satisfying the conclusions of (a)–(c) is a $\mathbb{C}$-vector space by some addition and scalar multiplication. Notably, we still have the hypotheses that $\int_{X\times Y} |u| \, d(\mu\times\nu) < \infty$ and $\int_{X\times Y} |v| \, d(\mu\times\nu) < \infty$.

- By writing $f = f^+ - f^-$ for $f^+, f^- \geq 0$, we will reduce to the case that f is nonnegative. Namely, achieving the result for the two functions

$$f^+(x) := \begin{cases} f(x) & \text{if } f(x) \geq 0, \\ 0 & \text{if } f(x) \leq 0, \end{cases} \qquad \text{and} \qquad f^-(x) := \begin{cases} -f(x) & \text{if } f(x) \leq 0, \\ 0 & \text{if } f(x) \geq 0, \end{cases}$$

  will achieve the result for f by summing.

Note that there is a technicality hidden in the above reasoning with linear combinations: for example, for the second reduction, even though we have the conclusion for f_x^+ and f_x^- are $\mathcal{N}$-measurable for all x, their difference might not be in $L^1(\nu)$ always. Well, we note that we can compute

$$\int_X \left(\int_Y f^+(x,y)\,d\nu(y) \right) d\mu(x) = \int_{X\times Y} f^+\,d(\mu\times\nu) < \infty,$$

so the inner function $x \mapsto \int_Y f_x^+\,d\nu(y)$ must be finite almost everywhere, or else this integral would be infinite! So we do indeed achieve that f_x^+ and f_x^- are in $L^1(\nu)$ almost everywhere, so their difference is in $L^1(\nu)$ almost everywhere. The argument for taking linear combinations in (b) is similar. ■

Let's see an example of why we want the hypothesis in Theorem 1.40.

Example 1.41. Set $X := \mathbb{N}$, and let μ and ν denote the counting measures on $\mathcal{M} = \mathcal{N} := \mathcal{P}(X)$. Note that $\mathcal{A}(X,X) = \mathcal{P}(X^2)$, so the product measure $\mu\times\nu$ is defined on all subsets; furthermore, we can see that the measure of a singleton is 1, so $\mu\times\nu$ is the counting measure. Then we define the function

$$f(x,y) := \begin{cases} +1 & \text{if } x = y, \\ -1 & \text{if } y = x+1, \\ 0 & \text{otherwise.} \end{cases}$$

For each x, we compute $\int_Y f(x,y)\,d\nu(y) = 0$ because each value of x has two values of y where $f(x,y)$ is nonzero. On the other hand, for each y, we compute $\int_X f(x,y)\,d\mu(x) = 0$ if $y > 0$ but is 0 if $y = 0$. The problem here is that $\int_{X\times Y} f(x,y)\,d(\mu\times\nu) = 0$.

Remark 1.42. According to Professor Christ, the above example is a "catastrophic failure" of a theorem rather than a "technical" one.

Remark 1.43. By induction, we are able to take products of any finite product of σ-finite measure spaces. Alternatively, one can redo the entire theory to do measurable rectangular prisms and so on. There are some extra checks here (e.g., does forming products associate meaningfully?), but it will work out in the end, essentially by the uniqueness of the construction provided by Theorem 1.11. Namely, up to the identification of products, we get the identification of the product σ-algebras and product measures because they should all agree on measurable rectangles, from which everything is generated.

1.5 January 26

Here we go.

1.5.1 Complete Measures

Recall the following definition.

Definition 1.44 (complete)**.** Fix a measure space $(X, \mathcal{M}, \mu)$. Then μ is *complete* if and only if it has all null sets: if $E \in \mathcal{M}$ has $\mu(E) = 0$, then any subset $F \subseteq R$ has $F \in \mathcal{M}$.

Non-Example 1.45. Let μ be the Lebesgue measure on the Borel algebra $\mathcal{M}$ of $\mathbb{R}$. Then μ is not complete: there are subsets of null sets which are not Borel. In fact, there are only $|\mathbb{R}|$ many Borel sets by a counting construction, but there are more null sets.

We now recall the following construction.

Proposition 1.46. Fix a measure space $(X, \mathcal{M}, \mu)$. Then there is a measure space $(X, \widetilde{\mathcal{M}}, \widetilde{\mu})$ such that $\widetilde{\mathcal{M}} \supseteq \mathcal{M}$, $\widetilde{\mu}$ extends μ, and $\widetilde{\mu}$ is complete.

Proof. We showed this in Math 202A. To sketch the idea, simply set $\widetilde{\mathcal{M}}$ as the union of elements in $\mathcal{M}$ with subsets of null sets, and define

$$\widetilde{\mu}(A \cup E) = \mu(A)$$

where $A \in \mathcal{M}$ and E is a subset of a null set. Checking that this is actually a measure space is annoying and hence omitted; checking that $\widetilde{\mu}$ is complete simply follows because $\widetilde{\mu}(A \cup E) = 0$ and $F \subseteq A \cup E$ implies that $\mu(A) = 0$ actually, so $A \cup E$ is a subset of a null set for μ still, so F is a subset of a null set for μ, so $F \in \widetilde{\mathcal{M}}$. ■

Remark 1.47. One can see that the above constructions the "minimal" completion in the sense that any other completion $(X, \widetilde{\mathcal{M}}', \widetilde{\mu}')$ has $\widetilde{\mathcal{M}} \subseteq \widetilde{\mathcal{M}}'$ and $\widetilde{\mu}'|_{\widetilde{\mathcal{M}}} = \widetilde{\mu}$.

We would like to examine completeness for our product measures. Sadly, in most cases, having complete metric spaces does not make the product measure complete.

Example 1.48. Let $(\mathbb{R}, \mathcal{L}, \mu)$ be the completion of the Borel Lebesgue measure. Then the product measure $\left(\mathbb{R}^2, \mathcal{L} \otimes \mathcal{L}, \mu \times \mu\right)$ fails to be complete. To see this, let $A \subseteq \mathbb{R}$ be a subset not in $\mathcal{L}$, and let $B \in \mathcal{L}$ be a nonempty null set. Then $E := A \times B$ is the desired set.

- Note $E \subseteq \mathbb{R} \times B$, and $\mathbb{R} \times B$ is a null set: note $\mu([-r, r] \times B) = 2r\mu(B) = 0$, so sending $r \to \infty$ shows $\mu(\mathbb{R} \times B) = 0$. Alternatively, we simply recall that it is convention that $\infty \times 0 = 0$ here.
- On the other hand, $E \notin \mathcal{L} \otimes \mathcal{L}$. Indeed, Theorem 1.25 would tell us that $E \in \mathcal{L} \otimes \mathcal{L}$ implies that $A = E_y$ is measurable for all $y \in B$, which is false.

To remedy our situation, we have the following result to recover Theorem 1.25.

Theorem 1.49. Fix σ-finite complete measure spaces $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$, and let $(X \times Y, \mathcal{L}, \lambda)$ denote the completion of $(X \times Y, \mathcal{M} \otimes \mathcal{N}, \mu \times \nu)$. Let $f \colon X \times Y \to [0, \infty]$ be $\mathcal{L}$-measurable.

(a) For μ-almost every x, the function $y \mapsto f(x, y)$ is $\mathcal{N}$-measurable.

(b) The function $x \mapsto \int_Y f(x, y)\, d\nu(y)$ is defined μ-almost everywhere, and it is $\mathcal{M}$-measurable.

(c) We have

$$\int_{X \times Y} f\, d(\mu \times \nu) = \int_X \int_Y f(x, y)\, d\nu(y)\, d\mu(x).$$

One can recover Theorem 1.40 in the analogous way.

We will not prove Theorem 1.49 in detail. The main point is to show the following.

Lemma 1.50. Fix everything as in Theorem 1.49. For $E \in \mathcal{L}$ with $\lambda(E) = 0$, we have $E_x \in \mathcal{N}$ for μ-almost every $x \in X$.

Indeed, as before, we can restrict to the case where $f = 1_E$, so this recovers (a), and then the arguments of Theorem 1.25 port over to prove (b) and (c) cleanly.

Proof. The main point is that we can find $A \supseteq E$ such that $A \in \mathcal{M} \otimes \mathcal{N}$. But then Theorem 1.25 reassures us that

$$(\mu \times \nu)(A) = \int_X \nu(A_x)\, d\mu(x) = 0,$$

so $\nu(A_x) = 0$ for μ-almost every $x \in X$. ■

1.5.2 Measuring Euclidean Spaces

For our discussion, we will want two pieces of notation.

Notation 1.51. Fix a nonnegative integer $d \geq 0$.

- For a topological space X, we define $\operatorname{Borel}(X)$ to be the σ-algebra of Borel subsets in X.
- We define $\mathcal{B}^d := \operatorname{Borel}(\mathbb{R})^d$.

There is some care one must take in our notation here, but not much care.

Proposition 1.52. For all nonnegative integers $d \geq 0$, we have $\mathcal{B}^d = \operatorname{Borel}\left(\mathbb{R}^d\right)$.

We will require the following lemma.

Lemma 1.53. Fix a positive integer $d \geq 1$, and let $U \subseteq \mathbb{R}^d$ be open. Then U is a disjoint union of countably many half-open cubes. Here, a "cube" is a product of intervals of the form

$$\prod_{i=1}^d [a_i, b_i).$$

Proof. For $n \geq 0$, let $\mathcal{D}_n$ denote the collection of "dyadic cubes" of the form

$$\prod_{i=1}^d [a_i, b_i)$$

where $b_i - a_i = 2^{-n}$ and $a_i \in 2^{-n}\mathbb{Z}$ for each i. We now let $\mathcal{C}$ denote the collection of cubes in some $\mathcal{D}_n$ contained in U. Because any point $x \in U$ is contained in some open ball $B(x, r)$ such that $B(x, 2r) \subseteq U$, we can find a cube in $\mathcal{D}_n$ for n large enough living inside $B(x, 2r)$ containing x.

The issue is now to make the cubes $\mathcal{C}$ disjoint. Well, define $\mathcal{C}'$ to be the subcollection of "maximal" cubes in the sense that $\prod_{i=1}^d [a_i, b_i)$ will be in $\mathcal{C}'$ if and only if $\prod_{i=1}^d [a_i, b_i + (b_i - a_i))$ is not in $\mathcal{C}$. Certainly $\mathcal{C}'$ still covers U, and its cubes are disjoint: certainly no two cubes in $\mathcal{C}'$ contain each other by construction, and dyadic cubes either contain each other or are disjoint. ■

1.6 January 29

Here we go.

1.6.1 More on Measuring Euclidean Spaces

We now prove the following statement from last class.

Proposition 1.52. For all nonnegative integers $d \geq 0$, we have $\mathcal{B}^d = \operatorname{Borel}\left(\mathbb{R}^d\right)$.

Proof. The case of $d = 0$ and $d = 1$ have no content. Now, in one direction, we see that $\operatorname{Borel}\left(\mathbb{R}^d\right) \subseteq \mathcal{B}^d$: note $\mathcal{B}^d$ is a σ-algebra containing the cubes of the form in Lemma 1.53, so it also contains open subsets of $\mathbb{R}^d$ by Lemma 1.53, so we conclude.

We now show that $\mathcal{B}^d \subseteq \operatorname{Borel}\left(\mathbb{R}^d\right)$. By the definition of $\mathcal{B}^d$, it is enough to show that $\operatorname{Borel}\left(\mathbb{R}^d\right)$ contains all measurable rectangles $A_1 \times \cdots \times A_d$ where $A_1 \in \operatorname{Borel}(\mathbb{R})$. We proceed inductively, claiming that if $A_i, A_{i+1}, \ldots, A_d$ are all open, then the entire product is Borel. For $i = 1$, there is nothing to do. Now, for the induction, suppose we have the claim for i, and we want the claim for $i + 1$. Well, fix everything except A_i, and we define

$$\mathcal{D} := \{B \subseteq \mathbb{R} : A_1 \times \cdots \times A_{i-1} \times B \times A_{i+1} \times \cdots \times A_d \text{ is Borel}\}.$$

Certainly if B is open, then $B \in \mathcal{D}$ by the induction. Additionally, arbitrary unions and intersections distribute over $\times$, so $\mathcal{D}$ is closed arbitrary unions and intersections. Lastly, if $B \in \mathcal{D}$, we see that

$$\begin{aligned}(A_1 \times \cdots \times A_{i-1} \times (\mathbb{R} \setminus B) \times A_{i+1} \times \cdots \times A_d) = (&A_1 \times \cdots \times A_{i-1} \times \mathbb{R} \times A_{i+1} \times \cdots \times A_d) \\ &\setminus (A_1 \times \cdots \times A_{i-1} \times B \times A_{i+1} \times \cdots \times A_d),\end{aligned}$$

and the right-hand side is the subtraction of two Borel sets, so $(\mathbb{R} \setminus B) \in \mathcal{D}$. Thus, $\mathcal{D}$ is a σ-algebra containing opens, so $\mathcal{D}$ contains $\operatorname{Borel}(\mathbb{R})$. ■

We now move towards some regularity conditions on our measures.

Definition 1.54 (regular)**.** Fix a topological space X and a measure μ on a σ-algebra $\mathcal{M}$ containing the Borel sets.

- μ is *outer regular* if and only if any $E \in \mathcal{M}$ has
$$\mu(E) = \inf_{\text{open } U \supseteq E} \mu(U).$$
- μ is *inner regular* if and only if any $E \in \mathcal{M}$ has
$$\mu(E) = \sup_{\text{compact } K \subseteq E} \mu(K).$$

Here is our result.

Theorem 1.55. Fix a nonnegative integer $d \geq 0$. Then the measures μ and $\widetilde{\mu}$ on $\mathbb{R}^d$ is outer regular.

Proof. The statement for $d = 0$ has no content. The outer regularity at $d = 1$ follows by its construction [Elb22, Lemma 6.15]; we will prove inner regularity from outer regularity momentarily. We proceed in steps.

1. We now show outer regularity for $d \geq 2$, using the $d = 1$ case. Take $E \in \mathcal{B}^d$, permissible by Proposition 1.52. If E has infinite measure, we can take $\mathbb{R}^d$ as the needed open set. Otherwise, we take $\varepsilon > 0$. By construction of μ, we get $\{A_i^j\}_{1 \leq i \leq d, 1 \leq j \leq n} \in \mathcal{B}^1$ of positive measure such that
$$\sum_{j=1}^n \mu(A_1^j) \cdots \mu(A_d^j) < \mu(E) + \varepsilon.$$

We now use outer regularity in dimension 1. For each A_i^j, we get some $U_i^j \supseteq A_i^j$ whose measure is within $\varepsilon \min\{A_i^j\} > 0$ of $\mu(A_i^j)$. Then we set U as the union of the $U_1^j \times \cdots \times U_d^j$ and find

$$\mu(U) \leq \sum_{j=1}^n \left(\mu(A_1^j) + \min\{A_k^\ell : k, \ell\}\varepsilon\right) \cdots \left(\mu(A_d^j) + \min\{A_k^\ell : k, \ell\}\varepsilon\right),$$

which we see is upper-bounded by

$$\underbrace{\sum_{j=1}^n \mu(A_1^j) \cdots \mu(A_d^j)}_{<\mu(E)+\varepsilon} + 2^n n\varepsilon \sum_{j=1}^n \mu(A_1^j) \cdots \mu(A_d^j)$$

when the ε_{ij} are chosen to be sufficiently small, and we are flagrantly collecting terms without explanation. Regardless, sending $\varepsilon \to 0^+$ recovers the result.

2. As an intermediate step, we note that we have the following form of outer regularity: for any $E \in \mathcal{B}^d$ of finite measure and $\varepsilon > 0$, we can find an open $U \supseteq E$ such that $\mu(U \setminus E) < \varepsilon$. In fact, we can even allow E to be of infinite measure: our measure is σ-finite, so we can write $E = \bigcup_{n=1}^\infty E_n$ where $E_n \in \mathcal{B}^d$ for each n and has finite measure. Then for any $\varepsilon > 0$, we find $U_n \supseteq E_n$ with $\mu(U_n) < \mu(E_n) + 2^{-n}\varepsilon$, and we see
$$\bigcup_{n=1}^\infty (U_n \setminus E_n) \supseteq U \setminus E,$$
where $U := \bigcup_{n=1}^\infty U_n$. Now, the left-hand side has measure bounded by $\sum_{n=1}^\infty 2^{-n}\varepsilon = \varepsilon$, as desired.

3. We now show inner regularity from outer regularity, for any $d \geq 1$. Fix $E \in \mathcal{B}^d$ and some $\varepsilon > 0$. We use $(-)^c$ to denote complement. Now, we are given some $U \supseteq E^c$ such that $\mu(U \setminus E^c) < \varepsilon$, but $U \setminus E^c = U \cap E = E \setminus U^c$, so $\mu(E \setminus U^c) < \varepsilon$. Now, we note that the closed set U^c satisfies
$$\mu(U^c) = \sup_{\text{compact } K \subseteq U^c} \mu(K)$$
because we can set $K_n := U^c \cap \overline{B(0,n)}$ to be compact subsets of U^c ascending to U^c, meaning $\mu(K_n) \to \mu(U^c)$ as $n \to \infty$. In particular, we can find n large enough so that $K_n \subseteq U^c$ has $\mu(E \setminus K_n) < 2\varepsilon$. Sending $\varepsilon \to 0^+$ completes the proof. ■

1.7 January 31

Ok let's begin.

Remark 1.56. To produce the Lebesgue measure on $\mathbb{R}^d$, one can imagine completing $(\mathbb{R}^d, \mathcal{L}^d, m^d)$ or completing $(\mathbb{R}^d, \mathcal{B}^d, m^d)$. Of course, one may just focus on showing that the σ-algebras are the same because then everything is the disjoint union of a null set and a measurable set.

For example, we note that $\mathcal{B}^1 \subseteq \mathcal{L}^1$, so completing makes $\widehat{\mathcal{B}^d} \subseteq \widehat{\mathcal{L}^d}$ by construction of our completion. For the reverse inclusion, by construction of our completion, it suffices to show that $\mathcal{L}^d \subseteq \widehat{\mathcal{B}^d}$. Looking at these as σ-algebras, it suffices to show that $\widehat{\mathcal{B}^d}$ contains measurable rectangles of $\mathcal{L}^1$-sets. Well, each set in $\mathcal{L}^1$ can be written as the union of a Borel set and a null set, so we can write the needed measurable rectangle as
$$(B_1 \cup N_1) \times (B_2 \cup N_2) \times \cdots \times (B_d \cup N_d).$$
Expanding out the product, the "leading term" $B_1 \times \cdots \times B_d$ is Borel, and then the remaining terms have null sets in them, so they are null sets. So the entire thing lives in $\mathcal{B}^d$.

1.7.1 Measuring Affine Maps

We will be interested in affine automorphisms of $\mathbb{R}^d$.

Definition 1.57 (affine)**.** An *affine map* $f\colon \mathbb{R}^d \to \mathbb{R}^d$ is one which can be written as

$$f(v) := Tv + a$$

where $T \in \mathrm{GL}_d(\mathbb{R})$ and $v_0 \in \mathbb{R}^d$.

Remark 1.58. Note that f is linear in its coordinates, so it is continuous. The inverse map is $v \mapsto T^{-1}(v - a)$, which is also affine, so f is in fact a homeomorphism and so sends open sets to open sets.

Let's quickly check that affine maps preserve Lebesgue sets.

Lemma 1.59. Fix a homeomorphism $h\colon X \to X$, where $(X, \mathcal{L}, \widehat{\mu})$ is a complete Borel measure space.

(a) h sends Borel sets to Borel sets.

(b) Suppose h sends Borel null sets to null sets. Then h preserves Lebesgue sets.

Proof. Here we go.

(a) Let $\mathcal{D}$ denote the collection of $E \subseteq X$ such that $h(E)$ is Borel. Well, h is an open map, so $\mathcal{D}$ contains open sets. Further, $\mathcal{D}$ is a σ-algebra because taking images preserves unions and complements because h is a bijection. Thus, $\mathcal{D}$ contains all Borel sets.

(b) Fix a Lebesgue set $B \cup N$ where N is a null set. Then $h(B \cup N) = h(B) \cup h(N)$ is the union of a Borel set $h(B)$ (by (a)) and a null set $h(N)$ by hypothesis. ∎

Remark 1.60. Note that an affine map $f\colon \mathbb{R}^d \to \mathbb{R}^d$ is a homeomorphism, so (a) above tells us that Borel sets get sent to Borel sets.

We can actually measure our images pretty well.

Proposition 1.61. Fix an affine map $f\colon \mathbb{R}^d \to \mathbb{R}^d$ of the form $f(v) := Tv + a$. For a Lebesgue set E, we have

$$\mu(f(E)) = |\det T|\, \mu(E).$$

Proof. By definition, we can decompose f into a translation map $\tau\colon v \mapsto v + a$ and a linear map $T\colon v \mapsto Tv$. It then suffices to check the result on translations and linear maps.

Well, for translations, we need $\mu(E + a) = \mu(E)$ for any Lebesgue set E. It suffices to do this for Borel sets E. Letting $\mathcal{D}$ denote the collection of Borel sets E with $\mu(E + a) = \mu(E)$, we note that $\mathcal{D}$ contains all cubes (compute the measure as a product of the side lengths via, say, Theorem 1.25), so $\mathcal{D}$ contains all open sets by Lemma 1.53. Further, we can see that $\mathcal{D}$ is a σ-algebra because it is closed under unions and complements because translation is a bijection, and μ preserves unions and complements (approximately speaking). (The complement argument needs to know that μ is σ-finite.)

For linear maps, we break down our maps even more. We can write any linear map T as a composition of maps of the following kinds.

- Permutations of coordinates: for $\sigma \in \mathrm{Sym}(\{1, \ldots, d\})$, we have the linear map $P_\sigma\colon (x_1, \ldots, x_d) \mapsto (x_{\sigma(1)}, \ldots, x_{\sigma(d)})$.

- Dilation: for $t \in \mathbb{R}^\times$, we have the linear map $D_t \colon (x_1, \ldots, x_d) \mapsto (x_1, \ldots, x_{d-1}, tx_d)$.
- Skew shifts: for $v \in \mathbb{R}^{d-1}$, we have the linear map $S_v \colon (x_1, \ldots, x_n) \mapsto (x_1, \ldots, x_{d-1}, x_d + (x_1, \ldots, x_{d-1}) \cdot v)$.

Gaussian elimination shows any linear map T is the composite of maps of the above form, so it suffices to take of matrices of the above form.

Well, for permutations, one expands out the integral and exchanges integrals via Theorem 1.25. For skew shifts, a trick is required. We use Theorem 1.25 (at the start and end) in order to write

$$\begin{aligned}
\mu(T(E)) &= \int_{(x_1,\ldots,x_{d-1})\in\mathbb{R}^{d-1}} \mu(T(E)_{(x_1,\ldots,x_{d-1})})\,d\mu(x_1,\ldots,x_{d-1}) \\
&= \int_{(x_1,\ldots,x_{d-1})\in\mathbb{R}^{d-1}} \mu(E_{(x_1,\ldots,x_{d-1}} + v\cdot(x_1,\ldots,x_{d-1}))\,d\mu(x_1,\ldots,x_{d-1}) \\
&\overset{*}{=} \int_{(x_1,\ldots,x_{d-1})\in\mathbb{R}^{d-1}} \mu(E_{(x_1,\ldots,x_{d-1})})\,d\mu(x_1,\ldots,x_{d-1}) \\
&= \mu(E),
\end{aligned}$$

as needed. Notably, $\overset{*}{=}$ used the fact that we already understand translations.

For dilations, we want to show that $\mu(T(E)) = |t|\,\mu(E)$. At $d = 1$, we note that the conclusion holds on open intervals by construction of the measure, so by taking finite unions, it holds on all open sets; then we can achieve the full conclusion at $d = 1$ by using Theorem 1.55. Now, for higher dimensions, we argue as above via Theorem 1.25 to note

$$\begin{aligned}
\mu(T(E)) &= \int_{(x_1,\ldots,x_{d-1})\in\mathbb{R}^{d-1}} \mu(T(E)_{(x_1,\ldots,x_{d-1})})\,d\mu(x_1,\ldots,x_{d-1}) \\
&= \int_{(x_1,\ldots,x_{d-1})\in\mathbb{R}^{d-1}} |t|\,\mu(E_{(x_1,\ldots,x_{d-1})})\,d\mu(x_1,\ldots,x_{d-1}) \\
&= |t|\,\mu(E),
\end{aligned}$$

as needed. ■

Remark 1.62. In fact, the above proof shows that measure-zero sets go to measure-sets under affine maps. In particular, affine maps send null sets to null sets.

1.8 February 2

We began class by finishing a proof from last class, into which I have edited directly.

1.8.1 Measuring Lipschitz Functions

We would like to understand functions with more curves than affine maps.

Definition 1.63 (Lipschitz). Fix an open subset $U \subseteq \mathbb{R}^d$, and let $f \colon U \to \mathbb{R}^k$ be a map. Then f is *Lipschitz* if and only if there is a finite positive real number A such that

$$|f(x) - f(y)| \le A\,|x - y|$$

for all $x, y \in U$. We define the *Lipschitz constant* $\|f\|_{\mathrm{Lip}}$ to be the infimum of all such possible A.

We need to know that Lipschitz functions send Borel sets to Borel sets.

Lemma 1.64. Fix a Lipschitz function $f\colon U \to \mathbb{R}^d$ where $U \subseteq \mathbb{R}^d$. Then for any Lebesgue measurable $E \subseteq U$, we have that $f(E)$ is Lebesgue measurable. In fact, null sets map to null sets.

Proof. Certainly if $K \subseteq U$ is compact, then $f(K)$ is compact (and in particular closed) and hence Borel and hence Lebesgue measurable.

Lastly, let E be an arbitrary Lebesgue set. By Theorem 1.55, we get a sequence of ascending compact sets $\{K_n\}_{n=1}^{\infty}$ such that $S \cup \bigcup_{n=1}^{\infty} K_n = E$ where S is a null set (and $S = \varnothing$ if E is actually Borel). Now, hitting this with f, we see that $f\left(\bigcup_{n=1}^{\infty} K_n\right) = \bigcup_{n=1}^{\infty} f(K_n)$ is the countable union of Borel sets, which is Borel and hence Lebesgue. Note that we have so far shown that E being Borel makes $f(E)$ measurable.

Lastly, we must show that $f(S)$ is a null set. As an intermediate step, we claim that there is a constant $C_f > 0$ such that

$$\mu(f(V)) \overset{?}{\leq} C_f \mu(V)$$

for any open $V \subseteq U$. To see why this is enough, note by Theorem 1.55, for any $\varepsilon > 0$, we can find an open $V \supseteq S$ such that $\mu(V) < \varepsilon$, so $\mu(f(S)) \leq \mu(f(V)) \leq C_f\varepsilon$. Sending $\varepsilon \to 0^+$ then shows $\mu(f(S)) = 0$.

It remains to show the claim. Well, use Lemma 1.53 to write V as the disjoint union of a countable collection of cubes $\{Q_n\}_{n\in\mathbb{N}}$, where Q_n has side length ℓ_n. Then

$$\mu(f(V)) \leq \sum_{n\in\mathbb{N}} \mu(f(Q_n)).$$

Now, the diameter of Q_n is $\ell_n\sqrt{d}$, so it is contained in a ball of radius equal to half that, so f will send this ball to a ball of radius $\frac{1}{2}\|f\|_{\text{Lip}}\,\ell_n\sqrt{d}$, which is contained in a cube of side length $\|f\|_{\text{Lip}}\,\ell_n\sqrt{d}$, which has measure $\left(\|f\|_{\text{Lip}}\,\ell_n\sqrt{d}\right)^d$. In total, we see

$$\mu(f(Q_n)) \leq \underbrace{\left(\|f\|_{\text{Lip}}\sqrt{d}\right)^d}_{C_f :=} \ell_n^d,$$

so by summing we have found our needed constant C_f. Importantly, C_f does not depend on V at all, so we are okay. ■

Anyway, here is our main result to change variables.

Definition 1.65 (C^1-diffeomorphism)**.** Fix open subsets $U, V \subseteq \mathbb{R}^d$. A map $f\colon U \to V$ is a C^1-*diffeomorphism* if and only if f is bijective, and f and f^{-1} are both C^1; i.e., f and f^{-1} are both continuously differentiable everywhere.

Remark 1.66. If f is a C^1-diffeomorphism, then $Df \circ Df^{-1} = \mathrm{id}$, so Df is actually invertible. In particular, the Jacobian is nonzero.

Remark 1.67. For example, the Inverse function theorem approximately says that a function f having invertible derivative at a point x implies that f becomes a C^1-diffeomorphism in a neighborhood of x.

Theorem 1.68. Fix open subsets $U, V \subseteq \mathbb{R}^d$, and fix a C^1-diffeomorphism $\Phi\colon U \to V$. Then the following hold.

(a) For each Lebesgue measurable $E \subseteq U$, the set $\Phi(E)$ is still Lebesgue measurable.

(b) For each Lebesgue measurable $E \subseteq U$, we have

$$\mu(\Phi(E)) = \int_U |J_\Phi(x)| \, d\mu(x),$$

where J_Φ is the Jacobian $\det D_\Phi$.

(c) For a Lebesgue measurable function $f\colon V \to \mathbb{C}$, we have

$$\int_V f \, d\mu = \int_U (f \circ \Phi) \, |J_\Phi| \, d\mu,$$

under hypotheses establishing the integrals make sense: either f is nonnegative, or f is in integrable, or $(f \circ \Phi) \, |J_\Phi|$ is integrable.

We will prove this result next class.

1.9 February 5

We began class by finishing a proof from last class, into which I have edited directly.

1.9.1 Change of Variables

Before doing anything, here are some lemmas.

Lemma 1.69. Fix a compact convex subset $K \subseteq \mathbb{R}^d$. Suppose $U \subseteq \mathbb{R}^d$ is an open subset containing K, and suppose $f\colon U \to \mathbb{R}^n$ is continuously differentiable. Then $f|_K$ is Lipschitz.

Proof. Fix $x, y \in K$. We would like to approximate $f(x) - f(y)$ by $Df(x)(y - x)$, for which we use the Mean value theorem.

To bring our target down to one dimension, we note

$$|f(y) - f(x)| = w \cdot (f(y) - f(x)),$$

where w is the unit vector in the direction of $f(y) - f(x)$. In order to bring our source down to one dimension, we define

$$g(t) := w \cdot f(x + t(y - x)),$$

which is well-defined because K is convex. Note g is continuously differentiable because it is the composition of continuously differentiable functions; in particular, we find $g'(t) = w \cdot Df(x + t(y - x))(y - x)$. Everything is now in one dimension, so the Mean value theorem provides $c \in (0, 1)$ such that

$$\frac{g(1) - g(0)}{1 - 0} = g'(c) = w \cdot Df(x + c(y - x))(y - x).$$

Thus, we may have the sequence of bounds

$$\begin{aligned}
|f(y) - f(x)| &= |g(1) - g(0)| \\
&= |w \cdot Df(x + c(y - x))(y - x)| \\
&\le |w| \cdot |Df(x + c(y - x))(y - x)| \\
&\le |y - x| \cdot \max_{z \in K} \|Df(z)\| .
\end{aligned}$$

Now, f is continuously differentiable, and K is compact, so Df is bounded on K. So we say that f is Lipschitz with Lipschitz constant bounded above by $\max_{z\in K}\|Df(z)\|$. ■

Now, the main content of Theorem 1.68 is in the following result.

Lemma 1.70. Fix a C^1-diffeomorphism $\Phi\colon U\to V$ where $U,V\subseteq\mathbb{R}^d$ is open. Further, let U' be an open subset with compact closure $\overline{U'}\subseteq U$ such that there is $\delta_0>0$ such that any $x\in U'$ has $d(x,U^c)\geq 2\delta_0$. Then we have

$$\mu(\varphi(U'))\leq\int_{U'}|J_\Phi|\,d\mu.$$

Remark 1.71. Here, $d(x,U^c)$ denotes the (infimum of the) distance from x to the set U^c. Because $\overline{U'}$ is compact, we note that we also get $d(x,U^c)\geq 2\delta_0$ for each $x\in\overline{U'}$ because $x\mapsto d(x,U^c)$ is a continuous function.

To prove the above lemma, we will want the following lemma.

Lemma 1.72. Fix a C^1-diffeomorphism $\Phi\colon U\to V$ where $U,V\subseteq\mathbb{R}^d$ is open. Further, let $U'\subseteq U$ be an open subset with compact closure such that there is $\delta>0$ such that any $x\in U'$ has $d(x,U^c)\geq 2\delta_0$. Then there is a decreasing "remainder" function R such that $R(t)\to 0$ as $t\to 0^+$ while the "Taylor remainder"

$$\mathcal{R}(x,u)\coloneqq\Phi(x+u)-\Phi(x)-D\Phi(x)(u)$$

satisfies $|\mathcal{R}(x,u)|\leq|u|\,R(|u|)$ for all $x\in U$ and $|u|\leq\delta_0$.

The content here is that the remainder function $\mathcal{R}$ (which ought to go to zero as $u\to 0$) is bounded in a way that does not depend on x.

Proof of Lemma 1.72. For u such that $|u|\leq\delta_0$, the construction of δ_0 implies that any $x\in U'$ has $x+u\in U$; in fact, we have $d(y,U^c)\geq\delta_0$ for all y of the form $x+u$. Notably,

$$\{y\in U: d(y,U^c)\geq\delta_0\}$$

is compact: it is closed because $y\mapsto d(y,U^c)$ is continuous, and it is bounded because it is a fixed distance away from the compact set $\overline{U'}$.

We now use the Mean value theorem to conclude. The proof is similar to Lemma 1.69. Fix x and u as in the statement. Then there is a unit vector w such that

$$w\cdot\mathcal{R}(x,u)=|\mathcal{R}(x,u)|.$$

This allows us to define

$$g(t)\coloneqq w\cdot\mathcal{R}(x,u),$$

which is now a function $g\colon[0,1]\to\mathbb{R}$ which is continuously differentiable because everything in sight is continuously differentiable. As such, the Mean value theorem provides $c\in(0,1)$ such that

$$g(1)-g(0)=g'(c)=w\cdot(D\Phi(x+cu)(u)-D\Phi(x)(u)).$$

Thus,

$$\begin{aligned}|\mathcal{R}(x,u)|&=|g(1)-g(0)|\\&\leq|w|\cdot|D\Phi(x+cu)(u)-D\Phi(x)(u)|\\&\leq\max_{c\in[0,1]}|D\Phi(x+cu)(u)-D\Phi(x)(u)|\\&\leq|u|\cdot\max_{\substack{x'\in\overline{U'}\\|u'|\leq|u|}}\|D\Phi(x'+u')-D\Phi(x')\|.\end{aligned}$$

The right-hand factor is our function R in terms of $|u|$, which is finite because we are taking the maximum of a continuous function on a compact set. Technically, we have not shown that R is decreasing, but we can make it decreasing by replacing R with $t \mapsto \sup\{R(s) : s \geq t\}$. ■

1.10 February 7

I have moved a proof to today's notes for continuity reasons.

1.10.1 More on Change of Variables

We are now ready to show Lemma 1.70.

Lemma 1.70. Fix a C^1-diffeomorphism $\Phi\colon U \to V$ where $U, V \subseteq \mathbb{R}^d$ is open. Further, let U' be an open subset with compact closure $\overline{U'} \subseteq U$ such that there is $\delta_0 > 0$ such that any $x \in U'$ has $d(x, U^c) \geq 2\delta_0$. Then we have

$$\mu(\varphi(U')) \leq \int_{U'} |J_\Phi| \, d\mu.$$

Proof. The point is to make everything as linear as possible. Fix $\delta > 0$ less than δ_0, which we will eventually send to 0^+. Now, by Lemma 1.53, we may divide U into countably many dyadic cubes $\{Q_j\}_{j\in\mathbb{N}}$, where Q_j has side-length ℓ_j; by possibly decomposing cubes finitely, we may assume that $\ell_j \leq \delta$ for each $j \in \mathbb{N}$. For our combinatorics, we let $c_j \in Q_j$ be the center of the cube so that any point in Q_j can be written as $c_j + u$ where $u = (u_1, \ldots, u_d)$ has $|u_i| \leq \frac{1}{2}\ell_j$ for each i.

Now, by Lemma 1.72, we may upper-bound

$$\Phi(c_j + u) = \Phi(c_j) + D\Phi(c_j)(u) + |u| \, R(|u|)$$

where R is conjured from Lemma 1.72. For our bounding, it will help to recognize that

$$|u| \, R(|u|) \leq \ell_j\sqrt{d}R(\ell_j\sqrt{d}) \leq \sqrt{d}R(\delta\sqrt{d})\ell_j.$$

To chart our progress, we note

$$\mu(\Phi(U')) \leq \sum_{j\in\mathbb{N}} \mu(\Phi(Q_j)),$$

and now Φ on Q_j is basically linear.

Indeed, set $T_j := D\Phi(c_j)$, and for $x = c_j + u$ in Q_j, we see

$$\underbrace{|\Phi(x) - \Phi(c_j) - T_j(u)|}_{\mathcal{R}(c_j,u):=} \leq |u| \, R(|u|)$$

for some remainder function R provided by Lemma 1.72. Now, to bound $\mu(\Phi(Q_j))$, we see that $\Phi(Q_j) \approx \Phi(c_j) + T_j(Q_j(0, \ell_j))$, where Q denotes the cube. Notably, we have some identity like $\Phi(c_j + u) = \Phi(c_j) + T_j\left(u + T_j^{-1}\mathcal{R}(c_j, u)\right)$, so we would like to bound $T_j^{-1}\mathcal{R}(c_j, u)$, which we do as

$$\left|T_j^{-1}\mathcal{R}(c_j, u)\right| \leq \|T_j\|^{-1} \cdot |u| \cdot R(|u|) \leq \|D\Phi(c_j)\|^{-1} \cdot \ell_j\sqrt{d} \cdot R(\delta\sqrt{d}).$$

Quickly, note that the function $x \mapsto D\Phi(x)^{-1}$ is continuous (on U) because Φ is continuously differentiable, so with $\overline{U'} \subseteq U$ is compact, we see that the function $x \mapsto D\Phi(x)^{-1}$ is upper-bounded on U by some C (which does not depend on δ!). So our bound becomes

$$\left|T_j^{-1}\mathcal{R}(c_j, u)\right| \leq C\ell_j\sqrt{d}R(\delta\sqrt{d}).$$

For brevity, set $\varepsilon_\delta := C\sqrt{d}R(\delta\sqrt{d})$. The point is that

$$u + T_j^{-1}\mathcal{R}(c_j, u) \in Q\left(0, \ell_j(1 + \varepsilon_\delta)\right),$$

so

$$\Phi(Q_j) \subseteq \Phi(c_j) + D\Phi(c_j)\left(Q\left(0, \ell_j(1+\varepsilon_\delta)\right)\right),$$

so

$$\mu(\Phi(Q_j)) \le |J_\Phi(c_j)| \underbrace{\ell_j^d}_{\mu(Q_j)} (1+\varepsilon_\delta)^d$$

by taking the measure of a cube under a linear transformation, so by summing over j, we achieve

$$\mu(\Phi(U')) \le (1+\varepsilon_\delta)^d \sum_{j\in\mathbb{N}} |J_\Phi(c_j)|\,\mu(Q_j).$$

It remains to relate the summation to the integral. Well, write

$$|J_\Phi(c_j)|\,\mu(Q_j) \le \int_{Q_j} |J_\Phi(x)|\,d\mu(x) + \int_{Q_j} \left||J_\Phi(x)| - |J_\Phi(c_j)|\right| d\mu(x),$$

and we see that we would like for $|J_\Phi(x) - J_\Phi(c_j)|$ to be small. It will turn out to be small, so we can upper-bound it by the maximum over any pairs (x, y) with distance $\delta\sqrt{d}$ apart, so we sum over all cubes to achieve

$$\mu(\Phi(U')) \le (1+\varepsilon_\delta)^d \left(\int_{U'} |J_\Phi|\,d\mu + \sup_{|x-y|\le\delta\sqrt{d}} \left||J_\Phi(x)| - |J_\Phi(c_j)|\right| \sum_{j\in\mathbb{N}} \mu(Q_j)\right).$$

Sending $\delta \to 0^+$ will send $\varepsilon_\delta \to 0^+$ and the supremum to 1 because that continuous function is uniformly continuous on the compact set $\overline{U'} \subseteq U$. ■

We now upgrade the lemma in various ways.

Lemma 1.73. Fix everything as in Lemma 1.70. For any Lebesgue measurable subset $E \subseteq U'$, we have

$$\mu(\Phi(E)) \le \int_E |J_\Phi|\,d\mu.$$

Proof. Fix $\varepsilon > 0$, and regularity in Theorem 1.55 promises some open $U'' \supseteq E$ contained in U' with $\mu(U'' \setminus E) < \varepsilon$. Then we rudely replace E with U'' and apply Lemma 1.70: note

$$\mu(\Phi(E)) \le \mu(\Phi(U'')) \le \int_{U''} |J_\Phi|\,d\mu = \int_E |J_\Phi|\,d\mu + \int_{U''\setminus E} |J_\Phi|\,d\mu.$$

So it remains to bound

$$\int_{U''\setminus E} |J_\Phi|\,d\mu \le \sup_{x\in\overline{U'}} |J_\Phi| \underbrace{\mu(U''\setminus E)}_{<\varepsilon},$$

which we see goes to 0 as $\varepsilon \to 0^+$. Note that the supremum exists because $\overline{U'}$ is compact. ■

We now upgrade away the U' entirely, which is mostly point-set topology.

Lemma 1.74. Fix a C^1-diffeomorphism $\Phi\colon U \to V$ where $U, V \subseteq \mathbb{R}^d$ is open. For any Lebesgue measurable $E \subseteq U$, we have

$$\mu(\Phi(E)) \le \int_E |J_\Phi|\,d\mu.$$

Proof. Fix a nonempty open subset $U \subseteq \mathbb{R}^d$. For each $n \in \mathbb{N}$, define $U_n \subseteq U$ to consist of the $x \in U$ such that $|x| < n$ and $\operatorname{dist}(x, U^c) < 1/n$. Then $\overline{U_n} \subseteq U$ and has a distance at most $1/n$ living inside U. So we may apply Lemma 1.73: set $E_n := E \cap U_n$, and we see

$$\mu(\Phi(E_n)) \le \int_{E_n} |J_\Phi| \, d\mu \le \int_E |J_\Phi| \, d\mu$$

for each n, and then sending $n \to \infty$ has $E = \bigcup_{n \in \mathbb{N}} E_n$ and so $\mu(\Phi(E_n)) \to \mu(\Phi(E))$, so we achieve the result. ■

We now upgrade from sets to functions.

Lemma 1.75. Fix a C^1-diffeomorphism $\Phi \colon U \to V$ where $U, V \subseteq \mathbb{R}^d$ is open. Given a measurable function $f \colon V \to [0, \infty]$, we have

$$\int_V f \, d\mu \le \int_U (f \circ \Phi) \, |J_\Phi| \, d\mu.$$

Proof. In the case where $f = 1_E$ for a Lebesgue measurable E, this result is just Lemma 1.74. Taking linear combinations and approximating below achieves the result for general measurable functions $f \colon V \to [0, \infty]$ via the Monotone convergence theorem [Elb22, Theorem 9.18]. ■

At long last, we produce equality.

Lemma 1.76. Fix a C^1-diffeomorphism $\Phi \colon U \to V$ where $U, V \subseteq \mathbb{R}^d$ is open. Given a measurable function $f \colon V \to [0, \infty]$, we have

$$\int_V f \, d\mu = \int_U (f \circ \Phi) \, |J_\Phi| \, d\mu.$$

Proof. Note $\le$ follows immediately from Lemma 1.75. For the other inequality, we set $\Psi := \Phi^{-1}$ and $g := (f \circ \Phi) \, |J_\Phi|$ so that $f := \left| J_\Phi \circ \Phi^{-1} \right| \left(g \circ \Phi^{-1} \right)$ and reapply Lemma 1.75 to g. ■

Theme 2
Banach Spaces

2.1 February 9

Today we move on to talk about Banach spaces. Throughout, $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ will be a field.

2.1.1 Banach Spaces

We begin with the definition of a normed vector space.

Definition 2.1 (norm)**.** Fix an $\mathbb{F}$-vector space V. Then a *norm* on V is a function $\|\cdot\| : V \to \mathbb{R}_{\geq 0}$ satisfying the following conditions.

- Positive-definite: $\|x\| = 0$ if and only if $x = 0$.
- Homogeneous: for $\lambda \in \mathbb{F}$, one had $\|\lambda x\| = |\lambda| \cdot \|x\|$.
- Triangle inequality: $\|x + y\| \leq \|x\| + \|y\|$.

We might say that the pair $(V, \|\cdot\|)$ is a *normed vector space*.

Notation 2.2. In the sequel, we may just say that V is a normed vector space, where we implicitly equip V with the norm $\|\cdot\|_V$.

Remark 2.3. Given a normed vector space $(V, \|\cdot\|)$. Then

$$d(x, y) := \|x - y\|$$

makes V into a metric space. Namely, d is a metric.

Definition 2.4 (Banach space)**.** A normed vector space $(V, \|\cdot\|)$ is a *Banach space* if and only if V is complete as a metric space.

Remark 2.5. Fix a normed vector space $(V, \|\cdot\|)$. Then we claim $|\|x\| - \|y\|| \leq \|x-y\|$. Indeed, the triangle inequality implies $\|x\| \leq \|x-y\| + \|y\|$, so

$$\|x\| - \|y\| \leq \|x-y\|.$$

Reversing x and y shows that $\|y\| - \|x\| \leq \|x-y\|$ as well, so the claim follows.

We will also want a notion of convergence.

Definition 2.6 (absolute convergence)**.** Fix a normed vector space $(V, \|\cdot\|)$. Then a sum $\sum_{i=1}^\infty v_i$ for vectors $\{v_i\}_{i=1}^\infty \subseteq V$ is absolutely convergent if and only if

$$\sum_{i=1}^\infty \|x_i\| < \infty.$$

This produces the following test of convergence.

Lemma 2.7. Fix a normed vector space $(V, \|\cdot\|)$. Then X is complete if and only if any absolutely convergent series converges (in V).

Proof. In one direction, any absolutely convergent series has partial sums which are Cauchy (by the absolute convergence), so it will converge in V.

In the other direction, suppose absolutely convergent series converge. Then suppose $\{v_n\}_{n\in\mathbb{N}}$ is a Cauchy sequence in V, which we would like to converge.

Remark 2.8. Here is an attempt which doesn't work: we might want to, define $w_n := w_n - w_{n-1}$ (take $w_0 = 0$), and we see that $\sum_{k=1}^n w_k = v_n$. So by hypothesis, it is enough to check that this series is absolutely convergent

$$\sum_{k=1}^\infty \|w_k\| < \infty,$$

but there is no reason to be true.

So we need to control how fast our series converges. Construct a strictly increasing sequence $\{N_k\}_{k=1}^\infty$ such that $m, n \geq N_k$ implies that $\|v_m - v_n\| < 2^{-k}$; for this sequence to be strictly increasing, it must be defined recursively. Thus, we define $w_k := y_{N_k}$ to be a subsequence of the $y_\bullet$, and it continues to be Cauchy. In fact, $\|w_k - w_{k-1}\| < 2^{-k-1}$ for each k, so the series

$$v_{N_n} := \sum_{k=1}^n w_k - w_{k-1}$$

is absolutely convergent. So a subsequence of $\{v_n\}_{n\in\mathbb{N}}$ converges, so our actual sequence converges by using the Cauchy condition. ■

Remark 2.9. Of course, there are convergent series which are not absolutely convergent: work in $\mathbb{R}$ with the usual norm, and take $v_n := (-1)^n/n$ for each $n \geq 1$.

2.1.2 A Little Linear Algebra

We will want the notion of a basis, which we now build.

Definition 2.10 (linearly independent)**.** Fix a subset S of a vector space V if and only if, for any $n \geq 1$ and distinct elements $\{v_1, \ldots, v_n\} \subseteq S$ has

$$a_1v_1 + \cdots + a_nv_n = 0$$

implies $a_1 = \cdots = a_n = 0$.

Definition 2.11 (finite dimensional)**.** A vector space V is *finite-dimensional* if and only if there is a finite subset $S \subseteq V$ such that any element of V can be written as a linear combination of elements of S.

Example 2.12. Let ℓ^∞ denote the set of bounded infinite sequences in $\mathbb{F}$, and we give it the norm

$$\|x\|_\infty := \sup_{n\in\mathbb{N}} |x_n| .$$

Then indicators $e_n \in \ell^\infty$ given by $(e_n)_i := 1_{n=i}$ are linearly independent.

It will turn out that subspaces of finite-dimensional normed vector spaces are closed, but this is not the case in general.

Example 2.13. Continue with ℓ^∞ with the norm $\|\cdot\|_\infty$. Consider V to be the set of finitely supported sequences, which we can see is a subspace. However, $\overline{V}$ contains the sequence v defined by $v_n := 1/n$. Indeed, for any $\varepsilon > 0$, we can find $N > 1/\varepsilon$ and then note that we can define $v' \in V$ by

$$v'_n := \begin{cases} 1/n & \text{if } n \leq N, \\ 0 & \text{if } n > N. \end{cases}$$

Then $\|v_n - v'_n\| < \varepsilon$. So we see $v \in \overline{V} \setminus V$.

But some parts of topology still make sense.

Remark 2.14. Fix a normed vector space $(V, \|\cdot\|)$. Then the open ball $B(0,1)$ is still open in V because it is just an open ball in the usual metric topology.

2.2 February 12

Let's talk about Banach spaces a little more.

2.2.1 Linear Maps

With our newfound topology, we want to control our linear maps.

Definition 2.15 (bounded)**.** Fix normed vector spaces $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$. Then a linear map $T\colon X \to Y$ is *bounded* if and only if there is a finite constant C such that

$$\|Tx\|_Y \leq C\|x\|_X .$$

We let $\mathcal{L}(X,Y)$ denote the space of bounded linear maps, and we define

$$\|T\|_{\mathcal{L}(X,Y)} := \sup_{x\in X\setminus\{0\}} \frac{\|Tx\|_Y}{\|x\|_X}.$$

Remark 2.16. We note $\|\cdot\|_{\mathcal{L}(X,Y)}$ is in fact a norm. Homogeneity has little content, and positive-definiteness holds because we are looking at a sup: if $\|T\|_{\mathcal{L}(X,Y)} = 0$, then we must have $\|Tx\|_Y = 0$ always, so $Tx = 0$ always. Lastly, for the triangle inequality, we note

$$\|T+S\|_{\mathcal{L}(X,Y)} = \sup_{x\in X\setminus\{0\}} \frac{\|Tx+Sx\|_Y}{\|x\|_X} \le \sup_{x\in X\setminus\{0\}} \frac{\|Tx\|_Y}{\|x\|_X} + \sup_{x\in X\setminus\{0\}} \frac{\|Sx\|_Y}{\|x\|_X} = \|T\|_{\mathcal{L}(X,Y)} + \|S\|_{\mathcal{L}(X,Y)}.$$

Here is the topological reason that we care.

Lemma 2.17. Fix normed vector spaces $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$. Then $T\colon X \to Y$ is bounded if and only if continuous.

Proof. For the forward direction, suppose T is bounded with constant $C > 0$. Then for any $\varepsilon > 0$, choose $\delta := \varepsilon/C$: for any $x, y \in X$, we find that $\|x-y\|_X < \delta$ implies that

$$\|Tx - Ty\|_Y \le C\,\|x-y\|_X < C\delta = \varepsilon,$$

so in fact T is uniformly continuous.

For the reverse direction, we will only use continuity of T at 0. Then there exists $\delta > 0$ such that $\|x\|_X < \delta$ implies that $\|Tx\|_Y < 1$, so any nonzero vector $x' \in X$ has

$$\frac{\|Tx'\|_Y}{\|x'\|_X} = \frac{\left\|T\left(\frac{\delta}{2\|x'\|_X}x'\right)\right\|_Y}{\left\|\frac{\delta}{2\|x'\|_X}x'\right\|_X} < \frac{1}{\delta/2} = \frac{2}{\delta}.$$

So we have a bound $\|T\|_{\mathcal{L}(X,Y)} < 2/\delta$, so T is bounded. ■

2.2.2 Equivalence of Norms

Let's try to classify our norms. Multiplying a norm by a scalar ought to be considered the same norm; generalizing this slightly produces the following definition.

Definition 2.18 (equivalent)**.** Fix a vector space X. Two norms $\|\cdot\|_1$ and $\|\cdot\|_2$ are *equivalent* if and only if there is a constant $C > 1$ such that

$$C^{-1} \le \frac{\|x\|_1}{\|x\|_2} \le C$$

for any nonzero $x \in X$.

Remark 2.19. Equivalence as above is in fact an equivalence relation. Reflexivity has no content (take $C = 1$), symmetry follows by taking the reciprocal of the given equation, and transitivity follows by multiplying the two constants together.

Example 2.20. The two norms $\|\cdot\|_2$ and $\|\cdot\|_\infty$ are equivalent on any finite-dimensional space $\mathbb{R}^d$. Namely, we see that

$$\|x\|_\infty \le \|x\|_2 = \sqrt{\sum_{k=1}^d x_k^2} \le \sqrt{d}\,\|x\|_\infty,$$

so equivalence follows.

Here is our main result on equivalence.

Proposition 2.21. Fix a finite-dimensional vector space X. Then any two norms $\|\cdot\|_1$ and $\|\cdot\|_2$ on X are equivalent.

Proof. We will show that any norm $\|\cdot\|_X$ on X is equivalent to a given one. Fix a basis $\{v_1, \ldots, v_n\}$ for our finite-dimensional vector space X. This basis defines an isomorphism $\varphi\colon \mathbb{F}^n \to X$ given by $(a_1, \ldots, a_n) \mapsto (a_1v_1 + \cdots + a_nv_n)$. We now define $\|\cdot\|_\varphi$ to be a norm on $\mathbb{F}^n$ defined by

$$\|x\|_\varphi := \left\|\sum_{i=1}^n x_iv_i\right\|.$$

We won't bother to check that $\|x\|_\varphi$ is a norm, which hold by passing the norm checks through the isomorphism φ of vector spaces.

We claim that $\|\cdot\|_\infty$ is continuous. Well, we compute

$$\|x\|_\varphi = \left\|\sum_{i=1}^n x_iv_i\right\| \le \sum_{i=1}^n |x_i| \cdot \|v_i\| \le \|x\|_\infty \underbrace{\sum_{i=1}^n \|v_i\|}_{A:=}.$$

Thus, we see that $\|\cdot\|_\varphi : \mathbb{F}^n \to \mathbb{R}_{\ge 0}$ is continuous: in fact, $\|\cdot\|_\varphi$ is Lipschitz continuous, where we compute

$$\left|\|x\|_\varphi - \|x'\|_\infty\right| \le \|\varphi(x) - \varphi(x')\| = F(x - x') \le A\|x - x'\|_\infty,$$

so $\|\cdot\|_\varphi$ is in fact continuous.

Thus, by continuity, upon restricting $\|\cdot\|_\varphi$ to the unit sphere S^{n-1}, we see that $\|\cdot\|_\varphi$ will achieve its minimum. But 0 is never achieved on S^{n-1}, so let δ be the minimum and find that

$$\delta \le \frac{\|x\|_\varphi}{\|x\|_\infty} \le A$$

for any $x \in S^{n-1}$, an inequality which extends to any nonzero $x \in X$ by scaling. Setting $C := \max\left\{\delta^{-1}, A\right\}$ to deduce that $\|\cdot\|$ is equivalent to $\|\cdot\|_\varphi$. Pushing through the isomorphism φ, we see that

$$C^{-1} \le \frac{\|x\|}{\|\varphi^{-1}(x)\|_\infty} \le C$$

for any nonzero $x \in X$. Thus, $\|\cdot\|$ is equivalent to the norm $x \mapsto \left\|\varphi^{-1}(x)\right\|_\infty$, but this latter norm is entirely independent of $\|\cdot\|$, so we have indeed shown that any norm on X is equivalent to a fixed one. ■

2.2.3 Topology on Normed Vector Spaces

Let's see some corollaries of Proposition 2.21.

Corollary 2.22. Fix a finite-dimensional normed vector space $(X, \|\cdot\|)$. Then X is complete.

Proof. Choose a basis, so we may let $\varphi\colon \mathbb{F}^n \to X$ be some isomorphism of vector spaces. Then $\|\cdot\|$ is equivalent to the norm $x \mapsto \left\|\varphi^{-1}(x)\right\|_\infty$ by Proposition 2.21, so φ upgrades to a homeomorphism. But then completeness of X follows from completeness of $\mathbb{F}^n$. ■

Corollary 2.23. Fix a normed vector space $(X, \|\cdot\|)$. Then any finite-dimensional subspace $V \subseteq X$ is closed.

Proof. We know that V is complete by Corollary 2.22. This is enough to see that V is closed: any convergent sequence $x_n \to x$ where $\{x_n\}_{n\in\mathbb{N}} \subseteq V$ will be Cauchy, but completeness of V implies that this Cauchy sequence has a limit $x' \in V$, but we must have $x = x'$ by the uniqueness of limits. ■

2.3 February 14

Let's talk about Banach spaces a little more.

2.3.1 Closed Unit Balls

We may want our closed balls to be compact, but we are out of luck.

Proposition 2.24. Fix a normed vector space $(X, \|\cdot\|)$. Then the closed unit ball

$$B \coloneqq \{x \in X : \|x\| \le 1\}$$

is compact if and only if X is finite-dimensional.

To prove this, we will want a lemma.

Definition 2.25. Fix a normed vector space $(X, \|\cdot\|)$. For $x \in X$ and subspace $V \subseteq X$, we define

$$d(x, V) \coloneqq \inf_{x \in V} \|x - v\| .$$

Note that this exists and is finite because V is nonempty (for example, $0 \in V$).

One of course has the bound $d(x, V) \le \|x\|$ because $0 \in V$. One cannot hope to do much better than this without explosion.

Lemma 2.26. Fix a normed vector space $(X, \|\cdot\|)$. Suppose that each $x \in X$ has $d(x, V) \le \frac{1}{2}\|x\|$. Then $\overline{V} = X$.

Let's see how this lemma implies Proposition 2.24.

Proof of Proposition 2.24 from Lemma 2.26. We will show that either B fails to be compact or X is not finite-dimensional. If $X = 0$, there is nothing to do; otherwise, for example, there is a nonzero vector, so scaling provides us with a unit vector.

We proceed with the following inductive process. Begin with some unit vector $x_1 \in X$. We then look for unit vectors x_2 such that $d(x_2, \operatorname{span}\{x_1\}) \ge \frac{1}{2}$; if no such vector exists, we terminate our process. We then continue this process inductively to produce a set B: given the finite set $\{x_1, \ldots, x_n\}$, we look for a unit vector x_{n+1} such that

$$d(x_{n+1}, \operatorname{span}\{x_1, \ldots, x_n\}) \ge \frac{1}{2}.$$

If no such vector exists, then the process terminates; otherwise, we add x_{n+1} to our finite set and continue.

We have two cases.

- If we ever terminate, then we trigger Lemma 2.26: we know there is a finite set $\{x_1, \ldots, x_n\}$ of unit vectors such that every unit vector x has $d(x, V) < 1/2$ where $V \coloneqq \operatorname{span}\{x_1, \ldots, x_n\}$. Scaling, we actually know that

  $$d(x, V) < \frac{1}{2}\|x\|$$

  for any $x \in X$: for nonzero x, we simply divide both sides of the above inequality by $\|x\|$. So we trigger Lemma 2.26, so $\overline{V} = X$. But V is finite-dimensional by construction, so $V \subseteq X$ is closed by Corollary 2.23, so $X = \overline{V} = V$ is finite-dimensional.

- Otherwise, we have an infinite set $\{x_1, x_2, \ldots\}$ of unit vectors such that $\|x_i - x_j\| \ge \frac{1}{2}$ for any indices $i < j$. In particular, B has a sequence with no convergent subsequence, so B fails to be compact. ■

We now prove Lemma 2.26.

Proof of Lemma 2.26. Fix $x \in X$. We want to show that $d(x, V) = 0$. In other words, for any $\varepsilon > 0$, we want $y \in V$ such that $\|x - y\| < \varepsilon$.

Well, suppose for the sake of contradiction that we have $\varepsilon > 0$ such that no $y \in V$ achieves $\|x - y\| < \varepsilon$, and we may assume that $\varepsilon := \inf\{\|x - y\| : y \in V\}$ is as small as possible. In particular, surely there will be some $y_1 \in V$ such that $\|x - y_1\| < \frac{3}{2}\varepsilon$, but then we can find $y_2 \in V$ such that

$$\|x - y_1 - y_2\| \leq \frac{1}{2} \|x - y_1\|,$$

so

$$\|x - (y_1 + y_2)\| \leq \frac{1}{2} \|x - y_1\| \leq \frac{3}{4}\varepsilon,$$

which is a contradiction to the construction of ε. ■

2.3.2 Functionals

It will be helpful to study linear maps to the ground field.

Definition 2.27 (functional)**.** Fix a normed vector space $(X, \|\cdot\|)$. Then a *bounded linear functional* is an element of $X^* := \mathcal{L}(X, \mathbb{F})$.

Example 2.28. Certainly X^* is nonempty because it has the zero element.

However, it is rather hard to get anything else in X^*.

Example 2.29. If X is finite-dimensional, then choose a basis $\{v_1, \ldots, v_n\}$. Then there are functionals

$$\sum_{i=1}^n a_i v_i \mapsto \sum_{i=1}^n c_i a_i$$

for any $(c_1, \ldots, c_n) \in \mathbb{F}^n$.

It will turn out that we can build lots of functionals, but it is not so obvious how to do this. We will have the following results.

Corollary 2.30. Fix a normed vector space $(X, \|\cdot\|)$. For any $x \in X$, there is a bounded linear functional $\ell \colon X \to \mathbb{F}$ such that $\ell(x) = \|x\|$ and $\|\ell\|_{X^*} = 1$.

Remark 2.31. Certainly $\ell(x) = \|x\|$ forces $\|\ell\|_{X^*} \geq 1$, so we are saying that ℓ is small away from x.

Corollary 2.32. Fix a normed vector space $(X, \|\cdot\|)$. Given distinct $x, y \in X$, there is a bounded linear functional $\ell \colon X \to \mathbb{F}$ such that $\ell(x) \neq \ell(y)$.

These results will arise as corollaries of the following result.

Theorem 2.33 (Hahn–Banach)**.** Fix a normed vector space $(X, \|\cdot\|)$, and let $V \subseteq X$ be a subspace. Given a bounded linear functional $\ell \colon V \to \mathbb{F}$, there is a bounded linear functional $L \colon X \to \mathbb{F}$ such that $L|_V = \ell$ and $\|L\| = \|\ell\|_{V^*}$

One even has the following extension for $\mathbb{R}$.

Definition 2.34. Fix an $\mathbb{R}$-vector space X. A function $p\colon X \to \mathbb{R}$ is a *sublinear functional* if and only if $p(x+y) \le p(x) + p(y)$ and $p(tx) = tp(x)$ for $t > 0$.

Example 2.35. Fix a normed vector space $(X, \|\cdot\|)$, and let $K \subseteq X$ be a closed convex subset, and suppose that K contains $B(0,\varepsilon)$ for some $\varepsilon > 0$. Then we define

$$p(x) \coloneqq \inf\left\{t > 0 : \frac{1}{t}x \in K\right\}.$$

It turns out that p is a sublinear functional. Notably, if K is not symmetric, then $p(x)$ need not equal $p(-x)$, so p need not be a norm.

Theorem 2.36 (Hahn–Banach)**.** Fix an $\mathbb{R}$-vector space X equipped with sublinear functional p, and let $V \subseteq X$ be a subspace. Given a linear functional $\ell\colon V \to \mathbb{F}$ such that $\ell \le p$ pointwise, there is a bounded linear functional $L\colon X \to \mathbb{F}$ such that $L|_V = \ell$ and $L \le p$ pointwise.

2.4 February 16

Here we go.

2.4.1 The Hahn–Banach Theorem

Our proofs of Theorems 2.33 and 2.36 will be by transfinite induction on V. Let's state the successor step.

Lemma 2.37. Fix an $\mathbb{R}$-vector space X equipped with sublinear functional p, and let $V \subseteq X$ be a subspace. Given a linear functional $\ell\colon V \to \mathbb{F}$ such that $\ell \le p$ pointwise and some $x \notin V$, there exists some linear $\widetilde{\ell}\colon (V + \mathbb{R}x) \to \mathbb{R}$ such that $\widetilde{\ell}|_V = \ell$ and $\widetilde{\ell} \le p$ pointwise.

Proof. Set $\widetilde{V} \coloneqq V + \mathbb{R}x$. Note that decomposition of an element of $\widetilde{V}$ into $v + rx$ where $v \in V$ and $r \in \mathbb{R}$ is unique: if $v + rx = v' + r'x$, then $(r - r')x = v' - v \in V$, so $r - r' = 0$, so $v' - v = 0$. So we simply define

$$\widetilde{\ell}(v + rx) \coloneqq \ell(v) + r\alpha$$

for some α to be determined later. For example, $\widetilde{\ell}$ is certainly linear by construction, and it restricts down to ℓ on V. So we are looking for α such that

$$t\alpha + \ell(v) \le p(tx + v) \qquad \text{and} \qquad -t\alpha + \ell(v) \le p(-tx + v)$$

for any $t \ge 0$ and $v \in V$. Certainly we may assume that $t > 0$ because $t = 0$ reduces down to ℓ, but then we can divide everything in sight by t so that it suffices for

$$\alpha + \ell(v) \le p(x + v) \qquad \text{and} \qquad -\alpha + \ell(v) \le p(-x + v)$$

for any $v \in V$. This rearranges to

$$\sup_{v \in V}(\ell(v) - p(-x + v)) \le \alpha \le \inf_{v \in V}(p(x + v) - \ell(v)).$$

So we can find the needed α if and only if

$$\sup_{v \in V}(\ell(v) - p(-x + v)) \le \inf_{v \in V}(p(x + v) - \ell(v)).$$

Fixing some vectors explicitly, it is enough to check that

$$\ell(u+v) = \ell(v) + \ell(u) \leq p(x+v) + p(-x+u)$$

for any $v, u \in V$. In particular, by hypothesis of ℓ, it is enough for $p(u+v) \leq p(x+v) + p(-x+u)$, but this follows from sublinearity of p by writing $u+v = (x+v) + (-x+u)$. ■

Remark 2.38. Replacing p with an actual norm $\|\cdot\|$ on X allows us to allow X to even by a $\mathbb{C}$-vector space, simply by repeating the proof verbatim.

We would now like to use Zorn's lemma to upgrade Lemma 2.37 to Theorem 2.36. We begin by stating Zorn's lemma in the form that we will use. This requires the notion of a partial order.

Definition 2.39 (partial order). A *partial order* on a set S is a binary relation $\leq$ satisfying the following.

- Reflexive: $x \leq x$ for all $x \in S$.
- Transitive: if $x \leq y$ and $y \leq z$, then $z \leq z$.
- Antisymmetric: if $x \leq y$ and $y \leq x$, then $x = y$.

We say that $\leq$ is *linearly ordered* (or is a *chain*) if and only if we satisfy the additional axiom of totality: $x \leq y$ or $y \leq x$ for all $x, y \in S$.

Example 2.40. Consider a set X. Then $\mathcal{P}(X)$ is partially ordered by $\subseteq$. It is not linear ordered by $\subseteq$.

And now here is Zorn's lemma.

Theorem 2.41. Suppose that $(S, \leq)$ is a nonempty partially ordered set such that any subchain $C \subseteq S$ is "bounded above" in the sense that there is some $y \in S$ such that $x \leq y$ for any $x \in C$. Then S has a maximal element; i.e., there is some $y \in S$ such that $x \leq y$ for any $x \in S$.

Proof. This is equivalent to the Axiom of choice. ■

To apply Zorn's lemma, here is the proof of Lemma 2.37.

Theorem 2.36 (Hahn–Banach)**.** Fix an $\mathbb{R}$-vector space X equipped with sublinear functional p, and let $V \subseteq X$ be a subspace. Given a linear functional $\ell \colon V \to \mathbb{F}$ such that $\ell \leq p$ pointwise, there is a bounded linear functional $L \colon X \to \mathbb{F}$ such that $L|_V = \ell$ and $L \leq p$ pointwise.

Proof. We will use Zorn's lemma. For this, we let our partially ordered set S consisting of all pairs (W, ℓ_W) where $W \subseteq X$ is a subspace and ℓ_W is a functional on W restricting to ℓ and has $\ell_W \leq p$ pointwise on W. Our partial order is given by $(W_1, \ell_1) \leq (W_2, \ell_2)$ if and only if (W_2, ℓ_2) "extends" (W_1, ℓ_1) in the sense that $W_1 \subseteq W_2$ and $\ell_1 = \ell_2|_{W_1}$.

Note S is nonempty because it has (V, ℓ) by construction. Now, to apply Theorem 2.41, we need to check that any chain $C \subseteq S$ has an upper bound. Well, define

$$W' := \bigcup_{(W, \ell_W) \in C} W,$$

and then we define $\ell' \colon W' \to \mathbb{R}$ by defining $\ell'(v)$ to be $\ell_W(v)$ where $v \in W$ for some $(W, \ell_W) \in C$. Notably, the choice of ℓ_W does not matter: if $v \in W_1 \cap W_2$ where $(W_1, \ell_1), (W_2, \ell_2) \in C$, then without loss of generality we can take $(W_1, \ell_1) \leq (W_2, \ell_2)$, implying $\ell_2(v) = \ell_1(v)$. Thus, ℓ' is well-defined, and one can check directly that it is linear on W'; similarly one can check that ℓ' extends ℓ and that $\ell' \le p$ pointwise. In conclusion, $(W', \ell') \in S$ and bounds C above.

Thus, Theorem 2.41 provides a maximal element (W, ℓ) of S. We claim that $W = X$, which will complete the proof. Well, if $W \subsetneq X$, then we can find $x \in X \setminus W$. But then Lemma 2.37 tells us that we can extend (W, ℓ) to some $(W + \mathbb{R}x, \widetilde{\ell})$ in S, and $W \subsetneq W + \mathbb{R}x$ contradicts the fact that (W, ℓ) is maximal. This completes the proof. ■

Remark 2.42. It turns out that Theorem 2.33 is not actually equivalent to the Axiom of choice.

Remark 2.43. The proof of Theorem 2.33 is similar, so we will omit it. Alternatively, one can apply Theorem 2.36 to $\operatorname{Re} \ell$, viewing X as an $\mathbb{R}$-vector space. Then one can recover ℓ from $\operatorname{Re} \ell$ via the identity

$$\ell(x) = \operatorname{Re} \ell(x) - i \operatorname{Re} \ell(ix),$$

so similarly one can upgrade the lifted functional $\operatorname{Re} \widetilde{\ell}$ provided by Theorem 2.36 to $\widetilde{\ell}$ via the above formula.

2.5 February 21

Here we go.

2.5.1 Baire's Theorem

Compactness is not going to be so helpful for Banach spaces, so we will need a different notion, which we will slowly build up to. To start off our story, we consider the following example.

Example 2.44. One can build an open dense subset covering $\mathbb{Q}$ but with arbitrarily small measure. Indeed, enumerate $\mathbb{Q}$ as $\{q_n\}_{n \in \mathbb{N}}$ and then set

$$U_\varepsilon := \left(q_n - \varepsilon 2^{-n}, q_n + \varepsilon 2^{-n}\right),$$

which we can compute has measure upper-bounded by $\sum_{n \in \mathbb{N}} 2\varepsilon 2^{-n} = 4\varepsilon$. Thus, for example,

$$A := \bigcap_{k \in \mathbb{N}} U_{1/k}$$

has measure zero but is dense and is the countable intersection of open sets.

One might be interested in what A looks like. It will turn out that A is uncountable, but of course it is rather hard to show there is a single element in $A \setminus \mathbb{Q}$.

Here is our theorem.

Theorem 2.45 (Baire). Fix a nonempty complete metric space X, and let $\{U_n\}_{n \in \mathbb{N}}$ be a collection of open dense subsets of X. Then

$$\bigcap_{n \in \mathbb{N}} U_n$$

is nonempty and in fact dense.

Proof. We will show density directly, so let V be some open subset of X which we would like $\bigcap_{n \in \mathbb{N}} U_n$ to intersect V nontrivially. Well, we proceed inductively. To begin, choose $x_0 \in U_0 \cap V$, and select $r_0 < 1$ such that $B(x_0, 3r_0) \subseteq U_1 \cap V$. Then given x_n and positive $r_n < r_{n-1}/2$ (where $r_{-1} = 2$), we select

$$x_{n+1} \in U_n \cap B(x_n, r_n)$$

and some positive $r_{n+1} < r_n/2$ such that $B(x_{n+1}, 3r_{n+1}) \subseteq U_n \cap B(x_n, r_n)$.

We now claim that $\{x_n\}_{n\in\mathbb{N}}$ is a Cauchy sequence, and its limit will be the desired point. Well, by an induction, we see $r_n < 2^{-n}$, so

$$d_X(x_n, x_{n+1}) < r_n < 2^{-n},$$

so for $m < n$, we see

$$d_X(x_m, x_n) \le \sum_{k=m}^{n-1} d_X(x_k, x_{k+1}) < \sum_{k=m}^{n-1} 2^{-k} < 2^{-m+1},$$

which does indeed vanish as $m \to \infty$.

Thus, by completeness, we get some $x \in X$ with $x_n \to x$ as $n \to \infty$. We claim that $x \in V \cap \bigcap_{n\in\mathbb{N}} U_n$, which will complete the proof. This will require some understanding of $d(x, x_m)$. Well, for each $m < n$, we know

$$d_X(x_m, x_n) < \sum_{k=m}^{n-1} d_X(x_k, x_{k+1}) < \sum_{k=m}^{\infty} r_k < 2r_m$$

by an argument similar to the above one, so we must have $d_X(x_m, x) \le 2r_m < 3r_m$, so $x \in B(x_m, 3r_m) \subseteq U_{m-1} \cap B(x_{m-1}, r_{m-1}$, where $B(x_{-1}, r_{-1})$ means V. So indeed x lives in every U_m and also in V, so we are done. ■

Remark 2.46. Completeness of X is necessary. For example, take $X = \mathbb{Q}$ with the standard metric and topology. But then we can enumerate $\mathbb{Q}$ as $\{q_n\}_{n\in\mathbb{N}}$ and define the open dense subset $U_n \coloneqq X \setminus \{q_n\}$, but

$$\bigcap_{n\in\mathbb{N}} U_n = \varnothing.$$

At this point, we are able to say that A of Example 2.44 is at least dense.

2.5.2 Nowhere Dense Sets

To see that A of Example 2.44 is fairly large, we want the following definitions.

Definition 2.47 (nowhere dense)**.** Fix a topological space X. Then a subset $E \subseteq X$ is *nowhere dense* if and only if $\overline{E}$ contains no nonempty subset.

Example 2.48. The Cantor set in $\mathbb{R}$ is nowhere dense.

Definition 2.49 (category)**.** Fix a topological space X. Then a subset $E \subseteq X$ is *of the first category* if and only if E is a countable union of nowhere dense sets.

Example 2.50. If X is countable, then singletons are certainly nowhere dense, so X is of the first category.

We are now allowed to extend Theorem 2.45 as follows.

Theorem 2.51 (Baire)**.** Fix a nonempty complete metric space X, and let $\{U_n\}_{n\in\mathbb{N}}$ be a collection of open dense subsets of X. Then

$$\bigcap_{n\in\mathbb{N}} U_n$$

is not of the first category.

Proof. Suppose for contradiction that we can find countably many nowhere dense subsets $\{E_k\}_{k\in\mathbb{N}}$ such that

$$\bigcap_{n\in\mathbb{N}} U_n \subseteq \bigcup_{k\in\mathbb{N}} E_k.$$

Rearranging, we see that

$$\left(\bigcap_{n\in\mathbb{N}} U_n\right) \cap \left(\bigcap_{k\in\mathbb{N}} X \setminus \overline{E_k}\right)$$

being empty. However, E_k being nowhere dense means that $X \setminus X \setminus \overline{E_k}$ is open and dense, so we have exhibited a countable union of open dense subsets of X with empty intersection. So we achieve contradiction from Theorem 2.45. ■

Example 2.52. If X is a complete metric space, one can take $U_n := X$ for all $n \in \mathbb{N}$, so X itself is not of the first category. Thus, for example, a complete metric space cannot be countable by comparing with Example 2.50.

The point of Theorem 2.51 is that A of Example 2.44 cannot be of the first category, so for example it cannot be countable.

2.5.3 The Open Mapping Theorem

Let's now apply our results for fun and profit. Here is our next statement.

Theorem 2.53 (Open mapping)**.** Fix a surjective bounded linear map $T\colon X \to Y$ of Banach spaces.

(a) Then T is open.

(b) There is $A > 0$ such that any $y \in Y$ has some $x \in X$ such that $T(x) = y$ and $\|x\|_X \le A\|y\|_Y$.

Remark 2.54. Theorem 2.53 is intended to recover some aspects of compactness. To see this, suppose $f\colon X \to Y$ is a continuous bijection of compact Hausdorff topological spaces. Then it turns out that f is open: it's enough to see that f is closed by taking complements, but for closed $K \subseteq X$, we see K is compact, so $f(K)$ is compact, so $f(K) \subseteq Y$ is closed.

Here's an application, explaining how to recover Remark 2.54.

Corollary 2.55. Fix a bijective bounded linear map $T\colon X \to Y$ of Banach spaces. Then T^{-1} is bounded.

Proof. Part (b) of Theorem 2.53 tells us that $\|T^{-1}(y)\|_X \le A\|\|$ (because $x = T^{-1}(y)$ is unique!), which is what we wanted. ■

2.6 February 23

Today we prove the Open mapping theorem.

2.6.1 Proof of the Open Mapping Theorem

Recall the statement.

Theorem 2.53 (Open mapping)**.** Fix a surjective bounded linear map $T\colon X \to Y$ of Banach spaces.

(a) Then T is open.

(b) There is $A > 0$ such that any $y \in Y$ has some $x \in X$ such that $T(x) = y$ and $\|x\|_X \le A\|y\|_Y$.

Proof. Let's focus on (b) for now. We proceed in steps; note that we may assume $y \neq 0$.

1. Consider $Y_n := \overline{T(B_X(0,n))}$ to be a subset of Y, and we note that

$$\bigcup_{n\in\mathbb{N}} Y_n = Y.$$

Thus, Theorem 2.51 implies there is some N such that Y_N contains some $B(\overline{y},\delta)$ for $\delta > 0$. Namely, the complements have empty intersection, so at least one of the $Y \setminus Y_\bullet$ must fail to be open and dense, so one of the $Y_\bullet$ will contain an open ball. So we get

$$\overline{T(B_X(0,N))} \supseteq B(\overline{y},\delta).$$

We now spend a couple steps upgrading this.

2. Now, choose some $y \in B_Y(0,\delta)$ and $\varepsilon > 0$. Then by construction, we may choose $x_1 \in X$ with $\|Tx_1 - \overline{y}\|_Y < \varepsilon$ and $\|x_1\|_X < N$ and $x_2 \in X$ such that $\|Tx_2 - (\overline{y}+y)\|_Y < \varepsilon$ and $\|x_2\|_X < N$. So setting $x := x_1 - x_2$, we find

$$\|Tx - y\|_Y \le \|Tx_1 - \overline{y}\|_Y + \|Tx_2 - (y+\overline{y})\|_Y < 2\varepsilon,$$

and $\|x\| < 2N$.

3. Continuing, set $C := 2N/\delta$. Then each nonzero $y \in Y$ and $\varepsilon > 0$ has some $x \in X$ such that $\|Tx - y\| < \varepsilon$ and $\|x\| < 2N/\delta$. Indeed, this is direct from the prior step: replace y with $y_0 := \frac{\delta}{2\|y\|}y$, conjuring x_0 with $\|x_0\|_X < 2N$ and $\|Tx_0 - y_0\| < \frac{\delta}{2\|y\|}\varepsilon$ via the prior step, and then scale x_0 back up to $x := \frac{2\|y\|}{\delta}$ to conclude.

4. We now complete the proof by a limiting procedure. To begin, we choose $x_1 \in X$ such that $\|x_1\|_X \le C\|y\|_Y$ and $\|Tx_1 - y\|_Y < 2^{-1}\|y\|$. Then given x_n, we choose $x_{n+1} \in X$ such that

$$\|Tx_{n+1} - (y - Tx_1 - \cdots - Tx_n)\|_Y < 2^{-n-1}\|y\|_Y$$

such that $\|x_{n+1}\|_X \le C\|y - Tx_1 - \cdots - Tx_n\|_Y$, and this upper bound is just $C2^{-n}\|y\|_Y$ by construction. So we may define

$$x := \sum_{k=1}^{\infty} x_k,$$

which converges because X is complete (and this series is absolutely convergent). Additionally, we see $\sum_{k=1}^{N} Tx_k \to y$ by construction, so continuity of T tells us $Tx = y$. Lastly, we bound $\|x\|_X$ by the above absolutely convergent series, which is bounded by $C\|y\|_Y \sum_{k=0}^{\infty} 2^{-k} = 2C\|y\|_Y$, as required.

It remains to show (a). Well, (b) provides a constant A such that $T(B_X(0,A))$ contains $B_Y(0,1)$. So we see $T(B_X(0,Ar)) \supseteq B_Y(0,r)$ for any $r > 0$. By translating, we see that any ball $T(B_X(x,\varepsilon))$ contains $B_Y(Tx,\varepsilon/A)$. Thus, for any open subset $U \subseteq X$, we ere able to conclude that $T(U) \subseteq Y$ is open: for any $y \in T(U)$, write $y = Tx$ for $x \in U$, but then there is $\varepsilon > 0$ such that $B(x,\varepsilon) \subseteq U$, so

$$B(y,\varepsilon/A) \subseteq T(U)$$

is an open neighborhood of y contained in $T(U)$. So $T(U)$ is in fact open. ■

Remark 2.56. The power of Theorem 2.53 is that it is able to provide a concrete bound "A" on solving the "system" of equations $Tx = y$.

2.6.2 The Closed Graph Theorem

Here is our next result.

Theorem 2.57 (Closed graph)**.** Fix a Banach spaces X and Y. Fix a linear map $T\colon X \to Y$. If T has a graph

$$\Gamma := \{(x, Tx) \in X \times Y : x \in X\}$$

which is closed in $X \times Y$, then T is bounded.

Remark 2.58. Again, the power of Theorem 2.57 is that we are able to get a concrete bound on the norm of T from a topological condition.

Wait, how do we give the product a topology? One could use the product topology, but we can see this on the level of norms as well.

Definition 2.59 (product)**.** Given normed vector spaces X and Y, then the *product* $X \times Y$ is a vector space with norm given by

$$\|(x, y)\|_{X\times Y} := \|x\|_X + \|y\|_Y .$$

Remark 2.60. If X and Y are both complete, then $X \times Y$ is also complete, which we can see by taking limits coordinate-wise.

Let's see an example of an unbounded operator.

Example 2.61. Let X be the collection of continuously differentiable functions $f\colon [0, 1] \to \mathbb{C}$, where we take $\|f\|_X := \sup\{f(x) : x \in [0, 1]\}$, which we can see is a norm. (Notably, $\|f\|_X = 0$ means that $f(x) = 0$ everywhere, so $f = 0$.) Then we let Y be the collection of continuous functions $f\colon [0, 1] \to \mathbb{C}$ with the same norm $\|f\|_Y := \sup\{f(x) : x \in [0, 1]\}$, and we define $D\colon X \to Y$ by $Df := f'$. Note D is well-defined and linear, and we see D is bounded is equivalent to requiring

$$\sup_{\substack{f\in X\\ f\neq 0}} \frac{\|Df\|_Y}{\|f\|_X} = \sup_{\substack{f\in X\\ f\neq 0}} \frac{\sup\{f'(x) : x \in [0, 1]\}}{\sup\{f(x) : x \in [0, 1]\}} < \infty,$$

but this is false; for example, we can consider the family $f_n(x) = \sin nx$ which has $\|f_n\|_X = 1$ but $\|Df_n\|_Y \to \infty$ as $n \to \infty$.

Example 2.62. We continue Example 2.61. Let Γ be the graph of D, but we note that Γ is closed! Unwinding definitions, we are being given a sequence of functions $\{f_n\}_{n\in\mathbb{N}} \subseteq X$ and g and h such that $f_n \to g$ uniformly and $f_n' \to h$ uniformly, and we must show that $g' = h$. Well, write

$$f_n(x) = f_n(0) + \int_0^x f_n'(t)\, dt$$

and take limits everywhere to see $g(x) = g(0) + \int_0^x h(t)\, dt$, which implies $g' = h$.

It might look like the previous example violates Theorem 2.57, but in fact X is not complete and hence not Banach.

Here is another example.

Example 2.63. Consider the function $f\colon \mathbb{R} \to \mathbb{R}$ by $f(x) := 1/x$ for nonzero x and $f(0) = 0$. Then the graph of $f(x)$ is $\{(0,0)\} \sqcup \{(x,y) : xy = 1\}$, which is a closed subset of $\mathbb{R}^2$. But f is certainly not continuous, which is not a violation of Theorem 2.57 because f is not linear!

Anyway, let's prove Theorem 2.57.

Proof of Theorem 2.57. Quickly, note that linearity of T means that the graph Γ is linear. Because Γ is closed, we actually see that Γ is a complete normed vector space, where the norm is given by $\|\cdot\|_{X\times Y}$. So we now define $S\colon \Gamma \to X$ by $S(x,y) := x$, which is linear (because look at it) and bounded because

$$\|S(x,y)\|_X = \|x\|_X \le \|x\|_X + \|y\|_Y = \|(x,y)\|_{X\times Y}.$$

Furthermore, S is a bijection, so in fact S^{-1} is bounded by Corollary 2.55, meaning that

$$\|(x,Tx)\|_{X\times Y} \le C\|x\|_X$$

for some absolute constant $C > 0$. This unwinds into exactly showing that T is bounded, as required. ■

2.7 February 28

Today we discuss the uniform boundedness principle.

2.7.1 The Uniform Boundedness Principle

Here is our statement.

Theorem 2.64 (Uniform boundedness principle)**.** Fix a Banach space X. Let $\{T_\alpha\}_{\alpha\in\kappa} \subseteq X^*$ be a nonempty collection of linear functionals. Suppose that the family $\{T_\alpha\}_{\alpha\in\kappa}$ is pointwise bounded; i.e., for each $x \in X$, we have

$$\sup_{\alpha\in\kappa} |T_\alpha(x)| < \infty.$$

Then in fact we are uniformly bounded: $\sup\{\|T_\alpha\|_{X^*} : \alpha \in \kappa\} < \infty$.

Here are some non-examples.

Example 2.65. Define a family $f_n\colon [0,1] \to \mathbb{R}$ (where $n \in \mathbb{Z}^+$) by

$$f_n(x) := \begin{cases} n^2x & \text{if } 0 \le x \le 1/n, \\ n^2(x-2/n) & \text{if } 1/n \le x \le 2/n, \\ 0 & \text{if } 2/n \le x \le 1. \end{cases}$$

Then $\{f_n\}_{n\in\mathbb{Z}^+}$ the family is pointwise bounded because $f_n(x) = 0$ for $n > 3/x$ (say), so $\{f_n(x)\}_{n\in\mathbb{N}}$. Simply put, we do not violate Theorem 2.64 because $[0,1]$ is not a Banach space, and our functions f_n are not linear.

Example 2.66. Define $X := \mathbb{C}^{\oplus\mathbb{N}}$ be an infinite direct sum of $\mathbb{C}$s. Explicitly, X consists of sequences which are eventually 0. We give X the norm

$$\|x\|_X := \max_{k\in\mathbb{Z}^+} \{|x_k|\}.$$

Now, define $T_n\colon X \to \mathbb{C}$ by $T_n(x) := nx_n$. Then the family $\{T_n\}_{n\in\mathbb{N}}$ is pointwise bounded: for any $x \in X$, there is some $N > 0$ such that $x_n = 0$ for $n > N$, so

$$\sup_{n\in\mathbb{N}} |T_n(x)| = \sup_{0\le n\le N} |T_n(x)| = \max_{0\le n\le N} |T_n(x)| < \infty.$$

However, we see that $\|T_n\|_{X^*} \ge n$ for each n by taking the sequence $x \in X$ given by $x_k := 1_{n=k}$, so this family is not uniformly bounded. (In fact, we see that $\|T_n\|_{X^*} \le n$ as well by bounding $|T_n(x)| / \|x\|_X = n|x_n| / \sup\{|x_k|\} \le n$.) The reason we do not violate Theorem 2.64 is that X is not complete and hence not Banach.

Anyway, let's prove Theorem 2.64.

Proof of Theorem 2.64. We proceed in steps. Let $\|\cdot\|$ be the norm on X.

1. For $M > 0$, define the subset $E_M \subseteq X$ by

$$E_M := \left\{x \in X : \sup_{\alpha\in\kappa} |T_\alpha(x)| \le M\right\}.$$

   Then we note that X is covered by the collection $\{E_M\}_{M\in\mathbb{Z}^+}$ because the family $\{T_\alpha\}_{\alpha\in\kappa}$ is pointwise bounded.

   Further, we note that E_M is closed for each $M > 0$. For this, we should show that E_M contains its limit points. Well, suppose that $\{x_n\}_{n\in\mathbb{N}} \subseteq X$ is a sequence in E_M converging to some $x \in X$, and we want to show that $x \in E_M$. To check this, fix any $\alpha \in \kappa$, and we must show that $|T_\alpha(x)| \le M$. For this, we bound

$$|T_\alpha(x)| = \left|\lim_{n\to\infty} T_\alpha(x_n)\right| = \lim_{n\to\infty} |T_\alpha(x_n)|,$$

   where the last equality is by continuity.

2. Thus, X is the union of the countably many closed sets $\{E_M\}_{M\in\mathbb{Z}^+}$, so Theorem 2.51 explains that there must be some $M \in \mathbb{Z}^+$ and $\delta > 0$ and $x_0 \in X$ such that $B(x, \delta) \subseteq E_M$.

   In fact, we claim that $B(0, \delta) \subseteq E_{2M}$. Indeed, if $\|x\| < \delta$, then define $x := (x_0 + x) - x_0$ so that any $\alpha \in \kappa$ has

$$|T_\alpha(x)| = |T_\alpha(x_0 + x)| - |T_\alpha(x_0)| \le 2M,$$

   so $x \in E_{2M}$.

3. We now do our bounding. Thus far we have shown that $\|x\| < \delta$ implies that $|T_\alpha(x)| \le 2M$ for any $\alpha \in \kappa$. But then we see we can write

$$\|T_\alpha\|_{X^*} = \sup_{\substack{x\in X\\ \|x\|=\delta/2}} \frac{|T_\alpha(x)|}{\|x\|} \le \frac{2M}{\delta/2},$$

   so we have achieved our uniform bound. ■

Remark 2.67. Later in life, we will use Theorem 2.64 do some Fourier analysis.

2.7.2 A Little on Bases

Let's talk a little about bases for vector spaces. Notably, we have never defined a "basis" for a Banach space, and we will avoid doing so ever. Instead, we will define bases with adjectives.

Definition 2.68 (Hamel basis)**.** A *Hamel basis* for a vector space X is a collection $\{v_\alpha\}_{\alpha\in\kappa}$ of vectors in X such that any $x \in X$ can be written uniquely as a linear combination

$$x = \sum_{\alpha\in\kappa} c_\alpha v_\alpha,$$

where the scalars $\{c_\alpha\}_{\alpha\in\kappa}$ are zero for all but finitely many α. (This sum certainly converges because it is actually a finite sum.)

Remark 2.69. Equivalently, we can ask for the linear combination to merely exist and for the subset to be linearly independent in the sense that having any linear combination

$$\sum_{\alpha\in\kappa} c_\alpha v_\alpha = 0$$

(where c_α is zero for all but finitely many α) will have $c_\alpha = 0$ for all α.

Remark 2.70. Any vector space has a Hamel basis by some argument with the Axiom of choice (e.g., via Zorn's lemma). In fact, any linearly independent collection of vectors can be extended to a Hamel basis. Rigorously, one builds a partially ordered set of linearly independent sets (ordered by inclusion), use Zorn's lemma to extract a maximal such set, and then show that a maximal linearly independent set is spanning.

Remark 2.71. Fix an infinite-dimensional normed vector space X over $\mathbb{F}$ and a nonzero normed vector space Y over $\mathbb{F}$. Then we can show that there is an unbounded linear functional $T\colon X \to Y$. Indeed, let $\{v_\alpha\}_{\alpha\in\kappa}$ be a Hamel basis, which is infinite because X is infinite-dimensional. Now, choose a nonzero vector $y \in Y$ and some unbounded sequence $\{c_\alpha\}_{\alpha\in\kappa} \subseteq \mathbb{F}$, and define

$$Tv_\alpha \coloneqq c_\alpha \|v_\alpha\|_X \, y$$

extended linearly to be a functional on X. (Explicitly, T sends $\sum_\alpha c_\alpha v_\alpha = c_\alpha \|v_\alpha\|_X \, y$, which is well-defined because we are working with a Hamel basis.) However, T is not bounded because

$$\|T\|_{X^*} \geq \sup_{\alpha\in\kappa} \frac{\|Tv_\alpha\|_Y}{\|v_\alpha\|_X} = \sup_{\alpha\in\kappa} |c_\alpha| \, \|y\|_Y = \infty.$$

Here are some more cautionary tales.

Remark 2.72. In general, closed subspaces of a Banach space do not have to have complements.

Remark 2.73. "Useful" bases for normed vector spaces X do not exist in general. Later in life, we will have Hilbert bases for some subclass of normed vector spaces.

Remark 2.74. Fix a normed vector space X. If $V \subseteq X$ is a closed subspace and a point $x \notin V$, there is no well-defined notion of "closest point." In finite dimensions, it need not be unique, and in general, it need not even exist.

2.8 March 1

We began class with a review of topological spaces. We just refer to [Elb22].

2.8.1 Weak Topologies

It will be helpful to have other topologies to work with on our normed vector spaces. We recall the following notion.

Definition 2.75 (weak topology)**.** Fix a nonempty set X, and let $\{f_\alpha\}_{\alpha\in\kappa}$ be a nonempty collection of maps $f_\alpha\colon X\to Y_\alpha$, where $\{Y_\alpha\}_{\alpha\in\kappa}$ is a collection of topological spaces. Then we define the *weak topology* on X to be the topology generated by the collection

$$\bigcup_{\alpha\in\kappa}\left\{f^{-1}(U_\alpha):\text{open } U_\alpha\subseteq Y_\alpha\right\}.$$

Remark 2.76. Notably, the given collection above is a subbasis because it contains $\varnothing=f_\alpha^{-1}(\varnothing)$ and $X=f_\alpha^{-1}(Y_\alpha)$. As such, we can give ourselves a basis by taking finite intersections of the sets in our given collection.

Remark 2.77. Observe that the weak topology on X promises that each map $f_\alpha\colon X\to Y_\alpha$ is continuous. Conversely, if $\mathcal{T}$ is a topology on X such that the maps $f_\alpha\colon X\to Y_\alpha$ are continuous, then $\mathcal{T}$ must contain the weak topology.

Remark 2.78. We have the following universal property: a map $g\colon Z\to X$ is continuous if and only if the composites $(f_\alpha\circ g)\colon X\to Y_\alpha$ are all continuous. The forward direction occurs by taking composition. For the backward direction, it suffices to check that any $g^{-1}\left(f_\alpha^{-1}(U_\alpha)\right)\subseteq Z$ is open for any $\alpha\in\kappa$ and open $U_\alpha\subseteq Y_\alpha$, which holds by the continuity of $f_\alpha\circ g$.

Example 2.79. One gives the infinite product $X:=\prod_{\alpha\in\kappa}Y_\alpha$ of topological spaces $\{Y_\alpha\}_{\alpha\in\kappa}$ the weak topology with respect to the projections $\operatorname{pr}_\alpha\colon X\to Y_\alpha$.

Remark 2.80. Let's explain closures quickly. Let $E\subseteq X$ be a subset as above. Then $x\in\overline{E}$ if and only if $B\cap E\neq\varnothing$ for any basic open neighborhood $B\subseteq X$ of x. In other words, for any finite collection $S\subseteq\kappa$ of indices and open subsets $U_\alpha\subseteq Y_\alpha$ for $\alpha\in S$, we have

$$B\cap\bigcap_{\alpha\in S}f_\alpha^{-1}U_\alpha\neq\varnothing.$$

Example 2.81. Consider $X:=\{0,1\}^{\mathbb{R}}$ endowed with the product topology. Set $x:=0\in X$ and E to be the collection of $y\in X$ such that $y_t=1$ or all but countably many $t\in\mathbb{R}$. Then actually $x\in\overline{E}$: indeed, a basic open subset of X containing x merely requires that some finite subset of indices $t\in\mathbb{R}$ vanish, which E does intersect.

However, there is no countable sequence $\{x_n\}_{n\in\mathbb{Z}^+}$ in E which converge to x. Indeed, by taking a union over all x_n, we see that there is a cocountable subset $T\subseteq\mathbb{R}$ such that $(x_n)_t=1$ for all $n\in\mathbb{Z}^+$ and $t\in T$. So our sequence cannot approach x (for example) by using a basic open subset of X requiring $y_t=0$ for some $t\in T$.

Here is our chief example of interest.

Definition 2.82. Fix a normed vector space $(X, \|\cdot\|)$. Then the *weak topology* on X is the weak topology obtained by requiring that all bounded linear functionals $f \in X^*$ are continuous.

Remark 2.83. The weak topology on X is Hausdorff: for any distinct $x, y \in X$, it is enough to show that there is a linear functional $f \in X^*$ such that $f(x) = f(y)$. If x and y are linearly independent, then define f by $f(x) = 1$ and extend f from $\operatorname{span}\{x\}$ to all X by Theorem 2.33. If they are not linearly dependent, then define f by $f(x) = 1$ and $f(y) = 0$ and extend f again via Theorem 2.33.

2.9 March 4

Here we go.

2.9.1 An Example Dual Space

We begin class with an example of convergence being strange. Let C_0 denote the collection of sequences $\{(x_n)_{n\in\mathbb{Z}^+}\}$ converging to 0, and we give C_0 the norm $\|\cdot\|_\infty$ defined by taking the maximum. One can see that C_0 is now a normed vector space (inherited from $L^\infty(\mathbb{N})$), and it is complete (one needs to check that $C_0 \subseteq L^\infty(\mathbb{N})$ is closed, which can be done via limit points).

Let's attempt to understand C_0^*. For this, we consider $L^1(\mathbb{N})$, which consists of sequences $(y_m)_m$ such that $\sum_{m\ge1} |y_m| < \infty$. Then $L^1(\mathbb{N})$ is a normed vector space with norm given by

$$\|y\|_1 := \sum_{m=1}^{\infty} |y_m| .$$

Now here is our duality check.

Lemma 2.84. The map $L_\bullet \colon L^1(\mathbb{N}) \to C_0^*$ defined by

$$L_y(x) := \sum_{m=1}^{\infty} x_m y_m$$

is an isomorphism.

Proof. Here are our many checks.

- Note $L_y(x) \in \mathbb{R}$: the series (absolutely) converges because

$$\sum_{m=1}^{\infty} |x_m y_m| \le \|x\|_\infty \sum_{m=1}^{\infty} |y_m| = \|x\|_\infty \|y\|_1 .$$

- Note L_y is linear for each $y \in L^1(\mathbb{N})$ by a direct check.
- In fact, note L_y is a bounded linear functional and hence in C_0^*; this finishes showing that $L_\bullet$ is well-defined. Indeed, we simply bound

$$\left| \sum_{m=1}^{\infty} x_m y_m \right| \le \|x\|_\infty \|y\|_1$$

  as above, meaning that $|L_y(x)| / \|x\|_\infty \le \|y\|_1$ for any nonzero $x \in C_0$.
- Note $L_\bullet$ is linear by a direct check.

- In fact, we claim that $\|L_y\|_{C_0^*} = \|y\|_{L^1}$ for any $y \in L^1(\mathbb{N})$; note that this shows that $L_\bullet$ is injective. Above we showed that $|L_y(x)| / \|x\|_\infty \le \|y\|_1$, so we get $\le$ immediately. For the other inequality, fix some $N > 0$ very large, and we define (x_n) to be

$$x_n := \begin{cases} +1 & \text{if } y_n \ge 0 \text{ and } n \le N, \\ -1 & \text{if } y_n \le 0 \text{ and } n \le N, \\ 0 & \text{if } n > N. \end{cases}$$

Then we see that $\|x\|_\infty = 1$

$$|L_y(x)| = \sum_{m=1}^{N} |y_m|$$

by construction, so $|L_y(x)| \to \|y\|_1$ as $N \to \infty$, enforcing the equality.

- Lastly, we must check surjectivity. This requires some work. Fix a linear functional $f \in C_0^*$, and we need to show that $f = L_y$ for some $y \in L^1(\mathbb{N})$. Well, simply define $y_n := f(e_n)$ where e_n is the n-indicator sequence $(e_n)_m := 1_{n=m}$. For example, this implies that

$$L_y(e_n) = f(e_n)$$

for each n. Then L_y and f agree on the subspace $C_{00} \subseteq C_0$ of sequences which are eventually 0. Note that we have not technically shown that $y \in L^1(\mathbb{N})$ yet from its construction. Well, for large N, define the sequence $(x_n)_n$ as in the previous point, and we see that

$$\sum_{m=1}^{N} |y_m| = |L_y(x)| = |f(x)| \le \|f\| \cdot \|x\| = \|f\|,$$

so sending $N \to \infty$ reveals that $y \in L^1(\mathbb{N})$. Thus, continuity of L_y (it is now known to be bounded) and density of C_{00} in C_0 (by definition of converging to 0) forces $L_y = f$ is forced. ■

Remark 2.85. Let's do a convergence check. Define the sequence $e_n \in C_0$ by $(e_n)_m := 1_{m=n}$, and it turns out that $e_n \to 0$ as $n \to \infty$ in the weak topology, even though this is notably false in C_0 because $\|x_m\|_\infty = 1$ for each m! In other words, we have $f(e_n) \to 0$ as $n \to \infty$ for any bounded linear functional $f \in C_0^*$. To see this, write $f = L_y$ for $y \in L^1(\mathbb{N})$ via Lemma 2.84, but now $f(e_n) = y_n$, which goes 0 as $n \to \infty$ because $y \in L^1(\mathbb{N})$.

Remark 2.86. In general, if $x_n \to x$ in the weak topology for C_0, then it turns out that $(x_n)_m \to x_m$ for any fixed m. Indeed, let $y = e_m$ be the m-indicator as above, and we must have $L_y(x_n) \to L_y(x)$ as $n \to \infty$, from which the claim follows.

However, the converse is not true: let $x_n := 2^n e_n$ be 2^n times the n-indicator. For this to converge to any x, the previous paragraph enforces $x = 0$, but this sequence does not converge to 0 by using the bounded linear functional L_y where $y_m := 1/2^m$ for each m. The problem here is that the norms are getting too large, and it will turn out that the converse holds as long as the norms are bounded (e.g., in the unit ball).

2.9.2 The Weak-$*$ Topology

Here is our definition.

Definition 2.87 (weak-$*$ topology). Fix a normed vector space $(X, \|\cdot\|)$. Then the *weak-$*$* topology on X^* is the weak topology obtained by requiring that the linear functionals $\operatorname{ev}_x \colon X^* \to \mathbb{F}$ defined by $\operatorname{ev}_x(f) := f(x)$ are continuous.

Remark 2.88. Note that ev_x is actually a bounded linear functional on X^*. Indeed, one computes that

$$\|\mathrm{ev}_x\|_{X^{**}} = \sup_{\substack{f\in X^* \\ f\neq 0}} \frac{|\mathrm{ev}_x(f)|}{\|f\|_{X^*}} = \sup_{\substack{f\in X^* \\ f\neq 0}} \frac{|f(x)|}{\|f\|_{X^*}} \le \|x\|_X ,$$

where the last inequality uses the definition of $\|f\|_{X^*}$. Thus, $\mathrm{ev}_x\colon X^* \to \mathbb{F}$ is already continuous, so the weak-$*$ topology is in fact coarser (has fewer open sets) than the usual topology.

Example 2.89. Given $g \in X^*$ and $\varepsilon > 0$ and $x_0 \in X$, the weak-$*$ topology has subbasic open subset given by

$$\{f \in X^* : |f(x_0) - g(x_0)| < \varepsilon\}.$$

Indeed, the point is that continuity of ev_{x_0} may as well be checked on basic open subsets of $\mathbb{F}$, and the above set is the pre-image of the open ball around $g(x_0)$ with radius $\varepsilon > 0$.

Example 2.90. We continue the discussion of section 2.9.1 to explicate the weak-$*$ topology on $C_0^* \cong L^1(\mathbb{N})$. This amounts to choosing some $z \in L^1(\mathbb{N})$ and $\varepsilon > 0$ and $x \in C_0$, for which our subbasic open subset is

$$\left\{ y \in L^1(\mathbb{N}) : \left| \sum_{m=1}^{\infty} (y_m - z_m)x_m \right| < \varepsilon \right\}.$$

2.10 March 6

Here we go.

2.10.1 Reflexive Banach Spaces

We begin with a quick digression, upgrading Remark 2.88.

Remark 2.91. We note that there is a canonical linear map $\mathrm{ev}_\bullet\colon X \to X^{**}$ which preserves norms. The map is defined by defining $\mathrm{ev}_x \in X^{**}$ as $\mathrm{ev}_x(f) := f(x)$ for $f \in X^*$. Note Remark 2.88 shows that $\mathrm{ev}_x \in X^{**}$, and we can see that $\mathrm{ev}_\bullet$ is linear. It remains to compute the norm of ev_x. For one inequality, we use the argument of Remark 2.88 to note

$$\|\mathrm{ev}_x\|_{X^*} = \sup_{\substack{f\in X^* \\ f\neq 0}} \frac{|f(x)|}{\|f\|_{X^*}}.$$

To show the equality, we need to exhibit a good f. Well, by Theorem 2.33, one gets $f \in X^*$ so that $\|f\|_{X^*} = 1$ and $|f(x)| = \|x\|_X$; explicitly, f is defined by extending $f\colon \mathbb{F}x \to \mathbb{F}$ defined by $f(x) := \|x\|_X$. So the above computation tells us that $\|\mathrm{ev}_x\|$ is at least $|f(x)| \,/\, \|f\|_{X^*} = \|x\|_X$, as required.

Remark 2.92. If X is finite-dimensional, then $\dim X^{**} = \dim X^* = \dim X$, so the injective linear map $\mathrm{ev}_\bullet\colon X \to X^{**}$ must actually be an isomorphism. For infinite-dimensional X, then $\mathrm{ev}_\bullet$ need not be an isomorphism.

In light of the above remark, we pick up the following definition.

Definition 2.93 (reflexive)**.** A Banach space X is *reflexive* if and only if the evaluation map $\mathrm{ev}_\bullet\colon X \to X^{**}$ is an isomorphism.

By Remark 2.91, being reflexive merely requires surjectivity.

Remark 2.94. On the homework, we will show that $\operatorname{im}\operatorname{ev}_\bullet$ is dense in X^{**} in the weak-$*$ topology.

Remark 2.95. One can show that the closed unit ball of a Banach space X is compact in the weak topology if and only if X is reflexive.

Non-Example 2.96. Section 2.9.1 tells us that C_0 is dual to $L^1(\mathbb{N})$. However, one can show that $L^1(\mathbb{N})$ is dual to $L^\infty(\mathbb{N})$, which is the space of bounded sequences, but the evaluation map $C_0 \to L^\infty(\mathbb{N})$ fails to be surjective (meaning C_0 is not reflexive!). For example, let $\mathcal{L} \subseteq L^1(\mathbb{N})$ of sequences with a limit; notably, the constant sequence of $(1, 1, \ldots)$ is in $\mathcal{L} \setminus C_0$, and one can use the Hahn–Banach theorem to bring this up to a linear functional on $L^1(\mathbb{N})$ not coming from C_0.

Non-Example 2.97. One can also see that $L^1(\mathbb{N})$ fails to be reflexive. One can again use the Hahn–Banach theorem or appeal to the following remark.

Remark 2.98. By functoriality, if X is reflexive, then X^* is also reflexive. We claim that the converse also holds! Suppose for the sake of contradiction that X^* is reflexive but X is not. Then $\operatorname{im}\operatorname{ev}_\bullet^X \subsetneq X^{**}$, so Theorem 2.33 permits us to find some bounded linear functional $\lambda \in X^{***}$ such that $\lambda \neq 0$ but $\lambda|_{\operatorname{im}\operatorname{ev}_\bullet^X} = 0$. However, we will show that $\lambda|_{\operatorname{im}\operatorname{ev}_\bullet^X} = 0$ implies $\lambda = 0$. Indeed, X^* is reflexive, so we are granted $\lambda_0 \in X^*$ such that $\lambda = \operatorname{ev}_{\lambda_0}^{X^*}$; it is enough now to show that λ_0 vanishes. Well, for each $x \in X$, we see that

$$\lambda_0(x) = \operatorname{ev}_{\lambda_0}^{X^*}(\operatorname{ev}_x) = \lambda(\operatorname{ev}_x) = 0.$$

2.10.2 Weak Compactness Results

Here is today's main result.

Theorem 2.99 (Helly)**.** Fix a separable Banach space $(X, \|\cdot\|)$. Then the closed unit ball of X^* is sequentially compact in the weak-$*$ topology. In other words, given a sequence of bounded linear functionals $\{f_n\}_{n\geq 1} \subseteq X^*$ such that $\|f_n\|_{X^*} \leq 1$ for each n, there is a subsequence $\{f_{n_k}\}_{k\geq 1}$ and $f \in X^*$ such that each $x \in X$ makes $f_n(x) \to f(x)$ as $n \to \infty$.

Proof. This is a Cantor diagonal argument. Choose a countable dense subset $\{x_n\}_{n\in\mathbb{N}}$. We now proceed with an inductive construction.

1. Because $|f_n(x_1)| \leq \|x_1\|_X$ for each n, we see that the sequence $\{f_n(x_1)\}_{n\in\mathbb{N}}$ is bounded, so we can find a subsequence $\{f_{n(i_1)}\}_{i_1\in\mathbb{N}}$ with a subsequence where the limit $\lim f_{n(i_1)}(x_1)$ exists.
2. Next up, we again pass to a subsequence where $f_{n(i_1)(i_2)}$ so that $\lim f_{n(i_1)(i_2)}(x_2)$ exists.

We have forced ourselves into the notation as above so that we can work with a subsequence $\{g_i\}_{i\geq 1}$ where $g_i := f_{n(i)\cdots(i)}$ repeated i times. As such, $g_i(x_n)$ converges as $i \to \infty$ for any n.

It remains to show that $\{g_i\}_{i\geq 1}$ actually converges to a function f in the closed unit ball. It suffices to show that this subsequence is Cauchy pointwise; then completeness of X^* tells us that our sequence converges to an actual bounded linear functional, and continuity of $\|\cdot\|_{X^*}$ means that we will land in the closed unit ball. Now, to show Cauchy, for each $x \in X$, choose $\varepsilon > 0$ and some x_j such that $\|x_j - x\| < \varepsilon$, and we find

$$\limsup_{m,n\to\infty} |g_m(x) - g_n(x)| \leq \limsup_{m,n\to\infty} |g_m(x_j) - g_n(x_j)| + 2\sup_{n\in\mathbb{N}} |g_n(x) - g_n(x_j)|$$

by some rearrangement. But g_n is in the closed unit ball, so the right-hand term is bounded above by 2ε. Additionally, the left term vanishes by construction of the $g_\bullet$, so are done. ■

Remark 2.100. Theorem 2.99 is useful because oftentimes one will construct a sequence of some functions (e.g., approximate solutions to a differential equation), and then we want to actually produce a limiting function. Then we see that it will merely remain to check some separability (which is the case when working in any Euclidean space) and a uniform boundedness (which is often not so bad).

Theorem 2.99 has the following analogous result without separability.

Theorem 2.101 (Alaoglu)**.** Let X be a Banach space. Then the closed unit ball of X^* is compact in the weak-$*$ topology.

Note that we are working with the dual space here the result is not true for the closed unit ball in the weak topology on X; see Remark 2.95.

It will be helpful to have the following extension result.

Lemma 2.102. Fix a normed vector space X, and let $B \subseteq X$ be the closed unit ball. Suppose $f\colon B \to \mathbb{F}$ is linear in the sense that $f(x+y) = f(x) + f(y)$ and $f(tz) = tf(z)$ whenever all inputs are in B. Then f has a unique linear extension $F\colon X \to \mathbb{F}$.

Proof. Define $F(x) := f(x/\|x\|)$. In fact, one finds that $F(x) = tf\left(\frac{1}{t}x\right)$ whenever $\left\|\frac{1}{t}x\right\| \leq 1$ by the given linearity; one can now check that F is linear. One can see that this construction of F is unique. ■

2.11 March 8

Today we complete our discussion of Banach spaces.

2.11.1 Alaoglu's Theorem

We are in the middle of proving the following result.

Theorem 2.101 (Alaoglu)**.** Let X be a Banach space. Then the closed unit ball of X^* is compact in the weak-$*$ topology.

Proof. Let $D := \{z \in \mathbb{F} : |z| \leq 1\}$ be the closed ball around 1, which we note is compact. Now, we set $Y := D^B$ to be the infinite product, which we note is compact by Tychonoff's theorem.

We will turn compactness of Y to compactness of our closed unit ball $B^* \subseteq X^*$. As such, we define $\varphi\colon B^* \to Y$ by

$$\varphi(f) := f|_B.$$

To see that $f|_B$ is in fact a map $B \to D$ note that $f \in B^*$ must have $|f(x)| \leq \|x\|$ for all $x \in X$, so in particular $|f(x)| \leq 1$ for $x \in B$.

We claim that φ is a homeomorphism of B^* (with the weak topology) onto a closed subset of Y. This will finish the proof because the image $\Lambda := \operatorname{im}\varphi$ will be compact in Y, which then pulls back to compactness for B^* via our homeomorphism. So it remains to show our claim.

- We claim that φ is injective. Indeed, suppose that $\varphi(f) = \varphi(g)$. Then $f|_B = g|_B$, so $f = g$ by Lemma 2.102.

  The point is that $\varphi\colon B^* \to \Lambda$ is now a bijection.

- We claim that $\varphi\colon B^* \to \Lambda$ is a homeomorphism. This comes down to understanding the open subsets of B^* and of Λ.

  For example, subbasic open subsets of Λ take the following form: choose some $x_0 \in B$ and $c \in \mathbb{F}$ and $\varepsilon > 0$, and we have a subbasic open subset of the form

  $$\{g \in \Lambda : |g(x_0) - c| < \varepsilon\}.$$

  Unraveling Λ, we see that we are looking at

  $$\{f|_B \in \Lambda : |f(x_0) - c| < \varepsilon\}.$$

  On the other hand, subbasic open subsets of B^* take the following form: choose some $x_0 \in X$ and $c \in \mathbb{F}$ and $\varepsilon > 0$, and we have a subbasic open subset of the form

  $$\{f \in B : |f(x_0) - c| < \varepsilon\}.$$

  (Namely, we are finding the subbasic open subset forced by continuity of $\operatorname{ev}_{x_0} : X^* \to \mathbb{F}$.) It may appear that we are gaining subbasic open subsets by allowing $x_0 \in X$ instead of restricting to $x_0 \in B$ as previously, but in fact we gain nothing: if $x_0 \notin B$, simply replace x_0 with $x_0/\|x_0\|$ and c with $c/\|x_0\|$ and ε with $\varepsilon/\|x_0\|$. So we see that our topologies are really the same, so we are done.

- We claim that $\Lambda \subseteq Y$ is closed. Well, choose some $g \in \overline{\Lambda}$, and we claim that $g(x+y) = g(x) + g(y)$ and $g(tz) = tg(z)$ whenever the relevant inputs live in B. This will imply that g comes from some linear functional on X by Lemma 2.102, and in fact $|g(x)| \le \|x\|$ is forced because $|g(x/\|x\|)| \le 1$, meaning g comes from some linear functional in B^*, as needed.

  So it remains to show the claim of the previous paragraph. We will show that $g(ax+by) = ag(x)+bg(y)$ whenever all inputs are in B. Well, note that

  $$V := \{f \in Y : f(ax+by) = af(x) + bf(y)\}$$

  is a closed subset of Y: the map $f \mapsto (f(x+y) - (f(x)+f(y)))$ is continuous as the linear combination of projections from Y, and V is the pre-image of this map from $\{0\}$, so V is closed. Additionally, $\Lambda \subseteq V$ because any element of Λ comes from a linear functional, so $\overline{\Lambda} \subseteq V$ follows. Thus, $g \in V$. ■

Remark 2.103. One can use nets to prove Theorem 2.101, but Professor Christ is quite proud of never using nets, so we will not.

2.11.2 Hilbert Spaces

We will want the notion of inner product spaces.

Definition 2.104 (inner product)**.** Fix an $\mathbb{F}$-vector space X. Then an *inner product* $\langle \cdot, \cdot \rangle$ on X is a mapping $X \times X \to \mathbb{C}$ satisfying the following.

- Additive: $\langle x+y, z\rangle = \langle x, z\rangle + \langle y, z\rangle$.
- Conjugate symmetry: $\langle x, y\rangle = \overline{\langle y, x\rangle}$.
- Positive definite: $\langle x, x\rangle \ge 0$ for each x and only zero when $x = 0$.
- Scalar multiplication: $\langle tx, y\rangle = t\langle x, y\rangle$.

Inner products produce norms.

Definition 2.105. Fix an inner product space X. Then the *norm* on X is given by

$$\|x\| := \langle x, x\rangle^{1/2}.$$

We will not bother to check that $\|\cdot\|$ is a norm, but it is. The most nontrivial check is the triangle inequality, which will be checked shortly in Proposition 3.3.

Remark 2.106. Note that

$$\|x-y\|^2 + \|x+y\|^2 = \langle x-y, x-y\rangle + \langle x+y, x+y\rangle \stackrel{*}{=} 2\langle x, x\rangle + 2\langle y, y\rangle = 2\|x\|^2 + 2\|y\|^2,$$

where $\stackrel{*}{=}$ is by direct expansion. This is called the "parallelogram law." It turns out that a norm satisfying the parallelogram law actually comes from an inner product. Explicitly, one can set

$$\langle x, y\rangle := \frac{\|x+y\|^2 - \|x-y\|^2}{4} + i\frac{\|ix-y\|^2 - \|ix+y\|}{4}$$

and check that the various identities hold by creative applications of the parallelogram law.

At long last, here is our definition.

Definition 2.107 (Hilbert space)**.** A normed vector space X is a *Hilbert space* if and only if is a complete inner product space (where the norm is induced by the inner product).

THEME 3

HILBERT SPACES

3.1 March 11

Here we go.

3.1.1 Bounding with Inner Products

Here are some fundamental bounding results.

Proposition 3.1 (Cauchy–Schwarz)**.** Fix an inner product space X. Then each $x, y \in X$ has

$$|\langle x, y\rangle| \leq \|x\| \cdot \|y\|.$$

Proof. Quickly, we take care of some special cases. If $\langle x, y\rangle = 0$, there is nothing to do. Let's also handle the case where $x = 0$. Then $\langle 0, y\rangle = 0$ because $\langle 0, y\rangle = \langle 0, y\rangle + \langle 0, y\rangle$ by bilinearity. So we conclude that $\langle 0, y\rangle \leq \|0\| \cdot \|y\|$.

Otherwise, by scaling everything via bilinearity, we may assume that $\|x\| = \|y\| = 1$. Then we find that

$$0 \leq \langle x + y, x + y\rangle = \|x\|^2 + \|y\|^2 + 2\operatorname{Re}\langle x, y\rangle$$

by a direct expansion using the bilinearity. We now twist to get rid of Re. Namely, choose $\theta \in \mathbb{R}$ so that $\left\langle e^{i\theta}x, y\right\rangle = -|\langle x, y\rangle|$, doable because $\left\langle e^{i\theta}x, y\right\rangle = e^{i\theta}\langle x, y\rangle$ allows us to adjust the angle of $\langle x, y\rangle$ until it is a negative real number. As such, rerunning the argument at the start of the paragraph, we are able to conclude that

$$0 \leq \|x\|^2 + \|y\|^2 - 2|\langle x, y\rangle|,$$

so

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

where we are using $\|x\| = \|y\| = 1$. ■

Remark 3.2. Proposition 3.1 tells us that the map $X \times X \to \mathbb{F}$ given by $(x, y) \mapsto \langle x, y\rangle$ is continuous. Indeed, it is enough to check that this linear functional is bounded, so we note that $\|(x, y)\| \leq 1$ means $\|x\|, \|y\| \leq 1$, so

$$|\langle x, y\rangle| \leq \|x\| \cdot \|y\| \leq 1$$

by Proposition 3.1.

Proposition 3.3 (Triangle inequality)**.** Fix an inner product space X. Then $\|x+y\| \le \|x\| + \|y\|$.

Proof. It suffices to check that the squares have the inequality, but

$$\|x+y\|^2 = \langle x+y, x+y\rangle = \|x\|^2 + \|y\|^2 + 2\operatorname{Re}\langle x, y\rangle.$$

Thus, Proposition 3.1 tells us that this is bounded above by $\|x\|^2 + \|y\|^2 + 2\|x\| \cdot \|y\|$, which is precisely $(\|x\| + \|y\|)^2$. ■

We now take a moment for an example.

Example 3.4. Let $L^2(\mathbb{N})$ denote the set of sequences $\{x_n\}_{n=1}^\infty$ such that $\sum_{n=1}^\infty |x_n|^2 < \infty$. To see that this is a vector space, note that

$$\sum_{n=1}^\infty |ax_n + by_n|^2 = |a|^2 \sum_{n=1}^\infty |x_n|^2 + |b|^2 \sum_{n=1}^\infty |y_n|^2 + |ab|^2 \sum_{n=1}^\infty 2|x_ny_n|^2$$

Now, $|x_ny_n| \le |x_n|^2 + |y_n|^2$, so the entire summation above absolutely converges. In fact, this same inequality tells us that

$$\langle x, y\rangle \coloneqq \sum_{n=1}^\infty x_n\overline{y_n}$$

absolutely converges, and so this formula provides us with an inner product on $L^2(\mathbb{N})$. For example, $\|x\|^2 = \langle x, x\rangle = \sum_{n=1}^\infty |x_n|^2$.

Remark 3.5. One can check that $L^2(\mathbb{N})$ is complete for its given norm, so in fact $L^2(\mathbb{N})$ is a Hilbert space. The point is that a Cauchy sequence $\{x_m\}_{m\in\mathbb{N}} \subseteq L^2(\mathbb{N})$ will converge component-wise because $\|x_m - x_{m'}\| \ge |(x_m)_n - (x_{m'})_n|^2$ for each n, so we can define the desired limit sequence x so that $(x_m)_n \to x_n$ as $m \to \infty$. Now the absolute convergence everywhere will tell us that $x \in L^2(\mathbb{N})$ and that $x_m \to x$ as $m \to \infty$.

3.1.2 Geometry in Hilbert Spaces

We now show the following nice geometric consequence: closest points to subspaces make sense.

Lemma 3.6. Fix a Hilbert space X. Given a closed subspace $V \subseteq X$ and $x \in X$ there is a unique $v \in V$ such that

$$\|x - v\| = \inf_{u\in V} \|x - u\|.$$

Proof. We begin by showing existence of v. Let D be the minimal distance $\inf\{\|x - v\| : v \in V\}$. Note that D is finite because V is nonempty (e.g., $0 \in V$). Using this infimum, we are provided a sequence $\{u_m\}_{m\in\mathbb{N}}$ such that $\|x - u_n\| \to D$ as $n \to \infty$.

By continuity of $\|\cdot\|$ (this is true even for normed vector spaces, or one can use Remark 3.2), it suffices to show that the sequence $\{u_m\}_{m\in\mathbb{N}}$ converges in V, so we want to show that this sequence is Cauchy (and the limit will land in V because V is closed). Well, for $m, n \in \mathbb{N}$, we use the Parallelogram law Remark 2.106 to write

$$\|u_m - u_n\|^2 = \|(u_m - x) - (u_n - x)\|^2 = 2\|u_m - x\|^2 + 2\|u_n - x\|^2 - \|u_m + u_n - 2x\|^2.$$

The first two terms will be small, and the last term is $4\left\|\frac{1}{2}(u_m+u_n)-x\right\| \geq 4D^2$ because $\frac{1}{2}(u_m+u_n)\in V$. Thus,

$$\limsup_{m,n\to\infty}\|u_m-u_n\|^2 = 2\limsup_{m,n\to\infty}\|u_m-x\|^2 + 2\limsup_{n\to\infty}\|u_n-x\|^2 - 4D^2 = 0$$

because $\|u_m-x\|\to D$ as $m\to\infty$ (and similar for n).

We now show that v is unique. Suppose that v_1 and v_2 both satisfy $\|x-v_1\|=\|x-v_2\|=D$. Then

$$\|v_1-v_2\|^2 \leq 2\|v_1-x\|^2 + 2\|v_2-x\|^2 - 4D^2 = 0$$

by the above computations, so $v_1=v_2$ is forced. ■

With a notion of closest, we now talk about orthogonality.

Definition 3.7 (orthogonal complement)**.** Fix a subset S of an inner product space X. Then we define the *orthogonal complement*

$$S^\perp := \{v\in X : \langle x,v\rangle = 0 \text{ for all } x\in S\}.$$

Remark 3.8. One can check that $S^\perp$ is always a subspace (by the bilinearity of the inner product). In fact, $S^\perp$ is closed: it is the pre-image of 0 along the continuous map $X\to\prod_{x\in S}\mathbb{F}$ defined by $v\mapsto\{\langle x,v\rangle\}_{x\in S}$. (This map is continuous because its coordinates are continuous: indeed, we want to check that the map $v\mapsto\langle x,v\rangle$ is continuous, which is true.)

As a sanity check, orthogonal complements produce unique decomposition.

Proposition 3.9. Fix a Hilbert space X and a closed subspace V. Then $V+V^\perp = X$. In fact, for any $x\in X$, there are unique $v\in V$ and $u\in V^\perp$ such that $x=v+u$.

Proof. We begin with existence. Let $v\in V$ be the closest vector to x, which exists by Lemma 3.6. We claim that $\langle x-v,w\rangle=0$ for all $w\in V$, which will complete the existence proof because then we may set $u:=x-v$. The point is to consider the function $F_w\colon\mathbb{R}\to\mathbb{C}$ defined by

$$F_w(t) := \|x-(v+tw)\|^2$$

for some fixed vector $w\in V$; this is a continuous function in t because $\|\cdot\|$ is continuous. Now, F_w is minimized at $t=0$ by assumption on v, so $\frac{d}{dt}F_w\big|_{t=0}=0$ provided we can show that this derivative makes sense. Well, we see that

$$F_w(t) = \|x-v\|^2 - 2t\operatorname{Re}\langle x-v,tw\rangle + t^2\|w\|^2,$$

which is a polynomial in t, so this is indeed differentiable at 0, whereupon we can compute its derivative to find that $\operatorname{Re}\langle x-v,w\rangle = 0$. If $\mathbb{F}=\mathbb{C}$, this is enough to conclude that $\langle x-v,w\rangle=0$, but in this case, we rerun the argument with w replaced by iw to see that $\operatorname{Re}\langle x-v,iw\rangle = -i\operatorname{Im}\langle x-v,w\rangle$ also vanishes, so $\langle x-v,w\rangle=0$.

It remains to check uniqueness. Suppose that $v,v'\in V$ and $u,u'\in V^\perp$ satisfies $v+u=v'+u'$. Then we want to check that $v=v'$ and $u=u'$. Well, we see that

$$w := v-v' = u'-u$$

lives in $V\cap V^\perp$, but then $\langle w,w\rangle=0$ by definition of $V^\perp$, so $\|w\|^2=0$, so $w=0$. ■

3.2 March 13

Here we go.

3.2.1 Functionals on Hilbert Spaces

We begin with a remark on inner product spaces.

Remark 3.10. Fix an inner product space X. Then the map $\varphi_y(x) := \langle x, y\rangle$ satisfies the following properties.

- Linear in x: we see $\varphi_y(ax + a'x') = a\langle x, y\rangle + a'\langle x', y\rangle = a\varphi_y(x) + a'\varphi_y(x')$.
- Conjugate-linear in y: we see $\varphi_{ay+a'y'}(x) = \langle x, ay + a'y'\rangle = \overline{a}\langle x, y\rangle + \overline{a'}\langle x, y'\rangle = (\overline{a}\varphi_y + \overline{a'}\varphi_{y'})(x)$.
- Bounded: by Proposition 3.1, we see

$$|f_y(x)| \le \|x\| \cdot \|y\|,$$

so $\|f_y\|_{X^*} \le \|y\|$. In fact, $f_y(y) = \|y\|$, so actually $\|f_y\|_{X^*} = \|y\|$ exactly.

So we get a conjugate linear isometry $\varphi_\bullet \colon X \to X^*$.

Being a Hilbert space lets us upgrade this map.

Proposition 3.11. Fix a Hilbert space X. Then the map $y \mapsto \varphi_y$ (where $\varphi_y \colon X \to \mathbb{R}$ is the map $\varphi_y(x) := \langle x, y\rangle$) is a surjection from X onto X^*.

Proof. Let f be a bounded linear functional on X. If $f = 0$, then $f = \varphi_0$, so we may assume that f is not identically zero. Because f is continuous, we note that

$$V := \ker f = \{x \in X : f(x) = 0\}$$

is a closed subspace of X not equal to X. As such, we can find an element of X not in V, and then Proposition 3.9 allows us to make this into a nonzero element $y \in V^\perp$. We are going to show that f is approximately φ_y.

Quickly, we claim that $\dim V^\perp = 1$ and actually spanned by y. Well, suppose that $x \in V^\perp$ and set $\alpha := f(x)/f(y)$ so that we would like to show $x = \alpha y$. Well, by rearranging, we know that

$$f(\alpha y - x) = 0,$$

so $\alpha y - x \in V$, so write $\alpha y = x + v$ for some $v \in V$, but then the uniqueness of Proposition 3.9 requires $(\alpha y, 0) = (x, v)$, so $x = \alpha y$.

We are now able to show that $f = \varphi_y$. For example, any $x \in X$ by Proposition 3.9 can be written uniquely as $\alpha y + v$ for some $\alpha \in \mathbb{F}$ and $v \in V$ so that

$$\varphi_y(\alpha y + v) = \langle \alpha y + v, y\rangle = \alpha \|y\|^2.$$

On the other hand, $f(x) = \alpha f(y)$, so

$$\frac{f_y(x)}{f(x)} = \frac{\|y\|^2}{f(y)}$$

for each (nonzero) x, so f and φ_y only differ by a fixed nonzero element of $\mathbb{F}$, so we can multiply y through by (the conjugate of) this element to complete. ■

Remark 3.12. One can use Proposition 3.11 to show that Hilbert spaces X are reflexive. In fact, this result tells us that X^* should be a Hilbert space because its norm will satisfy the Parallelogram law. Now checking surjectivity of $\mathrm{ev}_\bullet$ is some purely formal composition of surjections.

3.2.2 Orthonormal Sets

Here is our definition.

Definition 3.13 (orthonormal)**.** Fix an inner product space X. Then a set $S \subseteq X$ is an *orthonormal set* if and only if

$$\langle v, w \rangle = 1_{v=w}$$

for all $v, w \in S$. We say that S is merely *orthogonal* if and only if distinct $v, w \in S$ have $\langle v, w \rangle = 0$.

Example 3.14. Consider the Hilbert space $X = L^2(\mathbb{N})$. Then the set $\{e_n\}_{n\in\mathbb{N}}$ where $(e_n)_i := 1_{i=n}$ is an orthonormal set.

One can define L^2 in more generality than we have done so.

Example 3.15. Fix a measure space $(X, \mathcal{M}, \mu)$. Then the vector space $L^2(X, \mathcal{M}, \mu)$ of measurable functions $f\colon X \to \mathbb{C}$ such that $\int_X f^2\,d\mu < \infty$ is almost an inner product space when given the inner product

$$\langle f, g \rangle = \int_X f\overline{g}\,d\mu.$$

The issue is that $\langle f, f \rangle = 0$ only implies that $\int_X |f|^2\,d\mu = 0$, meaning that f merely vanishes almost everywhere. One fixes this by defining $\ell^2(X, \mathcal{M}, \mu)$ to be the vector space of equivalence classes of measurable functions.

We would like to work with infinite sums of our bases, but this requires a reasonable notion of convergence.

Definition 3.16. Fix a set S of nonnegative real numbers. Then we define the sum $\sum_{s\in S} S$ as

$$\sup_{\text{finite } S'\subseteq S} \sum_{s'\in S'} s'$$

Remark 3.17. If S is countable, then this is just usual convergence because absolute convergence permits us to sum in any order anyway.

Remark 3.18. If $\sum_{s\in S} s$ is actually finite, then S can only have countably many nonzero real numbers. Indeed, suppose the summation is finite. Then for each positive integer N, we see that

$$\sum_{s\in S} s \geq \sum_{\substack{s\in S\\ s\geq 1/N}} s \geq \frac{1}{N}\#\{s \in S : s \geq 1/N\},$$

so $\#\{s \in S : s \geq 1/N\}$ is finite for all N. Sending $N \to \infty$ tells us that the nonzero elements of S live in a countable union of finite sets, so the nonzero elements of S form a countable set.

So far we have discussed convergence for sums of real numbers. We now pass to vectors.

Proposition 3.19 (Bessel's inequality)**.** Fix an orthonormal subset S of an inner product space X. Then

$$\sum_{s\in S} |\langle x, s \rangle|^2 \leq \|x\|^2$$

for any $x \in X$.

Proof. By the definition of our summation, it is enough to show that

$$\sum_{s'\in S'} |\langle x, s\rangle|^2 \le \|x\|^2$$

for any finite subset $S' \subseteq S$. Thus, we may assume that S is finite.

The trick now is to take the projection onto the span of S. Set

$$y := \sum_{s\in S} \langle x, s\rangle s,$$

which makes sense as a sum because now S is finite. We claim that $\langle x - y, y\rangle = 0$. This will complete the proof because then $\|x\|^2 = \|x-y\|^2 + \|y\|^2$, and $\|y\|^2$ is the sum $\sum_{s\in S} |\langle x, s\rangle|^2$ because the $s \in S$ are mutually orthogonal.

So it remains to show the claim. This is purely formal. Indeed,

$$\begin{aligned}
\langle x - y, y\rangle &= \langle x, y\rangle - \langle y, y\rangle \\
&= \sum_{s\in S} \langle x, \langle x, s\rangle s\rangle + \sum_{s\in S} \langle x, s\rangle \overline{\langle x, s\rangle} \langle s, s\rangle \\
&= \sum_{s\in S} \overline{\langle x, s\rangle} \langle x, s\rangle + \sum_{s\in S} \langle x, s\rangle \overline{\langle x, s\rangle} \\
&= 0,
\end{aligned}$$

where we have repeatedly used the orthogonality of S. ■

Lemma 3.20. Fix a countable orthonormal subset S of a Hilbert space X. Given scalars $\{t_s\}_{s\in S}$ in $\mathbb{F}$ such that $\sum_{s\in S} |t_s|^2 < \infty$, then

$$\sum_{s\in S} t_s s$$

converges in X.

Proof. One might complain that the last summation we wrote down might depend on the order of summation. We will show that it does not. If S is finite, there is nothing to do, so we may assume that S is enumerated as $\{u_n\}_{n\in\mathbb{N}}$, and we set $t_k := t_{u_k}$.

Now, because X is complete, it suffices to show that our partial sums form a Cauchy sequence, meaning that we would like to show that

$$\lim_{m,n\to\infty} \left\| \sum_{k=m}^{n} t_k u_k \right\| = 0.$$

But now this is a finite sum, so we may write

$$\left\| \sum_{k=m}^{n} t_k u_k \right\| = \sum_{k=m}^{n} |t_k|^2 \le \sum_{k=m}^{\infty} |t_k|^2,$$

and the right-hand side vanishes as $m \to \infty$ because $\sum_k |t_k|^2$ converges, so we are done.

Lastly, we check that the order of the summation does not depend on our enumeration. Well, a different enumeration amounts to choosing a permutation $\sigma\colon \mathbb{N} \to \mathbb{N}$, and then we want to show that

$$\lim_{n\to\infty} \left\| \sum_{k=1}^{n} t_k u_k - \sum_{k=1}^{n} t_{\sigma(k)} u_{\sigma(k)} \right\|.$$

Well, for any $\varepsilon > 0$, choose n large enough so that $\sum_{k>n} |t_k|^2 < \varepsilon$ and $\sum_{k>n} \left|t_{\sigma(k)}\right|^2 < \varepsilon$. Then taking N large enough so that $\{\sigma(1), \ldots, \sigma(N)\}$ contains $\{1, \ldots, n\}$ tells us that the norm

$$\left\|\sum_{k=1}^{N} t_k u_k - \sum_{k=1}^{N} t_{\sigma(k)} u_{\sigma(k)}\right\| \le \left\|\sum_{k=n+1}^{N} t_k u_k\right\| + \left\|\sum_{k=n+1}^{N} t_{\sigma(k)} u_{\sigma(k)}\right\| < 2\varepsilon.$$

Sending $N \to \infty$ permits $\varepsilon \to 0^+$ and hence completes the proof. ■

Remark 3.21. This proof actually shows that

$$\left\|\sum_{n=1}^{\infty} t n u_n\right\|^2 = \sum_{n=1}^{\infty} |t_n|^2$$

under our hypotheses, by continuity of $\|\cdot\|$.

Remark 3.22. One might expect the hypothesis on the scalars to be $\sum_s |t_s| \le 1$ so that $\sum_s \|t_n u_n\| < \infty$. Our lemma is stronger though. Indeed, note that $\sum_s |t_s|^2 < \infty$ is a weaker condition than $\sum_s |t_s| < \infty$: if $\sum_s |t_s| < \infty$, then the sequence $\{t_s\}$ is bounded by some T, so we may write

$$\sum_{s\in S} |t_s|^2 \le T \sum_{s\in S} |t_s| < \infty.$$

In fact, the sum of squares converging is strictly weaker: one has $\sum_n 1/n^2 < \infty$ but $\sum_n 1/n = \infty$.

Remark 3.23. To further motivate our hypothesis on the scalars, note that if $\sum_s t_s s$ has any hope of convergence, then

$$\left\|\sum_{s\in S} t_s s\right\| = \sum_{s\in S} |t_s|^2$$

had better converge.

3.3 March 15

Here we go.

3.3.1 Orthonormal Basis

We now move towards our definition of orthonormal basis.

Definition 3.24 (complete)**.** Fix a Hilbert space X. Then an orthonormal subset $S \subseteq X$ is *complete* if and only if each $x \in X$ has $\langle x, s\rangle = 0$ for all $s \in S$ if and only if $x = 0$.

This is some sort of analogue of spanning. So here is our definition of basis.

Definition 3.25 (orthornomal basis)**.** Fix a Hilbert space X. Then a subset $S \subseteq X$ is an *orthonormal basis* if and only if it is orthonormal and complete.

Let's actually check that "complete" reasonably means spanning.

Lemma 3.26. Fix an orthonormal basis S of a Hilbert space X. Then for each $x \in X$, we have

$$x = \sum_{s \in S} \langle x, s \rangle s.$$

In particular, the sum converges.

Proof. We begin by noting that

$$\|x\|^2 \le \sum_{s \in S} |\langle x, s \rangle|^2,$$

from Proposition 3.19; in particular, the sum converges, so Remark 3.18 reassures us that only countably many of these terms are nonzero. Anyway, the point is that Lemma 3.20 tells us that we may set

$$y := \sum_{s \in S} \langle x, s \rangle s$$

to be a convergent series. Now, by continuity (noting that this is in fact a countable sum and then using sequence continuity), we see

$$\langle y, s' \rangle = \left\langle \sum_{s \in S} \langle x, s \rangle s, s' \right\rangle = \sum_{s \in S} \langle x, s \rangle \langle s, s' \rangle = \langle x, s' \rangle.$$

Thus, completeness of S tells us that having $\langle x - y, s \rangle = 0$ for all $s \in S$ requires $x = y$. ■

Remark 3.27. Now, Remark 3.21 combined with our equality implies

$$\|x\|^2 = \sum_{s \in S} |\langle x, s \rangle|^2.$$

And now we check that these bases exist.

Theorem 3.28. Any Hilbert space X has an orthonormal basis S.

Proof. We use Zorn's lemma. Let $\mathcal{S}$ denote the collection of orthonormal subsets of X, ordered by inclusion. Certainly $\mathcal{S}$ is nonempty because it contains $\varnothing$, and unions of orthonormal subsets are orthonormal, so $\mathcal{S}$ has an upper bound for any nonempty chain. Thus, Zorn's lemma provides a maximal orthonormal subset $S \subseteq X$. To see that S is complete, for any vector $x \in X$, if $\langle x, s \rangle = 0$ for all $s \in S$, then either $x = 0$ or $S \cup \{x / \|x\|\}$ is a larger orthonormal set than S, the latter of which cannot occur. So we must instead have $x = 0$. ■

It is useful to have countable bases, so we pick up the following proposition.

Proposition 3.29. Fix a Hilbert space X. Then the following are equivalent.

(a) X is separable.

(b) X has a countable orthonormal basis.

(c) Every orthonormal basis for X is countable.

(d) X has an isometric isomorphism to $L^2(\mathbb{N})$ or $L^2([n])$ for finite n.

Proof. We won't show all these in detail, but let's show some of them. For example, (c) implies (b) with no content by Theorem 3.28. To see that (a) implies (c), suppose for the sake of contraposition that X has an uncountable orthonormal basis $S \subseteq X$. Then note that $\|s - s'\| = \sqrt{2}$ for all distinct $s, s' \in S$, so the open subsets

$$B_s := B(s, \sqrt{2}/2)$$

are all nonempty and disjoint. So any dense subset of X must be uncountable to hit each of the B_s.

Continuing, to see that (b) implies (a), one fixes a countable orthonormal basis S, and then our countable dense subset is

$$\mathbb{Q}S := \left\{ \sum_{s \in S} q_s s : \{s : q_s \neq 0\} \text{ is finite} \right\}.$$

To see that this countable set is dense, the point is that any $x \in X$ can be written as $\sum_{s \in S} x_s s$ for real numbers x_s satisfying $\sum_{s \in S} |x_s|^2 = 0$. Then for any $\varepsilon > 0$, we see $\mathbb{Q}S \cap B(x, \varepsilon)$ is nonempty by cutting off the sum $\sum_{s \in S} |x_s|^2$ and then approximating the rest of the sum arbitrarily well with rationals.

Lastly, certainly (d) implies (a) because $L^2(\mathbb{N})$ is separable. And for the other implication, note that (b) implies (d) by choosing a countable orthonormal basis and mapping the two orthonormal bases to each other. ■

Remark 3.30. For (d), note that preserving the norm implies preserving the inner product because the inner product can be recovered from the norm (see Remark 2.106).

Remark 3.31. Here's a miscellaneous remark: given a bounded linear operator $T : X \to Y$ of normed vector spaces, then there is a notion of "transpose" $T^\intercal : Y^* \to X^*$ given by $T^* f := f \circ T$.

3.3.2 Overview of Lebesgue Spaces

We now discuss some important examples of normed vector spaces. Here is our definition.

Definition 3.32. Fix a measure space $(X, \mathcal{M}, \mu)$. Then for finite $p > 0$, the associated *Lebesgue space* is $\mathcal{L}^p(X)$, defined as the space of measurable functions $f : X \to \mathbb{C}$ such that

$$\int_X |f|^p \, d\mu < \infty.$$

It is not totally clear that $\mathcal{L}^p(X)$ is a $\mathbb{C}$-vector space with semi-norm given by

$$\|f\|_p := \left(\int_X |f|^p \, d\mu \right)^{1/p}.$$

(Here, the $(\cdot)^{1/p}$ is taken to ensure homogeneity of the norm.) However, we can see that $\|f\|_p = 0$ occurs whenever f vanishes almost everywhere, so $\|\cdot\|_p$ fails to be a norm on $\mathcal{L}^p$. To fix this, we have the following.

Definition 3.33. Fix a measure space $(X, \mathcal{M}, \mu)$. Then we define the equivalence relation $\equiv$ on measurable functions $X \to \mathbb{C}$ by $f \equiv g$ if and only if $f = g$ almost everywhere. Then we define L^p as $\mathcal{L}^p / \equiv$.

Remark 3.34. Let's quickly argue that $\equiv$ is an equivalence relation. Reflexivity and symmetry have little content. Lastly, if $f \equiv g$ and $g \equiv h$, then $f - g$ and $g - h$ both vanish outside possibly different null sets, but they will both vanish outside the union of those two null sets (which is still a null set!), so their sum $f - h$ continues to vanish outside a null set, meaning $f \equiv h$.

Once $\mathcal{L}^p(X)$ is a vector space, we see that $L^p(X)$ continues to be a vector space because it is the quotient of $\mathcal{L}^p(X)$ by the subspace consisting of measurable functions $f\colon X \to \mathbb{C}$ vanishing almost everywhere. We argued earlier that $\|f\|_p$ only depends on the class in $L^p(X)$, so $\|\cdot\|_p$ goes down to a semi-norm on $L^p(X)$. We will later show this semi-norm is in fact a norm, and that $L^p(X)$ becomes a Banach space with this norm.

3.4 March 18

We continue our discussion of L^p spaces.

3.4.1 Checks for Lebesgue Spaces

We begin running our checks on L^p. The most nontrivial check is the triangle inequality.

Proposition 3.35 (Minkowski's inequality)**.** Fix a measure space $(X, \mathcal{M}, \mu)$ and $p \geq 1$. Then for all $f, g \in \mathcal{L}^p(X)$, we have

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

The proof of this will require Hölder's inequality.

Proposition 3.36 (Hölder's inequality)**.** Fix a measure space $(X, \mathcal{M}, \mu)$ and real numbers $p, q \in (1, \infty)$ such that $\frac{1}{p} + \frac{1}{q} = 1$. Given $f \in \mathcal{L}^p(X)$ and $g \in \mathcal{L}^q(X)$, then $fg \in \mathcal{L}^1(X)$ and

$$\|fg\|_1 \leq \|f\|_p \|g\|_q .$$

Remark 3.37. The q here is called the "exponent conjugate to p." Notably, we can solve

$$q := \frac{p}{p-1},$$

verifying that $p > 1$ implies $q > 1$. Notably, if $p = 2$, then $q = 2$.

Our proofs will use the following geometric fact about convexity.

Lemma 3.38. Given real numbers $u, v \in \mathbb{R}$ and some $t \in [0,1]$, we have

$$e^{tu+(1-t)v} \leq te^u + (1-t)e^v.$$

Proof. The point is that the exponential function $\exp$ has positive second derivative, so it is convex up. To be more rigorous, we write

$$f(t) := te^u + (1-t)e^v - e^{tu+(1-t)v}.$$

We would like to show that $f(t) \geq 0$ on $[0,1]$. Well, we note that $f''(t) = -(v-u)^2 e^{tu+(1-t)v}$ is always negative, and $f(0) = f(1) = 0$, so on the interval $[0,1]$ our function must be increasing and then decreasing but always above $f(0) = f(1) = 0$, so we are done. Rigorously, if there is a point $p \in [0,1]$ with $f(p) < 0$, then the Mean value theorem tells us that there are $0 < p_- < p < p_+ < 1$ such that $f'(p_-) < 0$ and $f'(p_+) > 0$, so the Mean value theorem again tells us that there is some q between p_- and p_+ such that $f''(q) > 0$, which is a contradiction! ■

We now prove Proposition 3.36.

Proof of Proposition 3.36. By replacing f and g with their absolute values (which can only make $\int_X fg\,d\mu$ larger and not change any other integrals), we may assume that $f, g \geq 0$. Additionally, if f or g vanish, there is nothing to prove, so we may assume that they are nonzero. Then by scaling f and g by their norms, we may assume that $\|f\|_p = \|g\|_p = 1$. We may get rid of any zero set of fg by restricting X (this does not change $\int_X fg\,d\mu$, and it can only make our right-hand side smaller), so we define $u := \log f$ and $v := \log g$, so

$$\int_X fg\,d\mu = \int_X e^{u+v}\,d\mu = \int_X e^{p^{-1}\cdot pu + q^{-1}\cdot qv}\,d\mu$$

which by Lemma 3.38 is bounded above by

$$\int_X \left(p^{-1}e^{pu} + q^{-1}e^{qv}\right)\,d\mu = p^{-1}\,\|f\|_p + q^{-1}\,\|g\|_q = 1,$$

as desired. ■

Remark 3.39. If $p = 1$, then the interpretation of Proposition 3.36 should take $q = \infty$, so we are asking for $f \in \mathcal{L}^1(X)$ and $g \in \mathcal{L}^\infty(X)$ implies $\|fg\|_1 \le \|f\|_1 \cdot \|g\|_\infty$. But this is immediate because one we may upper-bound

$$\int_X |fg|\,d\mu \le \|g\|_\infty \int_X |f|\,d\mu = \|g\|_\infty \cdot \|f\|_1\,.$$

At long last, we are in a position to prove Proposition 3.35.

Proof of Proposition 3.35. Quickly, note that $p = 1$ is simply the triangle inequality, so we may assume that $p > 1$. Now, we are trying to show that

$$\int_X |f+g|^p\,d\mu \le \left(\|f\|_p + \|g\|_p\right)^p.$$

By replacing f and g with their absolute values (which does not change $\|f\|_p$ and $\|g\|_p$ but can make the above integral larger), it suffices to assume that $f, g \ge 0$.

We now employ a trick. Expand

$$\int_X (f+g)^p\,d\mu = \int_X f(f+g)^{p-1}\,d\mu + \int_X g(f+g)^{p-1}\,d\mu.$$

We would now like to use Hölder's inequality. Well, setting $q := \frac{p}{p-1}$ so that $\frac{1}{p} + \frac{1}{q} = 1$, we would like to know that $(f+g)^{p-1}$ lives in $\mathcal{L}^q(X)$. Well, we compute

$$\int_X (f+g)^{(p-1)q}\,d\mu = \int_X (f+g)^p\,d\mu = \|f+g\|_p^p < \infty.$$

As such, we may use Proposition 3.36 so that

$$\|f+g\|_p^p \le \|f\|_p \left\|(f+g)^{p-1}\right\|_q + \|g\|_p \left\|(f+g)^{p-1}\right\|_q,$$

which the previous computation tells us is

$$\|f+g\|_p^p \le (\|f\|_p + \|g\|_p)\,\|f+g\|_p^{p-1}\,.$$

If $\|f+g\|_p = 0$, then $f+g$ vanishes almost everywhere because $f, g \ge 0$ already, so there is nothing to prove. Otherwise, we may now cancel $\|f+g\|_p^{p-1}$ on both sides to conclude. ■

Remark 3.40. The nonnegative function $\|\cdot\|_p$ on $\mathcal{L}^p(X)$ now satisfies the triangle inequality. It certainly satisfies $\|tf\|_p = t\,\|f\|_p$ by a direct expansion, so $\|\cdot\|_p$ becomes a semi-norm on $\mathcal{L}^p(X)$, and it immediately descends to a semi-norm on $\mathcal{L}^p(X)$. In fact, we descend to a full norm on $L^p(X)$: if $\|f\|_p = 0$, then $|f|$ must vanish almost everywhere, so $f = 0$ in $L^p(X)$.

And we now check that $L^p(X)$ is a Banach space.

Theorem 3.41. Fix a measure space $(X, \mathcal{M}, \mu)$ and $p \geq 1$. Then $L^p(X)$ is a Banach space.

Proof. By Remark 3.40, it remains to show that $L^p(X)$ is complete. Turning the Cauchy condition into a series, it suffices to check that any series of functions $\sum_{n=1}^\infty f_n$ converges (for the norm $\|\cdot\|_p$) provided that

$$\sum_{n=1}^{\infty} \|f_n\|_p < \infty.$$

Of course, we would like to take the function $f := \sum_{n=1}^\infty f_n$, but to even define this function, we must know that f converges pointwise almost everywhere. Well, define

$$h_N(x) := \sum_{n=1}^{N} |f_n|(x)$$

so that $\|h_N\|_p \leq \sum_{n=1}^N \|f_n\|_p$ for all $N > 0$ by the triangle inequality. Thus, we see that $h_1 \leq h_2 \leq \cdots$ is a series of functions defined almost everywhere, and their values are bounded above for each $x \in X$, so Monotone convergence provides a limiting function h defined almost everywhere. Notably, Monotone convergence tells us that

$$\int_X h^p \, d\mu = \lim_{N\to\infty} \int_X h_N^p \, d\mu < \infty.$$

So we see that the series $f := \sum_{n=1}^\infty f_n$ absolutely converges almost everywhere (indeed, it absolutely converges to h), so f is actually defined as a function almost everywhere. Further, $|f| \leq |h|$, so Fatou's lemma [Elb22, Lemma 9.39] tells us that

$$\|f\|_p^p = \int_X |f|^p \, d\mu \leq \int_X \lim_{N\to\infty} \left| \sum_{n=1}^N f_n \right|^p \, d\mu \leq \liminf_{N\to\infty} \int_X \left| \sum_{n=1}^N f_n \right|^p \, d\mu = \liminf_{N\to\infty} \|h_N\|^p \leq \|h\|_p^p < \infty,$$

so $f \in L^p(X)$. (The second to last inequality uses the Monotone convergence theorem.)

We now complete the proof by showing that the series $\sum_{n=1}^\infty f_n$ converges to f with respect to the norm $\|\cdot\|_p$. We will use the Dominated convergence theorem [Elb22, Theorem 9.14]. Note that both $\left|\sum_{n=1}^N f_n\right|$ and $|f|$ are dominated by h, so their difference is dominated by $2h$, where $\|h\|_p^p < \infty$ can actually behave as a dominating function, so $\sum_{n=1}^N f_n \to f$ in norm. ■

3.5 March 20

We began class by completing the proof of Theorem 3.41.

Remark 3.42. It is possible to have a sequence of functions $\{f_n\}_{n\in\mathbb{N}}$ in $L^p(X)$ which go to 0 pointwise but do not go to 0 in $L^p(X)$. For example, set $f_n := 1_{[n,n+1]}$, so $\|f_n\|_p = 1$, so $f_n \in L^p(X)$. But of course $\|f_n - 0\|_p = 1$ does not go to 0 as $n \to \infty$.

Example 3.43. We show that the triangle inequality fails in $L^p(X)$ for $0 < p < 1$. Indeed, simply set $f := 1_{[0,1]}$ and $g := 1_{[3,4]}$, so $\|f+g\|_p = 2^{1/p}$ while $\|f\|_p = \|g\|_p = 1$.

Remark 3.44. For $0 < p < 1$, one does have

$$\int_X |f+g|^p \, d\mu \leq \int_X |f|^p \, d\mu + \int_X |g|^p \, d\mu,$$

so $L^p(X)$ does become a metric on a vector space via $d(f,g) := \int_X |f-g|^p \, d\mu$, though it will not be a normed vector space.

Remark 3.45. The notion of convergence is a bit tricky. Given a sequence of functions $\{f_n\}_{n\in\mathbb{N}}$ in $L^p(X)$ on a measure space $(X,\mathcal{M},\mu)$, we have the following notions of convergence to a function f.

- Perhaps $f_n \to f$ pointwise for all $x \in X$ or pointwise almost everywhere.
- Perhaps $f_n \to f$ in norm for $\|\cdot\|_p$.
- Perhaps $f_n \to f$ uniformly or uniformly almost everywhere.
- Perhaps $f_n \to f$ almost uniformly, meaning that any $\varepsilon > 0$ has some $E \subseteq X$ such that $\mu(X \setminus E) < \varepsilon$ and $f_n \to f$ uniformly on E.
- Perhaps $f_n \to f$ in measure, meaning that any $\varepsilon > 0$ has $\mu(\{x \in X : |f_n(x) - f(x)| > \varepsilon\}) \to 0$ as $n \to \infty$. See [Elb22, Example 8.11] for a nontrivial example.

Recall Egorov's theorem [Elb22, Theorem 9.13] tells us that $f_n \to f$ in measure in a finite measure space implies that any $\varepsilon > 0$ has some $E \subseteq X$ such that $\mu(E \setminus X) < \varepsilon$ and $f_n \to f$ almost uniformly on $E \setminus X$. Note then that almost uniform convergence provides a limiting function for a subsequence.

3.5.1 The Dual Space of $L^p(X)$

Here is our statement.

Theorem 3.46. Fix $p \in [1,\infty)$, and let $q := \frac{p}{p-1}$ be the conjugate exponent. Further, fix a σ-finite measure space $(X,\mathcal{M},\mu)$.

(a) If $p > 1$, then $L^p(X)^*$ is naturally isomorphic to $L^q(X)$ as Banach spaces.

(b) If $p = 1$, then $L^1(X)^*$ is naturally isomorphic to the space of equivalence classes of essentially bounded functions $f \colon X \to \mathbb{C}$.

Wait, what does essentially bounded mean?

Definition 3.47 (essentially bounded)**.** Fix a measure space $(X,\mathcal{M},\mu)$. A measurable function $f \colon X \to \mathbb{C}$ is *essentially bounded* if and only if there is some $M > 0$ such that

$$\mu(\{x : |f(x)| > M\}) = 0.$$

We denote the vector space of such functions by $\mathcal{L}^\infty(X)$ and their vector space of equivalence classes by $L^\infty(X)$.

Remark 3.48. We note that $\mathcal{L}^\infty(X)$ has a semi-norm given by

$$\|f\|_\infty := \inf\{A \ge 0 : \mu(\{x \in X : |f(x)| > A\}) = 0\}.$$

Quickly, note $\mu(\{x \in X : |f(x)| > \|f\|_\infty\}) = 0$ by some countable union argument of As approaching $\|f\|_\infty$. Now, we omit most of the checks that this is actually a semi-norm. As an example check, we note that $\|f+g\|_\infty \le \|f\|_\infty + \|g\|_\infty$ because

$$\begin{aligned}\mu(\{x \in X : |f(x)+g(x)| > \|f\|_\infty + \|g\|_\infty\}) &\le \mu(\{x \in X : |f(x)| > \|f\|_\infty\}) \\ &\quad + \mu(\{x \in X : |g(x)| > \|f\|\, g\infty\}).\end{aligned}$$

Remark 3.49. We also remark that $L^\infty(X)$ is complete, thus making $L^\infty(X)$ into a Banach space. It suffices to show that an infinite series $\sum_n f_n$ converges when $\sum_n \|f_n\|_\infty < \infty$. By adjusting equivalence classes, we may assume that the $f_\bullet$s are bounded by $\|f_n\|_\infty$. Then the infinite series is absolutely convergent and in fact uniformly so, so we get to define f by $f(x) := \sum_n f_n(x)$, and we get our uniform convergence for $\|\cdot\|_\infty$ by writing out the convergence.

We now prove Theorem 3.46.

Proof of Theorem 3.46. We begin with the proof of (a). Let's start by describing the map $T_\bullet \colon L^q(X) \to L^p(X)^*$. Well, for $g \in L^q(X)$, define $T_g \in L^p(X)^*$ by

$$T_g(f) := \int_X fg\,d\mu.$$

Note that $fg \in L^1(X)$ by Proposition 3.36, so $T_g(f)$ is a well-defined quantity (and note that it does not depend on the choice of equivalence class for either f or g). We also take a moment to note that T_g is linear in f because multiplication is linear, and in fact $T_\bullet \colon L^q(X) \to L^p(X)^*$ is again linear because multiplication is linear.

To continue we claim that $T_\bullet$ preserves norms. For one inequality, note Proposition 3.36 tells us $|T_g(f)| \le \|f\|_p \cdot \|g\|_q$, so

$$\|T_g\|_{L^p(X)^*} \le \|g\|_q\,.$$

For the equality, we need to choose some functions f making $T_g(f)$ large. Well, for each $n \in \mathbb{N}$, set $E_n := \{x : 1/n \le |g(x)| \le n\}$, and define $f_n := 1_{E_n} \cdot \overline{g}\,|g|^{q-2}$. To check that $f_n \in L^p(X)$, we compute

$$\|f_n\|_p = \left(\int_{E_n} |g|^q\,d\mu\right)^{(q-1)/q} < \infty.$$

Further, we see

$$T_g(f) = \int_{E_n} |g|^q\,d\mu,$$

so

$$\limsup_{n\to\infty} \frac{T_g(f_n)}{\|f_n\|_p} = \limsup_{n\to\infty} \left(\int_{E_n} |g|^q\right)^{1/q} = \|g\|_q\,, \tag{3.1}$$

providing our lower bound on $\|T_g\|_{L^p(X)^*}$. Note we did not actually need E_n in the above argument, but we will reuse this calculation later when we don't know that $g \in L^q(X)$ a priori.

Taking stock, currently we know that $T_\bullet \colon L^q(X) \to L^p(X)^*$ is an isometric embedding of Banach spaces. We also remark on the side that this map is natural in X, but we will not show it. It remains to check that $T_\bullet$ is surjective.

We begin by using σ-finiteness to reduce the surjectivity check to the case where X is a finite measure space. Write $X = \bigsqcup_{n=1}^\infty$ where the X_n are finite measure spaces. Now, if we are granted the finite measure space situation, the function $T|_{L^p(X_n)}$ gives us some $g_n \in L^q(X_n)$ such that

$$T(f) = \int_X g_n f\,d\mu.$$

Now, a direct argument as in (3.1) allows us to check that $\sum_{n=1}^\infty g_n$ is in $L^q(X)$, and then we can check that $T = T_g$ holds on simple functions, so it will hold everywhere because simple functions are dense in $L^q(X)$.

Thus, for the rest of the proof, we may assume that X is a finite measure space. We proceed in steps.

1. We would like to realize T as an integral, so define a measure λ on X by

$$\lambda(E) := T(1_E).$$

Note that this is well-defined for measurable sets E: X is finite, so E is finite, so 1_E is in $L^p(X)$, so $T(1_E)$ makes sense. We claim that λ is a measure, for which we need to check that λ is countably additive. Well, write $E = \bigsqcup_{n\in\mathbb{N}} E_n$ as a disjoint union of countable measurable sets, and we want

$$T(1_E) = \lambda(E) \stackrel{?}{=} \sum_{n=1}^{\infty} \lambda(E_n) = \sum_{n=1}^{\infty} T(1_{E_n}) \stackrel{*}{=} T\Bigg(\sum_{n=1}^{\infty} 1_{E_n}\Bigg).$$

Here, $\stackrel{*}{=}$ will hold because continuity of T allows it to move outside the summation. In particular, we would like to show that the partial sums $\sum_{n=1}^{N} 1_{E_n}$ approaches 1_E as $N \to \infty$, but we need this convergence to hold in $L^p(X)$! In other words, we want to look at

$$\left\| 1_E - \sum_{n=1}^{N} 1_{E_n} \right\|_p^p = \int_X 1_{E \setminus \bigsqcup_{n=1}^{N} E_n}\, d\mu = \mu\Bigg(\bigsqcup_{n>N} E_n\Bigg).$$

Because we are in a finite measure space, we know that this right-hand side vanishes as $N \to \infty$.

2. We now apply the Radon–Nikodym theorem to write λ as an integral. For this, we want to check that $\lambda \ll \mu$. Well, this means that we want to check that $\mu(E) = 0$ implies that $T(1_E) = \lambda(E) = 0$. The point is that $\mu(1_E) = 0$ implies that $\|1_E\|_p = 0$, but then $T(1_E) = 0$ because $1_E = 0$ in $L^p(X)$! Thus, the Radon–Nikodym theorem provides $g \in L^1(X)$ such that

$$\lambda(E) = \int_E g\, d\mu$$

for any measurable E.

It remains to check that $T = T_g$. We do this in steps.

3. For simple functions $f = \sum_{i=1}^{n} c_i 1_{E_i}$ where the $E_\bullet$ are disjoint measurable functions, so

$$T(f) = \sum_{i=1}^{n} c_i \lambda(E_i) = \int_X \Bigg(\sum_{i=1}^{n} c_i 1_{E_i}\Bigg) g\, d\mu = \int_X fg\, d\mu.$$

So we get $T(f) = T_g(f)$.

4. For bounded functions f, we know that f is the uniform limit of simple functions $\{f_n\}_{n\in\mathbb{N}}$: indeed by taking linear combinations, we may assume that f is real-valued, and then we partition the image of f (which is a bounded set) into very small intervals of length $1/N$, taking the preimage of these intervals to build our $E_i := f^{-1}([i/N, (i+1)/N))$. In particular, we may assume that $|f(x) - f_n(x)| < \frac{1}{n}$ for all x. But then we see that $f_n \to f$ even in $L^p(X)$ (note $p \geq 1$), so $T(f_n) \to T(f)$, and we even have $T_g(f_n) \to T_g(f)$ by the $1/n$ bound and finiteness of our measure space. So we complete by passing to the limit.

5. We show that $g \in L^q(X)$ for $p > 1$. For each n, define

$$G_n(x) := \begin{cases} \overline{g(x)}\,|g(x)|^{q-2} & \text{if } |g(x)| \leq n, \\ 0 & \text{otherwise.} \end{cases}$$

Now, the function $G_n(x)$ is always bounded, so we get $T(G_n) = T_g(G_n)$ by the previous step. The point is that an argument identical to (3.1) shows that

$$\|g\|_q = \limsup_{n\to\infty} \frac{\left|\int_X G_n\, d\mu\right|}{\|G_n\|_p} = \limsup_{n\to\infty} \frac{|T(G_n)|}{\|G_n\|_p} \leq \|T\|_{L^p(X)^*}$$

is finite.

6. We show that $g \in L^\infty(X)$ for $p = 1$. If f is bounded, we note that we will have
$$\left|\int_X fg\,d\mu\right| \le \|T\|_{L^1(X)^*} \|f\|_1$$
by bounding g by its maximum. Now, for each measurable set E, we set
$$f(x) := \begin{cases} 1_E(x) \cdot \overline{g}(x)/g(x) & \text{if } g(x) \ne 0, \\ 0 & \text{if } g(x) = 0, \end{cases}$$
whereupon we see that
$$\int_E |g|\,d\mu = \left|\int_E |g|\,d\mu\right| = \left|\int_X fg\,d\mu\right| = |T(f)| \le \|T\|_{L^1(X)^*} \|f\|_1 \le T_{L^1(X)^*}\mu(E).$$
Now, this implies that $|g| \le T_{L^1(X)^*}$ almost everywhere: if we violated this for strict equality on a set of positive measure E, we could plug in E above for contradiction. So $g \in L^\infty(X)$ follows.
7. Now, we know that $g \in L^q(X)$, so T_g fully makes sense. We have $T = T_g$ on simple functions by a previous step, and simple functions are dense in $L^p(X)$, so we get the equality everywhere by continuity. ■

Remark 3.50. The result of (a) is true even without σ-finite hypotheses, though (b) does require this.

Remark 3.51. For $p \in [1, \infty)$, one finds that $L^p\left(\mathbb{R}^d\right)$ is a separable space. Indeed, $L^p\left(\mathbb{R}^d\right)$ contains as a dense subset the $\mathbb{R}$-linear combinations of indicators of measurable sets. By approximating measurable sets close enough, we may assume that we are looking at $\mathbb{R}$-linear combinations of boxes. But then we may assume that the coefficients and endpoints of the boxes are all rational by another approximation argument, so we are done.

3.6 March 22

We spent class by completing the proof of Theorem 3.46.

3.6.1 Chebyshev's Inequality

We close class by proving an inequality. Here is our result.

Proposition 3.52 (Chebyshev's inequality)**.** Fix a measure space $(X, \mathcal{M}, \mu)$ and some real number $p \ge 1$. Then for any $f \in L^p(X)$ and $t > 0$, we have
$$\mu(\{x \in X : |f(x)| > t\}) \le t^{-p} \|f\|_p^p.$$

Proof. We would like to show that
$$\int_X |f|^p\,d\mu = \|f\|_p^p \stackrel{?}{\ge} t^p \mu(\{x \in X : |f(x)| > t\}).$$
Well, one simply notes that the integral on the left can be restricted to
$$E_t := X \setminus \{x \in X : |f(x)| \le t\},$$
and then bound $|f| \ge t$ on E_t (by definition), which is precisely the right-hand side. ■

Remark 3.53. One might complain that $\mu(\{x \in X : |f(x)| > t\})$ is a much simpler object that $\|f\|_p^p$. We work with $\|f\|_p$ instead of objects like this because $\|\cdot\|_p$ has a triangle inequality.

Theme 4
Riesz Representation

4.1 April 1

We began class with some remarks on $L^\infty(X)$, which I have just edited into last class's notes.

4.1.1 The Moral

Let X be a compact Hausdorff topological space X. We refer to [Elb22] and in particular subsection 4.2.1 for some facts about compact Hausdorff spaces.

Notation 4.1. Fix a compact Hausdorff topological space X. Then $C(X)$ is the $\mathbb{C}$-vector space of continuous functions $X \to \mathbb{C}$.

Remark 4.2. We note $C(X)$ can be upgraded to a normed vector space with the norm

$$\|f\|_{C(X)} := \sup_{x \in X} |f(x)|.$$

We will omit the checks that $\|\cdot\|_{C(X)}$ is actually a norm, though we note that it is basically $\|\cdot\|_\infty$. Because $C(X)$ can be checked to be complete with respect this norm (convergence for $\|\cdot\|_{C(X)}$ means uniform convergence, which preserves continuity), we see that $C(X)$ is a Banach space.

The Riesz representation theorem is basically the statement that $C(X)^*$ is exactly the space of complex Borel measures on X. (Notably, complex Borel measures on X form a $\mathbb{C}$-vector space.) This claim is not literally true, but we will be able to make it true enough.

4.1.2 Locally Compact Spaces

It will be helpful to have some understanding of locally compact spaces. We borrow the following exposition from [Elb22, subsection 4.8.1].

Definition 4.3 (Locally compact). A topological space $(X, \mathcal{T})$ is *locally compact* if and only if each point $x \in X$ has some open subset $U \in \mathcal{T}$ containing x such that $\overline{U}$ is compact.

Example 4.4. The set of real numbers $\mathbb{R}$ with the usual topology is locally compact. Indeed, any $x \in \mathbb{R}$ has the open neighborhood $(x-1, x+1)$ with closure $[x-1, x+1]$, and $[x-1, x+1]$ is compact.

Example 4.5. For the same reason, the space $[a, b)$ is also locally compact.

Example 4.6. Finite-dimensional Euclidean spaces V are locally compact spaces; for example, they can be embedded into the compact projective space $\mathbb{P}V$. However, infinite-dimensional Banach spaces fail to be locally compact. Indeed, up to shrinking, this would imply that sufficiently small open balls are compact, which is false.

Here are some helpful facts.

Lemma 4.7. Fix a locally compact Hausdorff space $(X, \mathcal{T})$. Then any $x \in X$ and open subset $U \in \mathcal{T}$ containing x has some open subset $U_x \subseteq X$ containing x such that $\overline{U_x}$ is compact and $\overline{U_x} \subseteq U$.

Proof. We begin by finding our promised U' containing x with $\overline{U'}$ compact. Thus, it suffices to find some open subset V containing x such that $\overline{V}$ is compact and $\overline{V} \subseteq U \cap U'$, but now we see that

$$\overline{U \cap U'} \subseteq \overline{U'}$$

is a closed subset of the compact space $\overline{U'}$ and therefore compact. In particular, we can replace U with $U \cap U'$ and assume that $\overline{U}$ is compact.

Now, let $\partial U := \overline{U} \setminus U$ be the boundary of U. Notably, ∂U is a closed subset of the compact space $\overline{U}$, so ∂U is compact. Because $\{x\}$ is a closed subset in U (note $X \setminus \{x\}$ is open, so $U \setminus \{x\}$ is open in the relative topology), the fact that compact Hausdorff spaces are normal grants open subsets U_x and U_∂ of $\overline{U}$ with $x \in U_x$ and $\partial U \subseteq U_\partial$.

Now, $U_x \subseteq \overline{U} \setminus U_\partial \subseteq \overline{U} \setminus \partial U$, so we see $\overline{U_x} \subseteq \overline{U} \setminus U_\partial$ because $\overline{U} \setminus U_\partial$ is a closed subset of $\overline{U}$. Further, $\overline{U_x}$ is a closed subset of a compact space $\overline{U}$, so $\overline{U_x}$ is compact, so we are done. ■

Remark 4.8. Lemma 4.7 basically says that open subspaces of locally compact Hausdorff spaces are locally compact.

We can extend the previous result past points to full compact sets.

Lemma 4.9. Fix a locally compact Hausdorff space $(X, \mathcal{T})$ and some compact subset $C \subseteq X$. Then any open subset U containing C has some open subset U_C containing C such that $\overline{U_C}$ is compact and $\overline{U_C} \subseteq U$.

Proof. We use Lemma 4.7. For each $x \in C$, find some U_x by Lemma 4.7 with U_x containing x with $\overline{U_x}$ compact and $\overline{U_x} \subseteq U$. Then we see that

$$C \subseteq \bigcup_{x \in C} U_x,$$

so we have provided an open cover for C, so we can choose finitely many $\{x_i\}_{i=1}^n \subseteq C$ with $U_i := U_{x_i}$ so that

$$C \subseteq \bigcup_{i=1}^n U_i \subseteq U_C.$$

Now, we see that

$$\overline{\bigcup_{i=1}^n U_i} = \bigcup_{i=1}^n \overline{U_i}$$

is a compact subset of U because being compact is closed under finite unions, so $\bigcup_{i=1}^n U_i$ is the required open subset. ■

4.2 April 3

Today we continue talking about locally compact Hausdorff spaces.

4.2.1 More on Locally Compact Hausdorff Spaces

Here is another lemma.

Lemma 4.10. Fix a locally compact Hausdorff space X. Let K be a compact subset contained in an open subset $U \subseteq X$. Then there is a real-valued continuous function $f\colon X \to [0,1]$ with compact support in U such that $1_K \le f$.

It will be helpful to have the following notation.

Definition 4.11. Fix a continuous function $f\colon X \to [0,1]$. Given an open subset $U \subseteq X$, we say that $f < U$ if and only if f has compact support contained in U.

Non-Example 4.12. Consider the function $f\colon \mathbb{R} \to [0,1]$ defined by $f(x) := \max\{0, 1-|x|\}$. Then $f^{-1}(\{x : x > 0\}) = (-1,1)$, so even though f has compact support, this compact support is not contained in $(-1,1)$, so $f \not< (-1,1)$.

We will use Lemma 4.9 to prove Lemma 4.10. One would like to use Urysohn's lemma directly, but this does not quite go through immediately because locally compact Hausdorff spaces need not be normal. To make sense of the previous sentence, we recall the following notions.

Definition 4.13 (normal)**.** A topological space X is *normal* if and only if disjoint closed subsets E and F have disjoint open neighborhoods U and V, respectively.

Example 4.14. A compact Hausdorff topological space X is normal by [Elb22, Proposition 4.18].

Theorem 4.15 (Urysohn's lemma)**.** Fix a normal topological space X. Given disjoint closed subsets $E, F \subseteq X$, there is a continuous function $f\colon X \to [0,1]$ such that $f|_E = 0$ and $f|_F = 1$.

Proof. See [Elb22, Theorem 3.8]. ■

We now prove Lemma 4.10.

Proof of Lemma 4.10. Lemma 4.9 provides an open subset $V \subseteq X$ such that $K \subseteq V \subseteq \overline{V} \subseteq U$ where $\overline{V}$ is compact. Now, $\overline{V}$ is a compact Hausdorff space, which is normal by Example 4.14, so Theorem 4.15 provides a continuous function $g\colon \overline{V} \to [0,1]$ such that $g|_K = 1$ and $g|_{\partial V} = 0$. This is not quite we want because we need a continuous function on all of X, so we rudely extend g to $\widetilde{g}\colon X \to [0,1]$ by

$$\widetilde{g}(x) := \begin{cases} g(x) & \text{if } x \in \overline{V}, \\ 0 & \text{if } x \notin \overline{V}. \end{cases}$$

Notably, we see that $1_K \le \widetilde{g}$ because $1_K \le g$, and the nonzero points of $\widetilde{g}$ are contained in $\overline{V}$, so the support of g is contained in $\overline{V} \subseteq U$. Thus, it only remains to show that $\widetilde{g}$ is continuous. We show continuity on the points of X by hand. For example, $\widetilde{g}$ is continuous on V because g is continuous on V. Further, $\widetilde{g}$ is continuous outside $\overline{V}$ because $\widetilde{g}$ vanishes there.

It remains to handle continuity at $x_0 \in \overline{V} \setminus V$. Well, we know that $\widetilde{g}(x_0) = g(x_0) = 0$. Fix some $\varepsilon > 0$, and we need an open neighborhood V' of x_0 such that $|\widetilde{g}(x)| < \varepsilon$ for $x \in V'$. Well, continuity of g provides some open neighborhood V' of x_0 such that $|g(x)| < \varepsilon$ for $x \in V' \cap \overline{V}$. But then we still have $|g(x)| = 0 < \varepsilon$ for $x \in V' \setminus \overline{V}$, so we are done. ■

We are now able to prove a version of partition of unity.

Theorem 4.16 (Partition of unity)**.** Fix a locally compact Hausdorff space X and a compact subspace K. Given a finite open cover $\mathcal{U}$ of K, there are continuous functions $X \to [0,1]$ denoted $\{f_U\}_{U\in\mathcal{U}}$ such that $f_U < U$ for each U and

$$\sum_{U\in\mathcal{U}} f_U|_K = 1.$$

Proof. We begin by claiming that we can find an open cover $\{V_U : U \in \mathcal{U}\}$ of K such that $V_U \subseteq U$ and $\overline{V}_U$ is compact for each $U \in \mathcal{U}$. To begin, each $x \in K$ is contained in some $U_x \in \mathcal{U}$, and then Lemma 4.7 provides an open neighborhood V_x with compact closure contained in U_x. Then we reduce this to a finite subcover $\{V_{x_i}\}_{i=1}^n$, and then

$$V_U \coloneqq \bigcup_{\substack{i=1\\ U_{x_i}=U}}^{n} V_{x_i}$$

makes $\{V_U : U \in \mathcal{U}\}$ an open cover of K satisfying the required hypotheses. (In particular, the closure is compact and contained in U because it is a subset of a finite union of compact sets contained in U.)

Now, we set $K_U \coloneqq \overline{V}_U$, which we know is compact and contained in V_U, and we see that Lemma 4.10 provides a continuous function $g_U \colon X \to [0,1]$ such that $g_U|_{K_U} = 1$ and has compact support contained in U for each U. To complete the proof, we need to normalize these functions. One try would be to divide out by the sum of the g_Us, but this does not quite work because this division has no reason to be continuous.

Well, we will simply smooth out the aforementioned normalization. Indeed, let δ be half the minimum of the $\sum_U g_U$ on K, and let V be the open subset of X on which $\sum_U g_U$ exceeds δ. In particular, $K \subseteq V$, so Lemma 4.10 provides a continuous $h \colon X \to [0,1]$ which is 1 on K and $h < V$. So we define

$$f_U(x) \coloneqq \begin{cases} h(x)g_U(x)/\left(\sum_{U'\in\mathcal{U}} g_{U'}(x)\right) & \text{if } x \in V, \\ 0 & \text{if } x \notin V. \end{cases}$$

These functions do in fact sum to 1 on K because h does not matter there. It remains to check continuity. If $x \in V$, then we see that the denominator is never zero, so this definition is in fact well-defined. Notably, this is continuous on V by definition, and it is continuous wherever h vanishes because then f_U is simply 0. So f_U is continuous on the open cover $\{V, X \setminus h^{-1}(\mathbb{R} \setminus \{0\})\}$ of X, so f_U is continuous. ■

4.3 April 5

We began class by proving Theorem 4.16.

4.3.1 Some More Spaces of Functions

Here are some more spaces of functions.

Definition 4.17. Fix a topological space X.

- $C(X)$ is the space of continuous functions $f \colon X \to \mathbb{C}$.
- $C_c(X)$ is the space of compactly supported continuous functions $f \colon X \to \mathbb{C}$.
- $C_0(X)$ is approximately the space of continuous functions f such that $f(x) \to 0$ as $x \to \infty$. Rigorously, we see that $f \in C_0(X)$ if and only if any $\varepsilon > 0$ has some compact set $K_\varepsilon \subseteq X$ containing

$$\{x \in X : |f(x)| \geq \varepsilon\}.$$

Remark 4.18. Note $C_c(X) \subseteq C_0(X) \subseteq C(X)$. The last inclusion has no content, and the first inclusion holds because $f \in C_c(X)$ has a single compact set K containing its support so that any $\varepsilon > 0$ has $|f|^{-1}((\varepsilon, \infty))$.

Remark 4.19. One might wonder why we are examining $C_0(X)$ instead of the space of bounded continuous function $X \to \mathbb{C}$. The reason is that the dual space of $C_0(X)$ is much better-behaved, as we will soon see.

These definitions allow us to make sense of a uniform norm.

Definition 4.20 (uniform norm)**.** Fix a topological space X. For $f \in C_0(X)$, we define

$$\|f\|_u := \sup_{x \in X} |f(x)| .$$

Remark 4.21. Let's check that $\|\cdot\|_u$ is finite on $C_0(X)$. Well, for $f \in C_0(X)$, we are granted a compact set K such that $|f(x)| < 1$ outside K. But then $\|f\|_u$ is bounded above by 1 outside K and has a maximum on K, so $\|f\|_u$ in total is bounded.

Remark 4.22. It is true that $\|\cdot\|_u$ becomes a norm on X, which we will not check too closely. Homogeneity has little content, and positive-definiteness holds because $\|f\|_u = 0$ would require that all values vanish. Lastly, the triangle inequality.

Example 4.23. Take $X := \mathbb{N}$. Recall $C_0(\mathbb{N})$ is the set of sequences $\{a_n\}_{n \in \mathbb{N}}$ which converge to 0.

Here are some fun checks.

Proposition 4.24. For any topological space X, the space $C_0(X)$ is a Banach space.

Proof. In light of Remark 4.22, we just have to check that $C_0(X)$ is complete. Well, suppose that $\{f_n\}_{n \in \mathbb{N}}$ is Cauchy for $\|\cdot\|_u$. So we are uniformly Cauchy, so our sequence converges to a continuous function $f \colon X \to \mathbb{C}$; in fact, this convergence is uniform.

It remains to check that $f \in C_0(X)$. Fix some $\varepsilon > 0$, and we need a compact set K such that $|f(x)| < \varepsilon$ outside K. Well, convergence in $\|\cdot\|_u$ grants some $n \in \mathbb{N}$ such that $\|f - f_n\|_u < \varepsilon/2$, and then $f_n \in C_0(X)$ promises some K such that $|f_n(x)| < \varepsilon$ outside K. Combining, we see that $|f(x)| < \varepsilon$ outside K, as needed. ■

Proposition 4.25. For any locally compact Hausdorff space X, the space $C_c(X)$ is dense in $C_0(X)$.

Proof. We use Lemma 4.10. Fix $f \in C_0(X)$ and $\varepsilon > 0$, and we need $g \in C_c(X)$ such that $\|f - g\|_u < \varepsilon$.

The point is to truncate f where it is small. Because $f \in C_0(X)$, there is a compact set $K \subseteq X$ such that $|f(x)| < \varepsilon$ outside K. Then Lemma 4.10 provides a continuous function $h \colon X \to [0,1]$ such that $h|_K = 1$ and h is compactly supported. So we set $g := fh$, and we find that $|f(x) - g(x)| = 0$ on K and $|f(x) - g(x)| = |f(x)| \cdot |1 - h(x)|$ is bounded by ε outside K. So g works. ■

Remark 4.26. In particular, if X is locally compact Hausdorff, then we see that $C_c(X)$ fails to be complete: its closure in $C(X)$ contains $C_0(X)$!

4.3.2 Functionals and Measures

In light of Remark 4.19, we will soon want to understand our dual space. We pick up the following definition.

Definition 4.27 (nonnegative)**.** Fix a topological space X. A linear functional $I\colon C_0(X) \to \mathbb{C}$ is *nonnegative* if and only if $I(f) \geq 0$ whenever $f\colon X \to [0, \infty)$.

We will learn that nonnegative bounded linear functionals give rise to nonnegative measures; loosening the nonnegative hypothesis gives rise to complex measures.

Example 4.28. Fix a nonnegative Borel measure μ on a topological space X, and suppose that μ is finite on compact sets. Then the map $C_c(X) \to \mathbb{R}$ given by

$$f \mapsto \int_X f \, d\mu$$

is a nonnegative functional on X. Notably, nonnegativity follows by the nonnegativity of μ.

Being finite on compact sets is something of an annoying condition, though it is true for all measures that come up in nature. So we pick up the following adjective to promise this.

Definition 4.29 (outer regular, inner regular, Radon)**.** Fix a Borel measure μ on a topological space X.

- μ is *outer regular* on $E \subseteq X$ if and only if $\mu(E) = \inf\{\mu(U) : U \supseteq E\}$.
- μ is *inner regular* on $E \subseteq X$ if and only if $\mu(E) = \sup\{\mu(K) : K \subseteq E\}$.
- μ is *Radon* if and only if μ is finite on compact sets, outer regular on all sets, and inner regular on open sets.

Bibliography

[Fol99] Gerald B. Folland. *Real analysis*. Second. Pure and Applied Mathematics (New York). Modern techniques and their applications, A Wiley-Interscience Publication. John Wiley & Sons, Inc., New York, 1999, pp. xvi+386. ISBN: 0-471-31716-0.

[Shu16] Neal Shusterman. *Scythe*. Arc of a Scythe. Simon & Schuster, 2016.

[Elb22] Nir Elber. *Introduction to Topology and Analysis*. 2022. URL: https://dfoiler.github.io/notes/202A/notes.pdf.

List of Definitions