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

$$\mathcal{M} \coloneqq \bigcap_{\mathcal{M}' \in \underline{\mathcal{M}}} \mathcal{M}'.$$

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

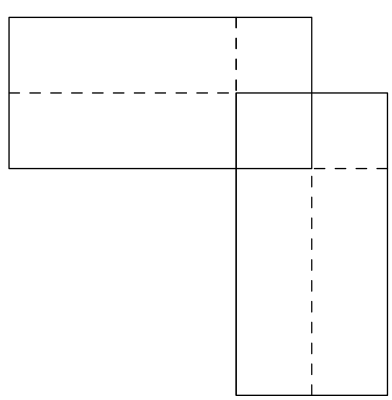
$$A_i = \bigsqcup_{\substack{\alpha \in \kappa \\ C_\alpha \subseteq A_i}} A_i.$$

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

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

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

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

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

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

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 $\frac{1}{2}\left\|f\right\|_{\text{Lip}}\ell_n\sqrt{d}$, which is contained in a cube of side length $\left\|f\right\|_{\text{Lip}}\ell_n\sqrt{d}$, which has measure $\left(\left\|f\right\|_{\text{Lip}}\ell_n\sqrt{d}\right)^d$. In total, we see

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

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

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

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

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

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 $\varepsilon_\delta \coloneqq C\sqrt{d} R(\delta \sqrt{d})$. 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

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

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

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

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

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

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

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

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

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

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

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

So it remains to bound

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

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

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

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

$$\mu(\Phi(E_n)) \leq \int_{E_n} |J_\Phi| \; d\mu \leq \int_E |J_\Phi| \; d\mu$$

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

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

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 $\leq$ 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*.

Remark 2.2. 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 2.3 (Banach space)**.** A normed vector space $(V, \|\cdot\|)$ is a *Banach space* if and only if V is complete as a metric space.

Remark 2.4. 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 2.5 (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 2.6. 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 2.7. 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 2.8. 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 2.9 (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 2.10 (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 2.11. 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 2.12. 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

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

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

But some parts of topology still make sense.

Remark 2.13. 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 2.14 (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

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

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 2.15. 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 2.16. 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 2.17 (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 2.18. 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 2.19. 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 2.20. 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

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

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 2.20.

Corollary 2.21. 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 2.20, so φ upgrades to a homeomorphism. But then completeness of X follows from completeness of $\mathbb{F}^n$. ■

Corollary 2.22. 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 2.21. 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 2.23. 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 2.24. 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 2.25. 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 2.23.

Proof of Proposition 2.23 from Lemma 2.25. 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 2.25: 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 2.25, so $\overline{V} = X$. But V is finite-dimensional by construction, so $V \subseteq X$ is closed by Corollary 2.22, 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 2.25.

Proof of Lemma 2.25. 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 2.26 (functional)**.** Fix a normed vector space $(X, \|\cdot\|)$. Then a *bounded linear functional* is an element of $X^* := \mathcal{L}(X, \mathbb{F})$.

Example 2.27. Certainly X^* is nonempty because it has the zero element.

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

Example 2.28. 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 2.29. 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 2.30. Certainly $\ell(x) = \|x\|$ forces $\|\ell\|_{X^*} \geq 1$, so we are saying that ℓ is small away from x.

Corollary 2.31. 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 2.32 (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 2.33. Fix an $\mathbb{R}$-vector space X. A function $p\colon X \to \mathbb{R}$ is a *sublinear functional* if and only if $p(x + y) \leq p(x) + p(y)$ and $p(tx) = tp(x)$ for $t > 0$.

Example 2.34. 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

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

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 2.35 (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.

2.4 February 16

Here we go.

2.4.1 The Hahn–Banach Theorem

Our proofs of Theorems 2.32 and 2.35 will be by transfinite induction on V. Let's state the successor step.

Lemma 2.36. 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 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 $\widetilde{\ell} \leq p$ pointwise.

Proof. Set $\widetilde{V} := V + \mathbb{R}x$. 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

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

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

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

for any $t \geq 0$ 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

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

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

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

So we can find the needed α if and only if

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

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 2.37. 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 2.36 to Theorem 2.35. We begin by stating Zorn's lemma in the form that we will use. This requires the notion of a partial order.

Definition 2.38 (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 2.39. 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 2.40. 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 2.36.

Theorem 2.35 (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 2.40, 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 $\ell' \leq p$ pointwise. In conclusion, $(W', \ell') \in S$ and bounds C above.

Thus, Theorem 2.40 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 2.36 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 2.41. It turns out that Theorem 2.32 is not actually equivalent to the Axiom of choice.

Remark 2.42. The proof of Theorem 2.32 is similar, so we will omit it. Alternatively, one can apply Theorem 2.35 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 2.35 to $\widetilde{\ell}$ via the above formula.

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