202A: Introduction to Topology and Analysis

Nir Elber

Fall 2022

Contents

Theme 1
Metric Spaces

My personal view on spaces is that every space I ever work with is either metrizable or is the Zariski topology.

—Evan Chen, [Che22]

1.1 August 24

Good morning everyone. This is my first class of the semester.

1.1.1 Administrative Notes

Here are some housekeeping remarks.

- The webpage for this class is `math.berkeley.edu/ rieffel/202AannF22.html`.
- The midterm date is negotiable. We will have a vote on Friday. The possible dates are Friday 14 October, Monday 17 October, or Wednesday 19 October.
- There will be no vote on the final exam. It is on 15 December at 7PM.
- Homework will be due Fridays by midnight, approximately every week.
- There is no particular text for this course, and any given text covers more than we have time for. That said, we will (very) loosely follow [Lan12], but it is helpful to have a number of different expositions around.
- Please wear a mask during lectures and office hours.

Here is a summary of the course.

- We will spend the next couple of lectures talking about metric spaces.
- We will then spend the first half of the course on general topology. The second half of the course will be on measure and integration.
- Throughout we will see a little on functional analysis.

1.1.2 Metric Spaces

Hopefully we remember something about metric spaces. Here's the definition.

Definition 1.1 (Metric)**.** A *metric* d on a set X is a function $d\colon X \times X \to \mathbb{R}_{\geq 0}$ satisfying the following rules for any $x, y, z \in X$.

(a) Zero: $d(x, x) = 0$.

(b) Zero: $d(x, y) = 0$ implies $x = y$.

(c) Symmetry: $d(x, y) = d(y, x)$.

(d) Triangle inequality: $d(x, y) + d(y, z) \geq d(x, z)$.

We call (X, d) a *metric space*.

Remark 1.2. It is occasionally helpful to think about a "reversed" triangle inequality: note $d(x, z) \leq d(x, y) + d(y, z)$ implies $d(x, z) - d(x, y) \leq d(y, z)$. Similarly, $d(x, y) - d(x, z) \leq d(y, z)$, so it follows

$$|d(y, x) - d(x, z)| \leq d(y, z).$$

We will want some "almost" metrics as well. Here are their names.

Definition 1.3 (Semi-metric)**.** A *semi-metric* d on a set X satisfies (a), (c), and (d) of Definition 1.1. We call (X, d) a *semi-metric space*.

Definition 1.4 (Extended metric)**.** An *extended metric* d on a set X is a function $d\colon X \times X \to \mathbb{R}^{\infty}_{\geq 0}$ satisfying (a)–(d) of Definition 1.1. We call (X, d) an *extended metric space*.

Intuitively, we might want extended metrics if we have points that we never want to be able to get to from other ones.

We can turn spaces with a semi-metric into a space with a metric.

Lemma 1.5. Fix a semi-metric space (X, d), and define the relation $\sim$ on X by $x \sim y$ if and only if $d(x, y) = 0$. Then $\sim$ is an equivalence relation.

Proof. We run these checks by hand. Fix any $x, y, z \in X$.

- Reflexive: $d(x, x) = 0$ means that $x \sim x$.
- Symmetry: if $x \sim y$, then $d(x, y) = 0$, so $d(y, x) = 0$, so $y \sim x$.
- Transitive: if $x \sim y$ and $y \sim z$, then

  $$0 \leq d(x, z) \leq d(x, y) + d(y, z) = 0,$$

  so $d(x, z) = 0$, so $x \sim z$. ■

As such, given a semi-metric space (X, d), we may look at the set of equivalence classes under $\sim$, which we will denote $X/{\sim}$.[1]

[1] The notation of $/{\sim}$ is intended to make us think of quotients.

Proposition 1.6. Fix a semi-metric space (X, d) and define $\sim$ as in Lemma 1.5. Then d naturally descends to a metric $\widetilde{d}$ on $X/{\sim}$.

Proof. Let $[x]$ denote the equivalence class of $x \in X$ under $\sim$. We claim that the function

$$\widetilde{d}([x],[y]) := d(x,y)$$

is a well-defined metric. We have the following checks; fix any $x, y, z \in X$.

- Well-defined: if $x \sim x'$ and $y \sim y'$, then note that
$$d(x,y) \le d(x,x') + d(x',y) = d(x',y) \le d(x',y') + d(y',y) = d(x',y').$$
By symmetry, we also have $d(x',y') \le d(x,y)$, so equality follows. So d does descent properly to the quotient $X/{\sim}$.
- Zero: note that $\widetilde{d}([x],[y]) = 0$ if and only if $d(x,y) = 0$ if and only if $x \sim y$ if and only if $[x] = [y]$.
- Symmetry: note that
$$\widetilde{d}([x],[y]) = d(x,y) = d(y,x) = \widetilde{d}([y],[x]).$$
- Triangle inequality: note that
$$\widetilde{d}([x],[z]) = d(x,z) \le d(x,y) + d(y,z) = \widetilde{d}([x],[y]) + \widetilde{d}([y],[z]),$$
which finishes. ■

Here are some examples of metric spaces.

Example 1.7. Given a connected graph $G = (V, E)$ with a weighting function $w \colon E \to \mathbb{R}_{\ge 0}$, we can build a metric as follows: define the "shortest-path" function $d \colon V \times V \to \mathbb{R}_{\ge 0}$ sending two vertices $v, w \in V$ to the length of the shortest path. If the graph G is not connected, we merely have an extended metric.

Example 1.8 (Euclidean metric)**.** The function $d \colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}_{\ge 0}$

$$d\big((x_1, \ldots, x_n), (y_1, \ldots, y_n)\big) := \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

is a metric.

Observe that it is not completely obvious that Example 1.8 satisfies the triangle inequality, but this will follow from the theory of the next subsections.

1.1.3 Norms on Vector Spaces

Norms provide convenient ways to build metrics.

Definition 1.9 (Norm)**.** Fix a vector space V over a normed field $(k, |\cdot|)$. A *norm* $\|\cdot\| \colon V \to \mathbb{R}_{\ge 0}$ is a function satisfying the following, for any $r \in \mathbb{R}$ and $v, w \in V$.

(a) Zero: $\|0\| = 0$.

(b) Zero: if $\|v\| = 0$, then $v = 0$.

(c) Scaling: $\|rv\| = |r| \cdot \|v\|$.

(d) Triangle inequality: $\|v + w\| \le \|v\| + \|w\|$.

Remark 1.10. We can probably work with a more general normed field instead of "merely" $\mathbb{R}$ or $\mathbb{C}$.

There is also an analogous notion of "semi-norm."

Definition 1.11 (Semi-norm)**.** Fix a vector space V over $\mathbb{R}$ or $\mathbb{C}$. A *semi-norm* $\|\cdot\|\colon V \to \mathbb{R}_{\geq 0}$ is a function satisfying (a), (c), and (d) of Definition 1.9.

And here is our result.

Proposition 1.12. Given a vector space V with a (semi-)norm $\|\cdot\| : V \to \mathbb{R}_{\geq 0}$, then the function

$$d(v, w) \coloneqq \|v - w\|$$

defines a (semi-)metric on V.

Proof. We run the checks directly. Let $x, y, z \in V$ be points. Quickly, we note that $d(x, y) = \|x - y\| \geq 0$ by hypothesis on $\|\cdot\|$.

(a) Zero: note that $d(x, x) = 0$ because $d(x, x) = \|x - x\| = \|0\| = 0$.

(b) Zero: if $d(x, y) = 0$, then $\|x - y\| = 0$, so $x - y = 0$, so $x = y$.

(c) Symmetry: note that

$$d(x, y) = \|x - y\| = |-1| \cdot \|y - x\| = 1 \cdot \|y - x\| = d(y, x).$$

(d) Triangle inequality: note that

$$d(x, z) = \|x - z\| = \|(x - y) + (y - z)\| \leq \|x - y\| + \|y - z\| = d(x, y) + d(y, z),$$

which finishes the check.

Thus, if $\|\cdot\|$ is a full norm, then d is a full metric. But if $\|\cdot\|$ is only a semi-norm satisfying (a), (c), and (d) of Definition 1.9, then the corresponding d only satisfies (a), (c), and (d) of Definition 1.1 and makes a semi-metric. ■

As an aside, we note that what's nice about semi-norms is that they will "algebraically" encode the equivalence relation of Lemma 1.5.

Proposition 1.13. Fix a vector space V over a normed field $(k, |\cdot|)$ and a semi-norm $\|\cdot\|$ on V. Then the set $N \coloneqq \{v \in V : \|v\| = 0\}$ is a subspace of V. In fact, the semi-norm $\|\cdot\|$ descends to a well-defined norm on V/N.

Proof. To show that $N \subseteq V$ is a subspace, we pick up $v, w \in N$ and scalars $r, s \in k$. Then we note

$$\|rv + sw\| \leq \|rv\| + \|sw\| = |r| \cdot \|v\| + |s| \cdot \|w\| = 0,$$

so it follows $\|rv + sw\| = 0$ and so $rv + sw \in N$.

It remains to descend $\|\cdot\|$ to V/N. Here are our checks; fix $v, w \in V$ and $r \in k$.

- Well-defined: if $v + N = w + N$, we need $\|v\| = \|w\|$. Well, $v + N = w + N$ tells us that there is some $z \in N$ with $v = w + z$ and so

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

  Similarly, $v + (-z) = w$ implies that $\|w\| \leq \|v\|$, so $\|v\| = \|w\|$ follows.

- Zero: note that $v+N=0+N$ implies that $\|v+N\|=\|0+N\|=\|0\|=0$.
- Zero: if $\|v+N\|=0$, then $\|v\|=0$, so $v\in N$, so $v+N=0+N$.
- Scaling: note $\|r(v+N)\|=\|rv+N\|=\|rv\|=|r|\cdot\|v\|=|r|\cdot\|v+N\|$.
- Triangle inequality: note $\|(v+N)+(w+N)\|=\|(v+w)+N\|=\|v+w\|\le\|v\|+\|w\|=\|v+N\|+\|w+N\|$.

■

Here are the usual examples.

Example 1.14. Set $V:=\mathbb{R}^n$ or $V:=\mathbb{C}^n$. Then the following are norms on V.

- $\|(x_1,\ldots,x_n)\|_2 := \left(\sum_{i=1}^n |x_i|^2\right)^{1/2}$.
- $\|(x_1,\ldots,x_n)\|_1 := \sum_{i=1}^n |x_i|$.

Here are some more esotetric examples.

Example 1.15. Set $V:=\mathbb{R}^n$ or $V:=\mathbb{C}^n$. Then

$$\|(x_1,\ldots,x_n)\|_\infty := \sup\{|x_1|,\ldots,|x_n|\}$$

provides a norm on V.

Example 1.16. Set $V:=\mathbb{R}^n$ or $V:=\mathbb{C}^n$. Then, given $p\ge 1$,

$$\|(x_1,\ldots,x_n)\|_p := \left(\sum_{i=1}^n |x_i|^p\right)^{1/p}$$

provides a norm on V.

Remark 1.17. Taking the limit as $p\to\infty$ of $\|f\|_p$ gives $\|f\|_\infty$. This justifies the notation.

Remark 1.18. Despite having lots of examples, all of these norms are equivalent in a topological sense.

These normed vector spaces actually allow us to define a metric on any subset.

Proposition 1.19. Given a metric space (X,d) and a subset $Y\subseteq X$, the restriction of d to $Y\times Y$ is a metric.

Proof. All the requirements for d on $Y\times Y$ are satisfied for any points in X, so we are done by doing no work. ■

Example 1.20. Any subset $X\subseteq\mathbb{R}^n$ has an induced metric by restricting the (say) Euclidean metric.

1.1.4 A Hint of L^p Spaces

Here is a more complicated example of a metric.

Example 1.21. Define $V := C([0,1])$ to be the $\mathbb{R}$-vector space of $\mathbb{R}$-valued (or $\mathbb{C}$-valued) continuous functions on $[0,1]$. The following are norms.

- $\|f\|_\infty := \sup\{|f(x)| : x \in [0,1]\}$.
- $\|f\|_1 := \int_0^1 |f(t)|\,dt$.
- $\|f\|_2 := \left(\int_0^1 |f(t)|^2\,dt\right)^{1/2}$.
- More generally, given $p \geq 1$
$$\|f\|_p := \left(\int_0^1 |f(t)|^p\,dt\right)^{1/p}.$$

These integrals are finite because $[0,1]$ is compact, forcing f to achieve a finite maximum on $[0,1]$.

Remark 1.22. We can tell the same story for $C(X)$, for any measurable compact space X.

Remark 1.23. Note the analogy of Example 1.21 with Example 1.16. To see this more rigorously, set X to be the finite set $\{1,\ldots,n\}$ so that $C(X) = \mathbb{R}^n$.

We should probably justify the claims of this subsection, so here is our result.

Proposition 1.24. Define $V := C([0,1])$ to be the vector space of $\mathbb{R}$-valued (or $\mathbb{C}$-valued) continuous functions on $[0,1]$. Then, given $p \geq 1$, the function $\|\cdot\|_p : C \to \mathbb{R}_{\geq 0}$ by
$$\|f\| := \left(\int_0^1 |f(t)|^p\,dt\right)^{1/p}$$
is a norm.

Proof. We run the checks directly.

- Zero: if $f = 0$, then of course $\int_0^1 |f(t)|^p\,dt = 0$.
- Zero: suppose that $f \in C([0,1])$ has $f(t_0) \neq 0$ for any $t_0 \in [0,1]$; set $y := f(t_0)$. Then $f^{-1}((y/2, 3y/2))$ is a nonempty open subset of X and hence contains a nonempty open interval (a,b) with $a < b$. As such,
$$\int_X |f(t)|^p\,dt \geq \int_a^b |f(t)|^p\,dt \geq \int_a^b |y/2|^p\,dt > 0,$$
so we are done.
- Scaling: given $f \in C([0,1])$ and a scalar r, we have
$$\|rf\| = \left(\int_0^1 |rf(t)|^p\,dt\right)^{1/p} = \left(|r|^p \int_0^1 |f(t)|^p\,dt\right)^{1/p} = |r| \cdot \|f\|.$$
- Triangle inequality: we borrow from [Tao09]. Given $f, g \in C([0,1])$, for psychological reasons we will assume that f and g are nonzero (else this is clear); then $\|f\|, \|g\| \neq 0$, so we may scale everything so that $\|f\| + \|g\| = 1$. In fact, we may again use scaling to find $a, b \in V$ such that
$$f = (1-\theta)a \qquad \text{and} \qquad g = \theta b$$

where $\theta \in (0,1)$ and $\|a\| = \|b\| = 1$. Now, the triangle inequality translates into showing

$$\int_0^1 |(1-\theta)a(t) + \theta b(t)|^p \, dt = \|(1-\theta)a + \theta b\|_p^p \stackrel{?}{\leq} \left(\|(1-\theta)a\|_p + \|\theta b\|_p\right)^p = 1.$$

Well, because $p \geq 1$, the function $t \mapsto t^p$ is convex, so we get to write

$$\int_0^1 |(1-\theta)a(t) + \theta b(t)|^p \, dt \leq (1-\theta)\int_0^1 |a(t)|^p \, dt + \theta \int_0^1 |b(t)|^p \, dt,$$

which is what we wanted.

The above checks complete the proof; note that the proof of the triangle inequality was nontrivial. ■

Remark 1.25. Now, to show Remark 1.23, replace all $\int_0^1$ with $\sum_{i=1}^n$ and adjust all the language accordingly. The point is that "integrating over $[0,1]$" is analogous to "integrating over $\{1,\ldots,n\}$." A more thorough understanding of measure theory will allow us to rigorize this.

Next class we will talk about completeness.

1.2 August 26

Today we're talking about completeness of metric spaces.

1.2.1 Isometries

In mathematics, we are interested in objects not in isolation but as they relate to each other. Namely, we are interested also in the maps between our objects.

The philosophy here comes from category theory, where one is really most interested in the "morphisms" between "objects" instead of the objects themselves. For concreteness, here is a definition of a category.

Definition 1.26 (Category). A *category* $\mathcal{C}$ consists of a class of objects $\operatorname{Ob}\mathcal{C}$ and class of morphisms $\operatorname{Mor}\mathcal{C}$ such that any two objects $A, B \in \operatorname{Ob}\mathcal{C}$ have a morphism class $\operatorname{Mor}(A,B)$. This data satisfy the following properties.

- Composition: given objects $A, B, C \in \operatorname{Ob}\mathcal{C}$, there is a binary composition operation

  $$\circ\colon \operatorname{Mor}(B,C) \times \operatorname{Mor}(A,B) \to \operatorname{Mor}(A,C).$$

  Explicitly, given $f \in \operatorname{Mor}(A,B)$ and $g \in \operatorname{Mor}(B,C)$, there is a composition $(g \circ f) \in \operatorname{Mor}(A,C)$.
- Given $A \in \operatorname{Ob}\mathcal{C}$, there is an identity morphism $\operatorname{id}_A \in \operatorname{Mor}(A,A)$.
- Identity: any $f \in \operatorname{Mor}(A,B)$ has $f \circ \operatorname{id}_A = f = \operatorname{id}_B \circ f$.
- Associativity: any $f \in \operatorname{Mor}(A,B)$ and $g \in \operatorname{Mor}(B,C)$ and $h \in \operatorname{Mor}(C,D)$ has $(h \circ g) \circ f = h \circ (g \circ f)$.

Example 1.27. There is a category of groups, where the morphisms are group homomorphisms. The identity function gives the identity morphism, and composition of functions gives the required composition.

For completeness, we check that composition is well-defined: given homomorphisms $f\colon A \to B$ and $g\colon B \to C$, we need $(g \circ f)\colon A \to C$ to be a group homomorphism. Well,

$$(g \circ f)(a \cdot a') = g(f(a \cdot a')) = g(f(a) \cdot f(a')) = g(f(a)) \cdot g(f(a')) = (g \circ f)(a) \cdot (g \circ f)(a').$$

In our discussion of metric spaces, there are many possible kinds of morphisms for us to consider. Here is the strongest type.

Definition 1.28 (Isometry)**.** Given metric spaces (X, d_X) and (Y, d_Y), an *isometry* is a function $f\colon X \to Y$ preserving the metric as

$$d_Y(f(x), f(x')) = d_X(x, x').$$

Example 1.29. The 90° rotation $r\colon \mathbb{R}^2 \to \mathbb{R}^2$ given by $r(x, y) \mapsto (y, -x)$ is an isometry, where $\mathbb{R}^2$ is given the Euclidean metric. Indeed, any $(x, y), (x', y') \in \mathbb{R}^2$ have

$$\begin{aligned} d\big(r(x, y), r(x', y')\big) &= d\big((y, -x), (y', -x')\big) \\ &= \sqrt{(y - y')^2 + (-x - -x')^2} \\ &= \sqrt{(x - x')^2 + (y - y')^2} \\ &= d\big((x, y), (x', y')\big). \end{aligned}$$

Notation 1.30. Fix two metric spaces (X, d_X) and (Y, d_Y). Given a function $f\colon X \to Y$ with extra structure respecting some aspect of the metric, we might write $f\colon (X, d_X) \to (Y, d_Y)$ to emphasize this.

To show that isometries are valid morphisms, we need to check that the identity function $\mathrm{id}_X\colon X \to X$ is an isometry (which of course it is) and that the composition of two isometries is an isometry. We check this last one in a quick lemma.

Lemma 1.31. Given two isometries $f\colon (X, d_X) \to (Y, d_Y)$ and $g\colon (Y, d_Y) \to (Z, d_Z)$, the composition $g \circ f$ is an isometry.

Proof. Well, any two points $x, x' \in X$ have

$$d_Z(g(f(x)), g(f(x'))) = d_Y(f(x), f(x')) = d_X(x, x'),$$

which is what we wanted. ■

One can restrict further to surjective isometries, where the main point is that (again) the composition of two surjective functions remains surjective. (Note that the identity is of course surjective.) The following is the reason why a surjective isometry is a good notion.

Lemma 1.32. A surjective isometry $f\colon (X, d_X) \to (Y, d_Y)$ is bijective, and its inverse function is also an isometry.

Proof. To see that f is bijective, we only need to know that f is injective. Well, given $x, x' \in X$, note that $f(x) = f(x')$ if and only if $d_Y(f(x), f(x')) = 0$ if and only if $d(x, x') = 0$ if and only if $x = x'$.[2]

Thus, f is indeed bijective; let $g\colon Y \to X$ be its inverse. We now need to show that g is an isometry. Well, given $y, y' \in Y$, we may find $x, x' \in X$ such that $f(x) = y$ and $f(x') = y'$. Then

$$d_X(g(y), g(y')) = d_X((g \circ f)(x), (g \circ f)(x')) = d_X(x, x') \stackrel{*}{=} d_Y(f(x), f(x')) = d_Y(y, y'),$$

where in $\stackrel{*}{=}$ we have used the fact that f is an isometry. ■

[2] In fact, this argument shows that all isometries are injective. We will shortly see that all actually Lipschitz continuous functions are injective.

Remark 1.33. The above result is somewhat subtle in its importance: the inverse function of a bijection is only an inverse in the category of sets. The above result is saying that this inverse morphism in the category of sets is lifting to an inverse morphism in the category of metric spaces with isometries as morphisms. In general, it is not always true that bijective morphisms are invertible, as we shall soon see.

1.2.2 Lipschitz Continuity

Isometries are somewhat restrictive, so we might weaken this as follows.

Definition 1.34 (Lipschitz continuous). Given metric spaces (X, d_X) and (Y, d_Y), a function $f\colon X \to Y$ is a *Lipschitz continuous* if and only if there is a constant $c \in \mathbb{R}$ such that

$$d_Y(f(x), f(x')) \le c \cdot d_X(x, x').$$

Remark 1.35. Equivalently, we are asking for the ratio

$$\frac{d_Y(f(x), f(x'))}{d_X(x, x')}$$

to be uniformly bounded above for all $x \neq x'$. Notably, the inequality is trivially satisfied whenever $x = x'$, or equivalently whenever $d(x, x') = 0$.

Example 1.36. Any isometry $f\colon (X, d_X) \to (Y, d_Y)$ is Lipschitz continuous: indeed, set $c := 1$ so that, for any $x, x' \in X$,

$$d_Y(f(x), f(x')) = d_X(x, x') \le 1 \cdot d_X(x, x').$$

Example 1.37. Provide $\mathbb{R}$ and $\mathbb{R}^2$ their usual Euclidean metrics. Then the projection $\pi\colon \mathbb{R}^2 \to \mathbb{R}$ by $\pi\colon (x, y) \mapsto x$ is Lipschitz continuous: indeed, set $c := 1$ so that, for any $(x, y), (x', y') \in \mathbb{R}^2$, we have

$$d_{\mathbb{R}^2}\big((x, y), (x', y')\big) = \sqrt{(x - x')^2 + (y - y')^2} \ge \sqrt{(x - x')^2} = d_{\mathbb{R}}(x, x') = d_{\mathbb{R}}\big(\pi((x, y)), \pi((x', y'))\big).$$

Example 1.38. Fix a normed vector space $(B, \|\cdot\|)$. We show the function $\|\cdot\| : B \to \mathbb{R}$ is Lipschitz continuous. Well, observe that $\|x\| \le \|x - y\| + \|y\|$, so by symmetry, it follows that

$$|\, \|x\| - \|y\| \,| \le \|x - y\|\,.$$

Again, one can see that the identity function $\mathrm{id}_X\colon (X, d_X) \to (X, d_X)$ is Lipschitz continuous (with $c := 1$), and here is our composition check.

Lemma 1.39. If $f\colon (X, d_X) \to (Y, d_Y)$ and $g\colon (Y, d_Y) \to (Z, d_Z)$ are Lipschitz continuous, then the composition $(g \circ f)\colon (X, d_X) \to (Z, d_Z)$ is also Lipschitz continuous.

Proof. We are given constants c and d such that any $x, x' \in X$ and $y, y' \in Y$ have

$$d_Y(f(x), f(x')) \le c \cdot d_X(x, x') \qquad \text{and} \qquad d_Z(g(y), g(y')) \le d \cdot d_Y(y, y').$$

As such, we use the constant cd to witness our Lipschitz continuity: any $x, x' \in X$ have

$$d_Z(g(f(x)), g(f(x'))) \le d \cdot d_Y(f(x), f(x')) \le cd \cdot d_X(x, x'),$$

which is what we wanted. ■

It will be shortly worth our time to talk about the constant c appearing in Definition 1.34.

Lemma 1.40. Fix a Lipschitz continuous function $f\colon (X, d_X) \to (Y, d_Y)$. Then there exists a constant c_f (possibly $-\infty$) such that any real number $c \geq c_f$ is equivalent to the following property: any $x, x' \in X$ have

$$d_Y(f(x), f(x')) \leq c \cdot d_X(x, x').$$

Proof. Let S denote the set of all constants c such that any $x, x' \in X$ have

$$d_Y(f(x), f(x')) \leq c \cdot d_X(x, x').$$

Equivalently, using Remark 1.35, S is the set of upper-bounds for

$$R := \left\{ \frac{d_Y(f(x), f(x'))}{d_X(x, x')} : x, x' \in X, x \neq x' \right\}.$$

Now, S is nonempty because f is Lipschitz continuity, so we set $c_f := \sup R$ to be the least upper bound for R—observe that $c_f = -\infty$ is permissible when X has one point. It is now pretty clear that $S = [c_f, \infty)$. ∎

Note that c_f the property stated in the lemma automatically implies that c_f is the least possible constant and is unique. Being least is immediate (by the backwards direction), and being unique follows from being least. So because we have some uniqueness, we get a definition.

Definition 1.41 (Lipschitz constant)**.** Given a Lipschitz continuous function $f\colon (X, d_X) \to (Y, d_Y)$, the *Lipschitz constant c_f for f* is the least real number c such that

$$d_Y(f(x), f(x')) \leq c \cdot d_X(x, x').$$

We could, as before, look at surjective Lipschitz continuous functions, but these need not be bijective anymore as shown by Example 1.37. What's worse is that, as warned possible in Remark 1.33, bijective Lipschitz continuous functions need not even have a Lipschitz continuous inverse.

Exercise 1.42. We exhibit a function between metric spaces which is bijective and Lipschitz continuous, but its inverse function is not Lipschitz continuous.

Proof. Set $X := (0, 1)$ and $Y := (1, \infty)$, both metric spaces with the Euclidean (subspace) metric, and set $f\colon (0, \infty) \to (0, \infty)$ by $f\colon x \mapsto 1/x$. Notably, $x \in X$ implies $f(x) \in Y$, and $y \in Y$ implies $f(y) \in X$.

- Note $f|_Y$ is bijective with inverse $f|_X$ because $f(f(x)) = f(1/x) = x$ for all $x \in (0, \infty)$.
- Note $f|_Y$ is Lipschitz continuous: set $c := 1$ and note that any $y, y' \in Y$ have

$$|f(y) - f(y')| = \left| \frac{1}{y} - \frac{1}{y'} \right| = \left| \frac{y - y'}{yy'} \right| \leq |y - y'|.$$

- But $f|_X$ is not Lipschitz continuous: suppose for contradiction that f_X is Lipschitz continuous, and use Lemma 1.40 to recover the needed constant c_0. Then set $c := \max\{c_0, 4\}$, which must also work as a constant, and set $x := 1/c$ and $x' := 1/(3c)$ so that

$$|f(x) - f(x')| = |c - 3c| = 2c > c \cdot |x - x'|.$$

This is a contradiction, so we are done. ∎

Remark 1.43 (Nir). In some sense, the problem here is that the definition of Lipschitz continuity allows $d_Y(f(x), f(x'))$ to be "too small," which permits the inverse function to have distances which blow up.

In light of Exercise 1.42, we introduce a new definition.

Definition 1.44 (Lipschitz isomorphism). Give metric spaces (X, d_X) and (Y, d_Y), a function $f\colon X \to Y$ is a *Lipschitz isomorphism* if and only if f is Lipschitz continuous and has an inverse function which is also Lipschitz continuous.

Remark 1.45. A good reason to care about this notion of continuity (and isomorphism) is that all normed $\mathbb{R}$-vector spaces of some finite dimension n are Lipschitz isomorphic.

1.2.3 Fun with Continuity

Here is yet a weaker notion of morphism.

Definition 1.46 (Uniformly continuous). Given metric spaces (X, d_X) and (Y, d_Y), a function $f\colon X \to Y$ is *uniformly continuous* if and only if every $\varepsilon > 0$ has some $\delta > 0$ such that

$$d_X(x, x') < \delta \implies d_Y(f(x), f(x')) < \varepsilon$$

for all $x, x' \in X$.

Example 1.47. Any Lipschitz continuous function $f\colon (X, d_X) \to (Y, d_Y)$ is also uniformly continuous: indeed, for any $\varepsilon > 0$, set $\delta := \max\{c_f, 1\}\varepsilon > 0$ (where c_f is the Lipschitz constant) so that

$$d_X(x, x') < \varepsilon \implies d_Y(f(x), f(x')) \le c_f \cdot d(x, x') < \delta.$$

Example 1.48. Give $[0, 1]$ the Euclidean (subspace) metric, and set $f\colon [0, 1] \to [0, 1]$ by $f(x) := \sqrt{x}$.

- Note f is uniformly continuous because it is continuous on a compact set.
- However, f is not Lipschitz continuous: for any constant $c > 0$, set $x = 1/(c + 1)^2$ and $x' = 0$ so that
$$\left|\frac{f(x) - f(x')}{x - x'}\right| = \left|\frac{1/(c+1)}{1/(c+1)^2}\right| = |c + 1| > c,$$
so Remark 1.35 tells us that we are not Lipschitz continuous.

By rearranging quantifiers, we get another useful (but weaker) notion.

Definition 1.49 (Continuous). Given metric spaces (X, d_X) and (Y, d_Y), a function $f\colon X \to Y$ is *continuous at $x \in X$* if and only if all $\varepsilon > 0$ have some $\delta_x > 0$ such that

$$d_X(x, x') < \delta_x \implies d_Y(f(x), f(x')) < \varepsilon.$$

Then f is *continuous* if and only if it is continuous at all $x \in X$.

Example 1.50. All uniformly continuous functions $f\colon (X,d_X)\to(Y,d_Y)$ are continuous. Indeed, at any $x_0\in X$ with $\varepsilon>0$, uniform continuity promises $\delta>0$ so that

$$|x-x'|<\delta \implies |f(x)-f(x')|<\varepsilon$$

for all $x,x'\in X$. Setting x' to x_0 recovers continuity.

Example 1.51. Give $\mathbb{R}$ the usual Euclidean metric, and set $f\colon \mathbb{R}\to\mathbb{R}$ by $f(x):=x^2$.

- Note $f(x)$ is continuous because it is a polynomial.
- However, $f(x)$ is not uniformly continuous: take $\varepsilon=1$. Now, for any $\delta>0$, set $x=1/\delta$ and $x'=1/\delta+\delta/2$ so that $|x-x'|<\delta$, but

$$|f(x)-f(x')|=\left(\frac{1}{\delta}+\frac{\delta}{2}\right)^2-\frac{1}{\delta^2}=1+\frac{\delta^2}{4}>\varepsilon.$$

As usual, the identity function is uniformly continuous and continuous (it's an isometry), and these continuities are preserved by composition. We will have a different way to see that continuous functions remain continuous under composition later, so for now we will focus on uniform continuity.

Lemma 1.52. Fix uniformly continuous morphisms $f\colon (X,d_X)\to(Y,d_Y)$ and $g\colon (Y,d_Y)\to(Z,d_Z)$. Then the function $(g\circ f)$ is uniformly continuous.

Proof. For any $\varepsilon>0$, the uniform continuity of g promises $\delta_g>0$ such that

$$d_Y(y,y')<\delta_g \implies d_Z(g(y),g(y'))<\varepsilon$$

for any $y,y'\in Y$. Continuing, the uniform continuity of f promises $\delta_f>0$ such that

$$d_X(x,x')<\delta_X \implies d_Y(f(x),f(x'))<\delta_Y \implies d_Z(g(f(x)),g(f(x')))<\varepsilon$$

for any $x,x'\in X$, which is what we wanted. ■

Remark 1.53. In some sense, isometries and Lipschitz continuous functions have their definition fundamentally interrelated with the metric. In contrast, the weaker notion of continuity will readily generalize to general topological spaces. Uniform continuity also generalizes to "uniformities," which is a different notion.

1.2.4 Convergent Sequences

To discuss completeness, we need to talk about convergence.

Definition 1.54 (Converge)**.** Fix a semi-metric d on a set X. A sequence of points $\{x_n\}_{n\in\mathbb{N}}\subseteq X$ *converges to* $x\in X$ if and only if, for any $\varepsilon>0$, we can find $N>0$ such that

$$n>N \implies d(x_n,x)<\varepsilon.$$

We might write this as "$x_n\to x$ as $n\to\infty$" or "$\lim_{n\to\infty}x_n=x$." In this event, we may say that the sequence $\{x_n\}_{n\in\mathbb{N}}$ *converges*, and its limit is x.

Remark 1.55 (Nir). As a sanity check, the limit of a sequence is unique if (X,d) is a metric space: if $x_n \to x$ and $x_n \to x'$ as $n \to \infty$, then any $\varepsilon > 0$ can find some large n so that $d(x_n,x), d(x_n,x') < \varepsilon/2$. As such,

$$d(x,x') < d(x_n,x) + d(x_n,x') = \varepsilon$$

for any $\varepsilon > 0$, so $d(x,x') = 0$ and thus $x = x'$ is forced.

Example 1.56. Given $x \in X$, define the sequence $\{x_n\}_{n\in\mathbb{N}}$ by $x_n := x$ for each n. Then $d(x_n,x) = 0$ for each n, so any $\varepsilon > 0$ may set $N = 0$ so that $n \geq N$ implies $d(x_n,x) < \varepsilon$. Thus, $x_n \to x$ as $n \to \infty$.

We have no reason yet to be convinced that any of our morphisms described previously are good notions, so let's start with continuity.

Lemma 1.57. Fix a continuous function between metric spaces $f\colon (X,d_X) \to (Y,d_Y)$. Then, if the sequence $\{x_n\}_{n\in\mathbb{N}} \subseteq X$ converges to $x \in X$, then the sequence $\{f(x_n)\}_{n\in\mathbb{N}} \subseteq Y$ converges to $f(x) \in Y$.

Proof. For any $\varepsilon > 0$,the continuity of f implies that we can find $\delta_x > 0$ so that

$$d_X(x_n,x) < \delta_x \implies d_Y(f(x_n),f(x)) < \varepsilon$$

for any x_n. But the fact that $x_n \to x$ as $n \to \infty$ means that there is $N > 0$ so that

$$n > N \implies d_X(x_n,x) < \delta_x \implies d_Y(f(x_n),f(x)) < \varepsilon,$$

so indeed, $f(x_n) \to f(x)$ as $n \to \infty$. ■

In fact, the converse also holds.

Lemma 1.58. Fix metric spaces (X,d_X) and (Y,d_Y), and fix a point $x \in X$. Then suppose a function $f\colon X \to Y$ satisfies that any convergent sequence $\{x_n\}_{n\in\mathbb{N}}$ with $x_n \to x$ as $n \to \infty$ has $f(x_n) \to f(x)$ as $n \to \infty$. Then f is continuous at x.

Proof. We proceed by contraposition. If f is not continuous at x, then any $n \in \mathbb{N}$ can find x_n such that $d_X(x,x_n) < 1/n$ even though $d_Y(f(x_n),f(x)) \geq 1$. In particular, $x_n \to x$ as $n \to \infty$ (for any ε, choose $N = 1/\varepsilon$), but the sequence $\{f(x_n)\}_{n\in\mathbb{N}}$ does not converge to $f(x)$ because no n has $d_Y(f(x),f(x_n)) < 1$. ■

We will want the following fact (much) later, but we prove it now while ideas are fresh.

Lemma 1.59. Fix a semi-norm $\|\cdot\|$ on a k-vector space V. Further, fix sequences $\{v_n\}_{n\in\mathbb{N}}$ and $\{w_n\}_{n\in\mathbb{N}}$ of vectors and two more vectors $v,w \in V$ such that $v_n \to v$ and $w_n \to w$ as $n \to \infty$.

(a) We have $v_n + w_n \to v + w$ as $n \to \infty$.

(b) For any scalar $a \in k$, we have $av_n \to av$ as $n \to \infty$.

Proof. Here we go. Let $|\cdot|$ denote the norm on k.

(a) For any $\varepsilon > 0$, having $v_n \to v$ promises N_v such that $n \geq N_v$ has

$$\|v - v_n\|_1 < \varepsilon/2.$$

Similarly, $w_n \to w$ promises N_w such that $n \geq N_w$ has

$$\|w - w_n\|_1 < \varepsilon/2.$$

As such, we set $N := \max\{N_v, N_w\}$ so that $n \geq N$ implies $n \geq N_v$ and $n \geq N_w$ and thus

$$\|(v+w)-(v_n+w_n)\|_1 \leq \|v-v_n\|_1 + \|w-w_n\|_1 < \frac{\varepsilon}{2}+\frac{\varepsilon}{2} = \varepsilon$$

by the triangle inequality.

(b) If $a=0$, then $av_n = av = 0$, so $av - av_n = 0$, so $\|av-av_n\|_1 = 0$. Thus, $av_n \to av$.
Otherwise, take $a \neq 0$ so that $|a|>0$. Now, having $v_n \to v$ promises N such that $n \geq N$ has

$$\|v-v_n\|_1 < \frac{\varepsilon}{|a|}.$$

Thus, $n \geq N$ has

$$\|av-av_n\|_1 = \|a(v-v_n)\|_1 \stackrel{*}{=} |a| \cdot \|v-v_n\|_1 < |a| \cdot \frac{\varepsilon}{|a|} = \varepsilon,$$

where $\stackrel{*}{=}$ is because $\|\cdot\|$ is a semi-norm. ■

1.2.5 Cauchy Sequences

We would like a notion of convergence which only uses data internal to the sequence, and this leads to the following definition.

Definition 1.60 (Cauchy)**.** Fix a semi-metric d on a set X. A sequence of points $\{x_n\}_{n\in\mathbb{N}} \subseteq X$ is a *Cauchy sequence* if and only if, for any $\varepsilon > 0$, we can find $N > 0$ such that

$$n, m > N \implies d(x_n, x_m) < \varepsilon.$$

Example 1.61. Given $x \in X$, define the sequence $\{x_n\}_{n\in\mathbb{N}}$ by $x_n := x$ for each n. Then $d(x_m, x_n) = 0$ for each n, so any $\varepsilon > 0$ may set $N = 0$ so that $n \geq N$ implies $d(x_m, x_n) < \varepsilon$. Thus, $\{x_n\}_{n\in\mathbb{N}}$ is Cauchy. More generally, we will see that convergent sequences are Cauchy in Lemma 1.64.

It would be rude if continuity was always the best kind of morphism, so this time around preserving Cauchy-ness requires something stronger.

Lemma 1.62. Fix a uniformly continuous function between metric spaces $f\colon (X, d_X) \to (Y, d_Y)$. Then, if the sequence $\{x_n\}_{n\in\mathbb{N}} \subseteq X$ is Cauchy, then the sequence $\{f(x_n)\}_{n\in\mathbb{N}} \subseteq Y$ is also Cauchy.

Proof. For any $\varepsilon > 0$, the uniform continuity of f promises $\delta > 0$ so that

$$d_X(x_n, x_m) < \delta \implies d(f(x_n), f(x_m)) < \varepsilon$$

for any x_n, x_m. However, the fact that $\{x_n\}_{n\in\mathbb{N}}$ is Cauchy promises N so that

$$n, m > N \implies d_X(x_n, x_m) < \delta \implies d(f(x_n), f(x_m)) < \varepsilon,$$

which is what we wanted. ■

Example 1.63. Continuous functions do not need to preserve Cauchy sequences: $f\colon (0,\infty) \to (0,\infty)$ by $f(x) := 1/x$ is continuous, and the sequence $\{1/n\}_{n\in\mathbb{N}} \subseteq (0,\infty)$ is Cauchy (it converges to 0 in $\mathbb{R}$) even though $\{f(1/n)\}_{n\in\mathbb{N}} = \{n\}_{n\in\mathbb{N}}$ certainly does not converge.

Anyway, it is quick to check that convergent sequences are Cauchy.

Lemma 1.64. Fix a metric space (X, d). Then all convergent sequences are Cauchy.

Proof. Suppose that the sequence $\{x_n\}_{n\in\mathbb{N}} \subseteq X$ converges to $x \in X$. Then, for any $\varepsilon > 0$, find N so that

$$d(x_n, x) < \varepsilon/2$$

for all $n > N$. Then any $n, m > N$ has

$$d(x_n, x_m) \le d(x_n, x) + d(x_m, x) < \varepsilon,$$

so the sequence $\{x_n\}_{n\in\mathbb{N}}$ is Cauchy. ■

As before, we will want the following fact later.

Lemma 1.65. Fix a semi-norm $\|\cdot\|$ on a k-vector space V. Further, fix Cauchy sequences $\{v_n\}_{n\in\mathbb{N}}$ and $\{w_n\}_{n\in\mathbb{N}}$ of vectors.

(a) The sequence $\{v_n + w_n\}_{n\in\mathbb{N}}$ is Cauchy.

(b) For any scalar $a \in k$, the sequence $\{av_n\}_{n\in\mathbb{N}}$ is Cauchy.

Proof. These proofs are essentially the same as Lemma 1.59. As usual, let $|\cdot|$ denote the norm on k.

(a) For any $\varepsilon > 0$, having $\{v_n\}_{n\in\mathbb{N}}$ Cauchy promises N_v such that $n \ge N_v$ has

$$\|v_m - v_n\|_1 < \varepsilon/2.$$

Similarly, $\{w_n\}_{n\in\mathbb{N}}$ Cauchy promises N_w such that $n \ge N_w$ has

$$\|w_m - w_n\|_1 < \varepsilon/2.$$

As such, we set $N := \max\{N_v, N_w\}$ so that $n \ge N$ implies $n \ge N_v$ and $n \ge N_w$ and thus

$$\|(v_m + w_m) - (v_n + w_n)\|_1 \le \|v_m - v_n\|_1 + \|w_m - w_n\|_1 < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

by the triangle inequality.

(b) If $a = 0$, then $av_n = av_m = 0$, so $av_m - av_n = 0$, so $\|av_m - av_n\|_1 = 0$. Thus, the sequence $\{av_m\}_{n\in\mathbb{N}}$ in Cauchy.

Otherwise, take $a \ne 0$ so that $|a| > 0$. Now, having $\{v_n\}_{n\in\mathbb{N}}$ Cauchy promises N such that $n \ge N$ has

$$\|v_m - v_n\|_1 < \frac{\varepsilon}{|a|}.$$

Thus, $n \ge N$ has

$$\|av_m - av_n\|_1 = \|a(v_m - v_n)\|_1 \overset{*}{=} |a| \cdot \|v_m - v_n\|_1 < |a| \cdot \frac{\varepsilon}{|a|} = \varepsilon,$$

where $\overset{*}{=}$ is because $\|\cdot\|$ is a semi-norm. ■

We in general hope that our Cauchy sequences converge. As such, we have the following definition.

Definition 1.66 (Complete)**.** A metric space (X, d) is *complete* if and only if every Cauchy sequence in X converges to a point in X.

We are sad when a metric space is not complete, so we hope to have a way to make it complete. The most natural way to do this is by using the notion of density.

Definition 1.67 (Dense)**.** Fix a metric space (X, d). Then $S \subseteq X$ is *dense* if and only if, given any $x \in X$ and $\varepsilon > 0$, we may find $x' \in S$ with $d(x, x') < \varepsilon$.

And here is our completion.

Definition 1.68 (Completion)**.** A *completion* of the metric space (X, d) is a metric space $(\overline{X}, \overline{d})$ equipped with an isometry $\iota\colon X \to \overline{X}$ such that $(\overline{X}, \overline{d})$ is complete and $\operatorname{im} \iota$ is dense in $\overline{X}$.

One can show that any metric space has a completion and that they are all isometric and therefore in some sense the same. We'll do these separately.

1.2.6 Existence of Completions

Let's start with existence.

Theorem 1.69. Any metric space (X, d) has a completion.

Proof. Let $\widetilde{X}$ denote the set of all Cauchy sequences in X. We hope to make $\widetilde{X}$ into our completion, but this requires a little care. To begin, we have the following lemma.

Lemma 1.70. Given a metric space (X, d) with two Cauchy sequences $\{x_n\}_{n\in\mathbb{N}}$ and $\{y_n\}_{n\in\mathbb{N}}$, then the sequence

$$\{d(x_n, y_n)\}_{n\in\mathbb{N}} \subseteq \mathbb{R}$$

converges.

Proof. Because $\mathbb{R}$ is a complete metric space, it suffices to show that the sequence $\{d(x_n, y_n)\}_{n\in\mathbb{N}}$ is Cauchy. Well, for any $\varepsilon > 0$, find a sufficiently large N so that

$$n, m > N \implies d(x_n, x_m), d(y_n, y_m) < \varepsilon/2.$$

Then any $n, m > N$ has

$$d(x_n, y_n) \le d(x_n, x_m) + d(x_m, y_m) + d(y_m, y_n) < \varepsilon + d(y_m, y_n),$$

and $d(x_m, y_m) < d(x_n, y_n) + \varepsilon$ as well by symmetry. It follows that any $n, m > N$ has

$$\left|d(x_n, y_n) - d(x_m, y_m)\right| < \varepsilon,$$

verifying that our sequence is Cauchy. ■

Remark 1.71. Here is a quick motivational remark for the definition of our metric below: if (X, d) is a metric space with $x_n \to x$ and $y_n \to y$ as $n \to \infty$, then we claim $d(x_n, y_n) \to d(x, y)$ as $n \to \infty$. Indeed, for any $\varepsilon > 0$, we can find N large enough so that $d(x_n, x), d(y_n, y) < \varepsilon/2$ for any $n > N$. As such,

$$d(x_n, y_n) \le d(x_n, x) + d(x, y) + d(y, y_n) < d(x, y) + \varepsilon.$$

By symmetry, we get $d(x, y) \le d(x_n, y_n) + \varepsilon$ as well, finishing.

Thus, we define $\widetilde{d}\colon \widetilde{X} \times \widetilde{X} \to \mathbb{R}_{\ge 0}$ by

$$\widetilde{d}\left(\{x_n\}, \{y_n\}\right) := \lim_{n\to\infty} d(x_n, y_n).$$

We claim that $\widetilde{d}$ is a semi-metric on $\widetilde{X}$. We have the following checks; fix Cauchy sequences $\{x_n\}, \{y_n\}, \{z_n\}$.

- Zero: note
$$\widetilde{d}(\{x_n\},\{x_n\}) = \lim_{n\to\infty} d(x_n,x_n) = 0.$$
- Symmetry: note
$$\widetilde{d}(\{x_n\},\{y_n\}) = \lim_{n\to\infty} d(x_n,y_n) = \lim_{n\to\infty} d(y_n,x_n) = \widetilde{d}(\{y_n\},\{x_n\}).$$
- Triangle inequality: note
$$\begin{aligned}\widetilde{d}(\{x_n\},\{y_n\}) + \widetilde{d}(\{y_n\},\{z_n\}) &= \lim_{n\to\infty} d(x_n,y_n) + \lim_{n\to\infty} d(y_n,z_n) \\ &= \lim_{n\to\infty} (d(x_n,y_n) + d(y_n,z_n)) \\ &\ge \lim_{n\to\infty} d(x_n,z_n) \\ &= \widetilde{d}(x_n,z_n),\end{aligned}$$
where we have implicitly used a number of limit laws.

So because $\widetilde{d}$ is a semi-metric, Proposition 1.6 tells us that $\widetilde{d}$ will descend naturally to a metric $\overline{d}$ on $\overline{X} \coloneqq \widetilde{X}/{\sim}$, where $\{x_n\} \sim \{y_n\}$ if and only if $\widetilde{d}(\{x_n\},\{y_n\}) = 0$. We will let $[\{x_n\}]$ denote the equivalence class of the Cauchy sequence $\{x_n\} \in \widetilde{X}$ in $\overline{X}$.

We now show that $(\overline{X},\overline{d})$ can be made into a completion for X.

- Given $x \in X$, note that the constant sequence $\{x\}$ is Cauchy (for any $\varepsilon > 0$, set $N = 0$), so we define $\iota\colon X \to \overline{X}$ by
$$\iota(x) \coloneqq [\{x\}].$$
To see that ι is an isometry, note any $x, x' \in X$ have
$$\overline{d}(\iota(x),\iota(x')) = \widetilde{d}(\{x\},\{y\}) = \lim_{n\to\infty} d(x,y) = d(x,y).$$
- We show that $\operatorname{im}\iota$ is dense in $\overline{X}$. Indeed, fix some $[\{x_n\}] \in \overline{X}$ and $\varepsilon > 0$. Then there is some N so that $n, m > N$ has
$$d(x_n,x_m) < \varepsilon/2.$$
Fixing a particular n_0 with $n_0 > N$, we set $x \coloneqq x_{n_0}$ so that
$$\overline{d}([\{x_n\}],\iota(x)) = \widetilde{d}(\{x_n\},x_{n_0}) = \lim_{n\to\infty} d(x_n,x_{n_0}).$$
Now, for $n > N$, we have $d(x_n,x_{n_0}) < \varepsilon/2$, so we conclude that this limit must be less than ε.
- We show that $(\overline{X},\overline{d})$ is a complete metric space. Fix a Cauchy sequence $\{\overline{x}_k\}$ in $\overline{X}$. To find the Cauchy sequence we are supposed to converge to, we use our density result: for each $k \in \mathbb{N}$, we can find $y_k \in X$ such that $\overline{d}(\overline{x}_k,\iota(y_k)) < 1/k$.
We claim that $\{y_k\}$ is Cauchy. Indeed, for any $\varepsilon > 0$, we can find N such that $k, \ell > N_0$ has
$$\overline{d}(\overline{x}_k,\overline{x}_\ell) < \varepsilon/3.$$
Then, setting $N \coloneqq \max\{3/\varepsilon, N_0\}$, we note that $k, \ell > N$ has
$$d(y_k,y_\ell) = \overline{d}(\iota(y_k),\iota(d_\ell)) \le \overline{d}(\overline{x}_k,\iota(y_k)) + \overline{d}(\overline{x}_\ell,\iota(y_\ell)) + \overline{d}(\overline{x}_k,\overline{x}_\ell) < \varepsilon.$$
Lastly, we claim that $\overline{x}_k \to [\{y_n\}]$ in $\overline{X}$. Indeed, for any $\varepsilon > 0$, find some sufficiently large N so that
$$k, \ell > N \implies d(y_k,y_\ell) < \varepsilon/2.$$
Then $k > \max\{N, 2/\varepsilon\}$ has
$$\overline{d}(\overline{x}_k,[\{y_n\}]) \le \overline{d}(\overline{x}_k,\iota(y_k)) + \overline{d}([\{y_n\}],\iota(y_k)) < \frac{\varepsilon}{2} + \lim_{n\to\infty} d(y_n,y_k).$$
Because $k > N$, we have $d(y_n,y_k) < \varepsilon/2$ for any $n > N$, so the entire right-hand side must be upper-bounded by ε. This finishes.

The above checks complete the proof. ■

Remark 1.72 (Nir)**.** One might complain that we used the completeness of $\mathbb{R}$ in this proof because one common way to construct the real numbers is as the completion of $\mathbb{Q}$ under the Euclidean metric. To remedy this, one ought to define the equivalence relation on Cauchy sequences more directly, saying that two Cauchy sequences $\{x_n\}_{n\in\mathbb{N}}$ and $\{y_n\}_{n\in\mathbb{N}}$ of real numbers are equivalent under $\sim$ if and only if

$$\lim_{n\to\infty} d_{\mathbb{R}}(x_n, y_n) = 0.$$

1.2.7 Uniqueness of Completions

We now show that any two completions of a metric space (X, d) are isometric, which is our uniqueness result. Here is the main intermediate result.

Lemma 1.73. Fix a metric space (X, d) and a completion $(\overline{X}, \overline{d})$ with its isometry $\iota\colon (X, d) \to (\overline{X}, \overline{d})$. Then, for any complete metric space (Y, d') and isometry $\varphi\colon (X, d) \to (Y, d')$, there is a unique isometry $\psi\colon (\overline{X}, \overline{d}) \to (Y, d')$ making the following diagram commute.

$$\begin{tikzcd} X \arrow[r, "\iota"] \arrow[dr, "\varphi"'] & \overline{X} \arrow[d, dashed, "\psi"] \\ & Y \end{tikzcd}$$

Proof. We start by showing the uniqueness of ψ. Well, for any $\overline{x} \in \overline{X}$, note that any $n \in \mathbb{N}$ allows us to find $x_n \in X$ with

$$\overline{d}(\overline{x}, \iota(x_n)) < 1/n$$

because $\operatorname{im}\iota$ is dense in $\overline{X}$. Now, we notice that $\iota(x_n) \to \overline{x}$ as $n \to \infty$ because any $\varepsilon > 0$ can set $N = 1/\varepsilon$. As such, we see that Lemma 1.57 applied to any possible $\psi\colon \overline{X} \to Y$ forces

$$\psi(\overline{x}) = \psi\Big(\lim_{n\to\infty} \iota(x_n)\Big) = \lim_{n\to\infty} \psi(\iota(x_n)) = \lim_{n\to\infty} \varphi(x_n).$$

Note that, a priori, we do not know if the sequence $\{\varphi(x_n)\}_{n\in\mathbb{N}}$ converges, but this argument tells us that it must; the limit is unique by Remark 1.55, so $\psi(\overline{x})$ is unique as well.

We now show that ψ exists. As before, any $\overline{x} \in \overline{X}$ can find a sequence $\{x_n\} \subseteq X$ such that $\iota(x_n) \to \overline{x}$ as $n \to \infty$. Thus, we note that $\{\varphi(x_n)\}$ is Cauchy by Lemma 1.62, so the completeness of Y gives it a limit; we set

$$\psi(\overline{x}) \coloneqq \lim_{n\to\infty} \varphi(x_n).$$

We have the following checks on ψ.

- Well-defined: if we have two sequences $\{x_n\}$ and $\{x'_n\}$ such that $\iota(x_n) \to x$ and $\iota(x'_n) \to x$ as $n \to \infty$, we need to show that

  $$\lim_{n\to\infty} \varphi(x_n) = \lim_{n\to\infty} \varphi(x'_n).$$

  For brevity, set y and y' to be the limits of $\{\varphi(x_n)\}$ and $\{\varphi(x'_n)\}$, respectively. Then, for any $\varepsilon > 0$, we note that there is a sufficiently large N such that

  $$n > N \implies d_Y(y, \varphi(x_n)), d_Y(y', \varphi(x'_n)) < \varepsilon/4.$$

  Further, we can make N even larger so that

  $$n > N \implies \overline{d}(\overline{x}, \iota(x_n)), \overline{d}(\overline{x}, \iota(x'_n)) < \varepsilon/4.$$

As such, any $n > N$ has

$$\begin{aligned} d_Y(y,y') &\le d_Y(y,\varphi(x_n)) + d_Y(\varphi(x_n),\varphi(x_n')) + d_Y(y',\varphi(x_n')) \\ &< \varepsilon/4 + d_X(x_n,x_n') + \varepsilon/4 \\ &= \varepsilon/2 + \overline{d}(\iota(x_n),\iota(x_n')) \\ &\le \varepsilon/2 + \overline{d}(\overline{x},\iota(x_n)) + \overline{d}(\overline{x},\iota(x_n')) \\ &< \varepsilon. \end{aligned}$$

It follows $d_Y(y,y') = 0$, so $y = y'$.

- Isometry: given $\overline{x},\overline{x}' \in \overline{X}$, find sequences $\{x_n\}$ and $\{x_n'\}$ in X so that $\iota(x_n) \to \overline{x}$ and $\iota(x_n') \to \overline{x}'$ as $n \to \infty$. Thus,

$$\begin{aligned} d_Y(\psi(\overline{x}),\psi(\overline{x}')) &= d_Y\Big(\lim_{n\to\infty}\varphi(x_n), \lim_{n\to\infty}\varphi(x_n')\Big) \\ &\overset{*}{=} \lim_{n\to\infty} d_Y(\varphi(x_n),\varphi(x_n')) \\ &= \lim_{n\to\infty} d(x_n,x_n') \\ &= \lim_{n\to\infty} \overline{d}(\iota(x_n),\iota(x_n')) \\ &= \overline{d}\Big(\lim_{n\to\infty}\iota(x_n), \lim_{n\to\infty}\iota(x_n')\Big) \\ &\overset{*}{=} \overline{d}(\overline{x},\overline{x}'), \end{aligned}$$

where we have used Remark 1.71 at the $\overset{*}{=}$.

- For any $x \in X$, we see that the (constant) Cauchy sequence $\{\iota(x)\}$ converges to $\iota(x)$, so

$$\psi(\iota(x)) = \lim_{n\to\infty}\varphi(x) = \varphi(x).$$

It follows $\psi \circ \iota = \varphi$.

Thus, we have finished establishing the existence of an isometry $\psi\colon \overline{X} \to Y$ such that $\varphi = \psi \circ \iota$. ■

Remark 1.74. One can also replace all isometries with uniformly continuous functions in the statement.

And here is our uniqueness result.

Theorem 1.75. Fix a metric space (X,d) and two completions $\iota\colon (X,d) \to (\overline{X},\overline{d})$ and $\iota'\colon (X,d) \to (\overline{X}',\overline{d}')$. Then there is a surjective isometry $\psi\colon (\overline{X},\overline{d}) \to (\overline{X}',\overline{d}')$.

Proof. Applying Lemma 1.73 twice, we get isometries $\psi\colon (\overline{X},\overline{d}) \to (\overline{X}',\overline{d}')$ and $\psi'\colon (\overline{X}',\overline{d}') \to (\overline{X},\overline{d})$ making the following diagrams commute.

$$\begin{tikzcd} X \arrow[r, "\iota"] \arrow[dr, "\iota'"'] & \overline{X} \arrow[d, "\psi"] \\ & \overline{X}' \end{tikzcd} \qquad \begin{tikzcd} X \arrow[r, "\iota'"] \arrow[dr, "\iota"'] & \overline{X}' \arrow[d, "\psi'"] \\ & \overline{X} \end{tikzcd}$$

In particular, we see that $\psi' \circ \psi$ makes the following diagram commute.

$$\begin{tikzcd} X \arrow[r, "\iota"] \arrow[dr, "\iota"'] & \overline{X} \arrow[d, "\psi'\circ\psi"] \\ & \overline{X} \end{tikzcd}$$

However, using Lemma 1.73 again, this isometry $\psi' \circ \psi$ is unique to make the diagram commute, and we could of course put the isometry $\mathrm{id}_{\overline{X}}$ here if we wanted to. Thus,

$$\psi' \circ \psi = \mathrm{id}_{\overline{X}}.$$

By symmetry, $\psi \circ \psi' = \mathrm{id}_{\overline{X'}}$, so we do see that ψ and ψ' are inverse isometries. This finishes the proof. ■

1.3 August 29

Good morning everyone.

1.3.1 Some Examples

Let's give some more examples of metric spaces. Let's start with spaces of continuous functions.

Definition 1.76. Given a (normed) topological field k, such as $\mathbb{R}$ or $\mathbb{C}$, we denote the k-vector space of k-valued continuous function from a topological space X as $C(X)$. By convention, we will take $k = \mathbb{C}$ unless otherwise specified.

And here are our two examples. The first is of a complete metric space.

Exercise 1.77. Give $V := C([0, 1])$ the uniform norm

$$\|f\|_\infty := \sup\{|f(t)| : t \in [0, 1]\}.$$

Then V is complete.

Proof. This is merely the statement that a sequence of continuous functions which are uniformly Cauchy will converge uniformly to a continuous function. We will prove this for completeness. Fix a sequence of continuous function $\{f_n\}_{n\in\mathbb{N}}$ which are Cauchy with respect to $\|\cdot\|_\infty$. In other words, for each $\varepsilon > 0$, there exists N_ε so that

$$n, m > N_\varepsilon \implies \|f_n - f_m\|_\infty < \varepsilon,$$

which means that $|f_n(t) - f_m(t)| < \varepsilon$ for all $t \in [0, 1]$.

In particular, for any fixed $t \in [0, 1]$, the sequence $\{f_n(t)\}_{n\in\mathbb{N}}$ is Cauchy in $\mathbb{R}$ (using the same N_ε), so we use the completeness of $\mathbb{R}$ to let this sequence converge to $f(t) \in \mathbb{R}$. We have the following checks.

- To see that $f_n \to f$ as $n \to \infty$ (under our metric), select any $\varepsilon > 0$, and then find N so that

  $$n, m > N \implies \|f_n - f_m\|_\infty < \varepsilon/3.$$

  Further, for any $t \in [0, 1]$, we see that we can find a large enough $n_t > N$ so that $|f(t) - f_{n_t}(t)| < \varepsilon/3$. But then $n > N$ has

  $$|f_n(t) - f(t)| \le |f_n(t) - f_{n_t}(t)| + |f_{n_t}(t) - f(t)| < 2\varepsilon/3,$$

  so $\|f - f_n\|_\infty \le 2\varepsilon/3 < \varepsilon$.

- To see that f is continuous, fix $t \in [0, 1]$ so that we want to show f is continuous at t. Well, for any $\varepsilon > 0$, find N large enough so that

  $$n, m > N \implies \|f_n - f_m\|_\infty < \varepsilon/4.$$

  Now, select $n_t > N$ large enough so that $|f(t) - f_{n_t}(t)| < \varepsilon/4$, and the continuity of f_{n_t} promises us $\delta > 0$ so that

  $$|t - t'| < \delta \implies |f_{n_t}(t) - f_{n_t}(t')| < \varepsilon/4.$$

In particular, for any t' with $|t-t'|<\delta$, find $n_{t'}>N$ large enough so that $|f(t')-f_{n_{t'}}(t')|<\varepsilon/4$, and then we see

$$|f(t)-f(t')|\le|f(t)-f_{n_t}(t)|+|f_{n_t}(t)-f_{n_t}(t')|+|f_{n_t}(t')-f_{n_{t'}}(t')|+|f_{n_{t'}}(t')-f(t')|<\varepsilon,$$

which is what we wanted. ■

The second example is the same space, but it is no longer complete.

Example 1.78. Fix $p\ge1$ finite. Give $V:=C([0,1])$ the L^p norm as

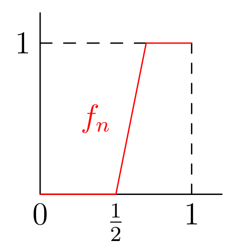
$$\|f\|_p:=\left(\int_0^1|f(t)|^p\,dt\right)^{1/p}.$$

Then V is not complete.

Proof. For each $n\ge2$, define f_n as the piecewise continuous function

$$f_n(t):=\begin{cases}0 & 0\le t\le\frac12,\\ n(t-\frac12) & \frac12\le t\le\frac12+\frac1n,\\ 1 & \frac12+\frac1n\le t\le1.\end{cases}$$

Here is the image.



The point is that f_n is trying to converge to a discontinuous function. To help us with the proof here, we pick up the following lemma.

Lemma 1.79. Fix $V:=C([0,1])$ and some finite $p\ge1$. If we have a convergent sequence $f_n\to f$ as $n\to\infty$ in the $\|\cdot\|_p$ metric, and $f_n(t)=g(t)$ for all sufficiently large n and $t\in U$ for some open $U\subseteq C([0,1])$, then $f|_U(t)=g(t)$.

Proof. Suppose for the sake of contradiction that we have $t_0\in U$ with $f(t_0)\ne g(t_0)$; we show that $\{f_n\}$ does not converge to f. Set $\varepsilon:=|f(t_0)-g(t_0)|$, which is nonzero. The continuity of $f-g$ now promises that there is $\delta>0$ for which

$$|t-t_0|<\delta\implies|(f-g)(t_0)-(f-g)(t)|<\varepsilon/2,$$

so in particular $|(f-g)(t)|\ge\varepsilon/2$. It follows that, for sufficiently large n, we have

$$\|f-f_n\|_p^p=\int_0^1|f(t)-f_n(t)|^p\,dt\ge\int_U|(f-g)(t)|\,dt\ge\int_{U\cap(t_0-\delta,t_0+\delta)}\frac{\varepsilon}{2}\,dt.$$

Because $U\cap(t_0-\delta,t_0+\delta)$ is open, it has nonzero measure, so this entire right-hand quantity is nonzero, thus violating that $f_n\to f$ as $n\to\infty$. ■

Now suppose for the sake of contradiction that $f_n \to f$ as $n \to \infty$ for some $f \in V$. Then, using $U = (0, 1/2)$, we conclude that $f(t) = 0$ for all $t \in (0, 1/2)$. Similarly, for any n, we set $U_n = (1/2 + 1/n, 1)$, so $f_m|_{U_n}$ returns 1 always for sufficiently large m; this then implies $f(t) = 1$ for any $t \in U_n$ for any n, so $f(t) = 1$ for any $t \in (1/2, 1)$.

However, the sequences $a_n := \frac{1}{2} - \frac{1}{n}$ and $b_n := \frac{1}{2} + \frac{1}{n}$ (for $n \geq 3$) have $a_n \to \frac{1}{2}$ and $b_n \to \frac{1}{2}$ both as $n \to \infty$ while the continuity of f would require

$$0 = \lim_{n\to\infty} f(a_n) = f(1/2) = \lim_{n\to\infty} f(b_n) = 1,$$

which is a contradiction. ■

Remark 1.80. In an attempt to make this metric space complete, we can try to specify which functions we want to look at, which motivates the theory of measure and integration.

Remark 1.81. The $\|\cdot\|_2$ norm on $C(X)$ for some (say) subset $X \subseteq \mathbb{R}$ with finite measure as coming from an inner product

$$\langle f, g\rangle := \int_X f(t)\overline{g(t)}\,dt.$$

When $\|\cdot\|_2$ is complete, we would then get a Hilbert space, which are very nice normed vector spaces, and we'll see more of them in Math 202B.

Remark 1.82 (Nir)**.** In contrast to the finite case, we see that the $\|\cdot\|_\infty$ norm induces a different (metric) topology on $C([0, 1])$ than the $\|\cdot\|_p$ norms with p finite because the former is complete while the latter are not. In fact, all the norms $\|\cdot\|_p$ induce different topologies on $C([0, 1])$.

Part I

Topology

THEME 2

BUILDING TOPOLOGIES

Sets are not doors.

—Munkres

2.1 August 29

We continue lecture by shifting to topology.

2.1.1 Metric Topology

We close our discussion of metric spaces with a taste of topology. Recall the following definition.

Definition 1.49 (Continuous)**.** Given metric spaces (X, d_X) and (Y, d_Y), a function $f\colon X \to Y$ is *continuous at* $x \in X$ if and only if all $\varepsilon > 0$ have some $\delta_x > 0$ such that

$$d_X(x, x') < \delta_x \implies d_Y(f(x), f(x')) < \varepsilon.$$

Then f is *continuous* if and only if it is continuous at all $x \in X$.

We are going to want to extend this definition to more general topological spaces. To step in that direction, we will want to talk about open sets, so we start with open balls.

Definition 2.1 (Ball)**.** Fix a metric space (X, d). Then the *open ball of radius* r *centered at* $x_0 \in X$ is

$$B(x_0, r) \coloneqq \{x \in X : d(x, x_0) < r\}.$$

The *closed ball* is $\overline{B(x_0, r)} \coloneqq \{x \in X : d(x, x_0) \leq r\}$.

We can now restate continuity as follows.

Definition 2.2 (Continuous)**.** Given metric spaces (X, d_X) and (Y, d_Y), a function $f\colon X \to Y$ is *continuous* at $x \in X$ if and only if, given any nonempty open ball $B(f(x_0), \varepsilon)$, there exists a nonempty open ball $B(x_0, \delta)$ such that

$$f(B(x_0, \delta)) \subseteq B(f(x_0), \varepsilon).$$

Namely, we've really only restated our inequalities.

To continue our generalization, we define the pre-image.

Definition 2.3 (Pre-image)**.** Fix a function $f\colon X \to Y$. Then we define the *pre-image* $f^{-1}\colon \mathcal{P}(Y) \to \mathcal{P}(X)$ by

$$f^{-1}(B) := \{x \in X : f(x) \in B\}.$$

Note that our pre-image notation matches with the notation of an inverse function. In general, no confusion will arise by confusing these two.

As such, let's restate continuity again: observe that $A \subseteq X$ and $B \subseteq Y$ has $f(A) \subseteq B$ if and only if all $a \in A$ have $f(a) \in B$ if and only if all $a \in A$ have $a \in f^{-1}(B)$ if and only if $A \subseteq f^{-1}(B)$.

Definition 2.4 (Continuous)**.** Given metric spaces (X, d_X) and (Y, d_Y), a function $f\colon X \to Y$ is *continuous* at $x \in X$ if and only if, given any nonempty open ball $B(f(x), \varepsilon)$, there exists a nonempty open ball $B(x, \delta)$ such that

$$B(x, \delta) \subseteq f^{-1}(B(f(x), \varepsilon)).$$

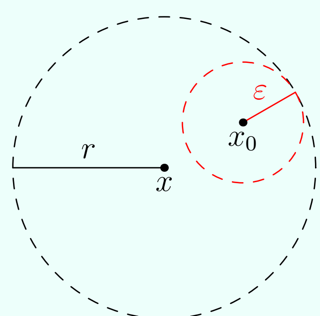
We defined open balls and promised open sets, so now let's define our open sets.

Definition 2.5 (Open set)**.** Fix a metric space (X, d). Then a subset $U \subseteq X$ is *open* if and only if, for each $x \in U$, there exists some $\varepsilon > 0$ such that $B(x, \varepsilon) \subseteq U$. In other words, each point in U has an open ball around it.

Example 2.6. Open balls are open sets. Indeed, given an open ball $B(x, r)$, note that any $x_0 \in B(x, r)$ has $d(x_0, x) < r$, so we take $\varepsilon := r - d(x_0, x)$. To see this works, observe $x' \in B(x_0, \varepsilon)$ will have

$$d(x', x) \leq d(x', x_0) + d(x_0, x) < \varepsilon + (r - \varepsilon) = r,$$

so $B(x_0, \varepsilon) \subseteq B(x, r)$ follows. Here is the image for what just happened.



And here is our definition of corresponding definition of continuity.

Lemma 2.7. Given metric spaces (X, d_X) and (Y, d_Y), a function $f\colon X \to Y$ is continuous at $x \in X$ if and only if, given any open set $U \subseteq Y$ with $f(x) \in U$, there is an open ball $B(x, \delta)$, such that

$$B(x, \delta) \subseteq f^{-1}(U).$$

Proof. Taking f to be continuous, note that we can find $\varepsilon > 0$ such that $B(f(x), \varepsilon) \subseteq U$ because U is open. Thus, continuity promises $\delta > 0$ such that

$$B(x, \delta) \subseteq f^{-1}(B(f(x), \varepsilon)) \subseteq f^{-1}(U).$$

Conversely, if f satisfies the conclusion of the statement, we can take $U = B(f(x), \varepsilon)$ for any $\varepsilon > 0$ by Example 2.6, and the conclusion promises $\delta > 0$ such that

$$B(x, \delta) \subseteq f^{-1}(U) = f^{-1}(B(f(x), \varepsilon)),$$

which is what we wanted. ■

It is cleaner to talk about the entire function being continuous instead of at a point.

Lemma 2.8. Given metric spaces (X, d_X) and (Y, d_Y), a function $f\colon X \to Y$ is continuous if and only if, given any open set $U \subseteq Y$ with $f(x) \in U$, the pre-image $f^{-1}(U)$ is open.

Proof. This is a matter of rearranging our quantifiers correctly. Lemma 2.7 tells us that, for all $x \in X$, all open $U \subseteq Y$ with $f(x) \in U$ has some $\delta > 0$ such that $B(x, \delta) \subseteq U$. Equivalently, for all open $U \subseteq Y$, any $x \in X$ with $x \in f^{-1}(U)$ has some $\delta > 0$ such that $B(x, \delta) \subseteq U$. But by definition of being open, we're just saying that all open $U \subseteq Y$ has $f^{-1}(U)$ also open. ■

So we have the following definition.

Definition 2.9 (Continuous). A function $f\colon X \to Y$ between metric spaces is *continuous* if and only if, for any open set $U \subseteq Y$, the pre-image $f^{-1}(U)$ is open.

The philosophy here is to try to understand open sets instead of trying to understand the metrics. This is the idea of topology.

2.1.2 Open Sets

Thus, we are motivated to understand open sets. Here are some basic properties.

Proposition 2.10. Fix a metric space (X, d), and let $\mathcal{T}$ be the collection of open sets.

(a) We have $X \in \mathcal{T}$ and $\varnothing \in \mathcal{T}$.

(b) Arbitrary union: given a collection $\mathcal{U} \subseteq \mathcal{T}$, the arbitrary union

$$\bigcup_{U \in \mathcal{U}} U$$

is open.

(c) Finite intersection: given a finite collection $\{U_1, \ldots, U_n\} \in \mathcal{T}$, we have

$$\bigcap_{i=1}^{n} U_i$$

is open.

Proof. We go in sequence.

(a) To show $X \in \mathcal{T}$, note that any $x \in X$ has $B(x, 1) \subseteq X$ by definition. To show $\varnothing \in \mathcal{T}$, note that any $x \in \varnothing$ has $B(x, 1) \subseteq \varnothing$ because there is no $x \in \varnothing$ at all.

(b) For any $x \in \bigcup_{U \in \mathcal{U}} U$, we have $x \in V$ for some particular $V \in \mathcal{U}$. Then the openness of V tells us we can find $\varepsilon > 0$ such that

$$B(x, \varepsilon) \subseteq V \subseteq \bigcup_{U \in \mathcal{U}} U,$$

which finishes.

(c) Fix x in the common intersection. Then, for any i, we have $x \in U_i$, so we have some $\varepsilon_i > 0$ such that $B(x, \varepsilon_i) \subseteq U$, and so we set

$$\varepsilon := \min_{1 \le i \le n} \varepsilon_i.$$

In particular, $\varepsilon > 0$ because n is finite, and we have

$$B(x, \varepsilon) \subseteq B(x, \varepsilon_i) \subseteq U_i$$

for each i, so $B(x, \varepsilon)$ is a subset of our intersection. ■

Remark 2.11. The arbitrary intersection of open sets need not be open: working in $\mathbb{R}$ with the usual metric,

$$\bigcap_{i=1}^{\infty} B(0, 1/n) = \{0\},$$

which is not open. (Namely, no $\varepsilon > 0$ has $B(x, \varepsilon) \subseteq \{0\}$.)

Motivated by Proposition 2.10, we have the following definition.

Definition 2.12 (Topology)**.** Fix a set X. Then a *topology* $\mathcal{T}$ on X is a collection of subsets $\mathcal{T} \subseteq \mathcal{P}(X)$ satisfying the following.

(a) We have $\varnothing \in \mathcal{T}$ and $X \in \mathcal{T}$.

(b) Arbitrary union: given a collection $\mathcal{U} \subseteq \mathcal{T}$, the arbitrary union $\bigcup_{U \in \mathcal{U}} U$ lives in $\mathcal{T}$.

(c) Finite intersection: given a finite collection $\{U_1, \ldots, U_n\} \subseteq \mathcal{T}$, the intersection $\bigcap_{i=1}^{n} U_i$ lives in $\mathcal{T}$.

We will say that the ordered pair $(X, \mathcal{T})$ is a *topological space*. We say that the sets in $\mathcal{T}$ are *open*.

Example 2.13. By Proposition 2.10, metric spaces with their open sets form a topological space.

Here are some more basic examples.

Definition 2.14 (Discrete topology)**.** Given a set X, the *discrete topology* is the topology $\mathcal{P}(X)$.

Definition 2.15 (Indiscrete topology)**.** Given a set X, the *indiscrete topology* is the topology $\{\varnothing, X\}$.

It is fairly routine to check that the above collections form topologies. In fact, they are closed under both arbitrary union and arbitrary intersection.

Remark 2.16. The discrete topology can be defined by the metric $d \colon X \times X \to \mathbb{R}_{\ge 0}$ by

$$d(x, x') := \begin{cases} 1 & x \ne x', \\ 0 & x = x'. \end{cases}$$

Indeed, for any $x \in X$, we see $B(x, 1/2) = \{x\}$, so any subset $U \subseteq X$ is the open set

$$U = \bigcup_{x \in U} \{x\} = \bigcup_{x \in U} B(x, 1/2).$$

Remark 2.17. If $\#X \geq 2$, the indiscrete topology cannot be given a metric. Indeed, find distinct points $a, b \in X$ and set $r := d(a,b)$, so $a \neq b$ implies $r > 0$. Now, $a \in B(a,r)$, but $b \notin B(a,r)$, so $B(a,r)$ is an open set distinct from both $\varnothing$ and X.

Remark 2.18. One can give topologies a partial order by inclusion. Then the discrete topology is the maximal one (definitionally, any topology is a subset of $\mathcal{P}(X)$), and the indiscrete topology is the minimal one (definitionally, any topology contains $\varnothing$ and X).

And so here is our general definition of continuity.

Definition 2.19 (Continuous)**.** Fix topological spaces $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$. Then a function $f\colon X \to Y$ is *continuous* if and only if, for any $U_Y \in \mathcal{T}_Y$, we have $f^{-1}(U_Y) \in \mathcal{T}_X$.

2.2 August 31

It is once again the morning.

2.2.1 Intersections of Topologies

We will want to have lots of topologies to work with. Here is a basic way to build them.

Proposition 2.20. Let X be a set, and pick up some collection of topologies $\{\mathcal{T}_\alpha\}_{\alpha\in\lambda}$. Then the intersection

$$\mathcal{T} := \bigcap_{\alpha\in\lambda} \mathcal{T}_\alpha$$

is also a topology on X.

Proof. This is mostly a matter of writing out the axioms.

(a) Note that $\varnothing, X \in \mathcal{T}_\alpha$ for each α, so $\varnothing, X \in \mathcal{T}$.

(b) Arbitrary union: given a collection $\mathcal{U} \subseteq \mathcal{T}$, we have $\mathcal{U} \subseteq \mathcal{T}_\alpha$ for each α, so $\bigcup_{U\in\mathcal{U}} U \in \mathcal{T}_\alpha$ for each α, so

$$\bigcup_{U\in\mathcal{U}} U \in \mathcal{T}$$

as well.

(c) Finite intersection: given a finite collection $\{U_1, \ldots, U_n\} \subseteq \mathcal{T}$, we have $\{U_1, \ldots, U_n\} \subseteq \mathcal{T}_\alpha$ for each α, so $\bigcap_{i=1}^n U_i \in \mathcal{T}_\alpha$ for each α, so

$$\bigcap_{i=1}^n U_i \in \mathcal{T}$$

follows. ■

Corollary 2.21. Fix a set X. Given a collection $\mathcal{S} \subseteq \mathcal{P}(X)$, there is a smallest topology $\mathcal{T}$ containing $\mathcal{S}$.

Proof. Certainly there is some topology containing $\mathcal{S}$, namely the discrete topology $\mathcal{P}(X)$. Thus, we can set our topology to be

$$\mathcal{T}(\mathcal{S}) := \bigcap_{\substack{\mathcal{T}\supseteq\mathcal{S} \\ \mathcal{T}\text{ a topology}}} \mathcal{T},$$

which is a topology (by Proposition 2.20) which contains $\mathcal{S}$ (because each topology in the intersection contains $\mathcal{S}$), and of course any topology $\mathcal{T}$ containing $\mathcal{S}$ will have $\mathcal{T}(\mathcal{S}) \subseteq \mathcal{T}$. ■

To codify this idea, we have the following idea.

Definition 2.22 (Generated topology)**.** Fix a set X. We say that a collection $\mathcal{S} \subseteq \mathcal{P}(X)$ *generates* its smallest topology $\mathcal{T}$. We will write $\mathcal{T}(\mathcal{S})$ for this topology.

Remark 2.23 (Nir)**.** The topology $\mathcal{T}(\mathcal{S})$ is unique. Indeed, suppose two topologies $\mathcal{T}$ and $\mathcal{T}'$ are minimal topologies containing $\mathcal{S}$. Then $\mathcal{T}\cap\mathcal{T}'$ is also a topology containing $\mathcal{S}$ by Proposition 2.20, but $\mathcal{T}\cap\mathcal{T}' \subseteq \mathcal{T},\mathcal{T}'$ forces $\mathcal{T} = \mathcal{T}\cap\mathcal{T}' = \mathcal{T}'$.

Remark 2.24 (Nir)**.** Given collections $\mathcal{S} \subseteq \mathcal{S}'$, then $\mathcal{T}(\mathcal{S}) \subseteq \mathcal{T}(\mathcal{S}')$. Indeed, we have

$$\mathcal{T}(\mathcal{S}) = \bigcap_{\substack{\mathcal{T}\supseteq\mathcal{S} \\ \mathcal{T}\text{ a topology}}} \mathcal{T} \subseteq \bigcap_{\substack{\mathcal{T}\supseteq\mathcal{S}' \\ \mathcal{T}_0\text{ a topology}}} \mathcal{T} = \mathcal{T}(\mathcal{S}').$$

Remark 2.25 (Nir)**.** If $\mathcal{T}$ is already a topology on X, then $\mathcal{T}(\mathcal{T}) = \mathcal{T}$. Indeed, of course $\mathcal{T} \subseteq \mathcal{T}(\mathcal{T})$, but then also

$$\mathcal{T}(\mathcal{T}) = \bigcap_{\substack{\mathcal{T}'\supseteq\mathcal{T} \\ \mathcal{T}'\text{ a topology}}} \mathcal{T}' \subseteq \mathcal{T}$$

because $\mathcal{T}$ is a topology containing $\mathcal{T}$.

2.2.2 Sub-bases

On the other side of things, we pick up the following definition.

Definition 2.26 (Sub-base)**.** Let $(X,\mathcal{T})$ be a topological space. A collection $\mathcal{S} \subseteq \mathcal{T}$ is a *sub-base* for $\mathcal{T}$ if and only if the following hold.

(a) $\mathcal{S}$ covers X, in that $X = \bigcup_{U\in\mathcal{S}} U$.

(b) $\mathcal{T}$ is generated by $\mathcal{S}$.

The point is that collections $\mathcal{S}$ are easy to find, so we have therefore found many topologies.

It will be useful to give a more concrete description of the topology generated by a collection $\mathcal{S}$. We start by taking finite intersections.

Lemma 2.27. Fix a set X and a collection $\mathcal{S} \subseteq \mathcal{P}(X)$ with $X = \bigcup_{U\in\mathcal{S}} U$. Then set

$$\mathcal{I}^{\mathcal{S}} := \left\{ \bigcap_{i=1}^n U_i : \{U_i\}_{i=1}^n \subseteq \mathcal{S} \right\}.$$

Then $\mathcal{S} \subseteq \mathcal{I}^{\mathcal{S}}$ and $\mathcal{I}^{\mathcal{S}}$ is closed under finite intersection. Further, the topology generated by $\mathcal{I}^{\mathcal{S}}$ is also the topology generated by $\mathcal{S}$.

Proof. We show the claims in sequence

- That $\{U\} \subseteq \mathcal{S}$ for any $U \in \mathcal{S}$ implies that $U \in \mathcal{I}^{\mathcal{S}}$ for any $U \in \mathcal{S}$, so $\mathcal{S} \subseteq \mathcal{I}^{\mathcal{S}}$ follows.
- To show $\mathcal{I}^{\mathcal{S}}$ is closed under finite intersection, pick up some finite collection $\{U_1, \ldots, U_n\} \subseteq \mathcal{I}^{\mathcal{S}}$. Then, for each i, we can find some finite collection $\mathcal{U}_i \subseteq \mathcal{S}$ such that

$$U_i = \bigcap_{V \in \mathcal{U}_i} V.$$

Setting $\mathcal{U} := \bigcup_{i=1}^n \mathcal{U}_i$, we see that $\mathcal{U}$ is finite and that

$$\bigcap_{i=1}^n U_i = \bigcap_{i=1}^n \bigcap_{V \in \mathcal{U}_i} V = \bigcap_{V \in \mathcal{U}} V$$

must live in $\mathcal{I}^{\mathcal{S}}$.

- Because $\mathcal{S} \subseteq \mathcal{I}^{\mathcal{S}}$, Remark 2.24 tells us $\mathcal{T}(\mathcal{S}) \subseteq \mathcal{T}\left(\mathcal{I}^{\mathcal{S}}\right)$. In the other direction, note that any finite collection $\{U_1, \ldots, U_n\} \subseteq \mathcal{S}$ also lives in $\mathcal{T}(\mathcal{S})$, so

$$\bigcap_{i=1}^n U_i \in \mathcal{T}(\mathcal{S}).$$

It follows $\mathcal{I}^{\mathcal{S}} \subseteq \mathcal{T}(\mathcal{S})$, so $\mathcal{T}\left(\mathcal{I}^{\mathcal{S}}\right) \subseteq \mathcal{T}(\mathcal{T}(\mathcal{S})) = \mathcal{T}(\mathcal{S})$ by Remark 2.25. ■

After taking finite intersections, we take arbitrary unions.

Lemma 2.28. Fix a set X and a collection $\mathcal{I} \subseteq \mathcal{P}(X)$ closed under finite intersection with $\bigcup_{U \in \mathcal{I}} U = X$. Then the collection of (arbitrary) unions of elements in $\mathcal{I}$, denoted

$$\mathcal{T} := \left\{ \bigcup_{U \in \mathcal{U}} U : \mathcal{U} \subseteq \mathcal{I} \right\},$$

is $\mathcal{T}(\mathcal{I})$.

Proof. If $\mathcal{T}'$ is a topology containing $\mathcal{I}$, then note any collection $\mathcal{U} \subseteq \mathcal{I}$ lives in $\mathcal{T}'$, so the arbitrary union

$$\bigcup_{U \in \mathcal{U}} U$$

lives in $\mathcal{T}'$. It follows that $\mathcal{T} \subseteq \mathcal{T}'$, so

$$\mathcal{T} \subseteq \bigcap_{\substack{\mathcal{T}' \supseteq \mathcal{T} \\ \mathcal{T}' \text{ a topology}}} \mathcal{T}' = \mathcal{T}(\mathcal{I}).$$

Thus, it remains to show that $\mathcal{T}$ is in fact a topology, which will imply from $\mathcal{I} \subseteq \mathcal{T}$ that $\mathcal{T}(\mathcal{I}) \subseteq \mathcal{T}(\mathcal{T}) = \mathcal{T}$ by Remark 2.24. Here are our checks.

- Setting $\mathcal{U} = \varnothing \subseteq \mathcal{I}$, we see that $\bigcup_{U \in \mathcal{U}} U = \varnothing$, so $\varnothing \in \mathcal{T}$. Also, by hypothesis, we have

$$X = \bigcup_{U \in \mathcal{I}} U \in \mathcal{T}.$$

- Arbitrary union: let $\mathcal{U} \subseteq \mathcal{T}$ be a subcollection. For any $U \in \mathcal{U}$, we can find a collection $\mathcal{V}_U \subseteq \mathcal{I}$ such that
$$U = \bigcup_{V \in \mathcal{V}_U} V.$$
Now, we set $\mathcal{V}$ to be the union of all the collections of $\mathcal{V}_U$ for each $U \in \mathcal{U}$, which is still contained in $\mathcal{I}$, so that
$$\bigcup_{U \in \mathcal{U}} U = \bigcup_{U \in \mathcal{U}} \bigcup_{V \in \mathcal{V}_U} V = \bigcup_{V \in \mathcal{V}} V \in \mathcal{T}.$$
- Finite intersection: by induction, it suffices to pick up two sets $U, V \in \mathcal{T}$ and show $U \cap V \in \mathcal{T}$. Well, we can find collections $\mathcal{U}, \mathcal{V} \subseteq \mathcal{I}$ such that
$$U = \bigcup_{U' \in \mathcal{U}} U' \qquad \text{and} \qquad V = \bigcup_{V' \in \mathcal{V}} V',$$
from which it follows (by distribution) that
$$U \cap V = \left(\bigcup_{U' \in \mathcal{U}} U' \right) \cap \left(\bigcup_{V' \in \mathcal{V}} V' \right) = \bigcup_{U' \in \mathcal{U}} \left(U' \cap \bigcup_{V' \in \mathcal{V}} V' \right) = \bigcup_{\substack{U' \in \mathcal{U} \\ V' \in \mathcal{V}}} (U' \cap V').$$
Now, $\mathcal{I}$ is closed under finite intersection, so $U' \cap V' \in \mathcal{I}$, so we have witnessed $U \cap V$ as an arbitrary union of elements of $\mathcal{I}$, so $U \cap V \in \mathcal{T}$ follows. ■

Corollary 2.29. Fix a set X and a collection $\mathcal{S} \subseteq \mathcal{P}(X)$ with $X = \bigcup_{U \in \mathcal{S}} U$. Letting $\mathcal{I}^{\mathcal{S}}$ be the collection of finite intersections of $\mathcal{S}$ and then $\mathcal{T}$ be the collection of arbitrary unions of $\mathcal{I}^{\mathcal{S}}$, we have that $\mathcal{T} = \mathcal{T}(\mathcal{S})$.

Proof. By Lemma 2.27, we have $\mathcal{T}(\mathcal{S}) = \mathcal{T}\left(\mathcal{I}^{\mathcal{S}}\right)$. Plugging $\mathcal{I}^{\mathcal{S}}$ into Lemma 2.28 (which applies because $\mathcal{I}^{\mathcal{S}}$ is closed under finite intersection and covers X because $\mathcal{S} \subseteq \mathcal{I}^{\mathcal{S}}$), we see that $\mathcal{T}\left(\mathcal{I}^{\mathcal{S}}\right) = \mathcal{T}$, finishing. ■

We quickly point out that the point of discussing sub-bases is that we will be allowed to check continuity on only a sub-base.

Lemma 2.30. Fix a topological space $(X, \mathcal{T}_X)$ and a set Y. Given a function $f \colon X \to Y$, the collection
$$\mathcal{T}(f) := \left\{ U \subseteq Y : f^{-1}(U) \in \mathcal{T}_X \right\}$$
forms a topology on Y.

Proof. Here are our checks.

- Note $f^{-1}(\varnothing) = \varnothing \in \mathcal{T}_X$, so $\varnothing \in \mathcal{T}(f)$. Also, $f^{-1}(Y) = X \in \mathcal{T}_X$, so $Y \in \mathcal{T}(f)$.
- Arbitrary union: given a collection $\mathcal{U} \subseteq \mathcal{T}(f)$, we see that
$$f^{-1}\left(\bigcup_{U \in \mathcal{U}} U \right) = \bigcup_{U \in \mathcal{U}} f^{-1}(U)$$
is a union of elements of $\mathcal{T}_X$ and therefore in $\mathcal{T}_X$. Thus, $\bigcup_{U \in \mathcal{U}} U \in \mathcal{T}(f)$.
- Finite intersection: this is identical to the previous check. Given a finite collection $\{U_1, \ldots, U_n\} \in \mathcal{T}(f)$, we see that
$$f^{-1}\left(\bigcap_{i=1}^{n} U_i \right) = \bigcap_{i=1}^{n} f^{-1}(U_i)$$
is a finite intersection of elements of $\mathcal{T}_X$ and therefore in $\mathcal{T}_X$. Thus, $\bigcap_{i=1}^{n} U_i \in \mathcal{T}(f)$. ■

Proposition 2.31. Fix topological spaces $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$, and let $\mathcal{S}$ be a sub-base for $\mathcal{T}_Y$. Then a function $f\colon X \to Y$ is continuous if and only if

$$f^{-1}(U) \in \mathcal{T}_X$$

for all $U \in \mathcal{S}$.

Proof. Certainly if f is continuous then the pre-image of any open set $U \in \mathcal{S} \subseteq \mathcal{T}_Y$ must be open. On the other hand, let $\mathcal{T}(f) \subseteq \mathcal{P}(Y)$ be the collection of subsets U for which $f^{-1}(U) \in \mathcal{T}_X$. This is a topology by Lemma 2.30, and it contains $\mathcal{S}$ by hypothesis, so it follows

$$\mathcal{T}_Y = \mathcal{T}(\mathcal{S}) \subseteq \mathcal{T}(f).$$

Thus, $f^{-1}(U) \in \mathcal{T}_X$ for any $U \in \mathcal{T}_Y$, so f is continuous. ■

2.2.3 Bases

Having defined a sub-base, we should be rightly upset that we have not defined a base.

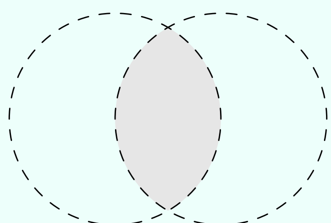
Definition 2.32 (Base)**.** Fix a set X. A collection $\mathcal{B} \subseteq \mathcal{P}(X)$ is a *base* (for a topology on X) if and only if the collection of arbitrary unions of $\mathcal{B}$ form a topology on X.

This definition is a little hard to access because we still don't have a good notion of what a topology is.

Example 2.33. Fix a set X. Given any collection $\mathcal{S} \subseteq \mathcal{P}(X)$, the collection of finite intersections $\mathcal{I}^{\mathcal{S}}$ is a base by Lemma 2.28.

However, in general we do not require a base to be closed under finite intersection.

Example 2.34. Fix a metric space (X, d). Then the collection of open balls $\mathcal{B}$ forms a topology by Example 2.13. Notably, the intersection of two open balls need not be an open ball, as follows.



Even though bases are not closed under finite intersection, we do have the following.

Proposition 2.35. Fix a set X and a collection $\mathcal{B} \subseteq \mathcal{P}(X)$. Then $\mathcal{B}$ is a base if and only if

(a) $X = \bigcup_{B \in \mathcal{B}} B$, and

(b) any $B_1, B_2 \in \mathcal{B}$ has some collection $\mathcal{U} \subseteq \mathcal{B}$ such that

$$B_1 \cap B_2 = \bigcup_{B \in \mathcal{U}} B.$$

Proof. In one direction, suppose that $\mathcal{B}$ is a base generating the topology $\mathcal{T}$.

(a) Because $X \in \mathcal{T}$, we see that X is the union of some subcollection $\mathcal{U} \subseteq \mathcal{B}$, so it follows

$$X = \bigcup_{U \in \mathcal{U}} U \subseteq \bigcup_{B \in \mathcal{B}} B \subseteq X.$$

(b) Given $B_1, B_2 \in \mathcal{B} \subseteq \mathcal{T}$, we see that $B_1 \cap B_2 \in \mathcal{T}$, so because $\mathcal{T}$ is made of arbitrary unions of $\mathcal{B}$, there is a collection $\mathcal{U} \subseteq \mathcal{B}$ such that

$$B_1 \cap B_2 = \bigcup_{B \in \mathcal{U}} B.$$

We now go in the other direction. Suppose $\mathcal{B}$ satisfies (a) and (b), and define

$$\mathcal{T} := \left\{ \bigcup_{U \in \mathcal{U}} U : \mathcal{U} \subseteq \mathcal{B} \right\}.$$

We now check that $\mathcal{T}$ is a topology.

- Using $\mathcal{U} = \varnothing \subseteq \mathcal{B}$, so we see that $\bigcup_{U \in \mathcal{U}} U = \varnothing$ is in $\mathcal{T}$. Also, by (a), we have

$$X = \bigcup_{B \in \mathcal{B}} B \in \mathcal{T}.$$

- Arbitrary union: this is the same as the check in Lemma 2.28. Given a collection $\mathcal{U} \subseteq \mathcal{T}$, each $U \in \mathcal{U}$ has some collection $\mathcal{V}_U \subseteq \mathcal{B}$ such that $\bigcup_{V \in \mathcal{V}_U} V = U$. Letting $\mathcal{V} \subseteq \mathcal{B}$ be the union of all the $\mathcal{V}_U$, we see

$$\bigcup_{U \in \mathcal{U}} U = \bigcup_{U \in \mathcal{U}} \bigcup_{V \in \mathcal{V}_U} V = \bigcup_{V \in \mathcal{V}} V$$

  lives in $\mathcal{T}$.

- Finite intersection: by induction, it suffices to pick up $U_1, U_2 \in \mathcal{T}$ and show $U_1 \cap U_2 \in \mathcal{T}$. Well, find $\mathcal{B}_1, \mathcal{B}_2 \subseteq \mathcal{B}$ such that

$$U_1 = \bigcup_{B_1 \in \mathcal{B}_1} B_1 \qquad \text{and} \qquad U_2 = \bigcup_{B_2 \in \mathcal{B}_2} B_2,$$

  which implies

$$U_1 \cap U_2 = \bigcup_{\substack{B_1 \in \mathcal{B}_1 \\ B_2 \in \mathcal{B}_2}} (B_1 \cap B_2).$$

  Now, (b) implies that $B_1 \cap B_2$ for any $B_1, B_2 \in \mathcal{B}$ is a union of elements in $\mathcal{B}$, so $B_1 \cap B_2 \in \mathcal{T}$. Thus, $U_1 \cap U_2$ is the arbitrary union of elements in $\mathcal{T}$, so $U_1 \cap U_2 \in \mathcal{T}$ by the previous check. ■

Remark 2.36 (Nir)**.** Careful readers might realize that we could rearrange the given exposition to show that, given a sub-base $\mathcal{S}$, the collection of finite intersections $\mathcal{I}^{\mathcal{S}}$ is a base instead of going through Lemma 2.28.

Remark 2.37. Of course, any base is also a sub-base. Notably, sub-bases only require that $X = \bigcup_{U \in \mathcal{S}} U$, which must be satisfied for bases.

Example 2.38. Set $X = \mathbb{R}$ with the usual topology $\mathcal{T}$. Then the collection $\mathcal{B}$ of open intervals (a, b) form a base for the usual topology (these are our open balls). In contrast, the collection

$$\mathcal{S} = \{(-\infty, a) : a \in \mathbb{R}\} \cup \{(a, \infty) : a \in \mathbb{R}\}$$

forms a sub-base for the usual topology. Namely, certainly $\mathcal{S} \subseteq \mathcal{T}$, and $\mathcal{B} \subseteq \mathcal{T}(\mathcal{S})$ because of the finite intersection $(-\infty, b) \cap (a, \infty) = (a, b)$ for any $a, b \in \mathbb{R}$. Namely, $\mathcal{T} = \mathcal{T}(\mathcal{B}) \subseteq \mathcal{T}(\mathcal{T}(\mathcal{S})) = \mathcal{T}(\mathcal{S})$ follows.

2.2.4 Induced Topologies

We start with the following motivating example.

Example 2.39. Fix a set X, and give it the discrete topology. Then, for any topological space $(Y, \mathcal{T}_Y)$, any function $f\colon X \to Y$ is continuous because the pre-image of any open subset $U_Y \subseteq Y$ is open in X.

In general, we might have some smallish collection of functions which we want to force to be continuous, so we might ask what topology is forced by their continuity.

Definition 2.40 (Induced topology)**.** Fix a set X and a collection of topologies $\{(Y_\alpha, \mathcal{T}_\alpha)\}_{\alpha\in\lambda}$ with some functions $f_\alpha\colon X \to Y_\alpha$ for each $\alpha \in \lambda$. Then

$$\bigcup_{\alpha\in\lambda} \left\{f_\alpha^{-1}(U_\alpha) : U_\alpha \in \mathcal{T}_\alpha\right\}$$

is a sub-base for an *induced topology*.

The one thing to check is that X belongs to the arbitrary unions of our collection, which is clear because $X = f_\alpha^{-1}(Y_\alpha)$.

Definition 2.41 (Relative topology)**.** Fix $(Y, \mathcal{T})$ a topological space. Then the *relative topology* for a subset $X \subseteq Y$ is the topology induced by the natural embedding $\iota\colon X \hookrightarrow Y$.

We have the following more concrete description.

Lemma 2.42. Fix $(Y, \mathcal{T}_Y)$ a topological space. Then the relative topology for a subset $X \subseteq Y$ consists of the subsets

$$\{X \cap U : U \in \mathcal{T}_Y\}.$$

Proof. Let $\iota\colon X \hookrightarrow Y$ be the natural embedding. Then we are given the sub-base

$$\mathcal{S} \coloneqq \left\{\iota^{-1}(U) : U \in \mathcal{T}_Y\right\}.$$

Now, $\iota^{-1}(U) = X \cap U$, and then we can check directly that this collection $\mathcal{S}$ gives a topology and finish by Remark 2.25. Here are the checks, which should be completely routine by now.

- Note $\varnothing \in \mathcal{T}_Y$ implies $\varnothing = X \cap \varnothing \in \mathcal{S}$. Also, $Y \in \mathcal{T}_Y$ implies $X = X \cap Y \in \mathcal{S}$.
- Arbitrary union: given a collection $\mathcal{U} \subseteq \mathcal{S}$, for each $U \in \mathcal{U}$ find $U_V \in \mathcal{T}_Y$ such that $U = X \cap U_V$. Then

  $$\bigcup_{U\in\mathcal{U}} = U = \bigcup_{U\in\mathcal{U}} X \cap U_V = X \cap \underbrace{\bigcup_{U\in\mathcal{U}} U_V}_{\in\mathcal{T}_Y}$$

  lives in $\mathcal{S}$.
- Finite intersection: given a finite collection $\{U_1, \ldots, U_n\} \subseteq \mathcal{S}$, find $V_i \in \mathcal{T}_Y$ such that $U_i = X \cap V_i$. Then

  $$\bigcap_{i=1}^n U_i = \bigcap_{i=1}^n (X \cap V_i) = X \cap \underbrace{\bigcap_{i=1}^n V_i}_{\in\mathcal{T}_Y}$$

  lives in $\mathcal{S}$. ■

2.3 September 2

There are no questions about anything.

2.3.1 Closed Sets

We begin, as always, with a definition.

Definition 2.43 (Closed). Fix a topological space $(X, \mathcal{T})$. A subset $V \subseteq X$ is *closed* if and only if $(X \setminus V) \in \mathcal{T}$.

Here are some basic properties.

Lemma 2.44. Fix a topological space $(X, \mathcal{T})$.

(a) The set $\varnothing$ and X are both closed.

(b) Arbitrary intersection: given a collection of closed sets $\mathcal{V}$, the intersection $\bigcap_{V \in \mathcal{V}} V$ is closed.

(c) Finite union: given a finite collection of closed sets $\{V_1, \ldots, V_n\}$, the union $\bigcup_{i=1}^n V_i$ is closed.

Proof. We proceed in sequence.

(a) Note that $X \setminus \varnothing = X$ and $X \setminus X = \varnothing$ are both open so $\varnothing$ and X are closed.

(b) Arbitrary intersection: observe that

$$X \Bigg\backslash \bigcap_{V \in \mathcal{V}} V = \bigcup_{V \in \mathcal{V}} (X \setminus V)$$

is an arbitrary union of open sets and therefore open. Thus, $\bigcap_{V \in \mathcal{V}} V$ is closed.

(c) Finite union: observe that

$$X \Bigg\backslash \bigcup_{i=1}^n V_i = \bigcap_{i=1}^n (X \setminus V_i)$$

is the finite intersection of open sets and therefore open. Thus, $\bigcup_{i=1}^n V_i$ is closed. ■

Remark 2.45. Observe that both X and $\varnothing$ are both open and closed. This is allowed.

Example 2.46. Fix a metric space (X, d). Then any closed ball $\overline{B(x_0, r)}$ is closed: we need to show

$$U := X \setminus \overline{B(x_0, r)} = \{x \in X : d(x, x_0) > r\}$$

is open. Well, for any $y \in U$, we see $d(y, x_0) > r$, so set $\varepsilon_y := d(y, x_0) - r$, so $y' \in B(y, \varepsilon_y)$ has $d(x_0, y') \geq d(x_0, y) - d(y, y') > r$. Thus, any $y \in U$ has $B(y, \varepsilon_y) \subseteq U$, finishing.

Remark 2.47. In $\mathbb{R}^2$ with the Euclidean metric,

$$\bigcup_{\varepsilon < 1}^{\infty} \overline{B(0, \varepsilon)} = \left\{x \in \mathbb{R}^2 : d(0, x) < \varepsilon \text{ for some } \varepsilon < 1\right\} = B(0, 1)$$

is not closed. Indeed, we need to show $U := X \setminus B(0, 1) = \left\{x \in \mathbb{R}^2 : d(0, x) \geq 1\right\}$ is not open. Well, note $(1, 0) \in U$, but any $\varepsilon > 0$ has $(1 - \varepsilon/2, 0) \in B((1, 0), \varepsilon)$ despite $(1 - \varepsilon/2, 0) \notin U$. Thus, U is not open.

Remark 2.48. One can define a topology by defining its closed sets to satisfy the axioms of Lemma 2.44. Then one defines the open sets as the complements of open sets.

Remark 2.49. Aligned with Remark 2.48, one can show that a function $f\colon (X, \mathcal{T}_X) \to (Y, \mathcal{T}_Y)$ is continuous if and only if $f^{-1}(V)$ is closed for all closed subsets $V \subseteq Y$.

- If f is continuous, then note any closed subset $V \subseteq Y$ has $Y \setminus V$ open, so $f^{-1}(Y \setminus V) = X \setminus f^{-1}(V)$ is open, so $f^{-1}(V)$ is closed.
- If f preserves closed sets, then any open subset $U \subseteq Y$ has $Y \setminus U$ closed, so $f^{-1}(Y \setminus U) = X \setminus f^{-1}(U)$ is closed, so $f^{-1}(U)$ is open.

In the case of metric spaces, we also have the following characterization of metric spaces.

Lemma 2.50. Fix a metric space (X, d) and $V \subseteq X$. The following are equivalent.

(a) V is closed.

(b) Any sequence $\{x_n\}_{n\in\mathbb{N}}$ in V which converges to a point $x \in X$ actually converges to $x \in V$.

Proof. In one direction, suppose V is closed, and suppose $x_n \to x$ as $n \to \infty$ with $x \notin V$. Then we show that some $n \in \mathbb{N}$ has $x_n \notin V$. Well, $x \in X \setminus V$, and $X \setminus V$ is open, so there is some $\varepsilon > 0$ with

$$B(x, \varepsilon) \subseteq X \setminus V.$$

However, $x_n \to x$ as $n \to \infty$ promises some large n such that $d(x, x_n) < \varepsilon$, implying that $x_n \in X \setminus V$ and so $x_n \notin V$.

In the other direction, suppose V is not closed. Then $X \setminus V$ is not open, so we can find $x \in X \setminus V$ for which there is no $\varepsilon > 0$ with $B(x, \varepsilon) \subseteq X \setminus V$. As such, $x \notin V$ but $B(x, 1/n) \cap V \neq \varnothing$ for all $n \in \mathbb{N}$, so just pick up some

$$x_n \in B(x, 1/n) \cap V$$

for each $n \in \mathbb{N}$. As such, $d(x, x_n) < 1/n$ for all $n \in \mathbb{N}$, so $x_n \to x$ as $n \to \infty$ (take $N = 1/\varepsilon$), and $x_n \in V$ for all $n \in \mathbb{N}$, but the limit x does not live in V. ■

Remark 2.51. The reason we are not generalizing the above lemma to arbitrary topological spaces is because we haven't generalized convergence yet.

Corollary 2.52. Fix a complete metric space (X, d). Then a closed subset $V \subseteq X$ given the restricted metric is also complete.

Proof. Suppose a sequence of points $\{x_n\}_{n\in\mathbb{N}}$ in V is Cauchy. Embedding back in X, this sequence is still Cauchy in X, so it has a limit $x \in X$. But Lemma 2.50 then promises $x \in V$, so $\{x_n\}_{n\in\mathbb{N}}$ does in fact have a limit x in V. ■

2.3.2 Closures

Given a general set, we can define the closure as follows.

Definition 2.53 (Closure)**.** Fix a topological space $(X, \mathcal{T})$. Given a subset $S \subseteq X$, we define the *closure* as

$$\overline{S} := \bigcap_{\substack{V \supseteq S \\ V \text{ closed}}} V.$$

Lemma 2.54. Fix a topological space $(X, \mathcal{T})$. Given a subset $S \subseteq X$, the closure $\overline{S}$ is the unique smallest closed set containing S.

Proof. Note that

$$\overline{S} := \bigcap_{\substack{V \supseteq S \\ V \text{ closed}}} V$$

is closed as the arbitrary intersection of closed sets, by Lemma 2.44. To see that $\overline{S}$ is a minimal such closed set, note that any closed V containing S must have $\overline{S} \subseteq V$ by definition of $\overline{S}$.

Lastly, to see that $\overline{S}$ is unique, note that if we have two minimal closed sets $\overline{S}_1$ and $\overline{S}_2$ containing S, then note $\overline{S}_1 \cap \overline{S}_2$ are both closed sets containing S by Lemma 2.44, so minimality forces $\overline{S}_1 = \overline{S}_1 \cap \overline{S}_2 = \overline{S}_2$. ■

Example 2.55. If $S \subseteq X$ is closed, then we see

$$S \subseteq \bigcap_{\substack{V \supseteq S \\ V \text{ closed}}} V \subseteq S$$

because S is a closed set containing S. Thus, $S = \overline{S}$.

Here is a more concrete way to work with the closure.

Lemma 2.56. Fix a topological space $(X, \mathcal{T})$ and a subset $A \subseteq X$. Then $x \in \overline{A}$ if and only if every open subset $U \subseteq X$ containing x has $U \cap A \neq \varnothing$.

Proof. In one direction, if there exists an open subset $U \subseteq X$ containing x such that $U \cap A \neq \varnothing$, then $A \subseteq X \setminus U$. By definition of the closure, it follows $\overline{A} \subseteq X \setminus U$, so $x \notin X \setminus U$ ensures $x \notin \overline{A}$.

In the other direction, suppose $x \notin \overline{A}$. Then $X \setminus \overline{A}$ is an open subset containing x (note $\overline{A}$ is closed by Lemma 2.54), and

$$A \cap (X \setminus \overline{A}) \subseteq \overline{A} \cap (X \setminus \overline{A}) = \varnothing,$$

so we have found an open set containing x disjoint from A. ■

With the notation, we note that we can move our notion of density from metric spaces to general topology.

Lemma 2.57. Fix a metric space (X, d). Then $S \subseteq X$ is dense if and only if $\overline{S} = X$.

Proof. In one direction, suppose that S is not dense in X, and we show $\overline{S} \subsetneq X$. Well, we are granted $x \in X$ and $\varepsilon > 0$ such that $S \cap B(x, \varepsilon) = \varnothing$, so $S \subseteq X \setminus B(x, \varepsilon)$. However, $X \setminus B(x, \varepsilon)$ is closed, so

$$\overline{S} \subseteq X \setminus B(x, \varepsilon) \subsetneq X,$$

as needed.

In the other direction, suppose $\overline{S} \subsetneq X$, and we show that S is not dense in X. Well, find $x \in X \setminus \overline{S}$. Because $X \setminus \overline{S}$ is open, we may find $\varepsilon > 0$ such that $B(x,\varepsilon) \subseteq X \setminus \overline{S}$, implying that

$$B(x,\varepsilon) \cap S \subseteq B(x,\varepsilon) \cap \overline{S} = \varnothing,$$

making S not dense in X. ■

Thus, we can generalize our definition as follows.

Definition 2.58 (Dense)**.** Fix a topological space $(X, \mathcal{T})$. Given subsets $A \subseteq B$, we say A is *dense* in B if and only if $B \subseteq \overline{A}$.

Remark 2.59. We are not requiring that B be closed for the definition of density. For example, $\mathbb{Q} \subseteq \mathbb{R}$ is dense in $\mathbb{Q}$.

2.3.3 The Product Topology

Let's see more examples of induced topologies. We start with the easiest example of the product topology.

Definition 2.60 (Product topology)**.** Fix topological spaces $(X_1, \mathcal{T}_1)$ and $(X_2, \mathcal{T}_2)$. The *product topology* on $X_1 \times X_2$ is the topology induced by the canonical projection mappings

$$\pi_1 \colon X_1 \times X_2 \to X_1 \qquad \text{and} \qquad \pi_2 \colon X_1 \times X_2 \to X_2.$$

We now give the following more concrete description of the product topology.

Lemma 2.61. Fix topological spaces $(X_1, \mathcal{T}_1)$ and $(X_2, \mathcal{T}_2)$. The product topology $\mathcal{T}$ on $X := X_1 \times X_2$ has a base given by

$$\mathcal{B} := \{U_1 \times U_2 : U_1 \in \mathcal{T}_1, U_2 \in \mathcal{T}_2\}.$$

Proof. The product topology is the minimal topology making $\pi_1 \colon X_1 \times X_2 \to X_1$ and $\pi_2 \colon X_1 \times X_2 \to X_2$ continuous. Namely, the product topology has a sub-base given by the sets

$$\pi_1^{-1}(U_1) = U_1 \times X_2 \qquad \text{and} \qquad \pi_2^{-1}(U_2) = X_1 \times U_2$$

for any $U_1 \in \mathcal{T}_1$ and $U_2 \in \mathcal{T}_2$. Using Example 2.33, we let $\mathcal{I}$ denote the finite intersections of these open sets and note $\mathcal{I}$ is a base for our topology.

Now, we finish by claiming $\mathcal{B} = \mathcal{I}$. On one hand, any $U_1 \times U_2 \in \mathcal{B}$ with $U_1 \in \mathcal{T}_1$ and $U_2 \in \mathcal{T}_2$ can be written as the finite intersection

$$U_1 \times U_2 = (U_1 \times X_2) \cap (X_1 \times U_2) = \pi_1^{-1}(U_1) \cap \pi_2^{-1}(U_2) \in \mathcal{I}.$$

On the other hand, pick finitely many sets of the form $\pi_1^{-1}(U_1)$ and $\pi_2^{-1}(U_2)$; dividing them into their classes, we can write our finite collection of sets as in $\{U_1^{(i)} \times X_2\}_{i=1}^m$ or $\{X_1 \times U_2^{(j)}\}_{j=1}^n$. Their intersection is

$$\left(\bigcap_{i=1}^m U_1^{(i)} \times X_2\right) \cap \left(\bigcap_{j=1}^n X_1 \times U_2^{(j)}\right) = \underbrace{\left(\bigcap_{i=1}^m U_1^{(i)}\right)}_{U_1 :=} \cap \underbrace{\left(\bigcap_{j=1}^n U_2^{(j)}\right)}_{U_2 :=}.$$

Now, $U_1 \subseteq X_1$ and $U_2 \subseteq X_2$ are finite intersection of open sets and therefore open, so our finite intersection takes the form $U_1 \times U_2$ and thus lives in $\mathcal{B}$. ■

Remark 2.62. Later in life we will discuss measurable sets, which are not quite topologies but will have similar ideas in spirit. For example, they will also care deeply about "rectangles."

We can define this more generally.

Definition 2.63 (Product topology)**.** Fix a collection of topological spaces $\{(X_\alpha, \mathcal{T}_\alpha)\}_{\alpha\in\lambda}$. The *product topology* on $X := \prod_{\alpha\in\lambda} X_\alpha$ is induced by the canonical projection maps

$$\pi_\alpha \colon X \to X_\alpha.$$

Here is our more concrete description.

Lemma 2.64. Fix a collection of topological spaces $\{(X_\alpha, \mathcal{T}_\alpha)\}_{\alpha\in\lambda}$. Then the product topology on $X := \prod_{\alpha\in\lambda} X_\alpha$ has a base

$$\mathcal{B} := \left\{ \prod_{\alpha\in\lambda} U_\alpha : U_\alpha \in \mathcal{T}_\alpha, U_\alpha = X_\alpha \text{ for all but finitely many } \alpha \right\}.$$

Proof. We are immediately given the sub-base of $\mathcal{S} := \{\pi_\alpha^{-1}(U_\alpha) : U_\alpha \in \mathcal{T}_\alpha\}$. Using Example 2.33, we let $\mathcal{I}$ denote the finite intersections of $\mathcal{S}$ so that $\mathcal{I}$ is a base for our product topology.

As before, we finish by claiming $\mathcal{I} = \mathcal{B}$. To stay organized, we proceed in steps.

- We show $\mathcal{B} \subseteq \mathcal{I}$. Namely, for any $\prod_{\alpha\in\lambda} U_\alpha$ in $\mathcal{B}$, we set $\lambda' := \{\alpha : U_\alpha \neq X_\alpha\}$, which we know must be finite. Then
$$\prod_{\alpha\in\lambda} U_\alpha = \bigcap_{\alpha\in\lambda} \pi^{-1}(U_\alpha) = \bigcap_{\alpha\in\lambda'} \pi_\alpha^{-1}(U_\alpha)$$
because $\pi^{-1}(X_\alpha) = X$. The right-hand side is indeed a finite intersection of elements of $\mathcal{S}$ and therefore in $\mathcal{I}$.

- We show $\mathcal{S} \subseteq \mathcal{B}$. For a given β and $U_\beta \in \mathcal{T}_\beta$, set $U_\alpha := X_\alpha$ for each $\alpha \neq \beta$. Then we see that
$$\pi_\beta^{-1}(U_\beta) = \prod_{\alpha\in\lambda} U_\alpha$$
is in $\mathcal{B}$ because $U_\alpha = X_\alpha$ for all but a single $\alpha \in \lambda$.

- We show $\mathcal{B}$ is closed under finite intersection. By induction, it suffices to pick up $U, U' \in \mathcal{B}$ and show that $U \cap U' \in \mathcal{B}$. Indeed, write
$$U = \prod_{\alpha\in\lambda} U_\alpha \qquad \text{and} \qquad U' = \prod_{\alpha\in\lambda} U'_\alpha,$$
where $\lambda_0 = \{\alpha : U_\alpha \neq X_\alpha\}$ and $\lambda'_0 = \{\alpha : U'_\alpha \neq X_\alpha\}$ are both finite. Then
$$U \cap U' = \prod_{\alpha\in\lambda} (U_\alpha \cap U'_\alpha),$$
and we have $U_\alpha \cap U'_\alpha = X_\alpha$ whenever $\alpha \notin (\lambda_0 \cup \lambda'_0)$, which is only finitely many exceptions because both λ_0 and λ_0 are finite.

- We show $\mathcal{I} \subseteq \mathcal{B}$. Indeed, $\mathcal{I}$ is made of the finite intersections of $\mathcal{S}$, and we see that $\mathcal{B}$ does indeed contain the finite intersections of $\mathcal{S}$ because $\mathcal{B}$ contains the finite intersections of itself, and $\mathcal{S} \subseteq \mathcal{B}$. ■

Remark 2.65. If λ is finite, then the arguments of Lemma 2.61 generalize to give the cleaner base

$$\left\{\prod_{\alpha\in\lambda} U_\alpha : U_\alpha \in \mathcal{T}_\alpha\right\}.$$

This also follows directly from Lemma 2.64, where we note that the "finitely many exceptions" actually permits all $\alpha \in \lambda$ to be an exception because λ is finite.

Example 2.66. Give $\{0,1\}$ the discrete topology. Then the space $X := \{0,1\}^{\mathbb{N}}$ given the product topology does not have

$$U := \prod_{n\in\mathbb{N}} \{0\}$$

open in X even though $\{0\} \subseteq \{0,1\}$ is always open. To see this, we note U has only a single element. On the other hand, for U to be open, Lemma 2.64 tells us U must contain a basis element B of the form

$$B := \prod_{n\in\mathbb{N}} U_n$$

where $U_n = \{0,1\}$ for all but finitely many n. However, B is infinite as the infinite product of sets containing more than 1 element, so $B \not\subseteq U$.

We quickly remark that the product topology satisfies the following universal property.

Lemma 2.67. Fix a collection of topological spaces $\{(X_\alpha, \mathcal{T}_\alpha)\}_{\alpha\in\lambda}$, and give the product $X := \prod_{\alpha\in\lambda} X_\alpha$ the projections $\pi_\alpha \colon X \to X_\alpha$ and the product topology $\mathcal{T}$. Given a topological space $(Y, \mathcal{T}_Y)$ and continuous maps $f_\alpha \colon Y \to X_\alpha$, there is a unique continuous map $f \colon Y \to X$ such that $f_\alpha = \pi_\alpha \circ f$ for each $\alpha \in \lambda$.

Proof. We show uniqueness and existence separately.

- Uniqueness: suppose both f and f' satisfy that $f_\alpha = \pi_\alpha \circ f = \pi_\alpha \circ f'$ for each $\alpha \in \lambda$. Then, for some $y \in Y$, we see that $f(y) = (x_\alpha)_{\alpha\in\lambda}$ and $f'(y) = (x'_\alpha)_{\alpha\in\lambda}$ have
$$x_\beta = (\pi_\beta \circ f)(y) = f_\beta(y) = (\pi_\beta \circ f')(y) = x'_\beta$$
for each $\beta \in \lambda$. So we conclude that $f(y) = f'(y)$ on all inputs. Observe that we have not used continuity anywhere.
- Existence: define $f \colon Y \to X$ by
$$f(y) := (f_\alpha(y))_{\alpha\in\lambda}.$$
We now need to check that f is continuous. By Proposition 2.31, it suffices to check this on the subbase of Lemma 2.64. In particular, pick up some finite $\lambda' \subseteq \lambda$ and set $U_\alpha \in \mathcal{T}_\alpha$ for each $\alpha \in \lambda$ while $U_\alpha = X_\alpha$ for $\alpha \notin \lambda'$. Then our basis element is
$$U := \prod_{\alpha\in\lambda} U_\alpha.$$
In particular,
$$\begin{aligned} f^{-1}(U) &= \{y \in Y : f_\alpha(y) \in U_\alpha \text{ for all } \alpha \in \lambda\} \\ &= \bigcap_{\alpha\in\lambda} f_\alpha^{-1}(U_\alpha) \\ &= \left(\bigcap_{\alpha\in\lambda'} f_\alpha^{-1}(U_\alpha)\right) \cap \left(\bigcap_{\alpha\notin\lambda'} f_\alpha^{-1}(\underbrace{U_\alpha}_{X_\alpha})\right), \end{aligned}$$

which is open because the left term is a finite intersection of open sets and the right term is just Y. ■

Corollary 2.68. Fix a collection of topological spaces $\{(X_\alpha, \mathcal{T}_\alpha)\}_{\alpha\in\lambda}$. Give the product $X := \prod_{\alpha\in\lambda} X_\alpha$ the projections $\pi_\alpha \colon X \to X_\alpha$ and the product topology $\mathcal{T}$. Given a topological space $(Y, \mathcal{T}_Y)$, a function $f \colon Y \to X$ is continuous if and only if the compositions $\pi_\alpha \circ f$ are continuous.

Proof. Certainly if f is continuous, then the continuity of π_α means that each $\pi_\alpha \circ f$ is continuous.

Conversely, set $f_\alpha := \pi_\alpha \circ f$ to be a continuous map $f_\alpha \colon Y \to X_\alpha$. Then Lemma 2.67 promises us a unique continuous map $\widetilde{f} \colon Y \to X$ such that

$$\pi_\alpha \circ \widetilde{f} = f_\alpha = \pi_\alpha \circ f.$$

However, the uniqueness proof of Lemma 2.67 showed that there is in fact one unique map of sets whose projections under π_α are f_α, so we conclude $f = \widetilde{f}$. Thus, f is continuous. ■

2.3.4 Comments on the Dual Space

Given a vector space V with a norm $\|\cdot\|$, we might be interested in the linear functionals on V, but because V is a metric space, we should actually be looking at the continuous linear functional. One can show (in Math 202B) that one has "plenty" of continuous linear functionals. Here is a lemma we will use a few times.

Lemma 2.69. Let $\|\cdot\|$ be a norm on an $\mathbb{R}$-vector space V. Then a linear functional $f \colon V \to \mathbb{R}$ is continuous if and only if there exists a real number $c > 0$ such that

$$|f(v)| \le c\,\|v\| \tag{2.1}$$

for all $v \in V$.

Proof. In one direction, suppose that we can find a real number $c > 0$ satisfying (2.1) for all $v \in V$. To show f is continuous, we use Lemma 1.58: suppose that we have a sequence $\{v_n\}_{n\in\mathbb{N}}$ such that $v_n \to v$ as $n \to \infty$. Then, for any $\varepsilon > 0$, find N such that $n > N$ implies

$$\|v - v_n\| < \varepsilon/c$$

so that

$$|f(v) - f(v_n)| \le c\,\|v - v_n\| < \varepsilon.$$

Conversely, suppose that f is continuous. Note that we don't have to worry about $v = 0$ because this gives equality. Now, we can find $\delta > 0$ such that $\|v\| < \delta$ implies $|f(v)| < 1$. It follows that any nonzero $v \in V$ will have

$$\left\|\frac{\delta}{2\,\|v\|}v\right\| < \delta,$$

so we see

$$|f(v)| = \frac{2\,\|v\|}{\delta}\left|f\left(\frac{\delta}{2\,\|v\|}v\right)\right| \le \frac{2}{\delta}\cdot\|v\|,$$

so $c := 2/\delta$ will do the trick. ■

Here is an example.

Exercise 2.70. Give $V := C([0,1])$ a p-norm $\|\cdot\|_p$ for some $p \ge 1$ or $p = \infty$. Then $g \in C([0,1])$ defines a continuous linear functional

$$\varphi_g \colon f \mapsto \int_0^1 f(t)g(t)\,dt.$$

Proof. To show φ_g is linear, pick up any $r_1, r_2 \in \mathbb{R}$ and $f_1, f_2 \in V$; then

$$\varphi_g(r_1f_1 + r_2f_2) = \int_0^1 (r_1f_1 + r_2f_2)(t)g(t)\,dt = r_1\int_0^1 f_1(t)g(t)\,dt + r_2\int_0^1 f_2(t)g(t)\,dt = r_1\varphi_g(f_1) + r_2\varphi_g(f_2).$$

Checking continuity is a little more involved. Note $|g|$ is a continuous function on a compact set $[0,1]$ and therefore has a maximum M. We now use Lemma 2.69; we have two cases.

- Suppose $p = \infty$. Then, for any $f \in V$, we see

  $$|\varphi_g(f)| = \left|\int_0^1 f(t)g(t)\,dt\right| \le M\int_0^1 |f(t)|\,dt \le M\,\|f\|_\infty\,,$$

  which finishes by Lemma 2.69.
- Suppose $p \ge 1$ is finite. To begin, we note

  $$|\varphi_g(f)| = \left|\int_0^1 f(t)g(t)\,dt\right| \le M\int_0^1 |f(t)|\,dt.$$

  Now, because the function $x \mapsto x^p$ is convex, we see that

  $$\left(\int_0^1 |f(t)|\,dt\right)^p \le \int_0^1 |f(t)|^p\,dt = \|f\|_p^p\,,$$

  so $|\varphi_g(f)| \le M\,\|f\|_p$. Lemma 2.69 finishes. ■

Even though the linear functionals we found were continuous for all $\|\cdot\|_p$, it is possible to find linear functionals continuous for some of our norms but not others.

Exercise 2.71. Fix $V := C([0,1])$, and select some $t_0 \in [0,1]$. Then

$$\varphi\colon f \mapsto f(t_0)$$

defines a linear functional on V which is continuous for $\|\cdot\|_\infty$ but not for $\|\cdot\|_p$ for any finite $p \ge 1$.

Proof. To see continuity with $\|\cdot\|_\infty$, we note that any $f \in V$ has

$$|\varphi(f)| = |f(t_0)| \le \|f\|_\infty\,,$$

so Lemma 2.69 finishes.

We now show that φ is not continuous for a fixed $\|\cdot\|_p$, where $p \ge 1$ is finite. Using Lemma 2.69, we just have to show that the ratio $|\varphi(v)|/\,\|v\|_p$ is unbounded for $v \in V$. For this, we define $f_c\colon [0,1] \to \mathbb{R}$ by

$$f(t) := \max\left\{0, c - c^{2p+1}(t-t_0)^2\right\}.$$

The idea here is that f has a sharp bump at t_0. Now, f is a continuous function on $[0,1]$ because it is the composition of continuous functions, so $f \in V$. We can compute

$$\|f\|_p = \left(\int_0^1 |f(t)|^p\,dt\right)^{1/p}.$$

Now, $f(t)$ will only be nonzero when $c - c^{2p+1}(t-t_0)^2 \ge 0$, which is equivalent to $t - t_0 \in (-c^{-p}, c^{-p})$, so we bound

$$\|f\|_p^p = \int_0^1 |f(t)|^p\,dt \le \int_{-c^{-p}}^{c^{-p}} \left(c - c^{2p+1}z^2\right)\,dz \le 2c^{1-p}.$$

Notably, as $c \to \infty$, we have that $\|f\|_p \le 2^{1/p} \cdot c^{1/p-1}$ is bounded, but $|\varphi(f)| = c$ grows unbounded. Thus, φ is discontinuous. ■

Remark 2.72. Now, we have exhibited many continuous functions

$$\varphi_g \colon C([0,1]) \to \mathbb{R},$$

so we can ask for the topology on $C([0,1])$ induced by these. It turns out that this induced topology is much weaker than any individual norm topology; this topology is often called the weak topology determined by $C([0,1])$.

Remark 2.73. By the end of the class, we will have a reasonable notion of the dual space of $\|\cdot\|_1$ and $\|\cdot\|_2$. The dual space for $\|\cdot\|_\infty$ will come up in Math 202B.

Remark 2.74. Still working with $C([0,1])$ given a specific norm $\|\cdot\|_p$, one can show that any $g \in C([0,1])$ has some $r_g \in \mathbb{R}$ with

$$\varphi_g(B(0,1)) \subseteq B(0,r_g).$$

It turns out to be helpful to be able to consider the product topology on the (very large) product

$$\prod_{g\in C([0,1])} B(0,r_g).$$

2.4 September 7

It's another day of sun.

2.4.1 Quotient Spaces

Here is a different way to induce a topology, the reverse of the induced topology.

Definition 2.75 (Final topology)**.** Fix a set Y and some topological spaces $\{(X_\alpha, \mathcal{T}_\alpha)\}_{\alpha\in\lambda}$. Given functions $f_\alpha \colon X_\alpha \to Y$, we define the *final topology* on Y to be the "strongest" (i.e., with the most open sets) making the f_α continuous.

Remark 2.76. Note that certainly some topology on Y exists making the f_α continuous because we can give Y the indiscrete topology, where $f_\alpha^{-1}(\varnothing) = \varnothing$ and $f_\alpha^{-1}(Y) = X_\alpha$ are open for each $\alpha \in \lambda$.

Here is a more concrete description.

Lemma 2.77. Fix a set Y and some topological spaces $\{(X_\alpha, \mathcal{T}_\alpha)\}_{\alpha\in\lambda}$, with functions $f_\alpha \colon X_\alpha \to Y$. Then the final topology is

$$\mathcal{T} \coloneqq \bigcap_{\alpha\in\lambda} \left\{S \subseteq Y : f_\alpha^{-1}(S) \in \mathcal{T}_\alpha\right\}.$$

Proof. Certainly each $\left\{S \subseteq Y : f_\alpha^{-1}(S) \in \mathcal{T}_\alpha\right\}$ is a topology by Lemma 2.30, as is their intersection by Proposition 2.20. Thus, $\mathcal{T}$ is a topology.

It remains to show that $\mathcal{T}$ is the strongest topology making each of the f_α continuous. Well, suppose $\mathcal{T}'$ is a topology making each of the f_α continuous. Then, for each $U \in \mathcal{T}'$, we have

$$f_\alpha^{-1}(U) \in \mathcal{T}_\alpha \text{ for each } \alpha \in \lambda,$$

so $U \in \mathcal{T}$ follows. Thus, $\mathcal{T}' \subseteq \mathcal{T}$. ■

We will be primarily interested in the case with just one function.

Remark 2.78. In the case of one function, which is Lemma 2.30, note that we might as well assume that $f\colon X \to Y$ is onto for otherwise we might as well just pass to the relative topology on $\operatorname{im} f$. To be explicit, we see $U \subseteq Y$ is open if and only if $f^{-1}(U)$ is open if and only if $f^{-1}(U \cap \operatorname{im} f)$ is open if and only if $U \cap \operatorname{im} f$ is open.

We are now ready to define the quotient space.

Lemma 2.79. Given sets $f\colon X \to Y$, there is an equivalence relation $\sim$ on X with $x \sim x'$ if and only if $f(x) = f(x')$.

Proof. We check the conditions one at a time. Find $x, x', x'' \in X$.

- Reflexive: note $f(x) = f(x)$, so $x \sim x$.
- Symmetric: if $x \sim x'$, then $f(x) = f(x')$, so $f(x') = f(x)$, so $x' \sim x$.
- Transitive: if $x \sim x'$ and $x' \sim x''$, then $f(x) = f(x') = f(x'')$, so $f(x) = f(x'')$, so $x \sim x''$. ■

With an equivalence relation, we may consider the set of equivalence classes $X/{\sim}$.

Remark 2.80. Conversely, given some partition $P \subseteq \mathcal{P}(X)$ of X, we can define $f\colon X \to P$ by $f\colon x \mapsto [x]$, where $[x] \in P$ is the element of P containing x. (Note $[x] \in P$ exists and is well-defined because P is a partition.) The point is that surjective functions give rise to equivalence relations, and equivalence relations give rise to surjective functions.

Anyway, here is our definition.

Definition 2.81 (Quotient topology)**.** Fix an equivalence relation $\sim$ on a set X with a topology $\mathcal{T}$. Then the *quotient topology* on $X/{\sim}$ is the final topology for the natural projection $X \twoheadrightarrow X/{\sim}$.

It turns out that we can talk about the quotient space by universal property as well.

Proposition 2.82. Fix an equivalence relation $\sim$ on a set X with a topology $\mathcal{T}$; let $\pi\colon X \twoheadrightarrow (X/{\sim})$ be the natural projection. Then, for any continuous map $f\colon X \to Z$ such that any $x \sim x'$ has $f(x) = f(x')$, there is a unique continuous map $\overline{f}\colon (X/{\sim}) \to Z$ such that

$$f = \overline{f} \circ \pi.$$

Proof. We show uniqueness and existence separately.

- Uniqueness: for any $[x] \in (X/{\sim})$, we see that we must have
$$\overline{f}([x]) = \overline{f}(\pi(x)) = f(x),$$
so $\overline{f}([x])$ is forced by our other data.
- Existence: for each $[x] \in (X/{\sim})$, define $\overline{f}([x]) := f(x)$. Note that this is well-defined: if $[x] = [x']$, then $x \sim x'$, so $f(x) = f(x')$ by hypothesis.
It remains to show that $\overline{f}$ is continuous. Well, for an open set $U \subseteq Z$, we note that
$$\overline{f}^{-1}(U) = \{[x] : \overline{f}([x]) \in U\} = \{[x] : f(x) \in U\} = \pi\left(f^{-1}(U)\right).$$
Now, $\pi^{-1}\left(\pi\left(f^{-1}(U)\right)\right) = f^{-1}(U)$ because $x \in \pi^{-1}\left(\pi\left(f^{-1}(U)\right)\right)$ if and only if $\pi(x) \in \pi\left(f^{-1}(U)\right)$, which is equivalent to there being $x' \in f^{-1}(U)$ with $\pi(x) = \pi(x')$, which is equivalent to there being x' with $x \sim x'$ while $f(x) = f(x') \in U$.
Thus, $\pi^{-1}\left(\pi\left(f^{-1}(U)\right)\right)$ is open, so it follows $\pi\left(f^{-1}(U)\right) \subseteq (X/{\sim})$ is open. ■

2.4.2 Homeomorphism

Homeomorphisms are isomorphisms in our category Top. To be technical, here is our definition.

Definition 2.83 (Homeomorphism)**.** A function $f\colon X \to Y$ between topological spaces $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$ is a *homeomorphism* if and only if f is continuous and has a continuous inverse. Formally, we require a continuous map $g\colon Y \to X$ such that

$$f \circ g = \mathrm{id}_Y \qquad \text{and} \qquad g \circ f = \mathrm{id}_X.$$

Warning 2.84. It is not enough for f to be continuous and bijective to be a homeomorphism. The hypothesis that the inverse function be continuous is necessary.

Remark 2.85. The definition above does not require that f be bijective, but this follows from f having an inverse.

Here are some examples.

Example 2.86. Fix a nonzero real number a and a real number b. Then the function $\varphi_{a,b}\colon \mathbb{R} \to \mathbb{R}$ by $\varphi_{a,b}(x) := ax + b$ is continuous: checking this on the subbase (which is enough by Proposition 2.31), we compute $\varphi_{a,b}^{-1}((c,d)) = ((c-b)/a, (d-b)/a)$. The inverse function is $\varphi_{1/a,-b/a}$—note $\varphi_{1/a,-b/a}(\varphi_{a,b}(x)) = \varphi_{a,b}(\varphi_{1/a,-b/a}(x)) = x$—which is continuous for the same reason, so this function $\varphi_{a,b}$ is a homeomorphism.

Lemma 2.87. Fix a homeomorphism $f\colon (X, \mathcal{T}_X) \to (Y, \mathcal{T}_Y)$. Further, for any subset $S \subseteq X$, give S and $f(S)$ their respective relative topologies. Then the restriction $f|_S\colon S \to f(S)$ is a homeomorphism.

Proof. For clarity, let $g\colon Y \to X$ be the inverse function for f; note that $g(f(S)) = \{g(f(x)) : x \in S\} = S$, so $g|_{f(S)}\colon f(S) \to S$. Observe that we still have $g(f(x)) = x$ and $f(g(y))$ for each $x \in X$ and $y \in Y$, so $f|_S$ and $g|_S$ are inverse functions by restricting these equations.

It remains to see that f and g are continuous. We will show that f is continuous, and g will follow by symmetry. Well, for an open subset $U \cap f(S) \subseteq f(S)$ (where $U \subseteq X$ is open), we see

$$f|_S^{-1}(U \cap f(S)) = \{x \in S : f(x) \in U \cap f(S)\} = S \cap \{x \in X : f(x) \in U\} \cap \{x \in S : f(x) \in f(S)\} = S \cap f^{-1}(U),$$

which is indeed open in the relative topology of S. ■

Example 2.88. Fix real numbers $b > a$. Continuing from Example 2.86, $\varphi_{a,b}\colon \mathbb{R} \to \mathbb{R}$ restricts by Lemma 2.87 to a homeomorphism

$$\varphi_{b-a,a}|_{[0,1]}\colon [0,1] \to [a,b].$$

Namely, $x \in [0,1]$ if and only if $0 \le x \le 1$ if and only if $a \le (b-a)x + a \le b$ if and only if $\varphi_{b-a,a}(x) \in [a,b]$.

Example 2.89. Give $\mathbb{R}$ the Euclidean topology, and let $\mathbb{R}_d$ be the real numbers with the discrete topology. Then the identity function $\iota\colon \mathbb{R}_d \to \mathbb{R}$ is continuous because all functions from the discrete topology are continuous. However, ι is its own inverse, and the inverse function

$$\pi\colon \mathbb{R} \to \mathbb{R}_d$$

(which is also the identity on $\mathbb{R}$) is not continuous. For example, $\pi^{-1}(\{0\}) = \{0\}$ is not open in $\mathbb{R}$ (by Remark 2.11) even though $\mathbb{R} \setminus \{0\} \subseteq \mathbb{R}_d$ is open.

Here are some more exotic examples.

Exercise 2.90. Give $X := [0,1]$ the subspace topology, and define the equivalence relation $\sim$ as having equivalence classes $\{0,1\}$ and $\{r\}$ for each $r \in (0,1)$. Then the quotient topology $X/\sim$ is homeomorphic to $S^1 \subseteq \mathbb{C}$.

Proof. We note that $\sim$ is an equivalence relation because its equivalence classes are a partition. Now, we define the maps

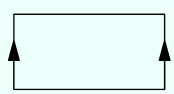
$$\begin{array}{ccc} (X/\sim) & \cong & S^1 \\ t & \mapsto & e^{2\pi i t} \\ \theta/2\pi & \leftarrow\!\shortmid & e^{i\theta} \end{array}$$

which we can see to be well-defined inverse. Note that $\mathbb{R} \to \mathbb{C}$ by $t \mapsto e^{it}$ is continuous by complex analysis (it's in fact holomorphic). Restricting, we get the continuous map $[0,1] \to S^1$, and then we can see that we can mod out by $0 \sim 1$ because they both go to the same place (using Proposition 2.82). One can check by hand that the inverse map is continuous, but we won't bother. ■

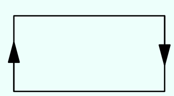
Remark 2.91 (Nir)**.** Here is a quick way to see that the inverse map is continuous: any continuous bijection $f\colon (X/\sim) \to S^1$ with $(X/\sim)$ compact—which is true because X is compact—and S^1 Hausdorff will send closed subsets $V \subseteq (X/\sim)$ (which are compact) to compact subsets of S^1 (which are closed). Thus, f is a closed map, so its inverse is continuous because f is bijective.

For the next few examples, we won't be very rigorous because we haven't provided good definitions of the relevant spaces.

Example 2.92. Give $X := [0,2] \times [0,1]$ the subspace topology, and define the equivalence relation $\sim$ as requiring $(0,r) \sim (2,r)$ only. Then X is homeomorphic to a circle by gluing its edges. One might draw X as follows.



Example 2.93. Continuing with the drawing style of Example 2.92, we have that



is the Möbius strip.

Remark 2.94. Note that these homeomorphisms do not care for the metric of our spaces. All that matters is the continuity.

Example 2.95. Let X be the unit sphere in $\mathbb{R}^3$ with the subspace topology, and define the equivalence relation on X by equivalence classes $\{v, -v\}$ for each $v \in X$. Then $X/\sim$ turns out to be $\mathbb{RP}^2$, which is hard to visualize.

2.4.3 Group Actions

A space might even have interesting homeomorphisms to itself.

Example 2.96. Fix a real number θ. The circle S^1 in $\mathbb{C}$ (given the subspace topology) has the rotation homeomorphism

$$r_\theta \colon e^{it} \mapsto e^{i(t+\theta)}.$$

Remark 2.97. In general, given a topological space $(X, \mathcal{T})$, we can make the group of homeomorphisms $\operatorname{Aut}(X)$ of homeomorphisms whose operation is composition.

This gives the following definition.

Definition 2.98 (Group action)**.** A *group action* by a group G on a topological space X is a group homomorphism

$$\varphi_\bullet \colon G \to \operatorname{Aut}(X).$$

Example 2.99. The group $\langle\sigma\rangle \simeq \mathbb{Z}/2\mathbb{Z}$ acts on a normed vector space $(V, \|\cdot\|)$ by sending σ^k to

$$\varphi_{\sigma^k} \cdot v \coloneqq (-1)^k v.$$

Notably, φ_{σ^k} is continuous and its own inverse for any k, so it is a homeomorphism. In fact, we can see directly that $\varphi_{\sigma^k} \circ \varphi_{\sigma^\ell} = \varphi_{\sigma^{k+\ell}}$.

Notably, with a group action comes a partition.

Definition 2.100 (Orbit)**.** Let G act on a topological space X by $\varphi_\bullet \colon G \to \operatorname{Aut}(X)$. Then the G*-orbit* Gx of a point $x \in x$ is the set

$$Gx \coloneqq \{\varphi_g(x) : g \in G\}.$$

We denote the set of all orbits $\mathcal{O}_x$ be X/G.

Remark 2.101. Note that the map $x \mapsto \mathcal{O}_x$ is a well-defined (surjective) map $X \to X/G$. In particular, we need to know that $x \in \mathcal{O}_{x'}$ implies that $\mathcal{O}_x = \mathcal{O}_{x'}$ so that there is exactly one orbit containing x. Well, $x \in \mathcal{O}_{x'}$ means we can find $g_0 \in G$ such that $x = \varphi_{g_0}(x')$, so

$$\mathcal{O}_x = \{\varphi_g(x) : g \in G\} = \{\varphi_g(\varphi_{g_0}(x')) : g \in G\} = \{\varphi_{gg_0}(x') : g \in G\} \subseteq \mathcal{O}_{x'}.$$

Conversely, we note that $x' = \varphi_{g_0^{-1}}(x)$, so $\mathcal{O}_{x'} \subseteq \mathcal{O}_x$ follows, giving equality.

Thus, the G-orbits partition X, so we can give the set X/G the quotient topology as the final topology of the natural projection $X \twoheadrightarrow X/G$.

Theme 3
Building Functions

I can assure you, at any rate, that my intentions are honourable and my results invariant, probably canonical, perhaps even functorial.

—Andre Weil, [Wei59]

3.1 September 9

The fun continues. The next problem set is going to be long but only in words, not in what we actually have to prove. We are being told not to be intimidated.

Remark 3.1. We are about to transition from making topologies to coming up with adjectives which will give "lots" of continuous maps to, say, the real numbers. A rigorization of this shall be provided shortly.

3.1.1 Normal Spaces

Last class we briefly mentioned the Hausdorff property.

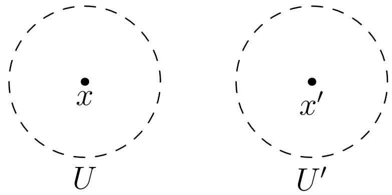
Definition 3.2 (Hausdroff)**.** Fix a topological space $(X,\mathcal{T})$. Then $(X,\mathcal{T})$ is *Hausdorff* if and only if, for any two distinct points $x,x'\in X$, there are disjoint open sets U and U' such that $x\in U$ and $x'\in U'$.

Example 3.3. A metric space (X,d) is Hausdorff. Indeed, given distinct points $x,x'\in X$, we have $d(x,x')>0$, so we set $r:=\frac{1}{2}d(x,x')$. Then $x\in B(x,r)$ and $x'\in B(x',r)$ (which are open sets by Example 2.6), we see $B(x,r)\cap B(x',r)=\varnothing$. Indeed, if we had $y\in B(x,r)\cap B(x',r)$, then we must have

$$d(x,x')\le d(x,y)+d(x',y)<2r=d(x,x'),$$

which is a contradiction.

Here is the image



Here is another adjective.

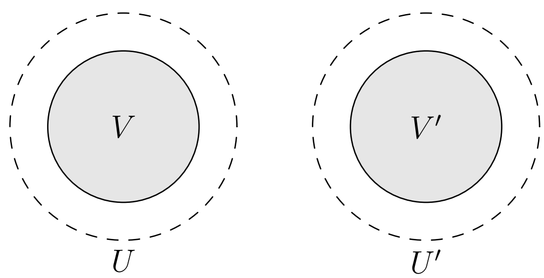
Definition 3.4 (Normal)**.** Fix a topological space $(X, \mathcal{T})$. Then $(X, \mathcal{T})$ is *Hausdorff* if and only if, for any two disjoint closed sets $V, V' \subseteq X$, there are disjoint open sets U and U' such that $V \in U$ and $V' \in U'$.

Remark 3.5. Intuitively, Hausdorff is approximately the normal property with singleton sets. In particular, some authors require "Hausdorff" in the definition of a normal space. We will not do this.

Example 3.6. Any set X given the indiscrete topology is normal. The problem here is that the only closed sets $\{\varnothing, X\}$, so the only possible pair of disjoint closed sets are $V_1 := \varnothing$ and $V_2 := \varnothing$, for which the open sets $U_1 := \varnothing$ and $U_2 := \varnothing$ are disjoint and cover these.

Example 3.7. A set X with more than 2 elements is normal, as shown in the previous example, but it is not Hausdorff. Namely, finding distinct points $x_1, x_2 \in X$, the only open subset of X containing x_1 or x_2 is X, so there are no disjoint open subsets U_1 containing x_1 and U_2 containing x_2.

Here is the image.



It is not completely obvious that metric spaces are normal, but we will see that they are.

Here is the main result for today.

Theorem 3.8 (Urysohn's lemma)**.** Fix a topological space $(X, \mathcal{T})$. If $(X, \mathcal{T})$ is normal, then for any disjoint closed subsets $V_0, V_1 \subseteq X$, there is a continuous function $f \colon X \to [0,1]$ such that $f(V_0) = \{0\}$ and $f(V_1) = \{1\}$.

So the point here is to realize Remark 3.1, where being normal is implying that we have "lots" of continuous functions.

Remark 3.9. Certainly if a topological space $(X, \mathcal{T})$ satisfies the conclusion of Theorem 3.8, then $(X, \mathcal{T})$ is normal. Indeed, for any disjoint closed subsets $V_0, V_1 \subseteq X$, pick up the promised continuous function f. Then

$$V_0 \subseteq f^{-1}((-1/2, 1/2)) \qquad \text{and} \qquad V_1 \subseteq f^{-1}((1/2, 3/2))$$

are disjoint open sets; namely, these are open because f is continuous, and they are disjoint because $f^{-1}((-1/2,1/2)) \cap f^{-1}((1/2,3/2)) = f^{-1}\big((-1/2,1/2) \cap (1/2,3/2)\big) = f^{-1}(\varnothing) = \varnothing$.

3.1.2 Urysohn's Lemma: Metric Spaces

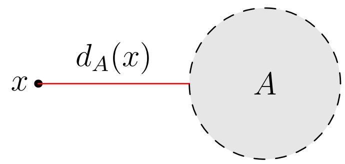
Let's see Theorem 3.8 for metric spaces, which will prove that metric spaces are normal by Remark 3.9. We pick up the following definition.

Definition 3.10. Fix a metric space (X, d). Then we define, for any $x \in X$ and nonempty subset $A \subseteq X$,

$$d_A(x) := \inf_{a \in A} d(x, a).$$

Remark 3.11. The infimum here exists because A is nonempty, so the set $\{d(x, a) : a \in A\}$ is nonempty (and bounded below by 0).

The image is that $d_A(x)$ is the distance from x to A.



We have the following continuity check.

Lemma 3.12. Fix a metric space (X, d). Then, for any nonempty subset $A \subseteq X$, the function $d_A \colon X \to \mathbb{R}$ is Lipschitz continuous.

Proof. Fix any $x, y \in X$. Then, for any given $a \in A$, we find that

$$d_A(x) \le d(x, a) \le d(x, y) + d(y, a).$$

Thus, $d_A(x) - d(x, y) \le d(y, a)$ for all $a \in A$, so we conclude that

$$d_A(x) - d(x, y) \le \inf_{a \in A} d(y, a) = d_A(y),$$

so $d_A(x) - d_A(y) \le d(x, y)$. By symmetry, we also have $d_A(y) - d_A(x) \le d(x, y)$, so it follows

$$\left|d_A(x) - d_A(y)\right| \le d(x, y),$$

which is what we need for our Lipschitz continuous. ■

As a sanity-check that this function behaves like it should, we pick up the following.

Lemma 3.13. Fix a metric space (X, d). Then, for any nonempty subset $A \subseteq X$, we have

$$d_A^{-1}(\{0\}) = \overline{A}.$$

Proof. Certainly $A \subseteq d_A^{-1}(\{0\})$ because $d_A(a) = 0$ for all $a \in A$. (In particular, $d_A(x) \ge 0$ everywhere, and $a \in A$ implies that $d_A(a) \le d(a, a) = 0$.) Because d_A is continuous by Lemma 3.12, we see $d_A^{-1}(\{0\})$ is closed, so containing A forces

$$\overline{A} \subseteq d_A^{-1}(\{0\}).$$

Conversely, suppose that $x \notin X \setminus \overline{A}$, and we show that $d_A(x) > 0$. Indeed, $X \setminus \overline{A}$ is open, so there is some open ball $B(x, \varepsilon)$ with $\varepsilon > 0$ such that $B(x, \varepsilon) \subseteq X \setminus \overline{A}$. It follows $B(x, \varepsilon) \cap \overline{A} = \varnothing$, so

$$d(a, x) \ge \varepsilon$$

for all $a \in A$. Thus, $d_A(x) \ge \varepsilon > 0$, so $d_A(x) \ne 0$. ■

Example 3.14. If $A \subseteq X$ is a dense subset, then $\overline{A} = X$, so $d_A \colon X \to \mathbb{R}$ is the constantly zero function.

Example 3.15. If $A \subseteq X$ is closed, then $\overline{A} = A$ by Example 2.55, so $d_A^{-1}(\{0\}) = A$. In other words, we have $x \in A$ if and only if $d_A(x) = 0$.

Let's now show Theorem 3.8 for metric spaces.

Proposition 3.16. Fix a metric space (X, d). For any disjoint closed subsets $V_0, V_1 \subseteq X$, there is a continuous function $f \colon X \to [0,1]$ such that $f(V_0) = \{0\}$ and $f(V_1) = \{1\}$.

Proof. The point is to use the Lipschitz continuous functions d_{V_0}, d_{V_1}. Then we define

$$f(x) := \frac{d_{V_0}(x)}{d_{V_0}(x) + d_{V_1}(x)}.$$

Note that defining $f \colon X \to \mathbb{R}$ does not have division-by-zero problems: because $d_{V_0}(x), d_{V_1}(x) \geq 0$, the only way to get zero in the denominator is by $d_{V_0}(x) = d_{V_1}(x) = 0$. However, this forces $x \in V_0 \cap V_1$ by Lemma 3.13 because V_0 and V_1 are closed, but in fact $V_0 \cap V_1 = \varnothing$.

We now run our checks on f.

- Because the quotient of two continuous functions is still continuous, we see that f is continuous.
- Using the fact that $d_A(x) \geq 0$ for any nonempty $A \subseteq X$ and $x \in X$, we find

  $$f(x) = \frac{d_{V_0}(x)}{d_{V_0}(x) + d_{V_1}(x)} \geq 0,$$

  and

  $$f(x) = 1 - \frac{d_{V_1}(x)}{d_{V_0}(x) + d_{V_1}(x)} \leq 1,$$

  so $\operatorname{im} f \subseteq [0,1]$.
- If $x \in V_0$, then $d_{V_0}(x) = 0$, so $f(x) = 0/(0 + d_{V_1}(x)) = 0$. If $x \in V_1$, then $d_{V_1}(x) = 0$, so $f(x) = d_{V_0}(x)/(d_{V_0}(x) + 0) = 1$. ■

And here is our check.

Corollary 3.17. Any metric space (X, d) is normal.

Proof. Plug Proposition 3.16 into Remark 3.9. ■

3.1.3 Urysohn's Lemma: The General Case

We will not prove the general case of Theorem 3.8 today, but we will make some progress. Here is a useful lemma.

Lemma 3.18. Fix a normal topological space $(X, \mathcal{T})$. Given a closed subset $V \subseteq X$ and an open subset $U_0 \subseteq X$ with $V \subseteq U_0$, there is an open set U such that

$$V \subseteq U \subseteq \overline{U} \subseteq U_0.$$

Proof. Because $V \subseteq U_0$, we define $V' := X \setminus U_0$, which is closed because U_0 is open. Further, $V' \subseteq X \setminus U_0 \subseteq X \setminus V$ forces $V \cap V' = \varnothing$. Thus, using the normality of $(X, \mathcal{T})$, we are promised disjoint open sets U and U' such that

$$V \subseteq U \qquad \text{and} \qquad V' \subseteq U'.$$

In particular, we see that

$$U \subseteq X \setminus U'$$

while $X \setminus U'$ is closed by definition. Thus, by definition of the closure, $\overline{U} \subseteq X \setminus U' \subseteq X \setminus V' = U_0$. This finishes the proof. ■

3.2 September 12

There are still no questions.

3.2.1 Urysohn's Lemma: The General Case

We continue the proof from last class.

> **Theorem 3.8** (Urysohn's lemma). Fix a topological space $(X, \mathcal{T})$. If $(X, \mathcal{T})$ is normal, then for any disjoint closed subsets $V_0, V_1 \subseteq X$, there is a continuous function $f : X \to [0,1]$ such that $f(V_0) = \{0\}$ and $f(V_1) = \{1\}$.

Proof. To begin, define $U_1 := X \setminus V_1$, which is open because V_1 is closed; notably $V_0 \subseteq U_1$. The idea here is that the points of U_1 will take value at most 1. Now, by Lemma 3.18, we find $U_{1/2}$ with

$$V_0 \subseteq U_{1/2} \subseteq \overline{U_{1/2}} \subseteq U_1.$$

Intuitively, we are going to let f take values at most $1/2$ on $U_{1/2}$. Using Lemma 3.18 again, we can find $U_{1/2}$ with

$$V_0 \subseteq U_{1/4} \subseteq \overline{U_{1/4}} \subseteq U_{1/2},$$

and now our function will take values at most $1/4$ on $U_{1/4}$. On the other side, we can use the containment $\overline{U_{1/2}} \subseteq U_1$ in Lemma 3.18 to find $U_{3/4}$ such that

$$\overline{U_{1/2}} \subseteq U_{3/4} \subseteq \overline{U_{3/4}} \subseteq U_1,$$

and here $U_{3/4}$ our function should take values less than $3/4$.

We can then continue the process for eights and then off to infinity. Let's describe what we have at the end of this inductive process. Set $\Delta := \{k/2^n : 0 < k \le 2^n\}$ to be the set of "dyadic" rationals in $(0,1]$; notably Δ is dense in $[0,1]$.[1] Then each $r \in \Delta$, we get an open set $U_r \subseteq X$. These have the following properties.

- Any $r, s \in \Delta$ with $r < s$ has $\overline{U_r} \subseteq U_s$.
- By construction $U_1 = X \setminus V_1$.
- Also, $V_0 \subseteq U_r$ for all $r \in \Delta$.

We now define

$$f(x) := \begin{cases} 1 & x \in V_1, \\ \inf\{r \in \Delta : x \in U_r\} & x \notin V_1, \end{cases}$$

where $x \notin V_1$ in the second case promises $x \in U_1$ so that the infimum in the second line makes sense. We now run the following checks on f.

[1] The fact we need is that $a, b \in [0,1]$ with $a < b$ have $r \in \Delta$ between them. Well, multiply $b - a$ by a suitably large power of 2 so that $2^n(b-a) > 1$, so there is an integer k in this interval between $2^n a$ and $2^n b$, so $a < k/2^n < b$.

- Note that $\operatorname{im} f(x) \subseteq \overline{\Delta} = [0,1]$.
- By the construction of these open sets, we have $f(x) = 1$ if $x \in V_1$.
- Further, $f(x) < r$ for all $r \in \Delta$ if $x \in V_0$, so $f(x) = 0$ for $x \in V_0$.
- It remains to check that f is continuous. For this, we use Proposition 2.31 to check the continuity on a subbase. Specifically, we use sets of the form $[0,a)$ and $(a,1]$ for $a \in (0,1)$. Indeed, note $[0,a) \cap (b,1] = (a,b)$, so intersections of these can give all open intervals strictly contained $[0,1]$; adding in the "open" intervals $[0,a)$ and $(a,0]$ make all the open intervals in $[0,1]$, which are a basis for our topology.

  We now proceed with our check; fix some $a \in (0,1)$.

  - Note that $x \in X$ has $f(x) < a$ if and only if there is some $r \in \Delta$ such that $f(x) < r < a$ (by density of Δ) if and only if there is some $r \in \Delta$ such that $x \in U_r$ and $r < a$ (by definition of the infimum). As such,
  $$f^{-1}([0,a)) = \bigcup_{r<a} U_r.$$
  - Note that $x \in X$ has $f(x) > a$ if and only if there is an $r, s \in \Delta$ with $f(x) > r > s > a$ (by density). It follows $x \notin U_r$, which contains $\overline{U_s}$, so $x \notin \overline{U_s}$ for some $s \in \Delta$ with $s > a$.
  On the other hand, $x \notin \overline{U_s}$ for some $s \in \Delta$ with $s > a$ implies that $x \notin U_r$ for any $r \in \Delta$ with $r > s > a$, so it follows $f(x) \geq s > a$.
  Thus, $f(x) > a$ if and only if $x \notin \overline{U_s}$ for $s \in \Delta$ with $s > a$, implying
  $$f^{-1}((a,1]) = \bigcup_{s>a} (X \setminus \overline{U_s}).$$

The above checks complete the proof. ■

Remark 3.19. We could not have f output to $\mathbb{Q} \cap [0,1]$ because we used the completeness of $\mathbb{R}$ in the construction of f.

Remark 3.20. It is somewhat noticeable that we have not discussed sequences at all in this class yet, even though they were featured prominently in metric space topology. The reason we have been avoiding them is that we prefer to use open sets and not points to study general topological spaces.

3.2.2 Bounded Functions

We are going to want a little functional analysis before we continue.

Definition 3.21 (Bounded)**.** Fix a metric space (X,d) and a nonempty set A. A subset $A \subseteq X$ is *bounded* if and only if there is an open ball $B(x,r)$ containing A. More generally, a function $f\colon A \to X$ is *bounded* if and only if $\operatorname{im} f \subseteq X$ is bounded, and we let $B(A,X)$ denote the set of all bounded functions $f\colon A \to X$.

We will be particularly interested in the case where X is a normed vector space.

The point of defining bounded functions is that we can provide them with a metric.

Definition 3.22 (Uniform metric)**.** Fix a nonempty set X and a metric space (Y,d). Then the *uniform metric* is the function $d_u\colon B(X,Y)^2 \to \mathbb{R}_{\geq 0}$ defined by
$$d_u(f,g) := \sup\{d(f(x),g(x)) : x \in X\}.$$

Lemma 3.23. Fix a set X and a metric space (Y, d). Then the uniform metric d_u on $B(X, Y)$ is a metric.

Proof. Here are our checks; fix $f, g, h \in B(X, Y)$.

- Well-defined: because f and g bounded, we can find open balls $B(a, r)$ and $B(b, s)$ containing $\operatorname{im} f$ and $\operatorname{im} g$ respectively. It follows that, for any $x \in X$, we have

$$d(f(x), g(x)) \leq d(f(x), a) + d(a, b) + d(b, g(x)) \leq r + d(a, b) + s,$$

  so the set $\{d(f(x), g(x)) : x \in X\}$ has an upper bound and hence a supremum.

- Nonnegative: fixing a particular $x \in X$, note $d_u(f, g) \geq d(f(x), g(x)) \geq 0$.

- Zero: note $d_u(f, f)$ is $\sup\{d(f(x), f(x)) : x \in X\} = \sup\{0 : x \in X\} = 0$.

- Zero: note $d_u(f, g) = 0$ implies that $\sup\{d(f(x), g(x)) : x \in X\} = 0$, so $d(f(x), g(x)) \leq 0$ for all $x \in X$, so $d(f(x), g(x)) = 0$ for all $x \in X$, so $f(x) = g(x)$ for all $x \in X$.

- Symmetric: note

$$d_u(f, g) = \sup\{d(f(x), g(x)) : x \in X\} = \sup\{d(g(x), f(x)) : x \in X\} = d_u(g, f).$$

- Triangle inequality: note that

$$d(f(x), h(x)) \leq d(f(x), g(x)) + d(g(x), h(x)) = d_u(f, g) + d_u(g, h)$$

  for all $x \in X$, so it follows $d_u(f, h) \leq d_u(f, g) + d_u(g, h)$ by taking the supremum. ■

Here is why we like this metric.

Proposition 3.24. Fix a set X and a complete metric space (Y, d). Then $B(X, Y)$ given the uniform metric is complete.

Proof. Fix a Cauchy sequence $\{f_n\}_{n \in \mathbb{N}}$ in $B(X, Y)$. Namely, for all $\varepsilon > 0$, there exists some N so that

$$n, m > N \implies d(f_n(x), f_m(x)) < \varepsilon$$

for all $x \in X$. In particular, fixing some particular $x \in X$, we see that $\{f_n(x)\}_{n \in \mathbb{N}}$ is a Cauchy sequence in Y, so the completeness of Y promises some limit $f(x)$.

It remains to check that the data of f assembles to a function $f \in B(X, Y)$. Well, any (fixed) $\varepsilon > 0$ promises an N so that $n, m > N$ forces $d(f_n(x), f_m(x)) < \varepsilon$ for all $x \in X$. Now, fixing some $x \in X$, any $\delta > 0$ has some N' large enough so that $m > N'$ has $d(f_m(x), f(x)) < \delta$, meaning that $n, m > \max\{N, N'\}$ gives

$$d(f_n(x), f(x)) \leq d(f_n(x), f_m(x)) + d(f_m(x), f(x)) < \varepsilon + \delta$$

for all $\delta > 0$. Thus, fixing some $n > N$, we see $d(f_n(x), f(x)) \leq \varepsilon$ for all $x \in X$.

To finish, we note $f_n \in B(X, Y)$ is bounded, so there is an open ball $B(a, r)$ containing $\operatorname{im} f_n$. Thus, for all $x \in X$,

$$d(a, f(x)) \leq d(a, f_n(x)) + d(f_n(x), f(x)) < r + \varepsilon,$$

so $\operatorname{im} f \subseteq B(a, r + \varepsilon)$. ■

We close with the following result.

Proposition 3.25. Fix a topological space $(X, \mathcal{T})$ and a metric space (Y, d). Let $B_c(X, Y) \subseteq B(X, Y)$ denote the metric subspace of bounded continuous functions $f\colon X \to Y$. Then $B_c(X, Y)$ is a closed subspace of $B(X, Y)$. In particular, if (Y, d) is complete, then $B_c(X, Y)$ is also complete.

Proof. Note that the second claim follows from the first claim by Corollary 2.52; thus, we focus on the first claim. For this, we use Lemma 2.50: fix a sequence $\{f_n\}_{n\in\mathbb{N}}$ of bounded continuous functions such that $f_n \to f$ as $n \to \infty$ where $f\colon X \to Y$ is just some bounded function. We need to show that f is continuous.

Well, fix an open set $U \subseteq Y$ so that we need to show $f^{-1}(U) \subseteq X$ is open. For this, we pick up any element $x \in f^{-1}(U)$, and we find an open neighborhood $U_x \subseteq f^{-1}(U)$ containing x; this will finish because it shows

$$f^{-1}(U) \subseteq \bigcup_{x\in U} U_x \subseteq f^{-1}(U),$$

so $f^{-1}(U)$ is the arbitrary union of open sets.

We now proceed with the proof directly.

1. Because $f(x) \in U$, and U is open, there is some $\varepsilon > 0$ such that $B(f(x), \varepsilon) \subseteq U$.

2. Because $\{f_n\}_{n\in\mathbb{N}}$ converges to f, there is a sufficiently large N so that $n > N$ has $d(f_n(y), f(y)) < \varepsilon/2$ for all $y \in X$. Fix some $n > N$.

3. Now, for all $y \in f_n^{-1}(B(f(x), \varepsilon/2))$, we see

$$d(f(y), f(x)) \le d(f(y), f_n(y)) + d(f_n(y), f(x)) < \varepsilon/2 + \varepsilon/2 = \varepsilon,$$

so $f(y) \in U$. As such, we see that $f_n^{-1}(B(f(x), \varepsilon/2))$ is open (because f_n is continuous), it contains x, and it is contained in $f^{-1}(U)$.

The above open neighborhood completes the proof of the first claim. ■

3.3 September 14

The march continues.

3.3.1 The Tietze Extension Theorem

Here is the main result for today.

Theorem 3.26 (Tietze extension)**.** Fix a normal topological space $(X, \mathcal{T})$, and give some closed subset $A \subseteq X$ the relative topology from X. Given a continuous function $f\colon A \to \mathbb{R}$, there exists a continuous function $\widetilde{f}\colon X \to \mathbb{R}$ such that $\widetilde{f}|_A = f$. In fact, if $\operatorname{im} f \subseteq [a, b]$, then we may enforce $\operatorname{im} \widetilde{f} \subseteq [a, b]$ as well.

This property is quite special to $\mathbb{R}$ shared by a few other spaces.

Example 3.27. Take $X := \overline{B(0, 1)} \subseteq \mathbb{R}^2$ given the relative topology, and let $A = \partial X$ be the boundary, which is the unit circle. Then the identity function $\mathrm{id}_A\colon A \to A$ does not extend continuously to a function $\widetilde{\mathrm{id}_A}\colon X \to A$. To see this rigorously, take a course in algebraic topology.

Example 3.28. Of course, any set Y given the indiscrete topology will be such that a continuous function $f\colon A \to Y$ can be extended to continuously to a function $\widetilde{f}\colon X \to Y$ because all functions to Y are continuous for free.

Remark 3.29. The condition of $\operatorname{im} f \subseteq [a,b]$ might as well be replaced by $\operatorname{im} f \subseteq [0,1]$ by using the homeomorphism $\mathbb{R} \to \mathbb{R}$ by $x \mapsto (x-a)/(b-b)$ which will send $[a,b]$ to $[0,1]$.

Here is a lemma which will help the proof of Theorem 3.26.

Lemma 3.30. Fix a normal topological space $(X,\mathcal{T})$, and give some closed subset $A \subseteq X$ the relative topology from X. Given a continuous function $f\colon A \to [0,r]$ (where $r > 0$), there exists a continuous function $g\colon X \to [0,r/3]$ such that

$$0 \le f(a) - g(a) \le 2r/3$$

for each $a \in A$.

Proof. Set $B := \{x \in A : f(x) \le r/3\} = f^{-1}([0,r/3])$ and $C := \{x \in A : f(x) \ge 2r/3\} = f^{-1}([(2r/3,r])$. Both $B, C \subseteq A$ and C are closed because they are the pre-image of closed subsets under $f\colon A \to \mathbb{R}$. In fact, by the relative topology, we can write $B = B' \cap A$ where $B' \subseteq X$ is closed. However, B' and A are both closed in X, so $B \subseteq X$ is closed. Similar holds for C.

Thus, so Urysohn's lemma provides (Theorem 3.8) a continuous function $g\colon X \to [0,1]$ such that $g|_B = 0$ and $g|_C = 1$. As such, we define $g\colon X \to [0,r/3]$ by

$$g(x) := (r/3) \cdot g(x),$$

which is still continuous because the map $x \mapsto (r/3)x$ is a homeomorphism $[0,1] \to [0,r/3]$ by Example 2.88. We can now see that g satisfies the needed properties. Fix some $a \in A$.

- If $a \in B$, then $g(a) = 0$ while $f(a) \le r/3$, so $0 \le f(a) - g(a) \le r/3$.
- If $a \in C$, then $g(a) = r/3$ while $f(a) \in [2r/3, r]$, so $0 \le f(a) - g(a) \le 2r/3$.
- Lastly, $a \notin B$ and $a \notin C$ means that $r/3 < f(a) < 2r/3$ while $0 \le g(a) \le r/3$, so it follows $0 \le f(a) - g(a) \le 2r/3$ still.

The above checks finish. ■

We now show the following special case of Theorem 3.26.

Proposition 3.31. Fix a normal topological space $(X,\mathcal{T})$, and give some closed subset $A \subseteq X$ the relative topology from X. Given a continuous function $f\colon A \to [0,1]$, there exists a continuous function $\widetilde{f}\colon X \to [0,1]$ such that $\widetilde{f}|_A = f$.

Proof. For brevity, define $\sigma := 2/3$. Taking $r = 1$ in Lemma 3.30, we get a function $g_1\colon X \to [0,1/3]$ with

$$0 \le f(a) - g_1(a) \le \sigma$$

for all $a \in A$, so define $\widetilde{f}_1 := g_1$. Next applying Lemma 3.30 to $(f - \widetilde{f}_1|_A)\colon A \to [0,\sigma]$ with $r = \sigma$, we get promised a function $g_2\colon X \to [0,\sigma/3]$ with

$$0 \le f(a) - \widetilde{f}_1(a) - g_2(a) \le \sigma^2$$

for any $a \in A$, so define $\widetilde{f}_2 := \widetilde{f}_1 + g_2$.

In general, suppose given a function $\widetilde{f}_n\colon X \to [0,1]$ with

$$0 \le f(a) - \widetilde{f}_n(a) \le \sigma^n$$

for $a \in A$, we can use Lemma 3.30 to $(f - \widetilde{f}_n|_A)\colon A \to [0,\sigma^n]$ to get a function $g_{n+1}\colon X \to [0,\sigma^n/3]$ with

$$0 \le f(a) - \widetilde{f}_n(a) - g_{n+1}(a) \le \sigma^{n+1}$$

for $a \in A$, allowing us to then set $\widetilde{f}_{n+1} := \widetilde{f}_n + g_{n+1}$.

Applying the above process inductively, we get a function

$$\widetilde{f}_n = \sum_{k=1}^{n} g_k$$

going to $[0,1]$ such that $\|g_k\|_\infty \le \sigma^{k-1}/3$ and $0 \le f(a) - \widetilde{f}_n(a) \le (2/3)^n$ for each $a \in A$ and $n \ge 1$. Notably, using the uniform metric d_u, we see that any $n \ge m$ has

$$d_u(\widetilde{f}_n, \widetilde{f}_m) = \sup_{x \in X} \left(\sum_{k=m+1}^{n} g_k(x) \right) \le \sum_{k=m+1}^{n} \frac{1}{3}\sigma^{k-1} \le \frac{\sigma^m}{3} \sum_{k=0}^{\infty} \sigma^k = \frac{\sigma^m}{3} \cdot \frac{1}{1-\sigma} = \left(\frac{2}{3}\right)^m,$$

which gets arbitrarily small. Thus, $\{\widetilde{f}_n\}_{n \in \mathbb{N}}$ is a Cauchy sequence: for any $\varepsilon > 0$, we can find N with $n > N$ having $(2/3)^n < \varepsilon$, meaning $n, m \ge N$ will have $d_u(\widetilde{f}_n, \widetilde{f}_m) < \varepsilon$. Now, because $[0,1] \subseteq \mathbb{R}$ is a closed subset of a complete metric space and hence complete by Corollary 2.52, the sequence $\{\widetilde{f}_n\}_{n \in \mathbb{N}}$ converges to a continuous function $\widetilde{f}\colon X \to [0,1]$ by Proposition 3.25.

It remains to check that $\widetilde{f}|_A = f$. Well, any $a \in A$ and $n \in \mathbb{N}$ have

$$|f(a) - \widetilde{f}(a)| \le |f(a) - \widetilde{f}_n(a)| + |\widetilde{f}_n(a) - \widetilde{f}(a)| \le \left(\frac{2}{3}\right)^n + |\widetilde{f}_n(a) - f(a)|.$$

Because $\widetilde{f}_n \to f$ as $n \to \infty$ under the metric d_u, we see that $|\widetilde{f}_n(a) - f(a)| \to 0$ as $n \to \infty$. Additionally, $(2/3)^n \to 0$ as $n \to \infty$, so the entire right-hand side goes to 0 as $n \to \infty$, meaning that $|f(a) - \widetilde{f}(a)| < \varepsilon$ for all $\varepsilon > 0$. Thus, $f(a) = \widetilde{f}(a)$ for each $a \in A$. ■

3.4 September 16

We continue the proof from last class.

3.4.1 The Tietze Extension Theorem: Proof

And here is the proof of the general case of Theorem 3.26.

> **Theorem 3.26 (Tietze extension).** Fix a normal topological space $(X, \mathcal{T})$, and give some closed subset $A \subseteq X$ the relative topology from X. Given a continuous function $f\colon A \to \mathbb{R}$, there exists a continuous function $\widetilde{f}\colon X \to \mathbb{R}$ such that $\widetilde{f}|_A = f$. In fact, if $\operatorname{im} f \subseteq [a,b]$, then we may enforce $\operatorname{im} \widetilde{f} \subseteq [a,b]$ as well.

Proof. Fix a continuous function $f\colon A \to \mathbb{R}$. Note that there is a homeomorphism $\varphi\colon \mathbb{R} \cong (-1,1)$, so we name composite

$$A \xrightarrow{f} \mathbb{R} \overset{\varphi}{\cong} (-1,1) \subseteq [0,1]$$

g and then extend it to a function $\widetilde{g}_0\colon X \to [-1,1]$ by Proposition 3.31. We would like to go back to $(-1,1)$ and then back to $\mathbb{R}$, but it is possible for $-1, 1 \in \operatorname{im} g_0$.

Isolating the problem, we set $B := \widetilde{g}_0^{-1}(\{-1,1\})$ and note that $A \cap (B_0 \cup B_1) = \varnothing$ because $\widetilde{g}_0(A) = g(A) \subseteq (-1,1)$. Now, by normality of X, we get promised by Theorem 3.8 a continuous function $\delta\colon X \to \mathbb{R}$ such that $\delta|_B = 0$ and $\delta|_A = 1$. Thus, we define

$$\widetilde{g}(x) := \delta(x)\widetilde{g}_0(x).$$

Notably, $\widetilde{g}|_A = \delta|_A \cdot \widetilde{g}_0|_A = 1 \cdot g = g$. But now $|\widetilde{g}(x)| = 1$ would force $|\widetilde{g}_0(x)| = 1$, but this implies $\delta(x) = 0$ by construction and so $\widetilde{g}(x) = 0$; thus, $\pm 1 \notin \operatorname{im} \widetilde{g}$, so we can pull back $\widetilde{g}$ through $\varphi\colon \mathbb{R} \cong (-1,1)$ to $\mathbb{R}$. ■

3.4.2 Existence of Completions, Again

We quickly provide another proof of the existence of completions. We begin with the following example.

Example 3.32. Given any topological space $(X, \mathcal{T})$, the metric space $(B_c(X, \mathbb{R}), d_u)$ of bounded continuous functions is complete by Proposition 3.25 because $\mathbb{R}$ is complete.

More generally, we will want to remember the following definition.

Definition 3.33 (Banach space)**.** A normed vector space $(V, \|\cdot\|)$ is a *Banach space* if and only if it is complete.

As such, we pick up the following tool.

Lemma 3.34. Fix an isometry $f: (X, d) \to (Y, d_Y)$ of metric spaces such that (Y, d_Y) is complete. Then $\overline{f(X)}$ equipped with the induced metric from Y is a complete metric space, and it is actually a completion of (X, d) when equipped with the natural embedding $\iota: X \to \overline{f(X)}$ from f.

Proof. For brevity, define $\overline{X} := \overline{f(X)}$ and set $\overline{d}$ to be the metric on $\overline{X}$ induced by (Y, d_Y). In particular, $\overline{X} \subseteq Y$ is a closed subset, and so $(\overline{X}, \overline{d})$ is complete by Corollary 2.52. Now, note that $\iota: (X, d) \to (\overline{X}, \overline{d})$ is an isometry because, for any $x, x' \in X$,

$$d(x, x') = d_Y(f(x), f(x')) = d_Y(\iota(x), \iota(x')) = \overline{d}(\iota(x), \iota(x'))$$

using our various restriction maps.

Lastly, we have to show that $\operatorname{im} \iota \subseteq \overline{X}$ is dense. Well, by Lemma 2.57, it suffices to note

$$\overline{\operatorname{im} \iota} = \overline{f(X)} = \overline{X},$$

which is what we wanted. ■

We are now ready to prove Theorem 1.69.

Theorem 1.69. Any metric space (X, d) has a completion.

Proof. Let our metric space be (X, d). For each $x \in X$, define $f_x(y) := d(x, y)$. To embed f_x into $B_c(X, \mathbb{R})$, we would need f_x to be bounded, but it need not be. To fix this, we choose a base-point $x_0 \in X$, and define

$$h_x := f_x - f_{x_0}.$$

In particular, any $y \in X$ will have $|h_x(y)| = |d(x, y) - d(x_0, y)| \le d(x, x_0)$, so h_x is bounded, and it is continuous as the sum of two continuous functions. More explicitly, for any $\varepsilon > 0$, take $\delta = \varepsilon$ so that $d(x_1, x_2) < \delta$ implies

$$|h_x(x_1) - h_x(x_2)| = |d(x, x_1) - d(x, x_2)| \le d(x_1, x_2) < \delta = \varepsilon.$$

We now need to show that the map $h_\bullet: (X, d) \to (B_c(X, \mathbb{R}), d_u)$ is an isometry. Indeed,

$$d_u(h_{x_1}, h_{x_2}) = \sup_{x \in X} \{h_{x_1}(x) - h_{x_2}(x)\} = \sup_{x \in X} \{d(x_1, x) - d(x_2, x)\}.$$

This is certainly upper-bounded by $d(x_1, x_2)$ by the triangle inequality, and we do achieve $d(x_1, x_2)$ at $x = x_2$ because $d(x_1, x_2) - d(x_2, x_2) = d(x_1, x_2)$. So indeed, $d_u(h_{x_1}, h_{x_2}) = d(x_1, x_2)$.

Thus, we have provided an isometry $h_\bullet: (X, d) \to (B_c(X, \mathbb{R}), d_u)$ from (X, d) to the complete metric space $(B_c(X, \mathbb{R}), d_u)$ (see Example 3.32), so $\overline{h_\bullet(X)}$ is a completion for (X, d) by Lemma 3.34. ■

Remark 3.35. Despite the above construction, it is actually fairly non-obvious what functions really are in $\overline{h_\bullet(X)}$.

Theme 4
Compactness

That something so small could be so beautiful.

—Anthony Doerr, [Doe14]

4.1 September 16

We continue the lecture, into compactness.

4.1.1 Compactness

The following is perhaps the most important definition in point-set topology.

Definition 4.1 (Open cover)**.** Fix a topological space $(X, \mathcal{T})$. An *open cover* of X is a collection $\mathcal{U} \subseteq \mathcal{T}$ of open sets such that

$$X = \bigcup_{U \in \mathcal{U}} U.$$

Definition 4.2 (Open subcover)**.** Fix a topological space $(X, \mathcal{T})$. An *(open) subcover* $\mathcal{U}'$ of an open cover $\mathcal{U}$ is an open cover $\mathcal{U}'$ of X such that $\mathcal{U}' \subseteq \mathcal{U}$.

And here is the relevant definition.

Definition 4.3 (Compact)**.** Fix a topological space $(X, \mathcal{T})$. We say that $(X, \mathcal{T})$ is *compact* if and only if every open cover of X has a finite subcover.

Example 4.4. The subset $[0, 1] \subseteq \mathbb{R}$ given the relative topology is compact.

In light of the previous example, it is helpful to extend our definition to subsets of a topological space.

Definition 4.5 (Compact)**.** Fix a topological space $(X, \mathcal{T})$. A subset $A \subseteq X$ is *compact* if and only if A is compact when given the relative topology from X.

Lemma 4.6. Fix a topological space $(X, \mathcal{T})$. Then A is compact if and only if any $\mathcal{U} \subseteq \mathcal{T}$ covering A has a finite subcover covering A.

Proof. The point is to use Lemma 2.42. In one direction, suppose A is compact. Then a cover $\{U_\alpha\}_{\alpha\in\lambda} \subseteq \mathcal{T}$ of A provides the open cover by

$$V_\alpha := A \cap U_\alpha$$

of A. Indeed, $A \cap U_\alpha \subseteq A$ is open, and $\bigcup_{\alpha\in\lambda} V_\alpha = \bigcup_{\alpha\in\lambda}(A \cap U_\alpha) = A$. Thus, compactness provides a finite subset $\lambda' \subseteq \lambda$ such that $\{V_\alpha\}_{\alpha\in\lambda'}$ still covers A, so

$$A = \bigcup_{\alpha\in\lambda'} (A \cap U_\alpha) \subseteq \bigcup_{\alpha\in\lambda'} U_\alpha,$$

meaning that the finite subcover $\{U_\alpha\}_{\alpha\in\lambda'} \subseteq \{U_\alpha\}_{\alpha\in\lambda}$ still covers A.

In the other direction, suppose that each open cover of A from $\mathcal{T}$ has a finite subcover. Now, give A some open cover $\{V_\alpha\}_{\alpha\in\lambda}$ from the relative topology on A. Each open subset V_α can be written as $U_\alpha \cap A$ where $U_\alpha \subseteq X$ is open by Lemma 2.42, so we define

$$\mathcal{U} := \{U_\alpha\}_{\alpha\in\lambda}.$$

Notably, $\bigcup_{\alpha\in\lambda} U_\alpha$ contains $\bigcup_{\alpha\in\lambda} V_\alpha$, which is A, so $\mathcal{U}$ covers A and hence has a finite subset $\lambda' \subseteq \lambda$ such that $\{U_\alpha\}_{\alpha\in\lambda'}$ covers A. But then

$$A = \bigcup_{\alpha\in\lambda'} (A \cap U_\alpha) = \bigcup_{\alpha\in\lambda'} V_\alpha,$$

so $\{V_\alpha\}_{\alpha\in\lambda'}$ provides a finite subcover of $\{V_\alpha\}_{\alpha\in\lambda}$. ■

In light of the above proof, it will be helpful to extend our notion of an open cover.

Notation 4.7. Given a topological space $(X, \mathcal{T})$, we will say that some open sets $\mathcal{U} \subseteq \mathcal{T}$ form an open cover for a subset $A \subseteq X$ if and only if

$$A \subseteq \bigcup_{U\in\mathcal{U}} U.$$

Remark 4.8. We will freely use Lemma 4.6 as a "definition" of compactness without reference.

Example 4.9. Given compact subsets $A_1, A_2 \subseteq X$ of a topological space $(X, \mathcal{T})$, we see that $A_1 \cup A_2$ is also compact. Indeed, given an open cover $\mathcal{U}$ of $A_1 \cup A_2$, we see that $\mathcal{U}$ is an open cover for both A_1 and A_2, so we can find our finite subcovers $\mathcal{U}_1 \subseteq \mathcal{U}$ and $\mathcal{U}_2 \subseteq \mathcal{U}$ by the compactness of A_1 and A_2, respectively. Thus, $\mathcal{U}_1 \cup \mathcal{U}_2 \subseteq \mathcal{U}$ is a finite collection covering A_1 and A_2 and therefore covering $A_1 \cup A_2$.

Here is a quick fact about compactness.

Lemma 4.10. Fix a compact topological space $(X, \mathcal{T})$. Then any closed subset $A \subseteq X$ is compact.

Proof. By Lemma 4.6, pick up an open cover $\mathcal{U}$ of A, and we would like to find a finite subcover. Then we set

$$\mathcal{V} := \mathcal{U} \cup \{X \setminus A\}.$$

Notably, $X \setminus A$ is open in X because A is closed, so we see

$$\bigcup_{U\in\mathcal{V}} U = (X \setminus A) \cup \bigcup_{U\in\mathcal{U}} U \supseteq (X \setminus A) \cup A = X,$$

so $\mathcal{V}$ is an open cover for X. As such, we can find a finite subcover $\mathcal{V}'$ for X, and we set $\mathcal{U}' := \mathcal{V} \cap \mathcal{U}$.

We claim that $\mathcal{U}'$ is a finite subcover of $\mathcal{U}$; indeed, $\mathcal{U}' \subseteq \mathcal{V}$ is finite, and $\mathcal{U}' \subseteq \mathcal{U}$ is a subset. It remains to check that $\mathcal{U}'$ covers A. Well, for any $a \in A$, we can find some $U' \in \mathcal{V}'$ containing a because $\mathcal{V}'$ covers X. However, $a \notin X \setminus A$, so $U' \neq X \setminus A$, so actually $U' \in \mathcal{U}'$. Thus,

$$A \subseteq \bigcup_{U \in \mathcal{U}'} U,$$

which is what we wanted. ■

Example 4.11. Give $X = \mathbb{R}$ the indiscrete topology. Then X has only two open sets, so any nonempty subset $S \subseteq X$ can only be covered by $\{X\}$, which is its own finite subcover. For example, $\{0\}$ is compact in X, but it is not closed because $\mathbb{R} \setminus \{0\}$ is not open.

4.2 September 19

There are questions today.

4.2.1 Compact Hausdorff Spaces

Last class we saw in Example 4.11 that compact subsets of a topological space need not be compact. It turns out that compact subsets of Hausdorff spaces are in fact closed. Let's see this.

Lemma 4.12. Fix a Hausdorff topological space $(X, \mathcal{T})$, and let $A \subseteq X$ be compact. Then, for any $x \notin A$, there are disjoint open sets U and V with $A \subseteq U$ and $x \in V$.

Proof. For each $y \in (X \setminus A)$, the Hausdorff condition promises disjoint open sets V_y and U_y such that $y \in U_y$ and $x \in V_y$. We would like to take the union of all the U_y and the intersection of all the U_x, but the arbitrary intersection of open sets need not be open.

To fix this, we note that $\{U_y\}_{y \in A}$ are some open sets which cover A, so the compactness of A allows us some finite subset $Y \subseteq A$ such that $\{U_y\}_{y \in Y}$ covers A. As such, we set

$$U := \bigcup_{y \in Y} U_y \qquad \text{and} \qquad V := \bigcap_{y \in Y} V_y.$$

Here are our checks.

- Both U and V are open because these are a finite union and a finite intersection of open sets, respectively.
- By construction of Y, we see that $A \subseteq U$.
- Note $x \in V_y$ for all $y \in Y \subseteq A$, so $x \in V$ as well.
- Lastly, we see that U and V are disjoint: for each $z \in U$, we can find some $y \in Y$ such that $z \in U_y$, but then $z \notin V_y$ by construction, so $z \notin V$. ■

Corollary 4.13. Fix a Hausdorff topological space $(X, \mathcal{T})$, and let $A \subseteq X$ be compact. Then A is closed.

Proof. For each $x \notin A$, Lemma 4.12 grants us an open subset V_x containing x which is disjoint from A. It follows $V_x \subseteq X \setminus A$, so we may say

$$(X \setminus A) \subseteq \bigcup_{y \in X \setminus A} V_y \subseteq \bigcup_{y \in X \setminus A} (X \setminus A) = X \setminus A,$$

so $X \setminus A = \bigcup_{y \in X \setminus A} V_y$ shows that $X \setminus A$ is a union of open sets and therefore open. It follows that A is closed. ■

Corollary 4.14. Fix a compact Hausdorff topological space $(X, \mathcal{T})$. Then all closed subsets $A \subseteq X$ and $x \notin A$ have disjoint open subsets U and V with $A \subseteq U$ and $x \in V$.

Proof. Lemma 4.10 says that A is compact, so Lemma 4.12 finishes. ■

The above property is useful enough to deserve a definition.

Definition 4.15 (Regular)**.** A topological space $(X, \mathcal{T})$ is *regular* if and only if each closed subset $A \subseteq X$ and $x \notin A$ have disjoint open subsets U and V with $A \subseteq U$ and $x \in V$.

Example 4.16. Every compact Hausdorff space is regular by Corollary 4.14.

Example 4.17. Any normal, Hausdorff space is regular. For example, metric spaces are regular.

In fact, compact Hausdorff spaces are not just regular but also normal.

Proposition 4.18. Fix a compact Hausdorff space $(X, \mathcal{T})$. Then $(X, \mathcal{T})$ is normal.

Proof. Fix disjoint closed subsets A and B. Then A and B are compact by Lemma 4.10.

Now, for any $y \in B$, we see $y \notin A$, so Lemma 4.12 grants us disjoint open subsets U_y and V_y such that U_y contains A and V_y contains y. As before, we see $\{V_y\}_{y \in B}$ forms an open cover of B, so the compactness of B promises a finite subset $Y \subseteq B$ such that $\{V_y\}_{y \in Y}$ still covers B. Thus, we set

$$U := \bigcap_{y \in Y} U_y \qquad \text{and} \qquad V := \bigcup_{y \in Y} V_y.$$

Here are our checks again.

- Note U is open as a finite intersection of open sets. Similarly, V is open as a union of open sets.
- By construction $A \subseteq U_y$ for each y, so $A \subseteq U$.
- By construction $\{V_y\}_{y \in Y}$ covers B, so $B \subseteq V$.
- Lastly, to see that U and V are disjoint, note that any $z \in V$ has $z \in V_y$ for some $y \in Y$, so $z \notin U_y$, so $z \notin U$. ■

4.2.2 Compact Images

We continue our fact-collection for compact spaces.

Lemma 4.19. Fix a continuous map $f\colon (X,\mathcal{T}_X)\to(Y,\mathcal{T}_Y)$. If $(X,\mathcal{T}_X)$ is compact, then $\operatorname{im} f\subseteq Y$ is also compact.

Proof. For psychological reasons, we may assume that $\operatorname{im} f = Y$, though we will not do this.

Suppose we have an open cover $\{V_\alpha\}_{\alpha\in\lambda}\subseteq\mathcal{T}_Y$ for $\operatorname{im} f$. Then we set

$$\mathcal{U}\coloneqq\left\{f^{-1}(V_\alpha)\right\}_{\alpha\in\lambda}$$

In particular, the continuity of f promises that everyone is $\mathcal{U}$ is open. We claim $\mathcal{U}$ covers X: for any $x\in X$, we see $f(x)\in\operatorname{im} f$, so $f(x)\in V_\alpha$ for some $\alpha\in\lambda$, so $x\in f^{-1}(V_\alpha)\in\mathcal{U}$.

Thus, the compactness of X promises a finite subset $\lambda'\subseteq\lambda$ so that $\left\{f^{-1}(V_\alpha)\right\}_{\alpha\in\lambda'}$ is still an open cover for X. Thus, we can see that the finite collection of open subsets

$$\{V_\alpha\}_{\alpha\in\lambda'}\subseteq\{V_\alpha\}_{\alpha\in\lambda}$$

still covers $\operatorname{im} f$. Indeed, for any $y\in\operatorname{im} f$, find $x\in X$ with $f(x)=y$, so place $x\in f^{-1}(V_\alpha)$ for some $\alpha\in\lambda'$, so $y\in V_\alpha$. ■

Corollary 4.20. Fix a continuous function $f\colon\mathbb{R}\to\mathbb{R}$. Then, for any closed interval $[a,b]$, f achieves its maximum on $[a,b]$.

Proof. Note that $f([a,b])$ is compact, and $\mathbb{R}$ is Hausdorff, so $f([a,b])$ is also closed. Further, $f([a,b])$ is bounded because it is compact. Thus, $f([a,b])$ has all of its limit points and in particular contains its supremum. ■

We take a moment to use this machinery to build an easier test for homeomorphisms; namely, we manifest Remark 2.91.

Proposition 4.21. Fix a compact topological space $(X,\mathcal{T}_X)$ and a Hausdorff topological space $(Y,\mathcal{T}_Y)$. Then any continuous bijection $f\colon X\to Y$ is a homeomorphism.

Proof. The bijectivity of f promises some inverse function $g\colon Y\to X$, which we need to show is continuous. Well, for an open subset $U\subseteq X$, we need to show that $g^{-1}(U)$ is open. But because g is the inverse of f, we see

$$g^{-1}(U)=\{y\in Y:g(y)\in U\}=\{f(x)\in Y:g(f(x))\in U\}=f(U),$$

so we need to show that $f(U)$ is open. Taking compliments, we set $A\coloneqq X\setminus U$ so that A is closed, and we will show that $f(A)$ is closed; this will finish because the bijectivity of f forces

$$f(U)=f(X\setminus A)=f(X)\setminus f(A)=Y\setminus f(A)$$

to be open.

We are now ready to finish the proof. Because $(X,\mathcal{T}_X)$ is compact, A being closed implies that A is compact by Lemma 4.10. It follows by Lemma 4.19 that $f(A)$ is compact, so because $(Y,\mathcal{T}_Y)$ is Hausdorff, we see Lemma 4.12 forces $f(A)$ to be closed. This finishes. ■

4.2.3 Compactness via Closed Sets

It will be helpful to be able to discuss compact sets in terms of closed sets.

Lemma 4.22. A set X is covered by a collection $\mathcal{S} \subseteq \mathcal{P}(X)$ if and only if

$$\bigcap_{S \in \mathcal{S}} (X \setminus S) = \varnothing.$$

Proof. Note

$$\bigcap_{S \in \mathcal{S}} (X \setminus S) = X \Bigg\backslash \bigcup_{S \in \mathcal{S}} X,$$

which is empty if and only if $\bigcup_{S \in \mathcal{S}} X = X$. ∎

Corollary 4.23. Fix a topological space $(X, \mathcal{T})$. Then $(X, \mathcal{T})$ is compact if and only if any collection of closed subsets $\mathcal{V}$ with $\bigcap_{V \in \mathcal{V}} V = \varnothing$ has some finite subcollection $\mathcal{V}' \subseteq \mathcal{V}$ with $\bigcap_{V \in \mathcal{V}'} V = \varnothing$.

Proof. If X is compact, then note any collection of closed subsets $\mathcal{V}$ with $\bigcap_{V \in \mathcal{V}} V = \varnothing$ has

$$X = X \Bigg\backslash \bigcap_{V \in \mathcal{V}} V = \bigcup_{V \in \mathcal{V}} (X \setminus V),$$

so $\mathcal{U} = \{(X \setminus V) : V \in \mathcal{V}\}$ is an open cover. Thus, we can find a finite subset $\mathcal{V}' \subseteq \mathcal{V}$ such that $\mathcal{U}' = \{(X \setminus V) : V \in \mathcal{V}'\}$ covers X, so it follows that $\bigcap_{V \in \mathcal{V}'} V = \varnothing$ by taking complements, as above.

Conversely, we show that X is compact. Well, pick up an open cover $\mathcal{U}$ of X. Then Lemma 4.22 says that $\mathcal{V} = \{(X \setminus V) : V \in \mathcal{V}\}$ has $\bigcap_{V \in \mathcal{V}} V = \varnothing$. By hypothesis on X, we get some finite subcollection $\mathcal{U}' \subseteq \mathcal{U}$ such that $\bigcap_{U \in \mathcal{U}'} (X \setminus U) = \varnothing$, so Lemma 4.22 says $\mathcal{U}'$ covers X. ∎

It will be useful to have some language to describe this.

Definition 4.24 (Finite intersection property)**.** Fix a set X. A collection $\mathcal{S} \subseteq \mathcal{P}(X)$ has the *finite intersection property* if and only if any nonempty finite subcollection $\mathcal{S}' \subseteq \mathcal{S}$ has

$$\bigcap_{S \in \mathcal{S}'} S \neq \varnothing.$$

In particular, we get the following.

Proposition 4.25. Fix a topological space $(X, \mathcal{T})$. Then $(X, \mathcal{T})$ is compact if and only if any collection $\mathcal{V}$ of closed subsets with the finite intersection property has

$$\bigcap_{V \in \mathcal{V}} V \neq \varnothing.$$

Proof. Applying contraposition to the conclusion, we are saying that any collection $\mathcal{V}$ with $\bigcap_{V \in \mathcal{V}} V = \varnothing$ has some finite subcollection $\mathcal{V}' \subseteq \mathcal{V}$ with $\bigcap_{V \in \mathcal{V}'} V = \varnothing$. This is equivalent to $(X, \mathcal{T})$ being compact by Corollary 4.23. ∎

Remark 4.26. It is somewhat important to notice that the proof of Proposition 4.25 does not require the Axiom of Choice to prove. It is purely moving around definitions cleverly.

4.3 September 21

Today we begin talking about Tychonoff's theorem.

4.3.1 Comments on Choice

Here is our main result for today.

Theorem 4.27 (Tychonoff)**.** Fix a collection $\{(X_\alpha, \mathcal{T}_\alpha)\}_{\alpha\in\lambda}$ of compact topological spaces, and give the product space $X := \prod_{\alpha\in\lambda} X_\alpha$ the product topology. Then X is compact.

Notably, we are not requiring the spaces X_α to be Hausdorff.

Warning 4.28. The proof of Theorem 4.27 will be the hardest part of this course.

Remark 4.29. The reason for Warning 4.28 is that we need to at least know that X is nonempty to say anything about X at all, and an arbitrary product being nonempty is equivalent to the Axiom of Choice. In fact, Theorem 4.27 (notably not assuming that the X_α are Hausdorff!) actually implies the Axiom of Choice, as shown by John Kelly.

To prepare ourselves, we will point out a few of the main ingredients we will use. We will use the Axiom of Choice, which we will go ahead and state now.

Axiom 4.30 (Choice)**.** Given a collection of nonempty sets $\{S_\alpha\}_{\alpha\in\lambda}$, the product $\prod_{\alpha\in\lambda} S_\alpha$ is nonempty.

We will also use Zorn's lemma. To state Zorn's lemma, we begin by defining a partially ordered set and its chains.

Definition 4.31 (Poset)**.** A *partially ordered set* or *poset* is a set P equipped with a reflexive, antisymmetric, and transitive relation $\le\ \subseteq P \times P$.

Example 4.32. Given a set X, the power set $\mathcal{P}(X)$ is a partially ordered set under inclusion $\subseteq$. Here are the checks.

- Reflexive: for $A \in \mathcal{P}(X)$, we see $A \subseteq A$.
- Antisymmetric: for $A, B \in \mathcal{P}(X)$, we see $A \subseteq B$ and $B \subseteq A$ implies $A = B$.
- Transitive: for $A, B, C \in \mathcal{P}(X)$, we see $A \subseteq B$ and $B \subseteq C$ implies $A \subseteq C$.

Replacing all the $\subseteq$s with $\supseteq$s shows that $\mathcal{P}(X)$ is also a partially ordered set under containment $\supseteq$.

Posets have very natural subposets.

Definition 4.33 (Subposet)**.** Given a partially ordered $(P, \le)$, a *subposet* is a subset $S \subseteq P$ equipped with the restricted partial order $\le \cap\, (S \times S)$.

All the checks for $(S, \le \cap\, (S \times S))$ being a partially ordered set are inherited directly from P, so the proof amounts to just writing them down.

Example 4.34. Given a topological space $(X, \mathcal{T})$, we see that $\mathcal{T}$ is a subposet of $\mathcal{P}(X)$, where $\mathcal{P}(X)$ can be given the partial order $\subseteq$ or $\supseteq$ from Example 4.32.

And here are our chains.

Definition 4.35 (Chain). Fix a partially ordered set $(P, \leq)$. Then a *chain* is a subset $C \subseteq P$ such that the subposet $(C, \leq)$ is totally ordered.

Zorn's lemma is interested in special kinds of partially ordered sets.

Definition 4.36 (Inductively ordered). A partially ordered set $(P, \leq)$ is *inductively ordered* if and only if every chain $C \subseteq P$ has an upper bound in P. In other words, there is an element $p \in P$ such that $c \leq p$ for all $c \in C$.

And here is Zorn's lemma.

Axiom 4.37 (Zorn's lemma). An inductively ordered partially ordered set $(P, \leq)$ has a maximal element.

Remark 4.38. It turns out that the Axiom of Choice (in the form of Zorn's lemma) is also equivalent to every vector space having a basis. (In one direction, given a vector space V, one can build a basis by taking a maximal linearly independent set of vectors in V.) One can get a feeling for the other direction because the $\mathbb{Q}$-vector space $\mathbb{R}$ doesn't have any "constructible" basis.

Remark 4.39. The fact that every (commutative) ring has a maximal ideal containing any given proper ideal is also equivalent to the Axiom of Choice (in the form of Zorn's lemma). Here are two examples.

- Given any set S, finding a maximal ideal of the ring $R := \mathbb{F}_2^S$ (whose operations are pointwise from $\mathbb{F}_2$) which contains the ideal $\mathbb{F}_2^{\oplus S}$ requires knowing that R is nonempty.
- The ring $R := C([0, \infty))$ of continuous $\mathbb{R}$-valued functions has the ideal

$$I := \left\{ f \in R : \lim_{x \to \infty} f(x) = 0 \right\}$$

  doesn't have any constructible maximal ideals containing it.

For our next example, we define a filter.

Definition 4.40 (Filter). Fix a set X. A *filter* $\mathcal{F}$ on X is a collection of nonempty subsets of X satisfying the following conditions.

(a) $\mathcal{F}$ is closed under finite intersection.

(b) If $A \in \mathcal{F}$ and $A \subseteq B \subseteq X$, then $B \in \mathcal{F}$.

Example 4.41. Given a topological space $(X, \mathcal{T})$ and a subset $A \subseteq X$, the subposet $\mathcal{T}$ of $(\mathcal{P}(X), \subseteq)$ has a filter $\mathcal{F}$ of all those open subsets containing A.

Example 4.42. Given a set X, the collection of subsets containing a given point $p \in X$ is a filter and in fact a "maximal" filter.

The point is that Zorn's lemma automatically promises us maximal filters, or "ultrafilters."

Example 4.43. Fix $X := [0, \infty)$. Then the collection $\mathcal{F}$ of the subsets of $A \subseteq X$ which contain $[n, \infty)$ for some integer n is a filter. However, there is no obvious maximal filter.

4.4 September 23

We continue discussing Tychonoff's theorem.

4.4.1 Tychonoff's Theorem

Here is our statement.

> **Theorem 4.27** (Tychonoff)**.** Fix a collection $\{(X_\alpha, \mathcal{T}_\alpha)\}_{\alpha\in\lambda}$ of compact topological spaces, and give the product space $X := \prod_{\alpha\in\lambda} X_\alpha$ the product topology. Then X is compact.

Proof. We will use Proposition 4.25. For each α, let $\pi_\alpha \colon X \to X_\alpha$ denote the canonical projection. Let $\mathcal{V}$ be a collection of closed subsets of X satisfying the finite intersection property, and we will show that $\bigcap_{V\in\mathcal{V}} V$ is nonempty. We proceed in steps.

1. The beginning of this proof does not use topology. Let $\Omega_{\mathcal{V}}$ be the collection of families of subsets $\mathcal{F}$ of X which contain $\mathcal{V}$ and have the finite intersection property. We claim that $W_{\mathcal{V}}$ is inductively ordered under $\supseteq$.

   Well, let $\Omega \subseteq \Omega_{\mathcal{V}}$ be some chain, and we define the collection
   $$\mathcal{U} := \bigcup_{\mathcal{F}\in\Omega} \mathcal{F},$$
   which we claim is the required upper bound for Ω. Of course, each $\mathcal{F}$ contains $\mathcal{V}$, and $\mathcal{U} \supseteq \mathcal{F}$ for each $\mathcal{F}$, so $\mathcal{U}$ both contains $\mathcal{V}$ and is an upper bound for Ω. It remains to show $\mathcal{U} \in \Omega_{\mathcal{V}}$, for which we need to show that $\mathcal{U}$ has the finite intersection property.

   For this, find some finite subcollection of nonempty subsets $\{A_k\}_{k=1}^n \subseteq \mathcal{U}$ which we would like to show have nonempty intersection. Now, for each k, there is some $\mathcal{F}_k \in \Omega$ containing A_k, by construction of $\mathcal{U}$ as the union over Ω. Because the number of subsets is finite, and because Ω is totally ordered, we may find the largest of the $\mathcal{F}_k$, which we call $\mathcal{F}$.

   Now, $\mathcal{F} \in \Omega \subseteq \Omega_{\mathcal{V}}$ must have the finite intersection property, so $\{A_k\}_{k=1}^n \subseteq \mathcal{F}$ forces
   $$\bigcap_{k=1}^n A_k \neq \varnothing,$$
   which is what we wanted. This completes the proof.

2. From the previous step, Zorn's lemma promises a maximal family $\mathcal{M}$. We claim that $\mathcal{M}$ is closed under taking finite intersections. Indeed, define $\mathcal{M}'$ as the set of all finite intersections of $\mathcal{M}$, and we will show that $\mathcal{M}' = \mathcal{M}$.

   Well, certainly $\mathcal{M} \subseteq \mathcal{M}'$ because intersections of exactly one set $F \in \mathcal{M}$ will just recover $F \in \mathcal{M}'$. Thus, if we can show $\mathcal{M}' \in \Omega_{\mathcal{V}}$, the desired equality $\mathcal{M}' = \mathcal{M}$ will follow by maximality.

   Certainly, we of course have $\mathcal{M} \supseteq \mathcal{V}$, so $\mathcal{M}' \supseteq \mathcal{V}$ as well. So to show $\mathcal{M}' \in \Omega_{\mathcal{V}}$, it remains to show the finite intersection property. Well, let $\{A_k\}_{k=1}^n \subseteq \mathcal{M}'$ be some finite subcollection of nonempty subsets, and we show their intersection is nonempty. By definition of $\mathcal{M}'$, each k lets us write
   $$A_k = \bigcap_{\ell=1}^{n_k} B_{k,\ell}$$
   for some subsets $B_{k,\ell} \in \mathcal{M}$; because $A_k \in \mathcal{M}'$ is nonempty, we see that $B_{k,\ell} \in \mathcal{M}$ is nonempty, so the finite intersection property on $\mathcal{M}$ tells us that
   $$\bigcap_{k=1}^n A_k = \bigcap_{k=1}^n \bigcap_{\ell=1}^{n_k} B_{k,\ell}$$
   is nonempty, which is what we wanted.

3. We claim that if a subset $B \subseteq X$ has $B \cap A \neq \varnothing$ for each $A \in \mathcal{M}$, then in fact $B \in \mathcal{M}$. Indeed, define $\mathcal{M}'' := \mathcal{M} \cup \{B\}$, and we show $\mathcal{M}'' = \mathcal{M}$.

   Certainly $\mathcal{M} \subseteq \mathcal{M}''$, so it is enough by maximality of $\mathcal{M}$ to show $\mathcal{M}'' \in \Omega_{\mathcal{V}}$. Certainly $\mathcal{V} \subseteq \mathcal{M} \subseteq \mathcal{M}''$, so it remains to show that $\mathcal{M}''$ satisfies the finite intersection property.

   For this, pick up some finite subcollection of nonempty subsets $\{A_k\}_{k=1}^n \subseteq \mathcal{M}''$, and we show their intersection is nonempty. If none of these subsets are B, then in fact $\{A_k\}_{k=1}^n \subseteq \mathcal{M}$, so the finite intersection property for $\mathcal{M}$ forces
$$\bigcap_{k=1}^n A_k \neq \varnothing.$$
   Otherwise, say $B = A_1$ without loss of generality. Then we may assume $B \neq A_k$ for each $k > 1$, so $A_k \in \mathcal{M}$ for each $k > 1$, so we note
$$\bigcap_{k=1}^n A_k = B \cap \bigcap_{k=2}^n A_k.$$
   However, $\mathcal{M}$ is closed under finite intersection, so in fact $\bigcap_{k=2}^n A_k \in \mathcal{M}$, and by the finite intersection property, we have that $\bigcap_{k=2}^n A_k$ is nonempty. Thus, by hypothesis on B, we see
$$B \cap \bigcap_{k=2}^n A_k \neq \varnothing,$$
   which is what we wanted.

4. We now begin touching our product. For given $\alpha \in \lambda$ and $\mathcal{F} \in \Omega_{\mathcal{V}}$, we claim that
$$\pi_\alpha(\mathcal{F}) := \{\pi_\alpha(A) : A \in \mathcal{F}\}$$
   satisfies the finite intersection property. Fix a finite subcollection of nonempty subsets $\{\pi_\alpha(A_k)\}_{k=1}^n$ of $\pi_\alpha(\mathcal{F})$, and we will show its intersection is nonempty. Then we must have A_k being nonempty for each k, so the finite intersection property on $\mathcal{F}$ forces
$$\bigcap_{k=1}^n A_k \neq \varnothing.$$
   Finding some a in this intersection, we see $\pi_\alpha(a) \in \pi_\alpha(A_k)$ for each k, so $\pi_\alpha(a)$ belongs in
$$\bigcap_{k=1}^n \pi_\alpha(A_k),$$
   thus making this intersection nonempty.

5. And now the topology begins. For given α, note that
$$\overline{\mathcal{M}}_\alpha := \{\overline{\pi_\alpha(A)} : A \in \mathcal{M}\}$$
   has the finite intersection property by the previous step. Namely, any finite subcollection of nonempty subsets $\{\pi_\alpha A_k\}_{k=1}^n$ has a nonempty intersection, so writing
$$\varnothing \neq \bigcap_{k=1}^n \pi_\alpha(A_k) \subseteq \bigcap_{k=1}^n \overline{\pi_\alpha(A_k)}$$
   gives what we want. However, X_α is compact (!), so Proposition 4.25 tells us that
$$\bigcap_{A \in \overline{\mathcal{M}}_\alpha} A \neq \varnothing.$$

6. Directly invoking the Axiom of Choice, we may find some $x_\alpha \in \bigcap_{A\in\overline{\mathcal{M}_\alpha}} A$ for each α. Set $x := (x_\alpha)_{\alpha\in\lambda}$ to be the corresponding element of X.

   We claim that each nonempty $A \in \mathcal{M}$ has $x \in \overline{A}$. By Lemma 2.56, it suffices to show that every open subset U containing x has nonempty intersection with A. Because each open subset U containing x has a(n open) basis set $B \subseteq U$ containing x, it suffices to check $B \cap A \neq \varnothing$ for basis elements, and $B \cap A \subseteq U \cap A$ will give the result.

   There are three steps. Observe that we must invoke the definition of the product topology on X to talk topologically about X, so we do so here.

   (a) We begin by checking this on the sub-base. For each $\alpha \in \lambda$, fix some sub-base element $\pi_\alpha^{-1}(U_\alpha)$ (where $U_\alpha \subseteq X_\alpha$ is open) containing x, and we claim

   $$\pi_\alpha^{-1}(U_\alpha) \cap A \stackrel{?}{\neq} \varnothing.$$

   Well, $x \in \pi_\alpha^{-1}(U_\alpha)$ requires $x_\alpha \in U_\alpha$, but $A \in \mathcal{M}$ forces $x_\alpha \in \overline{\pi_\alpha(A)}$. Thus, $U_\alpha \cap \pi_\alpha(A) \neq \varnothing$ by Lemma 2.56, so there is some $a \in A$ with $\pi_\alpha(a) \in U_\alpha$, so $\pi_\alpha^{-1}(U_\alpha) \cap A$ is in fact nonempty.

   (b) We show that each basis set containing x lives in $\mathcal{M}$. Part (a) above added to item 3 directly shows that every sub-base open set containing x lives in $\mathcal{M}$. Thus, item 2 tells us that any finite intersection of sub-basic sets containing x live in $\mathcal{M}$ as well, but these are exactly the basic sets containing x. (Namely, any basic set is the intersection of sub-basic sets, and x living in the basic set forces x to still live in those sub-basic sets.)

   (c) It follows from the finite intersection property for $\mathcal{M}$ that any basic set B containing x has $B \in \mathcal{M}$ and therefore $A \cap B \neq \varnothing$ because A is nonempty.

   The above steps finish this part.

7. We finish the proof. Any $V \in \mathcal{V}$ is closed and has $V \in \mathcal{M}$. By the above point, we see $x \in \overline{V}$, so $x \in V$ by Example 2.55, so we have exhibited

   $$\bigcap_{V\in\mathcal{V}} V \neq \varnothing.$$

The above steps have showed that $\bigcap_{V\in\mathcal{V}} \mathcal{V} \neq \varnothing$ from $\mathcal{V}$ having the finite intersection property, so we conclude that X is compact by Proposition 4.25. ■

4.5 September 26

We begin class by finishing the proof of Tychonoff's theorem (Theorem 4.27). I have gone ahead and just edited Friday's lecture for continuity.

4.5.1 Remarks on Tychonoff's Theorem

Here are some remarks.

Remark 4.44. Intuitively, the maximal element $\mathcal{M}$ is constructed in order to become some filter focused around the single point x. Similar to maximal ideals corresponding to points, adding in all the "maximality" constraints for $\mathcal{M}$ hones in our focus to the single constructed point x.

Remark 4.45. Here is an application of Theorem 4.27. One can show that any normed vector space $(V, \|\cdot\|)$ has "lots" of continuous functionals by extending those found on a finite-dimensional subspace; let V' be the complete normed vector space of continuous linear functionals. (The norm of some $v' \in V'$ is its Lipschitz constant, using Lemma 2.69.) Fixing the unit ball B of V', one can give V' the weakest topology making all the linear functionals "from V" continuous (this is the weak-* topology), which one can show is both Hausdorff and compact (!). This is the Banach–Alaoglu theorem, and it follows from Theorem 4.27 by showing the space we want is a closed subspace of the compact space

$$\prod_{v\in V}[-\|v\|, \|v\|].$$

Remark 4.46 (β-compactification)**.** Let $A := C([0,\infty))$ be the space of bounded continuous function $[0,\infty) \to \mathbb{R}$, which we can see directly is an $\mathbb{R}$-algebra by taking r to the constant function r. Let A' be the set of continuous functions $A \to \mathbb{R}$. Notably, any $x \in \mathbb{R}$ gives a continuous ring homomorphism $A \to \mathbb{R}$ by $f \mapsto f(x)$, so we let Y be the set of all homomorphisms $A \to \mathbb{R}$. Again, A' is compact using the weak-* topology, and so Y as a closed subset of A' can be given a compact topology. Then one can show that A is homeomorphic to $C(Y)$.

4.5.2 Tychonoff's Theorem and Choice

We now show that Tychonoff's theorem implies the Axiom of Choice.

Theorem 4.47 (Kelley)**.** Tychonoff's theorem implies the Axiom of Choice.

Proof. Assume Theorem 4.27 is true. Let $\{X_\alpha\}_{\alpha\in\lambda}$ be a collection of nonempty sets. We want to show that

$$X := \prod_{\alpha\in\lambda} X_\alpha$$

is nonempty.

The trick is to enlarge the X_α to be able to give them a suitable topology. Choose some (set) ω which does not live in $\bigcup_\alpha X_\alpha$; for example, setting ω to be equal to this set will do (using the Axiom of Foundation). Then we set

$$Y_\alpha := X_\alpha \cup \{\omega\},$$

which we give the topology $\mathcal{T}_\alpha := \{Y_\alpha, \varnothing, X_\alpha, \{\omega\}\}$. We quickly check that this is a topology.

- Note $\varnothing$ and Y_α are open.
- Arbitrary union: let $\mathcal{U} \subseteq \mathcal{T}_\alpha$ be a collection. Note that $\mathcal{U}$ is necessarily finite, so it suffices by induction to show that $U \cup U' \in \mathcal{T}_\alpha$ for any $U, U' \in \mathcal{T}_\alpha$. We have the following cases.
  - If $U = \varnothing$ or $U' = \varnothing$, then we get $U \cup U' \in \{U, U'\} \subseteq \mathcal{T}_\alpha$.
  - If $U = Y_\alpha$ or $U' = Y_\alpha$, then $U \cup U' = Y_\alpha \in \mathcal{T}_\alpha$.
  - Note $U = U'$ gives $U \cup U' = U \in \mathcal{T}_\alpha$.
  - We have left to deal with $\{U, U'\} \subseteq \{X_\alpha, \{\omega\}\}$ where U and U' are distinct, which means we just have to check $\mathcal{T}_\alpha \cup \{\omega\} = Y_\alpha$ is open.
- Finite intersection: note that $U \in \mathcal{T}_\alpha$ implies $Y_\alpha \setminus U \in \mathcal{T}_\alpha$ because $Y_\alpha \setminus \{\omega\} = X_\alpha$ and $Y_\alpha \setminus \varnothing = Y_\alpha$, and the other checks follow. Thus, we note any finite collection $\mathcal{U} \subseteq \mathcal{T}_\alpha$ has

$$Y_\alpha \Bigg\backslash \bigcap_{U\in\mathcal{U}} U = \bigcup_{\alpha\in\lambda}(Y_\alpha \setminus U)$$

  is a union of open sets and hence open. It follows that our intersection also lives in $\mathcal{T}_\alpha$.

Additionally, because $\mathcal{T}_\alpha$ has only finitely many sets, the space $(Y_\alpha, \mathcal{T}_\alpha)$ is compact: any subcollection of $\mathcal{T}_\alpha$ is finite, so all open covers of Y_α are automatically finite. It follows that the product

$$Y := \prod_{\alpha\in\lambda} Y_\alpha$$

is compact by applying Theorem 4.27 (!).

We will now extract out our element of X using compactness of Y via Proposition 4.25. Let $\pi_\alpha \colon Y \to Y_\alpha$ be the canonical projection. Note that $Y_\alpha \setminus X_\alpha = \{\omega\}$ is open in Y_α, so $X_\alpha \subseteq Y_\alpha$ is closed, so $V_\alpha := \pi_\alpha^{-1}(X_\alpha)$ is a closed subset of Y by the continuity of π_α (using Remark 2.49).

We now claim that the closed sets $\{V_\alpha\}_{\alpha\in\lambda}$ satisfy the finite intersection property: given a finite subcollection $\{V_{\alpha_i}\}_{i=1}^n$, one may finitely (!) choose a point $x_{\alpha_i} \in X_{\alpha_i}$. So we define

$$y_\alpha := \begin{cases} x_{\alpha_i} & \alpha \in \{\alpha_1, \ldots, \alpha_n\}, \\ \omega & \alpha \notin \{\alpha_1, \ldots, \alpha_n\}, \end{cases}$$

so the point $(y_\alpha)_{\alpha\in\lambda} \in Y$ has $\pi_{\alpha_i}(y) \in X_{\alpha_i}$ for each α_i, so $y_{\alpha_i} \in \pi_{\alpha_i}^{-1}(X_\alpha) = V_\alpha$, so

$$y \in \bigcap_{i=1}^n V_{\alpha_i}.$$

So we have verified the finite intersection property.

It follows from Proposition 4.25 that we can find

$$y \in \bigcap_{\alpha\in\lambda} V_\alpha.$$

However, this implies that each $\alpha \in \lambda$ has $y \in V_\alpha$ and so $\pi_\alpha(y) \in X_\alpha$. It follows that

$$y \in \prod_{\alpha\in\lambda} X_\alpha,$$

which finishes the proof. ■

Remark 4.48. The topology on Y_α need not be Hausdorff, so we needed Theorem 4.27 to allow non-Hausdorff spaces.

4.6 September 28

Today we discuss compactness for metric spaces.

4.6.1 Totally Bounded Spaces

Here is a quick lemma.

Lemma 4.49. Fix a compact metric space (X, d). For any $\varepsilon > 0$, there are finitely many points $\{x_i\}_{i=1}^n$ such that

$$X = \bigcup_{i=1}^n B(x_i, \varepsilon).$$

Proof. Note that of course

$$X = \bigcup_{x \in X} \{x\} \subseteq \bigcup_{x \in X} B(x, \varepsilon) = X,$$

so $\{B(x,\varepsilon)\}_{x \in X}$ is an open cover for X (see Example 2.6). The result follows by extracting a finite subcover. ■

This is a pretty nice finiteness property for a metric space to have, so we give it a name.

Definition 4.50 (Totally bounded)**.** Fix a metric space (X, d). A subset $A \subseteq X$ is *totally bounded* if and only if any $\varepsilon > 0$ has a finite set $\{x_i\}_{i=1}^n \subseteq A$ for which

$$A \subseteq \bigcup_{i=1}^n B(x_i, \varepsilon).$$

If X is totally bounded, we say that (X, d) is totally bounded.

Example 4.51. Any compact metric space is totally bounded by Lemma 4.49.

It's going to turn out that totally bounded is pretty close to compactness. Here is a quick sanity check.

Lemma 4.52. A totally bounded metric space (X, d), and $A \subseteq X$, then A with the induced metric is totally bounded.

Proof. For any $\varepsilon > 0$, we see that there is a finite set $S \subseteq X$ for which

$$X = \bigcup_{x \in S}^{n} B(x, \varepsilon)$$

because (X, d) is totally bounded. Now, let $T \subseteq S$ be the subset for which $B(x, \varepsilon) \cap A \neq \varnothing$ for each $x \in S$, and we then find some $y_x \in B(x, \varepsilon) \cap A$ for each $x \in T$. We now claim that

$$A \subseteq \bigcup_{x \in T}^{m} B(y_x, \varepsilon),$$

which will finish the proof. Indeed, if $a \in A$, then $a \in X$, so we can find some $x_0 \in S$ with $a \in B(x_0, \varepsilon/2)$. It follows that

$$d(a, y_{x_0}) \leq d(a, x_0) + d(x_0, y_{x_0}) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

so we get

$$a \in B(y_{x_0}, \varepsilon) \subseteq \bigcup_{x \in T} B(y_x, \varepsilon),$$

which is what we wanted. ■

Lemma 4.53. Fix a metric space (X, d) and a subset $A \subseteq X$ which is totally bounded. Then $\overline{A}$ is also totally bounded.

Proof. Fix any $\varepsilon > 0$. Because A is totally bounded, we may find $\{a_i\}_{i=1}^n \subseteq A$ for which

$$A \subseteq \bigcup_{i=1}^n B(a_i, \varepsilon/2).$$

We now claim that

$$\overline{A} \stackrel{?}{\subseteq} \bigcup_{i=1}^{n} B(a_i, \varepsilon),$$

which will finish the proof. Indeed, if $x \in \overline{A}$, then Lemma 2.56 tells us that $B(x, \varepsilon/2) \cap \overline{A}$ is nonempty, so place $a \in A \cap B(x, \varepsilon/2)$. By hypothesis on the a_i, there exists some a_i such that $a \in B(a_i, \varepsilon/2)$ as well, so

$$d(x, a_i) \le d(x, a) + d(a, a_i) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

so

$$x \in B(a_i, \varepsilon) \subseteq \bigcup_{i=1}^{n} B(a_i, \varepsilon).$$

The claim follows. ■

4.6.2 Nets

It will be beneficial to us to be able to talk about nets for convergence instead of just sequences.

Definition 4.54 (Directed set)**.** A partially ordered set Λ is a *directed set* if and only if any $a, b \in \Lambda$ have some $c \in \Lambda$ for which $c \ge a, b$.

Example 4.55. Any totally ordered set is a directed set. In particular, any $a, b \in \Lambda$ will have $a \ge b$ or $b \ge a$, so we just set c to be the larger of the two.

Definition 4.56 (Net)**.** Fix a topological space $(X, \mathcal{T})$. Given a directed set Λ, a net is a Λ-indexed sequence $\{x_\alpha\}_{\alpha \in \Lambda}$ in X.

Definition 4.57 (Cluster point)**.** Fix a topological space $(X, \mathcal{T})$ and a net $\{x_\alpha\}_{\alpha \in \Lambda}$. Then $x \in X$ is a *cluster point* if and only if, for any open subset U containing x and $\alpha \in \Lambda$, there is some $\alpha' > \alpha$ for which $x_{\alpha'} \in U$.

Remark 4.58. Fix a metric space (X, d). Then a cluster point x of a Cauchy sequence $\{x_n\}_{n \in \mathbb{N}}$ in X is in fact a limit point. Indeed, for any $\varepsilon > 0$, find some N_1 for which $m, n \ge N_1$ has $d(x_m, x_n) < \varepsilon/2$. Additionally, being a cluster point means there is N_2 with $d(x, x_{N_2}) < \varepsilon/2$. Thus, setting $N := \max\{N_1, N_2\}$, any $n > \max\{N_1, N_2\}$ has

$$d(x_n, x) \le d(x_n, x_{N_2}) + d(x_{N_2}, x) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Here is the application to metric spaces.

Proposition 4.59. Fix a compact topological space $(X, \mathcal{T})$. Then any net $\{x_\alpha\}_{\alpha \in \Lambda}$ has a cluster point.

Proof. Define

$$A_\alpha := \{x_\beta : \beta > \alpha\}.$$

Observe $\beta \ge \alpha$ implies $A_\beta \subseteq A_\alpha$, so $A_\beta \subseteq \overline{A_\alpha}$, so $\overline{A_\beta} \subseteq \overline{A_\alpha}$.

Additionally, we note that any finite subset of the A_α have a nonempty intersection. Indeed, for any finite $S \subseteq \Lambda$, inductively applying the fact that Λ is a directed set promises us some $\omega \in \Lambda$ with $\omega \ge \alpha$ for each $\alpha \in S$. It follows that $x_\omega \in A_\alpha$ for each $\alpha \in S$, so

$$\bigcap_{n \in S} A_n,$$

contains x_ω and hence is not empty.

Now, because $A_\alpha \subseteq \overline{A_\alpha}$, we see that the $\overline{A_\alpha}$ also have the finite intersection property: for any finite $S \subseteq \Lambda$, see

$$\varnothing \neq \bigcap_{\alpha \in S} A_\alpha \subseteq \bigcap_{\alpha \in S} \overline{A_\alpha},$$

But now the $\overline{A_\alpha}$ are closed, so the compactness of X (!) tells us that there is an element

$$x \in \bigcap_{\alpha \in \Lambda} \overline{A_\alpha}$$

by Proposition 4.25.

It remains to check that x is a cluster point. Indeed, for any open set U containing x, we see that $x \in \overline{A_\alpha}$ and so $U \cap A_\alpha \neq \varnothing$ for each α by Lemma 2.56. As such, for any $\alpha \in \Lambda$, we are being promised $U \cap A_\alpha \neq \varnothing$, so there is x_β with $\beta \geq \alpha$ with $x_\beta \in U$. This finishes. ■

Corollary 4.60. Any compact metric space (X, d) is complete.

Proof. Fix a Cauchy sequence $\{x_n\}_{n\in\mathbb{N}}$ of X. Because X is compact as a topological space, Proposition 4.59 promises us some cluster point $x \in X$. But then x is our limit point by Remark 4.58. ■

4.6.3 A "Metric" Completeness

Here is our capstone result: a converse for Lemma 4.49 combined with Corollary 4.60.

Theorem 4.61. Fix a metric space (X, d). If X is complete and totally bounded, then X is compact.

Proof. Suppose that X is not compact and totally bounded. We show that X is not complete. Because X is not compact, we can find an open cover $\mathcal{U}$ of X with no finite subcover.

Notice that, for any fixed $\varepsilon > 0$, being totally bounded means we can find some finite $S \subseteq X$ for which X is covered by the $\{B(x, \varepsilon)\}_{x \in S}$. If it were the case that each $x \in S$ has $B(x, \varepsilon)$ covered by some finite cover $\{U_{x,i}\}_{i=1}^{n_x} \in \mathcal{U}$, then we could write

$$X \subseteq \bigcup_{x \in S} B(x, \varepsilon) \subseteq \bigcup_{x \in S} \left(\bigcup_{i=1}^{n_x} U_{x,i} \right),$$

giving our finite subcover of $\mathcal{U}$. However, this violates the fact that $\mathcal{U}$ has no finite subcover, so there must be some $x \in S$ not covered by any finite subset of $\mathcal{U}$.

We can run the above argument starting with $\varepsilon = 1/2$ and find our x_1. Then we replace X with $B(x_1, 1/2)$ where $B(x_1, 1/2)$ has no finite subcover by $\mathcal{U}$, so running the argument with $\varepsilon = 1/2^2$ on the totally bounded space $B(x_1, 1/2)$ grants us $x_2 \in B(x_1, 1/2)$ such that $B(x_2, 1/2^2)$ still has no finite subcover by $\mathcal{U}$. Going again, we run the argument with $\varepsilon = 1/2^3$ on the totally bounded space $B(x_2, 1/2^2)$, so we get a totally bounded ball $B(x_3, 1/2^3)$ with no finite subcover by $\mathcal{U}$.

We can continue this process inductively, which gives a sequence $\{x_n\}_{n\in\mathbb{N}}$ such that each $n \in \mathbb{N}$ has $B(x_n, 1/2^n)$ with no finite cover by $\mathcal{U}$ and

$$d(x_n, x_{n+1}) \leq 1/2^n.$$

A standard argument shows that $\{x_n\}_{n\in\mathbb{N}}$ is a Cauchy sequence.[1] To finish the proof, we claim that it has no limit point.

[1] For any $m \geq n$, note that $d(x_m, x_n) \leq \sum_{k=n}^{m-1} d(x_{k+1}, x_k) \leq \sum_{k=n}^{m-1} 1/2^k < \sum_{k=n}^{\infty} 1/2^k = 1/2^{n-1}$. Namely, we see that $m, n \to \infty$ makes $d(x_m, x_n) \to 0$.

Indeed, suppose for the sake of contradiction that $x_n \to x$ as $n \to \infty$. Then we find some $U \in \mathcal{U}$ containing x, and by definition of a set being open, we can find some open ball $B(x, \varepsilon)$ contained in U. We now find some n large enough so that $1/2^n < \varepsilon/2$ and $d(x_n, x) < \varepsilon/2$ so that any $y \in B(x_n, 1/2^n)$ has

$$d(x, y) \le d(x, x_n) + d(x_n, y) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

so $y \in B(x, \varepsilon)$. It follows $B(x_n, 1/2^n) \subseteq B(x, \varepsilon) \subseteq U$, which is a contradiction to the construction of $B(x_n, 1/2^n)$. This completes the proof. ■

Corollary 4.62. Fix a complete metric space (X, d). Then a subset $A \subseteq X$ is compact if and only if A is closed and totally bounded.

Proof. In the forward direction, if A is compact, then A is totally bounded by Lemma 4.49, and A is closed by Corollary 4.13 because (X, d) is a metric space and thus Hausdorff. In the reverse direction, if A is closed and totally bounded, then A is complete by Corollary 2.52 and therefore compact by Theorem 4.61. ■

4.7 September 30

There are no questions.

4.7.1 Totally Bounded for Function Spaces

We continue our discussion of compactness in metric spaces. Fix a topological space $(X, \mathcal{T})$ and a metric space (M, d) so that we can give the space of bounded continuous functions $B_c(X, M)$ the uniform metric d_u by Proposition 3.25. We would like to understand the compact subset of $B_c(X, M)$, so Corollary 4.62 tells us that we are really interested in totally bounded subsets, and we'll take the closure afterward to get our compact sets.

Here are a few lemmas.

Lemma 4.63. Fix a topological space $(X, \mathcal{T})$ and a metric space (M, d) so that we can give the space of bounded continuous functions $B_c(X, M)$ the uniform metric d_u. Fixing a totally bounded subset $\mathcal{F} \subseteq B_c(X, M)$, the set

$$\{s(x) : s \in \mathcal{F}\}$$

totally bounded for any fixed $x \in X$.

Proof. For any $\varepsilon > 0$ has a finite set $\{f_1, \ldots, f_n\} \subseteq \mathcal{F}$ so that $\mathcal{F}$ is covered by the $B(s_i, \varepsilon/2)$. This is equivalent to saying that any $f \in \mathcal{F}$ has some f_i with

$$d(f(x), f_i(x)) \le \varepsilon/2$$

for all $x \in X$, so

$$\{f(x) : s \in \mathcal{F}\} \subseteq \bigcup_{i=1}^n B(f_i(x), \varepsilon),$$

so the claim follows. ■

Lemma 4.63 motivates the following definition.

Definition 4.64 (Pointwise totally bounded)**.** Fix topological spaces $(X, \mathcal{T}_X)$ and a metric space (M, d), and let $\mathcal{F}$ be a family of continuous functions $f \colon X \to M$. Then $\mathcal{F}$ is *pointwise totally bounded* if and only if any $x \in \mathcal{F}$ makes the set

$$\{f(x) : f \in \mathcal{F}\}$$

totally bounded.

Example 4.65. By Lemma 4.66, any totally bounded subset of $B_c(X, M)$ is pointwise totally bounded.

Lemma 4.66. Fix a topological space $(X, \mathcal{T})$ and a metric space (M, d) so that we can give the space of bounded continuous functions $B_c(X, M)$ the uniform metric d_u. Fixing a totally bounded subset $\mathcal{F} \subseteq B_c(X, M)$ and a point $x \in X$, any $\varepsilon > 0$ has some open subset $U \subseteq X$ containing x such that

$$d(f(x), f(y)) < \varepsilon$$

for any $y \in X$ and $f \in \mathcal{F}$.

Proof. Fix any $\varepsilon > 0$ and use our totally boundedness to extract $\{f_1, \ldots, f_n\} \subseteq \mathcal{F}$ such that the $B(f_i, \varepsilon/3)$ cover $\mathcal{F}$. Now, for any $f \in \mathcal{F}$, find some f_i with $d(f_i, f) < \varepsilon/3$, we see that any $y \in \mathcal{F}$ can write

$$d(f(x), f(y)) \le d(f(x), f_i(x)) + d(f_i(x), f_i(y)) + d(f_i(y), f(y)) \le 2\varepsilon/3 + d(f_i(x), f_i(y)).$$

Now, by the continuity of f_i, we see that there is an open subset U_i containing x such that $y \in U_j$ implies $d(f_i(x), f_i(y)) < \varepsilon/3$, so $d(f(x), f(y)) < \varepsilon$ follows.

We now let f vary, which allows the U_i to vary. Defining

$$U := \bigcap_{i=1}^{n} U_i,$$

we see U is an open subset of X containing x, and each $y \in U$ has $d(f(x), f(y)) < \varepsilon$ for any (!) $f \in \mathcal{F}$. ■

Lemma 4.66 motivates the following definition.

Definition 4.67 (Equicontinuous)**.** Fix topological spaces $(X, \mathcal{T}_X)$ and a metric space (M, d), and let $\mathcal{F}$ be a family of continuous functions $f \colon X \to M$. We say that the family $\mathcal{F}$ is *equicontinuous* at some $x \in X$ if and only if any $\varepsilon > 0$ has some open subset $U \subseteq X$ such that $y \in U$ has

$$d(f(y), f(x)) < \varepsilon$$

for all $f \in \mathcal{F}$. The entire family $\mathcal{F}$ is *equicontinuous* if any only if it is equicontinuous at all $x \in X$.

Example 4.68. By Lemma 4.66, any totally bounded subset of $B_c(X, M)$ is equicontinuous.

4.7.2 Arzelá–Ascoli's Theorem

We might hope for a converse of our given lemmas. Here is the result.

Theorem 4.69 (Arzelá–Ascoli)**.** Fix a compact topological space $(X, \mathcal{T})$ and a metric space (M, d) so that we can give the space of bounded continuous functions $B_c(X, M)$ the uniform metric d_u. Then any equicontinuous and pointwise totally bounded family $\mathcal{F} \subseteq B_c(X, M)$ is totally bounded.

Proof. Fix some $\varepsilon > 0$ so that we want to cover $\mathcal{F}$ with finitely balls of radius $\varepsilon > 0$.

The point is to use compactness on the equicontinuous statement. Indeed, for any $x \in X$, we are promised an open subset $U_x \subseteq X$ such that any $y \in U_x$ and $f \in \mathcal{F}$ has $d(f(x), f(y)) < \varepsilon/4$. However, this means

$$X \subseteq \bigcup_{x \in X} U_x$$

gives us an open cover of X, so compactness tells us that there is some finite sequence of points $\{x_i\}_{i=1}^n$ such that the $U_i := U_{x_i}$ cover X.

Now, fixing any particular i, we use the pointwise totally bounded condition to note

$$\{f(x_i) : f \in \mathcal{F}\}$$

is totally bounded, so we get a finite subset $S_i \subseteq \mathcal{F}$ such that

$$\{f(x_i) : f \in \mathcal{F}\} \subseteq \bigcup_{f \in S_i} B(f(x_i), \varepsilon/4).$$

We now define S as the union of all the S_i, which is finite as the finite union of finite sets.

To finish the proof, we will need to do a little bookkeeping. Let Ψ denote the set of functions from $\{1, \ldots, n\}$ to S so that we can set

$$\mathcal{F}_\psi := \{f \in \mathcal{F} : f(x_i) \in B(\psi(i), \varepsilon/4) \text{ for each } 1 \leq i \leq n\}$$

for any $\psi \in \Psi$. By construction, the $\mathcal{F}_\psi$ cover $\mathcal{F}$: fix some $f \in \mathcal{F}$. Note that any $y \in U_i$ implies that $d(f(y), f(x_i)) \leq \varepsilon/4$ by construction of U_i, and for any given $f(x_i)$, there is an element $s_i \in S_i$ such that $d(f(x_i), s_i) < \varepsilon/4$ by construction of the $S_i \subseteq S$. Defining ψ by $\psi(i) := s_i$ for this particular s_i, we see that any $y \in U_i$ for any i has

$$d(f(y), f(x_i)) \leq d(f(y), s_i) + d(s_i, f(x_i)) \leq \varepsilon/2.$$

Letting i vary, we recall that the U_i cover X, so we have found $\psi \in \Psi$ with $f \in \mathcal{F}_\psi$, which is what we wanted.

We will finish upon showing that $\mathcal{F}_\psi$ has diameter less than ε. Well, for any $f, g \in \mathcal{F}_\psi$, we need to show that $d_u(f, g) < \varepsilon$. Well, fix any $x \in X$ and find some j with $x \in U_j$. Then we see

$$d(f(x), g(x)) \leq d(f(x), f(x_j)) + d(f(x_j), g(x_j)) + d(g(x_j), g(x)) \leq \varepsilon/2 + d(f(x_j), g(x_j)).$$

Now, by construction of ψ, we see

$$d(f(x_j), g(x_j)) \leq d(f(x_j), \psi(j)) + d(\psi(j), g(x_j)) < \varepsilon/2,$$

so we see that $d(f(x), g(x)) < \varepsilon$ in total. It follows $\|f - g\|_\infty \leq \varepsilon$, so, say, dividing all εs by two will give $\mathcal{F}_\psi$ all with radius less than ε. ■

4.8 October 3

It's spooky season. We begin class by finishing the proof of Theorem 4.69. I have edited the proof from yesterday for continuity reasons.

4.8.1 Locally Compact Spaces

Here is our definition.

> **Definition 4.70** (Locally compact). A topological space $(X, \mathcal{T})$ is *locally compact* if and only if each point $x \in X$ has some open subset $U \in \mathcal{T}$ containing x such that $\overline{U}$ is compact.

Example 4.71. The set of real numbers $\mathbb{R}$ with the usual topology is locally compact. Indeed, any $x \in \mathbb{R}$ has the open neighborhood $(x-1, x+1)$ with closure $[x-1, x+1]$, and $[x-1, x+1]$ is compact.

Example 4.72. For the same reason, the space $[a, b)$ is also locally compact.

Remark 4.73. Even though compact Hausdorff spaces are normal (by Proposition 4.18), locally compact Hausdorff spaces do not have to be.

For today, we are going to look at only locally compact Hausdorff spaces.

Lemma 4.74. Fix a locally compact Hausdorff space $(X, \mathcal{T})$. Then any $x \in X$ and open subset $U \in \mathcal{T}$ containing x has some open subset $U_x \subseteq X$ containing x such that $\overline{U_x}$ is compact and $\overline{U_x} \subseteq U$.

Proof. We begin by finding our promised U' containing x with $\overline{U'}$ compact. Thus, it suffices to find some open subset V containing x such that $\overline{V}$ is compact and $\overline{V} \subseteq U \cap U'$, but now we see that

$$\overline{U \cap U'} \subseteq \overline{U'}$$

is a closed subset of the compact space $\overline{U'}$ and therefore compact by Lemma 4.10. In particular, we can replace U with $U \cap U'$ and assume that $\overline{U}$ is compact.

Now, let $\partial U := \overline{U} \setminus U$ be the boundary of U. Notably, ∂U is a closed subset of the compact space $\overline{U}$, so ∂U is compact by Lemma 4.10. Because $\{x\}$ is a closed subset in U (note $X \setminus \{x\}$ is open, so $U \setminus \{x\}$ is open in the relative topology), the fact that compact Hausdorff spaces are normal (Proposition 4.18) grants open subsets U_x and U_∂ of $\overline{U}$ with $x \in U_x$ and $\partial U \subseteq U_\partial$.

Now, $U_x \subseteq \overline{U} \setminus U_\partial \subseteq \overline{U} \setminus \partial U$, so we see $\overline{U_x} \subseteq \overline{U} \setminus U_\partial$ because $\overline{U} \setminus U_\partial$ is a closed subset of $\overline{U}$. Further, $\overline{U_x}$ is a closed subset of a compact space $\overline{U}$, so $\overline{U_x}$ is compact by Lemma 4.10, so we are done. ■

Remark 4.75. Lemma 4.74 basically says that open subspaces of locally compact Hausdorff spaces are locally compact.

We can extend the previous result past points to full compact sets.

Proposition 4.76. Fix a locally compact Hausdorff space $(X, \mathcal{T})$ and some compact subset $C \subseteq X$. Then any open subset U containing C has some open subset U_C containing C such that $\overline{U_C}$ is compact and $\overline{U_C} \subseteq U$.

Proof. We use Lemma 4.74. For each $x \in C$, find some U_x by Lemma 4.74 with U_x containing x with $\overline{U_x}$ compact and $\overline{U_x} \subseteq U$. Then we see that

$$C \subseteq \bigcup_{x \in C} U_x,$$

so we have provided an open cover for C, so we can choose finitely many $\{x_i\}_{i=1}^n \subseteq C$ with $U_i := U_{x_i}$ so that

$$C \subseteq \bigcup_{i=1}^n U_i \subseteq U_C.$$

Now, we see that

$$\overline{\bigcup_{i=1}^n U_i} = \bigcup_{i=1}^n \overline{U_i}$$

is a compact subset of U because being compact is closed under finite unions (by inductively applying Example 4.9), so $\bigcup_{i=1}^n U_i$ is the required open subset. ■

4.8.2 Supports

A nice thing about locally compact Hausdorff spaces is that they let us talk about supports.

Definition 4.77 (Support)**.** Fix a locally compact Hausdorff space $(X,\mathcal{T})$ and a normed vector space $(V,\|\cdot\|)$. Then the *support* of a continuous function $f\colon X\to V$ is

$$\operatorname{supp} f := \overline{\{x\in X : f(x)\neq 0\}}.$$

Notably, $\{x\in X : f(x)\neq 0\} = f^{-1}(V\setminus\{0\})$ is the pre-image of an open subset and is therefore open by the continuity of f. In particular, normed vector spaces are metric spaces and therefore Hausdorff, so $\{0\}\subseteq V$ is in fact a closed subset.

Here are some quick checks about the support.

Lemma 4.78. Fix a locally compact Hausdorff space $(X,\mathcal{T})$ and a normed k-vector space $(V,\|\cdot\|)$. Then, given two continuous functions $f,g\in C(X,V)$ and $a,b\in k$, we have that

$$\operatorname{supp}(af+bg)\subseteq(\operatorname{supp} f\cup\operatorname{supp} g)$$

Proof. Because $\operatorname{supp} f\cup\operatorname{supp} g$ is the union of two closed sets, it's closed, so it suffices by definition of the closure to show that

$$\{x\in X : (af+bg)(x)\neq 0\}\overset{?}{\subseteq}(\operatorname{supp} f\cup\operatorname{supp} g).$$

Well, if $f(x)=0$ and $g(x)=0$, then we see $(af+bg)(x)=af(x)+bg(x)=0$, so $x\notin\operatorname{supp}(af+bg)$. Applying contraposition, we see $x\in\operatorname{supp}(af+bg)$ implies $f(x)\neq 0$ or $g(x)\neq 0$, so $x\in\operatorname{supp} f$ or $x\in\operatorname{supp} g$. ■

Lemma 4.79. Fix a locally compact Hausdorff space $(X,\mathcal{T})$ and a normed k-algebra $(R,\|\cdot\|)$. Then, given two continuous functions $f,g\in C(X,R)$ and $a,b\in k$, we have that

$$\operatorname{supp} fg\subseteq(\operatorname{supp} f\cap\operatorname{supp} g)$$

Proof. Again, because $\operatorname{supp} f\cap\operatorname{supp} g$ is the intersection of closed sets, it's closed, so it suffices to show that

$$\{x\in X : (fg)(x)\neq 0\}\overset{?}{\subseteq}(\operatorname{supp} f\cap\operatorname{supp} g).$$

Well, if $f(x)=0$ or $g(x)=0$, then $(fg)(x)=f(x)g(x)=0$. Thus, by contraposition, if $(fg)(x)=0$, then $f(x)\neq 0$ and $g(x)\neq 0$, so $x\in(\operatorname{supp} f\cap\operatorname{supp} g)$. ■

We tend to like small things, so here are our small functions.

Definition 4.80 (Compact support)**.** Fix a locally compact Hausdorff space $(X,\mathcal{T})$ and a normed vector space $(V,\|\cdot\|)$. A continuous function $f\colon X\to V$ has *compact support* if and only if its support is compact. We let $C_c(X,V)$ denote the continuous functions of compact support.

Here is a quick sanity check.

Lemma 4.81. Fix a locally compact Hausdorff space $(X,\mathcal{T})$ and a normed $k$vector space $(V,\|\cdot\|)$. Then $C_c(X,V)$ is a k-subspace of $C(X,V)$. If V is a normed k-algebra, then $C_c(X,V)$ is a k-subalgebra.

Proof. We have the following checks.

- Zero: note that the zero function $z\colon X \to V$ by $z(x) = 0$ for all $x \in X$ has
$$\{x \in X : z(x) \neq 0\} = \varnothing.$$
The closure of the empty set is still empty (certainly $\overline{\varnothing} \subseteq \varnothing$ by definition of the closure), so we conclude that $\operatorname{supp} z = \varnothing$. Now, $\varnothing$ is compact because any open cover can take the empty subcover, which is certainly finite. Thus, $z \in C_c(X, C)$.
- Linear combination: given $f, g \in C_c(X, V)$ and $a, b \in k$, we see from Lemma 4.78 that $\operatorname{supp}(af + bg)$ is a closed subset of $\operatorname{supp} f \cup \operatorname{supp} g$. However, $\operatorname{supp} f \cup \operatorname{supp} g$ is the union of two compact sets and therefore compact by Example 4.9, so $\operatorname{supp}(af + bg)$ is a closed subset of a compact space and hence compact by Lemma 4.10.
- Multiplication: given $f, g \in C_c(X, V)$, we see from Lemma 4.79 that $\operatorname{supp}(fg)$ is a closed subset of
$$\operatorname{supp} f \cap \operatorname{supp} g \subseteq \operatorname{supp} f.$$
However, $\operatorname{supp} f$ is compact, so $\operatorname{supp} fg$ is a closed subset of a compact space and hence compact by Lemma 4.10.

The first two checks tell us that we have a subspace, and the last check uses the algebra structure to get a subalgebra. ■

Of course, we would like to know that there are a nontrivial number of functions of compact support, so here we go.

Proposition 4.82. Fix a locally compact Hausdorff space $(X, \mathcal{T})$ and a normed vector space $(V, \|\cdot\|)$. For any compact subset $C \subseteq X$ and open subset $U \subseteq X$ containing C, there is a continuous function $f\colon X \to \mathbb{R}$ of compact support such that $f|_C = 1$ and $f|_{X \setminus U} = 0$.

Proof. The point is to apply Theorem 3.8. By Proposition 4.76, we may find an open subset V containing C such that $\overline{V}$ is compact and $\overline{V} \subseteq U$. Then we see C and $\overline{V} \setminus V$ are disjoint closed subsets of $\overline{V}$—note C is closed because X is Hausdorff, using Corollary 4.13.

Thus, because $\overline{V}$ is a normal space (it's compact and Hausdorff, so Proposition 4.18 applies), we are promised a continuous function $f_{\overline{V}}\colon \overline{V} \to \mathbb{R}$ such that $f_{\overline{V}}|_C = 1$ and $f_{\overline{V}}|_{\overline{V} \setminus V} = 0$. We now extend $\overline{V}$ to all of X by
$$f(x) := \begin{cases} f_{\overline{V}}(x) & x \in \overline{V}, \\ 0 & x \notin \overline{V}. \end{cases}$$
Indeed, if $x \in C$, we see $x \in \overline{V}$, so $f(x) = 1$; similarly, if $x \notin U$, then $x \notin \overline{V}$ and so $f(x) = 0$. Lastly, to see that f is continuous, we pick up some open closed $W \subseteq V$; we have the following cases.

- If $0 \notin W$, then we see that $f(x) \in W$ forces $x \in \overline{V}$, so
$$f^{-1}(W) = f_{\overline{V}}^{-1}(W)$$
is a closed subset of $\overline{V}$ by the continuity of $f_{\overline{V}}$. Closed subsets of closed subspaces are still closed, though, so we see that $f^{-1}(W)$ is closed in X.
- If $0 \in W$, then we do casework. If $x \in \overline{V}$, then actually $x \in f_{\overline{V}}^{-1}(W)$, which is closed in $\overline{V}$ and hence closed in X by continuity of $f_{\overline{V}}$. Otherwise, $x \notin \overline{V}$, but then we see that $x \in X \setminus V$ as well; conversely, if $x \in \overline{V} \setminus V$, then either $x \in \overline{V}$ and so $f_{\overline{V}}(x) = 0 \in W$, or $x \notin \overline{V}$ and so $f(x) = 0 \in W$.
In total, we see
$$f^{-1}(\{0\}) = f_{\overline{V}}(W) \cup (X \setminus V)$$
is the union of closed sets and thus closed.

Lastly, we see that $\operatorname{supp} f \subseteq \overline{V}$, so $\operatorname{supp} f$ is a closed subset of a compact set, so we conclude $\operatorname{supp} f$ is compact by Lemma 4.10, so f has compact support. ■

Remark 4.83. By Proposition 3.25, the space of bounded continuous functions $X \to \mathbb{R}$ is complete under $\|\cdot\|_\infty$. We note that $C_c(X, \mathbb{R})$ is a subalgebra, but it is not a closed subset. It turns out that its closure is $C_\infty(X, \mathbb{R})$, which is the space of functions which vanish at infinity: namely, for any $\varepsilon > 0$, there is a compact set $C \subseteq X$ such that $|f(x)| \leq \varepsilon$ for each $x \notin C$.

Part II

Measure Theory

THEME 5

DEFINING MEASURES

One fish, two fish, red fish, blue fish.

—Dr. Suess, [Gei60]

5.1 October 5

We begin today by making some motivating remarks on C^*-algebras and the like. I hope it's not important because I didn't understand it very well.

5.1.1 Evaluation Maps

For this subsection, we will want to work with the fields $\mathbb{R}$ and $\mathbb{C}$ at the same time, so we pick up the following definition.

Definition 5.1. An *archimedean field* is either $\mathbb{R}$ or $\mathbb{C}$.

We now recall the following piece of notation, which we will state in the case we now care about.

Notation 5.2. Fix an archimedean field k and a compact Hausdorff space X. Then we let $C(X)$ denote the continuous functions $X \to k$.

Remark 5.3. Note that $C(X)$ is a k-subalgebra of k^X because the constantly one function is continuous and the sum and product of two continuous functions is still continuous.

It will turn out that $C(X)$ can tell us a lot about X. For example, homomorphisms we can use X to build homomorphisms $C(X) \to k$.

Example 5.4. Given any $x \in X$, the function $\operatorname{ev}_x \colon C(X) \to k$ by $f \mapsto f(x)$ is a homomorphism. To see that this is a homomorphism, note that $\operatorname{ev}_x(1) = 1$, and $\operatorname{ev}_x(f+g) = (f+g)(x) = f(x) + g(x)$, and $\operatorname{ev}_x(fg)(x) = (fg)(x) = f(x)g(x)$.

In fact, these are all the homomorphisms!

Theorem 5.5. Fix an archimedean field k and a compact Hausdorff space X. Then all homomorphisms $C(X) \to k$ take the form ev_x for some $x \in X$.

Proof. Fix some homomorphism $\varphi\colon C(X) \to k$, and suppose for the sake of contradiction that $\varphi := \mathrm{ev}_x$ for each $x \in X$. To relate our geometry and our algebra, we will use the fact that the "algebraic" set $k \setminus \{0\}$ is open.

Now, we can find $f_x \in C(X)$ with $\varphi(f_x) \neq f_x(x)$ for each $x \in X$. However, $(f_x - \varphi(f_x)1_X)\colon X \to k$ is a continuous function, so

$$U_x := \{y \in X : (f_x - \varphi(f_x)1_X)(y) \neq 0\}$$

is the preimage of the open subset $k \setminus \{0\}$ through the continuous function $(f_x - \varphi(f_x)1_X)$.

Further, $x \in U_x$ because $(\varphi(f_x)1_X)(x) = \varphi(f_x) \neq f_x(x)$, so the open sets $\{U_x\}_{x\in X}$ produce an open cover of X, so we can finitely many of these points in $\{x_1, \ldots, x_n\}$ so that the open sets $U_i := U_{x_i}$ cover X. Thus, the function

$$f := \sum_{i=1}^{n} (f_x - \varphi(f_x)1_X)^2$$

is nonzero everywhere and thus a unit in $C(X)$. On the other hand, $\varphi(f_x - \varphi(f_x)1_X) = \varphi(f_x) - \varphi(f_x)\varphi(1_X) = 0$, so summing gives $\varphi(f) = 0$, which is a contradiction because ring homomorphisms send units to units! ■

This motivates us to work in a little more generality: fix a local field k. Given a k-algebra A, set $A^* := \mathrm{Hom}_k(A, k)$, and given $a \in A$, define the homomorphism $\mathrm{ev}_a\colon A^* \to k$ by $f \mapsto f(a)$. Then, by convention, we will give A^* the weakest topology making the $a \mapsto \mathrm{ev}_a$ continuous.

Example 5.6. Using the k-algebra $A = C(X)$, the map $X \to A^*$ by $x \mapsto \mathrm{ev}_x$ is a homeomorphism. Namely, this is a bijection by Theorem 5.5, and it is a homeomorphism essentially by the definition of the topology on A^*.

The point of the above example is that the algebra $C(X)$ and its evaluation maps are able to fully recover the topological space X!

5.1.2 The Gelfand–Naimark Theorem

By adding in a little more data, we can read even more information off $C(X)$.

Remark 5.7. With $k = \mathbb{C}$, note that complex conjugation extends to a function $C(X) \to C(X)$ by $f \mapsto \overline{f}$. Then one can check that

$$\left\|\overline{f} \cdot f\right\|_\infty = \|f\|_\infty^2 .$$

In fact, the converse is true!

Theorem 5.8 (Gelfand–Naimark)**.** Suppose that A is a commutative Banach $\mathbb{R}$-algebra or $\mathbb{C}$-algebra equipped with an involution $a \mapsto a^*$ such that $\|aa^*\| = \|a\|^2$. Then there is an isomorphism

$$A \simeq C(A^*).$$

In particular, all of these Banach algebras come from a topological space!

Example 5.9. When X is locally compact, set $C_\infty(X)$ to be the set of continuous functions $X \to k$ which vanish at infinity. Even though $C_\infty(X)$ has no multiplicative unit, it is still the case that $C_\infty(A^*) \cong C_\infty(X)$, and in fact $A^* \cong X$. Not having a unit turns out to not be a problem because we can have a function be 1 over a large interval, which is topologically close enough to a unit.

Example 5.10. In contrast, the bounded continuous functions $A := C_b(X)$ have $A \cong C(A^*)$ still, even though A^* is compact. This is weird: the embedding $X \hookrightarrow A^*$ is going to have elements not live in the image, but the elements outside the image require the Axiom of Choice to see.

The above example is why we prefer to work with $C_\infty(X)$ when we talk about locally compact spaces X.

Before jumping into measure theory, we will want to pick up the following definition.

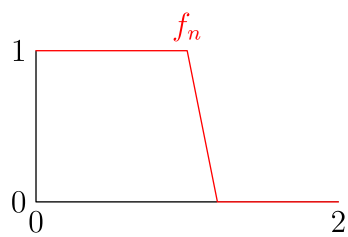
Definition 5.11. A *Hilbert space* is a complete inner product $\mathbb{R}$- or $\mathbb{C}$-vector space.

Example 5.12. Given a Hilbert space H, the set of linear operators $B(H)$ on H has a conjugation again, giving us an involution $T \mapsto \overline{T}$. One still has $\|T\overline{T}\| = \overline{T}^2$, so Theorem 5.8 applies, and we can think about these spaces as spaces of functions.

The above example will generalize to the study of C^* algebras, but we won't discuss this further.

5.1.3 Finitely Additive Measures

We begin with a motivating example. Consider the set of functions $f_n\colon [0,2] \to \mathbb{R}$, given by the following image.



More precisely, we can write

$$f_n(x) := \min\{1, \max\{0, 1 - n(x-1)\}\}.$$

These functions are all continuous by definition, but we can also give them a piecewise definition as

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

In particular, we can see that $f_n \to 1_{[0,1]}$ as $n \to \infty$ with respect to the $\|\cdot\|_p$ norm for $p \in [1, \infty)$: the error is

$$\left\|1_{[0,1]} - f_n\right\|_p^p = \int_0^2 |1_{[0,1/2]}(t) - f_n(t)|^p\,dt = \int_1^{1+1/n} |f_n(t)|^p\,dt \le \int_1^{1+1/n} dt = 1/n,$$

which goes to 0 as $n \to \infty$. Namely, to complete the set of our continuous functions $C([0,1])$ equipped with $\|\cdot\|_p$ for $p \in [1,\infty)$, we need to add in these indicator functions. Nonetheless, we just integrated over $1_{[0,1]}$ just fine above, so we will want to build a class of functions which includes $1_{[0,1]}$ both for completeness reasons but also for integration reasons.

It turns out that not all sets should be able to be integrated over; this leads to the notion of measurable sets. So we will have some collection of subsets $\mathcal{R} \subseteq \mathcal{P}(\mathbb{R})$ and then some measuring function $\mu\colon \mathcal{R} \to [0,\infty]$ (we must allow infinity!). Let's discuss what we want to be true of μ.

- Additivity: if $E, F \in \mathcal{R}$ are disjoint, then $E \sqcup F$ should be in $\mathcal{R}$, and we had better have $\mu(E \sqcup F) = \mu(E) + \mu(F)$. Namely, the sum of the sizes of two disjoint sets had better just be size of the disjoint union.

- Splitting: if $E, F \in \mathcal{R}$ with $F \subseteq E$, then we want (from the above)

$$\mu(E) = \mu(E \cap F) + \mu(E \setminus F),$$

where the idea is that we can look at just the size of $E \cap F$ and $E \setminus F$ individually to more locally compute our sizes.

We can view the above rules as first dictating what sets should be measured at all. As such, we have the following definition.

Definition 5.13 (Ring)**.** Fix a set X. A ring is a nonempty collection $\mathcal{R} \subseteq \mathcal{P}(X)$ with the following properties.

- Union: if $E, F \in \mathcal{R}$, then $E \cup F \in \mathcal{R}$.
- Subtraction: if $E, F \in R$, then $E \setminus F \in \mathcal{R}$.

Example 5.14. Of course, the full collection $\mathcal{P}(X)$ is a ring. More generally, given a subset $S \subseteq X$, the collection of subsets of S is a ring: if E, F are subsets of S, we see $E \cup F$ and $E \setminus F$ are both subsets of F.

Example 5.15. Of course, $\{\varnothing\}$ is a ring.

Example 5.16. The set of all finite subsets of X is a ring. Indeed, if $E, F \subseteq X$ are finite, then both $E \cup F$ and $E \setminus F$ are finite as well.

Remark 5.17. Fix a ring $\mathcal{R}$ and some $E, F \in \mathcal{R}$. Note that $E \cap F = E \setminus (E \setminus F)$, so $E \cap F \in \mathcal{R}$ as well.

Remark 5.18. Given a ring $\mathcal{R}$, we note that $\varnothing \in \mathcal{R}$: we know there is some $E \in \mathcal{R}$, so it follows $\varnothing = E \setminus E \in \mathcal{R}$.

Adding in the desired properties for our μ, we can now define "small" measures.

Definition 5.19 (Finitely additive measure)**.** Fix a set X and ring $\mathcal{R} \subseteq \mathcal{P}(X)$. Then a *finitely additive measure* is a function $\mu\colon \mathcal{R} \to [0, \infty]$ such that any disjoint $E, F \in \mathcal{R}$ have

$$\mu(E \sqcup F) = \mu(E) + \mu(F)$$

Remark 5.20. Note that $\mu(\varnothing) = \mu(\varnothing \sqcup \varnothing) = 2\mu(\varnothing)$, so it follows $\mu(\varnothing) = 0$.

Remark 5.21. Note that being finitely additive tells us that $E \subseteq F$ implies $E = F \sqcup (E \setminus F)$ because an element of E is either in F or not in F. Thus, we see $\mu(E) = \mu(F) + \mu(E \setminus F)$, so if $\mu(F) < \infty$, we may write $\mu(E \setminus F) = \mu(E) - \mu(F)$.

It turns out that being finitely additive is not good enough.

Example 5.22. We use the usual measure μ on $\mathbb{R}$. Fix a sequence of disjoint intervals $\{E_i\}_{i=1}^{\infty}$ in $[0,1]$, and we see that we should have

$$\sum_{i=1}^{\infty} \mu(E_i) < \infty.$$

Defining $F_n := \bigsqcup_{i \le n} E_i$ and $F := \bigsqcup_{i=1}^{\infty} E_i$, we see that the characteristic functions 1_{F_n} is a Cauchy sequence converging to 1_F, but we don't immediately have access to 1_F because it's an infinite union!

So next class we will discuss how adding a countably additive condition will help us.

5.2 October 7

We continue our discussion into measure theory.

5.2.1 σ-Things

Motivated by Example 5.22, we see that we want to be able to measure countable unions. As such, we have the following definitions.

Definition 5.23 (σ-ring)**.** Fix a set X. Then a ring $\mathcal{R} \subseteq \mathcal{P}(X)$ is a σ*-ring* if and only if R is closed under countable unions.

Remark 5.24. As in Remark 5.17, we note σ-rings $\mathcal{S}$ have countable intersections. Fix some $\{E_i\}_{i=1}^{\infty} \subseteq \mathcal{S}$. Then we note

$$E_1 \Bigg\backslash \bigcup_{i=1}^{\infty}(E_1 \setminus E_i) = \bigcap_{i=1}^{\infty}(E_1 \setminus (E_1 \setminus E_i)) = \bigcap_{i=1}^{\infty}(E_1 \cap E_i) = E_1 \cap \bigcap_{i=1}^{\infty} E_i = \bigcap_{i=1}^{\infty} E_i$$

lives in $\mathcal{S}$, finishing.

Definition 5.25 (σ-algebra)**.** Fix a set X. Then a ring $\mathcal{R} \subseteq \mathcal{P}(X)$ is a σ*-algebra* if and only if R is a σ-ring and contains X.

Example 5.26. Given a set X, we see $\mathcal{P}(X)$ is a σ-ring because a countable union of subsets of X is still a subset of X. Further, $\mathcal{P}(X)$ is a σ-algebra because $X \in \mathcal{P}(X)$.

Example 5.27. Fix a set X. Then the collection $\mathcal{S} \subseteq \mathcal{P}(X)$ of countable subsets of X is a σ-ring; here are our checks.

- Countable union: suppose $\{E_i\}_{i=1}^{\infty} \subseteq \mathcal{S}$. Then

  $$\bigcup_{i=1}^{\infty} E_i$$

  is the countable union of countable subsets of X and therefore countable. It follows that $\bigcup_{i=1}^{\infty} E_i \in \mathcal{S}$.

- Subtraction: if $E, F \in \mathcal{S}$, then E and F are both countable, so $E \setminus F \subseteq E$ is still countable, so $E \setminus F \in \mathcal{S}$.

Notably, if X itself is not an uncountable set, then $X \notin \mathcal{S}$, so $\mathcal{S}$ is not a σ-algebra.

As usual, we may give the collection of all σ-rings (and σ-algebras) the subposet structure coming from inclusion on $\mathcal{P}(X)$. For example, $\mathcal{P}(X)$ is the largest collection in $\mathcal{P}(X)$ and is thus the largest σ-ring and also the largest σ-algebra.

Analogous to our discussion of topologies in Proposition 2.20, we pick up the following lemma to make our σ-rings smaller.

Lemma 5.28. Fix a set X, and fix a collection Σ of rings, σ-rings, or σ-algebras. Then

$$\mathcal{S} := \bigcap_{\mathcal{R}\in\Sigma} \mathcal{R}$$

is another ring, σ-ring, or σ-algebra, respectively.

Proof. We show the axioms get inherited individually.

(a) Suppose that each $\mathcal{R} \in \Sigma$ is closed under finite unions. Then for any $E, F \in \mathcal{S}$, we see $E, F \in \mathcal{R}$ for each $\mathcal{R} \in \Sigma$, so $E \cup F \in \mathcal{R}$ for each $\mathcal{R} \in \Sigma$, so $E \cup F \in \mathcal{S}$.

(b) Suppose that each $\mathcal{R} \in \Sigma$ is closed under subtraction. Then for any $E, F \in \mathcal{S}$, we see $E, F \in \mathcal{R}$ for each $\mathcal{R} \in \Sigma$, so $E \setminus F \in \mathcal{R}$ for each $\mathcal{R} \in \Sigma$, so $E \setminus F \in \mathcal{S}$.

(c) Suppose that each $\mathcal{R} \in \Sigma$ is closed under countable union. Then for any countable collection $\{E_i\}_{i=1}^\infty \in \mathcal{S}$, we see $\{E_i\}_{i=1}^\infty \in \mathcal{R}$ for each $\mathcal{R} \in \Sigma$, so $\bigcup_{i=1}^\infty E_i \setminus F \in \mathcal{R}$ for each $\mathcal{R} \in \Sigma$, so $\bigcup_{i=1}^\infty E_i \in \mathcal{S}$.

(d) Suppose that each $\mathcal{R} \in \Sigma$ contains X. Then $X \in \mathcal{S}$.

The above checks complete the proof. For example, if Σ contains σ-rings, then checks (a)–(c) show $\mathcal{S}$ is still a σ-ring. ■

Corollary 5.29. Fix a set X and a collection $\mathcal{C} \subseteq \mathcal{P}(X)$. Then there is a unique smallest ring, σ-ring, or σ-algebra containing $\mathcal{C}$.

Proof. Let Σ denote the collection of all rings, σ-rings, or σ-algebras containing $\mathcal{C}$. We want to show that Σ contains a unique minimum element. Well, we set

$$\mathcal{S} := \bigcap_{\mathcal{R}\in\Sigma} \mathcal{R}.$$

Notably, $\mathcal{S} \in \Sigma$ by Lemma 5.28, and $\mathcal{S}$ is its minimum somewhat directly: for any $\mathcal{R} \in \Sigma$, we have $\mathcal{S} \subseteq \mathcal{R}$ by construction of $\mathcal{S}$. ■

This gives us the following definition.

Definition 5.30 (σ-ring generated by)**.** Fix a set X. Then give a collection $\mathcal{C}$, we let $\mathcal{S}(\mathcal{C})$ denote the *σ-ring generated by* $\mathcal{C}$, as conjured by Corollary 5.29.

There are analogous definitions for ring and σ-algebra, but we won't state them explicitly.

Remark 5.31. As usual, we note that $\mathcal{C} \subseteq \mathcal{C}'$ implies $\mathcal{S}(\mathcal{C}) \subseteq \mathcal{S}(\mathcal{C}')$ because $\mathcal{S}(\mathcal{C}')$ is a σ-ring containing $\mathcal{C}$.

Remark 5.32. Also as usual, if $\mathcal{S}$ is already a σ-ring, then $\mathcal{S}(\mathcal{S}) = \mathcal{S}$. Of course, $\mathcal{S} \subseteq \mathcal{S}(\mathcal{S})$, but also $\mathcal{S}$ is a σ-ring containing $\mathcal{S}$, so $\mathcal{S}(\mathcal{S}) \subseteq \mathcal{S}$ follows.

Example 5.33. Fix a set X. We claim σ-ring generated by the collection $\mathcal{F}$ finite subsets of X is the σ-ring $\mathcal{S}$ of countable subsets of X. Certainly $\mathcal{S}(\mathcal{F}) \subseteq \mathcal{S}$ because $\mathcal{S}$ is a σ-ring by Example 5.27. On the other hand, any countable subset $E \subseteq X$ has

$$E = \bigcup_{x \in E} \{x\}$$

while $\{x\} \in \mathcal{F} \subseteq \mathcal{S}(\mathcal{F})$ and therefore $E \in \mathcal{S}(\mathcal{F})$. Thus, $\mathcal{S} \subseteq \mathcal{S}(\mathcal{F})$.

5.2.2 Measures

We are now ready to define measures.

Definition 5.34 (Countably additive)**.** Fix a set X and a collection of subsets $\mathcal{C} \subseteq \mathcal{P}(X)$. A function $\mu\colon \mathcal{C} \to [0, \infty]$ is *countably additive* if and only if any pairwise disjoint subcollection $\{E_i\}_{i=1}^\infty \subseteq \mathcal{C}$ with $\bigsqcup_{i=1}^\infty E_i \in \mathcal{C}$ has

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

Notably, we are allowed to have the right-hand side diverge to ∞ if the left-hand side is ∞.

Remark 5.35. In general, it is pretty difficult to actually show that a function is countably additive, but one can take advantage of the fact that

$$\bigsqcup_{i=1}^\infty E_i$$

might not actually be in $\mathcal{C}$.

And here is our definition.

Definition 5.36 (Measure)**.** Fix a set X and σ-ring $\mathcal{S}$. Then a *measure* on $\mathcal{S}$ is a function $\mu\colon \mathcal{S} \to [0, \infty]$ which is countably additive.

Remark 5.37. Note that the countable unions of sets in $\mathcal{S}$ to check the countably additive condition are always in $\mathcal{S}$ because $\mathcal{S}$ is a σ-ring. Namely, the trick suggested in Remark 5.35 doesn't help us.

Remark 5.38. In general, it is not a good idea to ask for unions larger than countable. Approximately speaking, we really want to have countable unions, but we need to be careful adding any other infinities. The main problem is that those infinite sums don't have easy notions of convergence. Even if we don't want to work with something like nets to allow larger convergences, then allowing arbitrary unions for $E \subseteq X$ gives

$$\mu(E) = \sum_{x \in X} \mu(\{x\}),$$

which intuitively should vanish if we make our points have measure 0.

Remark 5.39. Fix a set X and measure $\mu\colon \mathcal{S} \to [0, \infty]$. If $\mu(\varnothing) < \infty$, then note $\varnothing = \bigsqcup_{i=1}^\infty \varnothing$ implies that the sum $\sum_{i=1}^\infty \mu(\varnothing) = \mu(\varnothing)$ converges, so $\mu(\varnothing) = 0$ is forced. Otherwise, if $\mu(\varnothing) = \infty$, then any $E \in \mathcal{S}$ has $E = E \sqcup \varnothing$, so $\mu(E) = \mu(E) + \mu(\varnothing) = \infty$.

Remark 5.40. If μ is a measure a σ-ring $\mathcal{S}$, then $\mu|_\mathcal{T}$ remains a measure on any σ-ring $\mathcal{T} \subseteq \mathcal{S}$. Indeed, any pairwise disjoint subcollection $\{T_i\}_{i=1}^\infty \subseteq \mathcal{T}$ also lives in $\mathcal{S}$, so we maintain having

$$\mu|_\mathcal{T}\left(\bigsqcup_{i=1}^\infty T_i\right) = \mu\left(\bigsqcup_{i=1}^\infty T_i\right) = \sum_{i=1}^\infty \mu(T_i) = \sum_{i=1}^\infty \mu|_\mathcal{T}(T_i).$$

Let's see some examples.

Exercise 5.41. More generally, fix a set X using the σ-ring $\mathcal{S} := \mathcal{P}(X)$ of countable subsets of X. For a function $f\colon X \to [0,\infty)$, we define

$$\mu_f(E) := \sum_{x\in E} f(x)$$

for each countable subset $E \subseteq X$. Then μ_f is a measure.

Proof. Note that the order of the sum over $x \in X$ doesn't matter because if the sum converges, then it absolutely converges because all the terms in the sum are positive. Now, to see that we have a measure, pick up some countably many pairwise disjoint countable subsets $\{E_i\}_{i=1}^\infty$ of X. Then

$$\mu_f\left(\bigsqcup_{i=1}^\infty E_i\right) = \sum_{x\in\bigsqcup_{i=1}^\infty E_i} f(x) \stackrel{*}{=} \sum_{i=1}^\infty \sum_{x\in E_i} f(x) = \sum_{i=1}^\infty \mu_f(E_i),$$

where $\stackrel{*}{=}$ holds because each $x \in \bigsqcup_{i=1}^\infty E_i$ lives in exactly one of the E_i. ■

Example 5.42. Fix a set X with σ-ring $\mathcal{S} := \mathcal{P}(X)$. Then we set $\mu(E) := \#E$ for each $E \subseteq X$; namely, if $\mu(E) = \infty$ if and only if E is infinite. We claim that μ is a measure: if $\{E_i\}_{i=1}^\infty$ is a countable collection of pairwise disjoint subsets of X, then it's a property of cardinality that the cardinality of the (disjoint) union is the sum of the cardinalities.

5.2.3 Premeasures

We are going to want to build measures, but this is somewhat difficult. So we begin with something a little weaker. We begin by weakening our rings.

Definition 5.43 (Prering)**.** Fix a set X. A *prering* of a set X is a nonempty collection $\mathcal{P} \subseteq \mathcal{P}(X)$ satisfying the following.

- Intersection: if $E, F \in \mathcal{P}$, then $E \cap F = \mathcal{P}$.
- Decomposition: if $E, F \in \mathcal{P}$, then we can write

  $$E \setminus F = \bigsqcup_{i=1}^n G_i$$

  for some finite disjoint union on the right-hand side with $G_i \in \mathcal{P}$ for each i.

Remark 5.44. Fix a prering $\mathcal{P}$. Note any $E \in \mathcal{P}$ has $E \setminus E = \varnothing$, so $\varnothing \in \mathcal{P}$ always because $\mathcal{P}$ is required to be nonempty.

And now here are our weaker measures.

Definition 5.45 (Premeasure)**.** Fix a set X and a prering $\mathcal{P} \subseteq \mathcal{P}(X)$. A *premeasure* on $\mathcal{P}$ is a countably additive function $\mu\colon \mathcal{P} \to [0, \infty]$.

It will turn out that premeasures on prering will give measures on the generated σ-ring. This is nicer because the countably additive condition might be easier to check on a prering, using ideas of Remark 5.35.

Here is our main example.

Exercise 5.46. Fix our set $X \coloneqq \mathbb{R}$, and let $\mathcal{P}$ be the collection of half-open intervals $[a, b)$ where $a, b \in \mathbb{R}$. Then $\mathcal{P}$ is a prering.

Proof. We begin by checking that $\mathcal{P}$ is a prering.

- Intersection: suppose that $[a, b), [a', b') \in \mathcal{P}$; without loss of generality, take $a \le a'$ so that $x \in [a, b) \cap [a', b')$ requires $a' \le x$. Now, note

$$\begin{aligned} [a, b) \cap [a', b') &= \{x \in \mathbb{R} : a \le x \text{ and } a' \le x \text{ and } x < b \text{ and } x < b'\} \\ &= [\max\{a, a'\}, \min\{b, b'\}). \end{aligned}$$

- Decomposition: suppose that $[a, b), [a', b') \in \mathcal{P}$. Now, note

$$\begin{aligned} [a, b) \setminus [a', b') &= \{x \in \mathbb{R} : a \le x \text{ and } x < b \text{ and } (a' > x \text{ or } x \ge b')\} \\ &= \{x \in \mathbb{R} : a \le x \text{ and } x < b \text{ and } x < a'\} \cup \{x \in \mathbb{R} : a \le x \text{ and } x < b \text{ and } b' \le x\} \\ &= [a, \min\{b, a'\}) \cup [\max\{a, b'\}, b). \end{aligned}$$

The above checks complete the proof. ■

Continuing from Exercise 5.46, it will turn out that the function $\mu\colon \mathcal{P} \to \mathbb{R}$ given by

$$\mu([a, b)) \coloneqq b - a$$

will give a premeasure, but we will not show this today. (We will say that one should use ideas of Exercise 5.46.) This is surprisingly annoying to prove.

Example 5.47. Give $\mathbb{Q} \cap [0, 1)$ an enumeration $\{q_k\}_{k \in \mathbb{N}}$. Then define the interval $F_k \coloneqq [q_k, q_{k+1}) \cup [q_{k+1}, q_k)$ and "disjoint-ize" these intervals by taking

$$E_k \coloneqq F_k \Big\backslash \bigcup_{\ell=1}^{k-1} F_\ell$$

and then decompose E_k into a finite disjoint union of G_ks so that the G_ks are now disjoint. Any proof that μ is a premeasure must account for pathologies like this.

5.3 October 10

The midterm exam is coming. It will cover topology things.

5.3.1 The Lebesgue Premeasure

We continue with our attempts to construct measures.

Proposition 5.48. Fix a left-continuous, increasing function $\alpha\colon \mathbb{R} \to \mathbb{R}$, and let $\mathcal{P} \subseteq \mathcal{P}(\mathbb{R})$ as the prering of half-open intervals $[a, b)$ for $a < b$. Then

$$\mu_\alpha([a, b)) := \alpha(b) - \alpha(a)$$

is a premeasure on $\mathcal{P}$.

Proof. Quickly, note that the fact that α is increasing implies that $\mu([a, b)) = \alpha(b) - \alpha(a) \geq 0$ for any $[a, b) \in \mathcal{P}$.

Fix some $[a, b) \in \mathcal{P}$ which has been decomposed into an infinite disjoint union

$$[a, b) = \bigsqcup_{i=1}^{\infty} [a_i, b_i).$$

We need to show that $\mu_\alpha([a, b))$ is the sum of all the $\mu_\alpha([a_i, b_i))$s. We will show our two inequalities separately.

- In the easy direction, we show $\sum_{i=1}^{\infty} \mu_\alpha([a_i, b_i)) \leq \mu_\alpha([a, b))$. It suffices to show that, for any $n \in \mathbb{N}$, we have
$$\sum_{i=1}^{n} \mu_\alpha([a_i, b_i)) \stackrel{?}{\leq} \mu_\alpha([a, b)),$$
which will finish by taking the limit as $n \to \infty$. Well, let $\sigma\colon \{1, \ldots, n\} \to \{1, \ldots, n\}$ be the permutation such that $a_{\sigma(1)} \leq a_{\sigma(2)} \leq \cdots \leq a_{\sigma(n)}$. Notably, $a_{\sigma(i)} \leq a_{\sigma(i+1)}$ implies that $b_{\sigma(i)} \leq a_{\sigma(i+1)}$ because $[a_{\sigma(i)}, b_{\sigma(i)}) \cap [a_{\sigma(i+1)}, b_{\sigma(i+1)}) = \varnothing$ requires $a_{\sigma(i+1)} \notin [a_{\sigma(i)}, b_{\sigma(i)})$.
Thus, $b_{\sigma(i)} \leq a_{\sigma(i+1)}$ implies $\alpha(b_{\sigma(i)}) \leq \alpha(a_{\sigma(i+1)})$, so
$$\sum_{i=1}^{n} \mu_\alpha([a_i, b_i)) = \sum_{i=1}^{n} \big(\alpha(b_{\sigma(i)}) - \alpha(a_{\sigma(i)})\big) = -\alpha(a_{\sigma(1)}) + \sum_{i=1}^{n-1} \big(- \alpha(a_{\sigma(i+1)}) + \alpha(b_{\sigma(i)})\big) + \alpha(b_{\sigma(n)})$$
has $-\alpha(a_{\sigma(i+1)}) + \alpha(b_{\sigma(i)}) \leq 0$ for each i. Finishing up,
$$\sum_{i=1}^{n} \mu_\alpha([a_i, b_i)) \leq -\alpha(a_{\sigma(1)}) + \alpha(b_{\sigma(n)}) \leq \alpha(b) - \alpha(a) = \mu_\alpha([a, b)),$$
where we have used $a \leq a_{\sigma(1)}$ and $b_{\sigma(n)} \leq b$ in our bounding.

- In the difficult direction, we show $\mu_\alpha([a, b)) \leq \sum_{i=1}^{\infty} \mu_\alpha([a_i, b_i))$. Fix any $\varepsilon > 0$, and we will actually show $\mu_\alpha([a, b)) \leq \sum_{i=1}^{\infty} \mu_\alpha([a_i, b_i)) + \varepsilon$, which will be enough upon sending $\varepsilon \to 0^+$.
To set up the proof, set $\varepsilon_i := \varepsilon/2^{i+1}$ so that
$$\sum_{i=1}^{\infty} \varepsilon_i = \sum_{i=1}^{\infty} \frac{\varepsilon}{2^{i+1}} = \frac{\varepsilon}{2} \sum_{i=1}^{\infty} \frac{1}{2^i} = \frac{\varepsilon}{2}.$$
(This is a surprise tool which will help us later.)
We now proceed in steps. The idea is to approximate all of our $[a_i, b_i)$ by open intervals to use compactness of closed intervals.
  1. Find some $b' \leq b$ such that $\alpha(b) - \varepsilon/2 \leq \alpha(b') \leq \alpha(b)$, using the left-continuity of α. Similarly, for each $i \in \mathbb{N}$, we may select $a_i' < a_i$ such that $\alpha(a_i) - \varepsilon_i \leq \alpha(a_i') \leq \alpha(a_i)$. Thus,
  $$\bigcup_{i=1}^{\infty} (a_i', b_i) \supseteq \bigcup_{i=1}^{\infty} [a_i, b_i) \supseteq [a, b) \supseteq [a, b'].$$

Thus, we have given $[a,b']$ a countable open cover! So compactness (!) provides us with a finite subcover given by indices $\{i_1,\ldots,i_n\}$. Letting N be the largest of the indices, then, we see that

$$\bigcup_{i=1}^{N}(a_i',b_i)\supseteq\bigcup_{k=1}^{n}(a_{i_k}',b_{i_k})\supseteq[a,b'].$$

2. We now inductively relabel our intervals. Some open interval must contain a, so we find $j_1\in\{1,\ldots,N\}$ so that $a\in(a_{j_1}',b_{j_1})$. If $b_{j_1}>b'$, then we are done because we have covered $[a,b']$. Otherwise, $b_{j_1}\in[a,b']$, so we find $j_2\in\{1,\ldots,N\}$ so that $b_{j_1}\in(a_{j_2}',b_{j_2})$. If $b_{j_2}>b'$, then we are done because we have covered $[a,b']$; otherwise we find j_3 and continue.

   The above inductive process must terminate because each of the j_i are distinct—at each point, b_{j_i} is strictly greater than all previous $b_{j_\bullet}$s—and we were already promised that the indices up to N will produce a finite subcover. So we have produced some open cover

$$\bigcup_{k=1}^{m}(a_{j_k}',b_{j_k})\supseteq[a,b'].$$

3. We are finally able to give the argument that everyone always wants to. Observe that

$$\sum_{k=1}^{m}\left(\alpha(b_{j_k})-\alpha(a_{j_k}')\right)=-\alpha(a_{j_1}')+\sum_{k=1}^{m-1}\left(\alpha(b_{j_k})-\alpha(a_{j_{k+1}}')\right)+\alpha(b_{j_m})$$

   by some re-indexing. However, $a_{j_{k+1}}'<b_{j_k}$, so $\alpha(b_{j_k})-\alpha(a_{j_{k+1}}')\geq0$ always, so

$$\sum_{k=1}^{m}\left(\alpha(b_{j_k})-\alpha(a_{j_k}')\right)\geq\alpha(b_{j_m})-\alpha(a_{j_1}')\geq\alpha(b')-\alpha(a),$$

   where at the end we have used the fact that $b_{j_m}\geq b'$ and $a'\geq a_{j_1}$. But now $\alpha(b')\geq\alpha(b)-\varepsilon/2$, so we get

$$\sum_{k=1}^{m}\left(\alpha(b_{j_k})-\alpha(a_{j_k}')\right)\geq\alpha(b)-\alpha(a)-\varepsilon/2. \tag{5.1}$$

4. Now, on the other side, we write

$$\sum_{k=1}^{m}\left(\alpha(b_{j_k})-\alpha(a_{j_k}')\right)\leq\sum_{k=1}^{m}\left(\alpha(b_{j_k})-\alpha(a_{j_k})+\varepsilon_{j_k}\right),$$

   using the fact that $\alpha(a_{j_k}')\geq\alpha(a_{j_k})-\varepsilon_{j_k}$. We can now just add in all the indices to get

$$\sum_{k=1}^{m}\left(\alpha(b_{j_k})-\alpha(a_{j_k}')\right)\leq\sum_{i=1}^{\infty}\left(\alpha(b_i)-\alpha(a_i)\right)+\sum_{i=1}^{\infty}\varepsilon_i\leq\sum_{i=1}^{\infty}\mu_\alpha([a_i,b_i))+\frac{\varepsilon}{2}. \tag{5.2}$$

5. In total, we combine (5.1) and (5.2) to get

$$\sum_{i=1}^{\infty}\mu_\alpha([a_i,b_i))+\varepsilon\geq\mu_\alpha([a,b)).$$

   Sending $\varepsilon\to0^+$ finishes the proof. ■

Remark 5.49. The "easy" part of the above proof works fine without using the completeness of $\mathbb{R}$, but it is very necessary for the harder part.

5.4 October 12

The midterm exam is still coming. It is closed-book. Only bring writing implements. He might ask for definitions, statements of theorems, proofs of theorems, and relatively quick applications of theorems.

5.4.1 Premeasure Subtraction

Last class, in Proposition 5.48, we showed that $\mu_\alpha\colon \mathcal{P} \to [0,\infty]$ gave a suitable premeasure. We are now going to embark on a somewhat long story to show that μ_α (and premeasures in general) can turn into a full measure.

To begin our journey, we pick up some annoying facts about prerings and premeasures.

Lemma 5.50. Fix a set X and a prering $\mathcal{P}$. For any set $E \in \mathcal{P}$ and any $\{E_i\}_{i=1}^m \subseteq \mathcal{P}$, there exist finitely many $\{F_j\}_{j=1}^n \subseteq P$ which are pairwise disjoint and satisfy

$$E \Big\backslash \bigcup_{i=1}^m E_i = \bigsqcup_{j=1}^n F_j.$$

Proof. We induct on m, using the prering condition. When $m = 0$, set $F_1 = E$, and there is nothing else to say.

Now suppose that we can write

$$E \Big\backslash \bigcup_{i=1}^m E_i = \bigsqcup_{j=1}^n F_j.$$

Picking up some other $E_{m+1} \in \mathcal{P}$, we note

$$E \Big\backslash \bigcup_{i=1}^{m+1} E_i = \left(E \Big\backslash \bigcup_{i=1}^m E_i\right) \setminus E_{m+1} = \left(\bigsqcup_{j=1}^m F_j\right) \cap (X \setminus E_{m+1}) \overset{*}{=} \bigcup_{j=1}^m (F_j \cap (X \setminus E_{m+1})) = \bigcup_{j=1}^m (F_j \setminus E_{m+1})$$

where we have used the distributivity of intersection over union in $\overset{*}{=}$. For each j, because $\mathcal{P}$ is a prering, we may find pairwise disjoint $\{G_{j,k}\}_{k=1}^{m_j} \subseteq \mathcal{P}$ such that

$$F_j \setminus E_{m+1} = \bigsqcup_{k=1}^{m_j} G_{j,k}$$

so that

$$E \Big\backslash \bigcup_{i=1}^{m+1} E_i = \bigcup_{j=1}^m \bigcup_{k=1}^{m_j} G_{j,k}.$$

We now claim that the $\{G_{j,k}\}$ are pairwise disjoint, which will finish the proof. Indeed, if we can find $x \in G_{j,k} \cap G_{j',k'}$, then $G_{j,k} \subseteq F_j$ and $G_{j',k'} \subseteq F_{j'}$ tells us $x \in F_j \cap F_{j'}$, so $j = j'$ because the $F_\bullet$ are pairwise disjoint. Thus, $x \in G_{j,k} \cap G_{j,k'}$ further implies $k = k'$ because the $G_{j,\bullet}$ are pairwise disjoint. So $(j,k) = (j',k')$, and we are done. ■

Lemma 5.51. Fix a prering $\mathcal{P}$ on X and a finitely additive function $\mu\colon \mathcal{P} \to [0,\infty]$. Given $E, F \in \mathcal{P}$, then $\mu(E) \ge \mu(E \cap F)$. In particular, if $E \supseteq F$, then $\mu(E) \ge \mu(F)$.

Proof. Note that an element of E is always exactly one of in F or not, so $E = (E \cap F) \sqcup (E \setminus F)$. Now, we use the prering condition on $\mathcal{P}$ to write

$$E \setminus F = \bigsqcup_{i=1}^n G_i$$

for some pairwise disjoint $G_1, \ldots, G_n \in \mathcal{P}$. We also note that $G_i \subseteq X \setminus F$ for each i, so $G_i \cap (E \cap F) = \varnothing$ for each i, so the sets $(E \cap F), G_1, \ldots, G_n$ are pairwise disjoint and grant

$$\mu(E) = \mu(E \cap F) + \sum_{i=1}^{n} \mu(G_i).$$

However, $\mu(G_i) \geq 0$ always, so the first assertion follows. The second assertion follows upon noticing $E \supseteq F$ implies $E \cap F = F$. ■

The above result motivates the following definition.

Definition 5.52 (Monotone)**.** Fix a collection $\mathcal{F}$ of subsets of a set X. A function $\mu \colon \mathcal{F} \to [0, \infty]$ is *monotone* if and only if any $E, F \in \mathcal{F}$ with $E \subseteq F$ have $\mu(E) \leq \mu(F)$.

Example 5.53. Finitely additive premeasures on prerings are monotone by Lemma 5.51.

5.4.2 Finite Subadditivty

We now pick up some subadditivity lemmas.

Lemma 5.54. Fix a prering $\mathcal{P}$ on X and a finitely additive function $\mu \colon \mathcal{P} \to [0, \infty]$. Given $E \in \mathcal{P}$ and some pairwise disjoint $\{E_i\}_{i=1}^{n} \subseteq \mathcal{P}$ such that $E_i \subseteq E$ for such i, we have

$$\sum_{i=1}^{n} \mu(E_i) \leq \mu(E).$$

Proof. By Lemma 5.50, we note that we may write

$$E \Big\backslash \bigcup_{i=1}^{n} E_i = \bigsqcup_{j=1}^{m} F_j$$

for pairwise disjoint $\{F_j\}_{j=1}^{m} \subseteq \mathcal{P}$. We now note that all the E_i and F_j are pairwise disjoint from each other: note that $E_i \cap E_j \neq \varnothing$ implies $i = j$ by hypothesis on the $E_\bullet$, and $F_i \cap F_j \neq \varnothing$ implies $i = j$ by hypothesis on the $F_\bullet$. Further, we note that $E_i \cap F_j \subseteq E_i \cap (E \setminus E_i) = \varnothing$ for each i and j, by construction of the F_j.

In total, we see that we have a disjoint union

$$E = \left(\bigsqcup_{i=1}^{n} E_i \right) \sqcup \left(\bigsqcup_{j=1}^{m} F_j \right),$$

so the finite additivity of μ tells us

$$\mu(E) = \sum_{i=1}^{n} \mu(E_i) + \sum_{j=1}^{m} \mu(F_j) \geq \sum_{i=1}^{n} \mu(E_i),$$

which is what we wanted. ■

Lemma 5.55. Fix a prering $\mathcal{P}$ on a set X and a finitely additive function $\mu \colon \mathcal{P} \to [0, \infty]$. Given $E \in \mathcal{P}$ and some $\{F_j\}_{j=1}^{m} \subseteq \mathcal{P}$ covering E, we have

$$\mu(E) \leq \sum_{j=1}^{m} \mu(F_j).$$

Proof. To begin, we note $E = \bigcup_{j=1}^m (E \cap F_j)$, so we note that it suffices for

$$\mu(E) \le \sum_{k=1}^m \mu(E \cap F_j),$$

which will finish because $\mu(E \cap F_j) \le \mu(F_j)$ for each j by Lemma 5.51. Thus, we just replace each F_j with $E \cap F_j$ so that $E = \bigcup_{j=1}^m F_j$.

Next, we force the F_j to be disjoint, using Lemma 5.50 to write

$$H_j := F_j \setminus \bigcup_{k=1}^{j-1} F_k = \bigsqcup_{k=1}^{n_j} G_{j,k}$$

where the $G_{j,k} \subseteq H_j$ live in $\mathcal{P}$ and are pairwise disjoint for each fixed j. Now, we note that each $x \in E$ will live in some F_j with least j, so $x \in H_j$ for this j, so the H_j cover E.

We now note that all the $G_{j,k}$ are disjoint. Indeed, if $x \in G_{j,k} \cap G_{j',k'}$, we see that $G_{j,k} \subseteq H_j$ and $G_{j',k'} \subseteq H_{j'}$, so $x \in H_j \subseteq H_{j'}$. If $j \ne j'$, say that $j < j'$ without loss of generality, so $x \in H_j \subseteq F_j$ while $x \in H_{j'}$ has $H_{j'}$ disjoint from F_j, so we have a contradiction. So instead we see $j = j'$, so $x \in G_{j,k} \cap G_{j,k'}$, and it follows that $k = k'$ because the $G_{j,\bullet}$ are disjoint.

In total, we see that

$$E = \bigsqcup_{j=1}^m \bigsqcup_{k=1}^{n_j} G_{j,k},$$

so the finitely additive condition tells us that

$$\mu(E) = \sum_{j=1}^m \sum_{k=1}^{n_k} \mu(G_{j,k}).$$

However, we note that the $G_{j,k}$ are disjoint for any fixed j and have $G_{j,k} \subseteq F_j$ for each k, so we see that

$$\sum_{k=1}^{n_k} \mu(G_{j,k}) \le \mu(F_j)$$

for each j by Lemma 5.54, so we conclude

$$\mu(E) = \sum_{j=1}^m \sum_{k=1}^{n_k} \mu(G_{j,k}) \le \sum_{j=1}^m \mu(F_j),$$

which is what we wanted. ■

THEME 6
BUILDING MEASURES

So the man gave him the bricks, and he built his house with them.

—Joseph Jacobs, "The Story of the Three Little Pigs" [Jac90]

6.1 October 14

We will probably still have homework next week, despite the midterm.

6.1.1 Countable Subadditivity

Continuing our story from last time, we pick up the following definition. The above result motivates the following definition.

Definition 6.1 (Countably subadditive)**.** Fix a set X and a collection $\mathcal{F} \subseteq \mathcal{P}(X)$. A function $\mu\colon \mathcal{F} \to [0,\infty]$ is *countably subadditive* if and only if

$$E \subseteq \bigcup_{i=1}^{\infty} E_i \implies \mu(E) \le \sum_{i=1}^{\infty} \mu(E_i)$$

for any $E \in \mathcal{F}$ and $\{E_i\}_{i=1}^{\infty} \subseteq \mathcal{F}$.

Lemma 6.2. Fix a prering $\mathcal{P}$ on a set X, and let μ be a premeasure on $\mathcal{P}$. Then μ is countably subadditive.

Proof. We repeat the proof of Lemma 5.55, essentially verbatim, replacing the bound m with ∞. Indeed, pick up any $E \in \mathcal{P}$ and some $\{F_j\}_{j=1}^{\infty} \subseteq \mathcal{P}$ with $E \subseteq \bigcup_{j=1}^{\infty} F_j$, and we want to show that

$$\mu(E) \le \sum_{j=1}^{\infty} \mu(F_j).$$

To begin, we note $E = \bigcup_{j=1}^{\infty} (E \cap F_j)$, so we note that it suffices for

$$\mu(E) \le \sum_{k=1}^{\infty} \mu(E \cap F_j),$$

which will finish because $\mu(E \cap F_j) \le \mu(F_j)$ for each j by Lemma 5.51. Thus, we just replace each F_j with $E \cap F_j$ so that $E = \bigcup_{j=1}^m F_j$.

Next, we force the F_j to be disjoint, using Lemma 5.50 to write

$$H_j := F_j \setminus \bigcup_{k=1}^{j-1} F_k = \bigsqcup_{k=1}^{n_j} G_{j,k}$$

where the $G_{j,k} \subseteq H_j$ live in $\mathcal{P}$ and are pairwise disjoint for each fixed j. Now, we note that each $x \in E$ will live in some F_j with least j, so $x \in H_j$ for this j, so the H_j cover E.

We now note that all the $G_{j,k}$ are disjoint. Indeed, if $x \in G_{j,k} \cap G_{j',k'}$, we see that $G_{j,k} \subseteq H_j$ and $G_{j',k'} \subseteq H_{j'}$, so $x \in H_j \subseteq H_{j'}$. If $j \ne j'$, say that $j < j'$ without loss of generality, so $x \in H_j \subseteq F_j$ while $x \in H_{j'}$ has $H_{j'}$ disjoint from F_j, so we have a contradiction. So instead we see $j = j'$, so $x \in G_{j,k} \cap G_{j,k'}$, and it follows that $k = k'$ because the $G_{j,\bullet}$ are disjoint.

In total, we see that

$$E = \bigsqcup_{j=1}^{\infty} \bigsqcup_{k=1}^{n_j} G_{j,k},$$

so the finitely additive condition tells us that

$$\mu(E) = \sum_{j=1}^{m} \sum_{k=1}^{n_k} \mu(G_{j,k}).$$

However, we note that the $G_{j,k}$ are disjoint for any fixed j and have $G_{j,k} \subseteq F_j$ for each k, so we see that

$$\sum_{k=1}^{n_k} \mu(G_{j,k}) \le \mu(F_j)$$

for each j by Lemma 5.54, so we conclude

$$\mu(E) = \sum_{j=1}^{\infty} \sum_{k=1}^{n_k} \mu(G_{j,k}) \le \sum_{j=1}^{m} \mu(F_j),$$

which is what we wanted. ■

6.1.2 Hereditary Rings

We continue trying to move from premeasures to measures. Our next step is to add in lots and lots of sets to our prering, which we will later filter out to get our actual measure.

Definition 6.3 (Hereditary σ-ring)**.** Fix a set X and nonempty family $\mathcal{F} \subseteq \mathcal{P}(X)$. Then the *hereditary σ-ring* $\mathcal{H}(F)$ generated by $\mathcal{F}$ consists of all subsets $E \subseteq X$ such that there exists a countable subcollection $\{F_i\}_{i=1}^{\infty} \subseteq \mathcal{F}$ such that

$$E \subseteq \bigcup_{i=1}^{\infty} E_i.$$

Remark 6.4. Because $\mathcal{F}$ is nonempty, find some $E \in \mathcal{F}$. Then $\varnothing \subseteq E \subseteq X$ tells us that $\varnothing \in \mathcal{H}(\mathcal{F})$.

Here's a quick sanity check.

Lemma 6.5. Fix a set X and a family $\mathcal{F} \subseteq \mathcal{P}(X)$. Then $\mathcal{H}(\mathcal{F})$ is a σ-ring.

Proof. Here are our checks.

- Union: suppose $\{E_i\}_{i=1}^\infty \subseteq \mathcal{H}(\mathcal{F})$. Then, for each i, we can write
$$E_i = \bigcup_{j=1}^\infty F_{ij}$$
where $F_{ij} \in \mathcal{F}$ for each j. So
$$\bigcup_{i=1}^\infty E_i = \bigcup_{i=1}^\infty \bigcup_{j=1}^\infty F_{ij}$$
shows that $\bigcup_{i=1}^\infty E_i$ is contained in a countable union of elements $F_{ij} \in \mathcal{F}$, so $\bigcup_{i=1}^\infty E_i \in \mathcal{H}(\mathcal{F})$.
- Subtraction: suppose $E, F \in \mathcal{H}(\mathcal{F})$. Indeed, we can write $E \subseteq \bigcup_{i=1}^\infty E_i$ for some $E_i \in \mathcal{F}$, so
$$E \setminus F \subseteq E \subseteq \bigcup_{i=1}^\infty E_i$$
has covered $E \setminus F$ by countably many elements E_i of $\mathcal{F}$, so we conclude $E \setminus F \in \mathcal{H}(\mathcal{F})$. ■

Example 6.6. Take $X = \mathbb{R}$ and $\mathcal{P}$ the prering from Exercise 5.46. But now we see that
$$\mathbb{R} = \bigcup_{i\in\mathbb{Z}} [i, i+1),$$
so any subset $E \subseteq \mathbb{R} \subseteq \bigcup_{i\in\mathbb{Z}}[i, i+1)$ is contained in a countable union of elements from $\mathcal{P}$. Thus, $\mathcal{H}(\mathcal{P}) = \mathcal{P}(\mathbb{R})$.

Example 6.7. Fix a set X and $\mathcal{P}$ the prering of finite set of X. Then any set in $\mathcal{H}(\mathcal{P})$ is countable as contained in a countable union of finite sets, and conversely any countable subset $E \subseteq X$ can write
$$E = \bigcup_{x\in E} \{x\}$$
to show that E is covered by countably many finite sets $\{x\} \in \mathcal{P}$. Thus, $\mathcal{H}(\mathcal{P})$ contains exactly the countable subsets of X.

It might feel like taking all the subsets of $\mathbb{R}$ makes us too big, but there are measures here anyway.

Example 6.8. Fix a set X and an element $x \in X$. Then we define the measure $\delta_x \colon \mathcal{P}(X) \to [0, \infty]$ by
$$\delta_x(E) := \begin{cases} 1 & x \in E, \\ 0 & x \notin E. \end{cases}$$
To see that this is a measure, fix disjoint $\{E_i\}_{i=1}^\infty$, and set $E := \bigsqcup_{i=1}^\infty E_i$. We have two cases.

- If $x \notin E_i$ for each i, then $x \notin E$, so $\delta_x(E) = 0 = \sum_{i=1}^\infty \delta_x(E_i)$.
- If $x \in E_{i_0}$ for some i_0, then note $x \in E_i$ for exactly one i because the E_i are disjoint. Also, $x \in E$ because $E_{i_0} \subseteq E$, so $\delta_x(E) = 1 = \delta_x(E_{i_0}) = \sum_{i=0}^\infty \mu(E_i)$.

The term "hereditary" comes from the following definition.

Definition 6.9 (Hereditary)**.** Fix a set X and nonempty family $\mathcal{G} \subseteq \mathcal{P}(X)$. Then $\mathcal{G}$ is *hereditary* if and only if $A \in \mathcal{G}$ and $A' \subseteq A$ implies $A' \in \mathcal{G}$.

Example 6.10. Indeed, given a collection $\mathcal{F} \subseteq \mathcal{P}(X)$, we can see that $\mathcal{H}(\mathcal{F})$ is hereditary. To see this, note $E \in \mathcal{H}(\mathcal{F})$ can be contained as $E \subseteq \bigcup_{i=1}^\infty E_i$ for some $\{E_i\}_{i=1}^\infty \subseteq \mathcal{F}$, but then any $E' \subseteq E$ has

$$E' \subseteq E \subseteq \bigcup_{i=1}^\infty E_i,$$

so E' is covered by a countable union of elements of $\mathcal{F}$, so $E' \in \mathcal{H}(\mathcal{F})$.

Remark 6.11. Note that the intersection of hereditary rings is still hereditary. Indeed, fixing our set X for hereditary rings $\{\mathcal{H}_\alpha\}_{\alpha\in\lambda}$ of X, we need to show

$$\mathcal{H} := \bigcap_{\alpha\in\lambda} \mathcal{H}_\alpha$$

is still hereditary. Well, for any $E \in \mathcal{H}$ and $E' \subseteq E$, we see $E \in \mathcal{H}_\alpha$ for each $\alpha \in \lambda$, so $E' \subseteq E$ forces $E' \in \mathcal{H}_\alpha$ for each $\alpha \in \lambda$, so actually $E' \in \mathcal{H}$.

Remark 6.12. Thus, we can see that the hereditary σ-ring $\mathcal{H}(\mathcal{F})$ generated by a family $\mathcal{F} \subseteq \mathcal{P}(X)$ is in fact the smallest hereditary σ-ring $\mathcal{H}$ containing $\mathcal{F}$, where $\mathcal{H}$ is the intersection of all the hereditary σ-rings containing $\mathcal{F}$. (Note $\mathcal{H}$ is hereditary by Remark 6.11 and a σ-ring by Lemma 5.28.)

- Certainly $\mathcal{H}(\mathcal{F})$ is a σ-ring by Lemma 6.5, and $\mathcal{H}(\mathcal{F})$ is hereditary by Example 6.10, so $\mathcal{H} \subseteq \mathcal{H}(\mathcal{F})$.
- Conversely, any $E \in \mathcal{H}(\mathcal{F})$ is contained in some countable union as $E \subseteq \bigcup_{i=1}^\infty E_i$ where $\{E_i\}_{i=1}^\infty \subseteq \mathcal{F}$. But then $E_i \in \mathcal{H}$ for each i, so $\bigcup_{i=1}^\infty E_i \in \mathcal{H}$ because $\mathcal{H}$ is a σ-ring, so $E \in \mathcal{H}$ because $\mathcal{H}$ is hereditary.

6.1.3 Outer Measures

We now have the following construction.

Notation 6.13. Fix a set X and nonempty family $\mathcal{F} \subseteq \mathcal{P}(X)$. Then give $\mu \colon \mathcal{F} \to [0,\infty]$, we will define $\mu^* \colon \mathcal{H}(\mathcal{F}) \to [0,\infty]$ by

$$\mu^*(E) := \inf\left\{\sum_{i=1}^\infty \mu(E_i) : \{E_i\}_{i=1}^\infty \subseteq \mathcal{F} \text{ and } E \subseteq \bigcup_{i=1}^\infty E_i\right\}.$$

Remark 6.14. Note that $E \in \mathcal{H}(\mathcal{F})$ tells us that the set we are taking the infimum of is in fact nonempty because $E \in \mathcal{H}(\mathcal{F})$ is contained in some countable collection of elements from $\mathcal{F}$. And in fact, for any subcollection $\{E_i\}_{i=1}^\infty \subseteq \mathcal{F}$ covering $E \in \mathcal{H}(\mathcal{F})$, we see that

$$\sum_{i=1}^\infty \mu(E_i) \geq 0$$

by definition of μ, so $\mu^*(E) \geq 0$ for any $E \in \mathcal{H}(\mathcal{F})$.

Here are some quick facts.

Lemma 6.15. Fix a set X and nonempty family $\mathcal{F} \subseteq \mathcal{P}(X)$. Further, fix some $\mu\colon \mathcal{F} \to [0,\infty]$. Then we have the following.

(a) $\mu^*(E) \le \mu(E)$ for any $E \in \mathcal{F}$.

(b) μ^* is monotone.

(c) μ^* is countably subadditive.

Proof. Here we go.

(a) Note that $\{E\} \subseteq \mathcal{F}$ covers E, so $\mu^*(E) \le \mu(E)$ follows.

(b) Suppose $E \subseteq F$ with $E, F \in \mathcal{H}(\mathcal{F})$. We need to show $\mu^*(E) \le \mu^*(F)$; certainly, if $\mu^*(F) = \infty$, then there is nothing to say. Otherwise, pick up any $\varepsilon > 0$, and we show

$$\mu^*(E) \le \mu^*(F) + \varepsilon,$$

which will be enough upon sending $\varepsilon \to 0^+$.

Now, the definition of $\mu^*(F)$ as an infimum promises some countable subcollection $\{F_i\}_{i=1}^\infty \subseteq \mathcal{F}$ covering F such that

$$\sum_{i=1}^\infty \mu(F_i) < \mu^*(F) + \varepsilon.$$

But now $E \subseteq F \subseteq \bigcup_{i=1}^\infty F_i$, so the definition of μ^* lets us conclude

$$\mu^*(E) \le \sum_{i=1}^\infty \mu(F_i) < \mu^*(F) + \varepsilon,$$

which finishes because we may now take $\varepsilon \to 0^+$.

(c) This requires some effort. Suppose that $A \in \mathcal{H}(\mathcal{F})$ and some $\{B_i\}_{i=1}^\infty \subseteq \mathcal{H}(\mathcal{F})$ covering A. We need to show that

$$\mu^*(A) \le \sum_{i=1}^\infty \mu^*(B_i).$$

Well, fix any $\varepsilon > 0$, and we will actually show that

$$\mu^*(A) \le \varepsilon + \sum_{i=1}^\infty \mu^*(B_i),$$

which will be enough upon sending $\varepsilon \to 0^+$. Certainly if $\mu^*(B_i)$ is infinite for any i, then there is nothing to say. Otherwise, each $\mu^*(B_i)$ is finite, so we may use the definition of μ^* as an infimum to find some countably subcollection $\{E_{ij}\}_{j=1}^\infty$ such that

$$B_i \subseteq \bigcup_{j=1}^\infty E_{ij} \qquad \text{and} \qquad \sum_{j=1}^\infty \mu(E_{ij}) \le \mu^*(B_i) + \frac{\varepsilon}{2^i}$$

because $\varepsilon/2^i > 0$ always. It follows that

$$A \subseteq \bigcup_{i=1}^\infty B_i \subseteq \bigcup_{i=1}^\infty \bigcup_{j=1}^\infty E_{ij},$$

so the definition of μ^* lets us say

$$\mu^*(A) \le \sum_{i=1}^\infty \sum_{j=1}^\infty \mu(E_{ij}) \le \sum_{i=1}^\infty \left(\frac{\varepsilon}{2^i} + \mu^*(B_i)\right) = \varepsilon + \sum_{i=1}^\infty \mu^*(B_i),$$

which is what we wanted. ■

It seems somewhat frustrating that we don't get equality in part (a) of Lemma 6.15, but we need a few more adjectives to make the proof go through.

Lemma 6.16. Fix a set X and a prering $\mathcal{P}$ on X equipped with a premeasure μ on $\mathcal{P}$. Then $\mu^*(E) = \mu(E)$ for any $E \in \mathcal{P}$.

Proof. Fix some $E \in \mathcal{F}$. Note that $\mu^*(E) \le \mu(E)$ already from Lemma 6.15, so we just need the other inequality. Well, for any cover

$$E \subseteq \bigcup_{i=1}^{\infty} E_i,$$

where $\{E_i\}_{i=1}^{\infty} \subseteq \mathcal{F}$, countable subadditivity tells us that

$$\mu(E) \le \sum_{i=1}^{\infty} \mu(E_i)$$

by Lemma 6.2. Thus, $\mu(E) \le \mu^*(E)$, which is what we wanted. ∎

The above results motivate the following definition.

Definition 6.17 (Outer measure)**.** Fix a set X and a hereditary σ-ring $\mathcal{H}$. An *outer measure* is a function $\mu^* \colon \mathcal{H} \to [0, \infty]$ which is monotone and countably subadditive.

Example 6.18. From Lemma 6.15, we note that if μ is a premeasure on a prering $\mathcal{P}$, then μ^* is an outer measure on the hereditary σ-ring $\mathcal{H}(\mathcal{P})$.

6.2 October 17

Please write neatly on the exam.

6.2.1 Restricting Outer Measures

Last time, in Example 6.18, we constructed an outer measure from a premeasure. We might hope that this outer measure is actually countably additive, thus giving us our measure, but most of the time it is not.

Instead, we are going to restrict our outer measure to some "σ-subring" which will then be a measure. The following definition is due to Carathéodory.

Definition 6.19. Fix a set X and a hereditary σ-ring $\mathcal{H}$ on X, and fix an outer measure $\nu \colon \mathcal{H} \to [0, \infty]$. Then a set $E \subseteq \mathcal{H}$ is *ν-measurable* if and only if

$$\nu(A) = \nu(A \cap E) + \nu(A \setminus E)$$

for any $A \in \mathcal{H}$. We will let $\mathcal{M}(\nu)$ denote the set of ν-measurable sets.

Remark 6.20. Because ν is already an outer measure, it is countably subadditive, so $\nu(A) \le \nu(A \cap E) + \nu(A \setminus E)$. Thus, we really only need to focus on proving

$$\nu(A) \ge \nu(A \cap E) + \nu(A \setminus E).$$

Here are the main results.

Theorem 6.21. Fix a set X and a hereditary σ-ring $\mathcal{H}$ on X, and fix an outer measure $\nu\colon \mathcal{H} \to [0,\infty]$. If nonempty, $\mathcal{M}(\nu)$ is a σ-ring, and $\nu|_{\mathcal{M}(\nu)}$ is a measure.

Remark 6.22. Later, we will also show that, give a premeasure μ on a prering $\mathcal{P}$, we will see $\mathcal{P} \subseteq \mathcal{M}(\mu^*)$, so $\mu^*|_{\mathcal{S}(\mathcal{P})}$ will be a measure on $\mathcal{S}(\mathcal{P})$ extending μ. We won't be precise about this until we need to, but we do want to see that we are close to the finish line.

Remark 6.23. It is indeed possible for $\mathcal{M}(\nu)$ to be empty. For example, the outer measure $\nu\colon \mathcal{P}(X) \to [0,\infty]$ by $\nu(E) := 1$ for any $E \subseteq X$ has no ν-measurable sets.

Let's begin our proof.

Proof of Theorem 6.21. We proceed in steps.

1. Finite union: given $E, F \in \mathcal{M}(\nu)$, we show $E \cup F \in \mathcal{M}(\nu)$. Well, for any $A \in \mathcal{H}$, we compute

$$\begin{aligned}\nu(A \cap (E \cup F)) + \nu(A \setminus (E \cup F)) &= \nu((A \cap E) \sqcup ((A \setminus E) \cap F)) + \nu((A \setminus E) \setminus F) \\ &\le \nu(A \cap E) + \nu((A \setminus E) \cap F) + \nu((A \setminus E) \setminus F),\end{aligned}$$

   where we have used subadditivity at the end.

   Because F is ν-measurable, the last two pieces become $\nu(A \setminus E)$, where we note $A \setminus E \subseteq E \in \mathcal{H}$ implies $A \setminus E \in \mathcal{H}$. Thus, because E is ν-measurable, this in total collapses down to $\nu(A)$, which is enough by Remark 6.20.

2. Subtraction: given $E, F \in \mathcal{M}(\nu)$, we show $E \setminus F \in \mathcal{M}(\nu)$. Well, for any $A \in \mathcal{H}$, we compute

$$\begin{aligned}\nu(A \cap (E \setminus F)) + \nu(A \setminus (E \setminus F)) &= \nu((A \cap E) \setminus F) + \nu((A \setminus E) \sqcup (A \cap E \cap F)) \\ &\le \nu((A \cap E) \setminus F) + \nu(A \setminus E) + \nu((A \cap E) \cap F),\end{aligned}$$

   where we have used subadditivity at the end.

   Now, because F is ν-measurable, we see $\nu((A\cap E)\setminus F)+\nu((A\cap E)\cap F) = \nu(A\cap E)$, where $A\cap E \subseteq E$ is in $\mathcal{H}$ because $E \in \mathcal{H}$. Thus, because E is ν-measurable, this in total collapses down to $\nu(A)$, which is enough by Remark 6.20.

3. Strong finitely additive: for any $A \in \mathcal{H}$ and disjoint $E, F \in \mathcal{M}(\nu)$, we claim

$$\nu(A \cap (E \sqcup F)) \stackrel{?}{=} \nu(A \cap E) + \nu(A \cap F).$$

   Well, because E is measurable, we note $A \cap (E \sqcup F) \subseteq A$ must live in $\mathcal{H}$ and so

$$\nu(A \cap (E \sqcup F)) = \nu(A \cap (E \sqcup F) \cap E) + \nu(A \cap (E \sqcup F) \setminus E) = \nu(A \cap E) + \nu(A \cap F),$$

   where the last equality has used the fact that $E \cap F = \varnothing$.

   By induction, for finitely many pairwise disjoint ν-measurable subsets $\{E_i\}_{i=1}^n \subseteq \mathcal{M}(\nu)$, we see

$$\nu\left(A \cap \bigsqcup_{i=1}^n E_i\right) = \sum_{i=1}^n \nu(A \cap E_i).$$

4. Finitely additive: we show $\nu|_{\mathcal{M}(\nu)}$ is finitely additive: for finitely many pairwise disjoint ν-measurable subsets $\{E_i\}_{i=1}^n E_i \in \mathcal{M}(\nu)$, we note $A := \bigcup_{i=1}^n E_i \in \mathcal{M}(\nu)$ because $\mathcal{M}(\nu)$ is a preserved by finite unions, so we set $A := \bigcup_{i=1}^n E_i$ to give

$$\nu\left(\bigsqcup_{i=1}^n E_i\right) = \nu\left(A \cap \bigsqcup_{i=1}^n E_i\right) = \sum_{i=1}^n \nu(A \cap E_i) = \sum_{i=1}^n \nu(E_i).$$

5. Countable union: given some countable subcollection $\{E_i\}_{i=1}^\infty \subseteq \mathcal{M}(\nu)$, and let F be their union. Now, we set
$$F_i := E_i \setminus \bigcup_{j<i} E_j$$
so that F is the union of the F_i (certainly $F_i \subseteq E_i \subseteq F$ for any i, and conversely any $x \in F$ is in some E_i, for i as small as possible, so $x \in F_i$), and the F_i are pairwise disjoint (if $i \neq j$, then $i < j$ without loss of generality, so $F_i \subseteq E_i \setminus F_j$ is disjoint from F_j). The point is that we are now dealing with pairwise disjoint subsets.

   We now need to show that F is ν-measurable. Well, fix any $A \in \mathcal{H}$. Then, for any n, we note that $\mathcal{M}(\nu)$ is already a ring, so $\bigsqcup_{i=1}^n F_i$ is in $\mathcal{M}(\nu)$ for any finite n, so
$$\nu(A) = \nu\left(A \cap \bigsqcup_{i=1}^n F_i\right) + \nu\left(A \setminus \bigsqcup_{i=1}^n F_i\right) = \sum_{i=1}^n \nu(A \cap F_i) + \nu\left(A \setminus \bigsqcup_{i=1}^n F_i\right),$$
where we have used finite additivity. Because ν is monotone, we may lower-bound this by
$$\nu(A) \geq \sum_{i=1}^n \nu(A \cap F_i) + \nu(A \setminus F)$$
for any n. Sending $n \to \infty$ now, we see
$$\nu(A) \geq \sum_{i=1}^\infty \nu(A \cap F_i) + \nu(A \setminus F),$$
but then countable subadditivity of ν kicks in and tells us that
$$\nu(A) \geq \nu(A \cap F) + \nu(A \setminus F),$$
so we are done by Remark 6.20.

6. Countably additive: we show $\nu|_{\mathcal{M}(\nu)}$ is countably additive. Well, given some countable pairwise disjoint collection of ν-measurable sets $\{E_i\}_{i=1}^n \subseteq \mathcal{M}(\nu)$, we see that the previous step has told us
$$\nu(A) \geq \sum_{i=1}^\infty \nu(A \cap E_i) + \nu\left(A \setminus \bigsqcup_{i=1}^n E_i\right) \geq \nu\left(A \cap \bigsqcup_{i=1}^n E_i\right) + \nu\left(A \setminus \bigsqcup_{i=1}^n E_i\right) \geq \nu(A).$$
But $\mathcal{M}(\nu)$ is a σ-ring, so may set $A := \bigsqcup_{i=1}^\infty E_i$ so that the above equalities actually read
$$\nu(A) = \sum_{i=1}^n \nu(E_i),$$
which is what we wanted. ■

The previous theorem has an annoying hypothesis that $\mathcal{M}(\nu)$ is nonempty. In the cases we're interested in, this is no issue.

Theorem 6.24. Fix a set X and a prering $\mathcal{P}$ on X equipped with a premeasure μ on $\mathcal{P}$. Then $\mathcal{P} \subseteq \mathcal{M}(\mu^*)$.

Proof. Fix some $E \in \mathcal{P}$, and we need to show that $E \in \mathcal{M}(\mu^*)$. By Remark 6.20, it suffices to pick up any $A \in \mathcal{H}(\mathcal{P})$ and show
$$\mu^*(A) \geq \mu^*(A \cap E) + \mu^*(A \setminus E).$$
If $\mu^*(A) = \infty$, then there is nothing to say. Otherwise, we have $\mu^*(A) < \infty$. As usual, fix some $\varepsilon > 0$, and we will show that $\mu^*(A \cap E) + \mu^*(A \setminus E) \leq \mu^*(A) + \varepsilon$.

Well, by definition of $\mu^*(A)$, we can find some collection $\{E_i\}_{i=1}^\infty \subseteq \mathcal{P}$ such that

$$\mu^*(A) + \varepsilon > \sum_{i=1}^\infty \mu(E_i).$$

We now decompose each $\mu(E_i)$. Note that $E_i = (E_i \cap E) \sqcup (E_i \setminus E)$ is a disjoint union because an element of E_i is exactly one of in E or not. Further, by the prering property, we can write

$$E \setminus E_i = \bigsqcup_{j=1}^{n_i} F_{ij}$$

for some disjoint $\{F_{ij}\}_{j=1}^\infty \subseteq \mathcal{P}$. In total, we see that $E_i = (E_i \cap E) \sqcup \bigsqcup_{j=1}^{n_i} F_{ij}$ is a disjoint union because $E_i \cap E$ is certainly disjoint from each of the $F_{ij} \subseteq E \setminus E_i$, and the F_{ij} are disjoint from each other by construction. In total, we use the countable additivity of μ to write

$$\mu^*(A) + \varepsilon > \sum_{i=1}^\infty \mu(E_i) \geq \sum_{i=1}^\infty \left(\mu(E_i \cap E) + \sum_{j=1}^{n_i} \mu(F_{ij}) \right) = \sum_{i=1}^\infty \mu(E_i \cap E) + \sum_{i=1}^\infty \sum_{j=1}^{n_i} \mu(F_{ij}).$$

Now, we note that $A \cap E \subseteq \bigcup_{i=1}^\infty (E_i \cap E)$ and $A \setminus E \subseteq \bigcup_{i=1}^\infty (E_i \setminus E) \subseteq \bigcup_{i=1}^\infty \bigcup_{j=1}^{n_i} F_{ij}$, so countable subadditivity of μ^* (by Lemma 6.15) lets us conclude

$$\mu^*(A) + \varepsilon > \mu(A \cap E) + \mu(A \setminus E).$$

Note that we have implicitly used the fact that $\mu^*|_{\mathcal{P}} = \mu$ from Lemma 6.16. Anyway, sending $\varepsilon \to 0^+$ completes the proof. ■

6.2.2 Completeness

We have a notion of completeness for our measures; here is the definition.

Definition 6.25 (Compelete)**.** Fix a set X and a family $\mathcal{F} \subseteq \mathcal{P}(X)$. Then a function $\nu \colon \mathcal{F} \to [0, \infty]$ is *complete* if and only if any $E \in \mathcal{F}$ with $F \subseteq E$ and $\nu(E) = 0$ must have $F \in \mathcal{F}$ and $\nu(F) = 0$.

Remark 6.26. If μ is a complete measure on a σ-ring $\mathcal{S}$ on X, then we claim that any countable subset in $\mathcal{S}$ has measure zero. Indeed, if $A \in \mathcal{S}$ is countable, we can write

$$\mu(A) = \mu\left(\bigsqcup_{a \in A} \{a\} \right) = \sum_{a \in A} \mu(\{a\}) = \sum_{a \in A} 0 = 0.$$

We will continue this after the midterm.

6.3 October 21

I did poorly on the midterm, and I'm too tired to be okay with it.

6.3.1 Miscellaneous Outer Measures

We quickly complete an example from last class.

Lemma 6.27. Fix an outer measure ν on a hereditary σ-ring $\mathcal{H}$. Then any set $E \in \mathcal{H}$ with $\nu(E) = 0$ is ν-measurable.

Proof. Fix any $A \in \mathcal{H}$. Because outer measures are monotone, we see

$$\nu(A \setminus E) + \nu(A \cap E) \leq \nu(A) + \nu(E) = \nu(A),$$

so we conclude that $\nu(A \setminus F) + \nu(A \cap F) = \nu(A)$ by Remark 6.20. ■

Lemma 6.28. If ν is an outer measure on a hereditary σ-ring $\mathcal{H}$, then $\nu|_{\mathcal{M}(\nu)}$ is complete when $\mathcal{M}(\nu)$ is nonempty.

Proof. Given $E \in \mathcal{M}(\nu)$ with $\nu(E) = 0$, we note that any $F \subseteq E$ is ν-measurable and has $\nu(F) = 0$. Well, we certainly have $0 \leq \nu(F)$, and then we see that $\nu(F) \leq \nu(E) = 0$ because ν is monotone, so we conclude that $\nu(F) = 0$. Thus, F is ν-measurable by Lemma 6.27. ■

We take a moment to acknowledge that our restricted outer measures are in fact extending our premeasures when appropriate.

Lemma 6.29. Fix a premeasure μ on a σ-ring $\mathcal{S}$ (viewed as a prering). Then, for any $B \in \mathcal{H}(\mathcal{S})$, there exists some $E \in \mathcal{S}$ such that $B \subseteq E$ and $\mu^*(B) = \mu(E)$.

Proof. Recall that

$$\mu^*(B) = \inf\left\{\sum_{i=1}^\infty \mu(E_i) : \{E_i\}_{i=1}^\infty \subseteq \mathcal{S} \text{ and } B \subseteq \bigcup_{i=1}^\infty E_i\right\}.$$

Notably, if we have some $\{E_i\}_{i=1}^\infty$ covering B, then we are told that $\mu^*(B) \leq \sum_{i=1}^\infty \mu(E_i)$, but in fact $\mathcal{S}$ being a σ-ring forces $E := \bigcup_{i=1}^\infty E_i$ to be in $\mathcal{S}$, so $B \subseteq E$ forces the stronger inequality

$$\mu^*(B) \leq \mu^*(E) = \mu(E) \leq \sum_{i=1}^\infty \mu(E_i).$$

Note that we have used countable subadditivity from Lemma 6.2 and $\mu^*(E) = \mu(E)$ from Lemma 6.16. It follows that $\inf\{\mu(E) : B \subseteq E\} \leq \mu^*(B)$. But of course $B \subseteq E$ forces $\mu^*(B) \leq \mu(E)$ from definition of μ^*, so in fact

$$\mu^*(B) = \inf\{\mu(E) : B \subseteq E\}. \tag{6.1}$$

It remains to show that this infimum is achievable. Certainly if $\mu^*(B) = \infty$, then any $E \in \mathcal{S}$ with $B \subseteq E$ will have $\mu^*(E) = \infty$, finishing.

Otherwise, take $\mu^*(B) < \infty$. From (6.1), we can a sequence $\{E_i\}_{i=1}^\infty \subseteq \mathcal{S}$ such that $E_i \subseteq B$ and $\mu(E_i) < \mu^*(B) + \frac{1}{i}$ for each i. We now define

$$E := \bigcap_{i=1}^\infty E_i,$$

which is an element of $\mathcal{S}$ by Remark 5.24. However, because μ^* is monotone by Lemma 6.15, we see that $B \subseteq E$ forces $\mu^*(B) \leq \mu^*(E)$ while $E \subseteq E_i$ for each i forces

$$\mu^*(E) \leq \mu^*(E_i) < \mu^*(B) + 1/i$$

for each i. Sending $i \to \infty$ recovers $\mu(E) = \mu^*(E) = \mu^*(B)$, where we have used Lemma 6.16 to finish. ■

6.3.2 Uniqueness of Extensions

It is not always true that the extension of a measure must be unique.

Example 6.30. Give an uncountable set X the discrete topology, and let $\mathcal{S}$ denote the σ-ring of countable sets. Then the zero function μ on S is a measure; however, we have the following two extensions ν to a measure on all of $\mathcal{P}(X)$.

- We could set $\nu(E) = 0$ for any uncountable E.
- We could set $\nu(E) = \infty$ for any uncountable E.

Example 6.31. Let $\mathcal{P}$ be the prering of right-half-open intervals of $\mathbb{R}$. Then the measure μ on $\mathcal{P}$ by $\mu([a, b)) = \infty$ for $a < b$ while $\mu(\varnothing) = 0$. Then here are two extensions of μ.

- We could set μ to be infinite for any nonempty subset of $\mathbb{R}$.
- We could set $\mu(E)$ be the counting measure on $\mathbb{R}$.

The issue in these examples is that there is too much allowed ∞. To deal with this, we have the following definition.

Definition 6.32 (σ-finite)**.** Fix a set X and a prering $\mathcal{P}$ on X. Then a premeasure μ on $\mathcal{P}$ is *σ-finite* if and only if $E \subseteq \mathcal{P}$ has some countable collection $\{E_i\}_{i=1}^\infty \subseteq \mathcal{P}$ with $E = \bigcup_{i=1}^\infty E_i$ and $\mu(E_i) < \infty$ for each i.

Remark 6.33. In the above situation, we note that any $E \in \mathcal{H}(\mathcal{P})$ can be covered by $\{G_i\}_{i=1}^\infty \subseteq \mathcal{P}$ with $\mu(G_i) < \infty$ for each i. Indeed, we can at least cover $E \in \mathcal{H}(\mathcal{P})$ by some $\{E_i\}_{i=1}^\infty \subseteq \mathcal{P}$, and then each E_i has a cover

$$E_i \subseteq \bigcup_{j=1}^\infty F_{ij}$$

where $\mu(F_{ij}) < \infty$ because μ is σ-finite. Reordering our countable union of countable unions covering E into some sequence $\{G_i\}_{i=1}^\infty$, we see $E \subseteq \bigcup_{i=1}^\infty G_i$ while $\mu(G_i) < \infty$.

It will now turn out that σ-finite things have unique extensions. Let's first see that our outer measure extension is special, though it need not be the only extension yet.

Lemma 6.34. Fix a set X and a prering $\mathcal{P}$ on X equipped with a premeasure μ. Then for any σ-ring containing $\mathcal{P}$ and contained in $\mathcal{M}(\mu^*)$, we have $\mu^*|_\mathcal{S}$ is the largest measure on $\mathcal{S}$ extending μ on $\mathcal{P}$. In other words, if ν is any measure extending $\mathcal{P}$ to $\mathcal{S}$, then $\nu(E) \le \mu^*(E)$ for any $E \in \mathcal{S}$.

Proof. Note that μ^* extends μ by Lemma 6.16.

Now, suppose that ν is a measure on $\mathcal{S}$ extending the premeasure ν on $\mathcal{P}$. Now, for any $G \in \mathcal{S}$ the fact that $\mathcal{S} \subseteq \mathcal{H}(\mathcal{P})$, we may find some $\{E_i\}_{i=1}^\infty$ contained in $\mathcal{P}$ covering G. This tells us

$$\nu(G) \le \sum_{i=1}^\infty \nu(E_j) = \sum_{i=1}^\infty \mu(E_i) = \sum_{i=1}^\infty \mu^*(E_i) = \sum_{i=1}^\infty \mu(E_i).$$

Taking the infimum allows us to conclude $\nu(G) \le \mu^*(G)$. ■

Now, here is our main result.

Theorem 6.35. Fix a set X and a prering $\mathcal{P}$ on X equipped with a σ-finite premeasure μ on X. Then, for some σ-ring $\mathcal{S} \subseteq \mathcal{M}(\mu^*)$, our $\mu^*|_{\mathcal{S}}$ is the unique extension of μ to a measure on $\mathcal{S}$.

Proof. Let ν be some measure on $\mathcal{S}$ extending μ. Note that we really only have one inequality here, thanks to Lemma 6.34. Anyway, we proceed in steps. Fix any $G \in \mathcal{S}$.

1. If $G \in \mathcal{S}$ has $\nu(G) = \infty$, then our bound $\mu^*(G) \geq \nu(G)$ from Lemma 6.34 forces equality. So we may now assume that $\nu(G) < \infty$.

2. Otherwise, we take $\nu(G) < \infty$. Suppose that $G \in \mathcal{S}$ and $G \subseteq E$ where $E \in \mathcal{P}$. Note that we at least still know $\nu(G) \leq \mu^*(G) \leq \mu^*(E) = \mu(E)$ because μ^* is monotone by Lemma 6.15. On the other hand, we know that $E \in \mathcal{P}$ is measurable by Theorem 6.24, so we can use additivity to write

$$\nu(E) = \nu(G) + \nu(E \setminus G) \leq \mu^*(G) + \mu^*(E \setminus G) = \mu^*(E) \stackrel{*}{=} \mu(E) = \nu(E).$$

   Note that we have used Lemma 6.16 in $\stackrel{*}{=}$. Thus, equalities must follow everywhere, so in particular $\nu(G) = \mu^*(G)$ is forced.

3. Lastly, we take $\nu(G) < \infty$ with any $G \in \mathcal{S}$. We now use the σ-finite hypothesis: by Remark 6.33, we may cover G with a countable collection $\{F_i\}_{i=1}^\infty \subseteq \mathcal{P}$ covering G such that $\nu(F_i) < \infty$ for each i.

   Now, as usual, we set

$$F_i' := F_i \Bigg\backslash \bigsqcup_{j<i} F_j$$

   so that the F_i' are now disjoint (if $i \neq i'$, say with $i < i'$ without loss of generality, then $F_{i'}' \subseteq F_{i'} \setminus F_i$) even though G is still covered by the F_i' (any $x \in G$ lives in some least F_i, so $x \in F_i'$ follows). Additionally, $F_i' \in \mathcal{S}$ by Remark 5.24, so the previous step tells us that $G \cap F_i' \subseteq F_i'$ implies $\nu(G \cap F_i') = \mu^*(G \cap F_i')$ and thus

$$\nu(G) = \sum_{i=1}^\infty \nu(G \cap F_i') = \sum_{i=1}^\infty \mu^*(G \cap F_i') = \mu^*(G)$$

   by using additivity. ■

6.4 October 24

The midterms were all graded. The mean was 15.68, and the standard deviation was 7.64. Roughly speaking, a score of 15 (and continuing to work at that level for the rest of the class) should roughly correspond to a B+.

We completed the proof of Theorem 6.35 from last class, but I have simply completed the proof there for continuity reasons.

6.4.1 Continuity Properties

Let's discuss continuity a little.

Proposition 6.36. Fix a σ-ring $\mathcal{S}$ on a set X equipped with a measure μ on $\mathcal{S}$. A collection $\{E_i\}_{i=1}^\infty \subseteq \mathcal{S}$ such that $E_n \subseteq E_{n+1}$ for each i will have

$$\lim_{n\to\infty} \mu(E_n) = \mu\left(\bigcup_{i=1}^\infty E_i\right).$$

Proof. Set $E := \bigcup_{i=1}^\infty E_i$, for brevity, and we define

$$F_i := E_i \setminus E_{i-1},$$

where $E_0 := \varnothing$. Note that the F_i are now pairwise disjoint: if $i \neq j$, then without loss of generality say $i < j$, so $F_i \subseteq E_i \subseteq E_{j-1}$ while $F_j = E_j \setminus E_{j-1}$ is disjoint from E_{j-1}, so $F_i \cap F_j = \varnothing$. Thus, we note that

$$E_n = \bigsqcup_{k=1}^n F_k.$$

Indeed, certainly each $k \leq n$ has $F_k \subseteq E_k \subseteq E_n$; and conversely any $x \in E_n$ belongs to some E_k with $k \leq n$ minimal, implying $x \notin F_{k-1}$ and so $x \in F_k$. In particular, we note that

$$E = \bigcup_{n=1}^\infty E_n = \bigcup_{n=1}^\infty \bigcup_{i=1}^n F_i = \bigcup_{i=1}^\infty \bigcup_{n \geq i} F_i = \bigcup_{i=1}^\infty F_i = \bigsqcup_{i=1}^\infty F_i$$

is still a disjoint union because the F_i are pairwise disjoint.

Thus, by the countable additivity of μ, we compute

$$\begin{aligned}
\lim_{n\to\infty} \mu(E_n) &= \lim_{n\to\infty} \mu\left(\bigsqcup_{k=1}^n F_k\right) \\
&= \lim_{n\to\infty} \sum_{k=1}^n \mu(F_k) \\
&= \sum_{k=1}^\infty \mu(F_k) \\
&= \mu\left(\bigsqcup_{k=1}^\infty F_k\right) \\
&= \mu(E),
\end{aligned}$$

which is what we wanted. ■

Corollary 6.37. Fix a σ-ring $\mathcal{S}$ on a set X equipped with a measure μ on $\mathcal{S}$. Suppose we have a collection $\{E_i\}_{i=1}^\infty \subseteq \mathcal{S}$ such that $\mu(E_1) < \infty$ and $E_n \supseteq E_{n+1}$ for each i. Then we have

$$\lim_{n\to\infty} \mu(E_n) = \mu\left(\bigcap_{i=1}^\infty E_i\right).$$

Proof. Set

$$E := \bigcap_{i=1}^\infty E_i.$$

Then we define $F_i := E_1 \setminus E_i$ so that

$$F := \bigcup_{i=1}^\infty F_i = \bigcup_{i=1}^\infty (E_1 \setminus E_i) = E_1 \Big\backslash \bigcap_{i=1}^\infty E_i = E_1 \setminus E.$$

On the other hand, we note $E_i \supseteq E_{i+1}$ implies $E_1 \setminus E_i \supseteq E_1 \setminus E_{i+1}$, so $F_i \subseteq F_{i+1}$.

Thus, applying Proposition 6.36, we see

$$\lim_{n\to\infty} \mu(F_n) = F.$$

Rearranging gets the needed result. However, we note that $\mu(E_i) \leq \mu(E_1) < \infty$ for each i because μ is monotone by Lemma 5.51, so we can say

$$\begin{aligned}\lim_{n\to\infty}\mu(E_n) &= \lim_{n\to\infty}\mu(E_1\setminus F_n)\\ &\overset{*}{=} \lim_{n\to\infty}(\mu(E_1)-\mu(F_n))\\ &= \mu(E_1)-\lim_{n\to\infty}\mu(F_n) = \mu(E_1)-\mu(F)\\ &\overset{*}{=} \mu(E_1\setminus F) = \mu(E),\end{aligned}$$

where we have used Remark 5.21 at each $\overset{*}{=}$. This finishes. ■

Remark 6.38. If we do not require $\mu(E_1) < \infty$, then the statement is false: set $\alpha(t) := t$ be an increasing, left-continuous function, and let μ be the corresponding measure coming as a restricted outer measure from the premeasure measure μ_α of Proposition 5.48.

Then set $E_i := [i, \infty)$, which is measurable by Theorem 6.24. Here are our checks.

- Note $\mu(E_i) = \infty$. Indeed, for any positive integer N, we note that $\mu(E_i) \geq \mu([i, i+N)) = N$ because μ is monotone by Lemma 5.51 and restricts properly by Lemma 6.16. It follows $\mu(E_i) > \infty$.
- On the other hand, note $\bigcap_{i=1}^\infty E_i = \varnothing$ because no real number is larger than every positive integer, and $\mu(\varnothing) = \mu([0,0)) = 0$ using Lemma 6.16.

6.4.2 Borel Measures

We take a moment to recognize that we've actually built a measure.

Definition 6.39 (Lebesgue–Stieltjes measure)**.** Let $\mathcal{P}$ be the prering of Exercise 5.46 and some increasing, left-continuous function $\alpha\colon \mathbb{R}\to\mathbb{R}$. Then the measure $\mu_\alpha^*|_{\mathcal{M}(\mu_\alpha^*)}$ restricted by Remark 5.40 from the premeasure of Proposition 5.48 is the *Lebesgue–Stieltjes measure*. The *Lebesgue measure* is the measure coming from $\alpha(t) = t$.

The measurable sets for each μ_α might be difficult to handle, so let's find some subsets which are always measurable.

Definition 6.40 (Borel set)**.** The σ-ring generated by the prering $\mathcal{P}$ of Exercise 5.46 is called the σ-ring of *Borel sets*. A measure on the Borel sets is called a *Borel measure*.

Let's go find some Borel sets.

Example 6.41. Here are some Borel sets of $\mathbb{R}$. Let $a, b \in \mathbb{R}$.

- Note $(-\infty, a) = \bigcup_{n=1}^\infty [a-n, a)$, so $(-\infty, a)$ is a Borel set.
- Note $(a, \infty) = \bigcup_{i=1}^\infty [a+1/i, a+i)$, so (a,∞) is a Borel set. Namely, if $b > a$, then there is some positive integer $i > \max\{1/(b-a), b-a\}$, so $b \in [a+1/i, a+i)$.
- From Remark 5.24, we note that $(a,b) = (-\infty, b) \cap (a, \infty)$ is a Borel set.

Exercise 6.42. Any open subset $U \subseteq \mathbb{R}$ is a Borel set.

Proof. If $U = \mathbb{R}$, there is nothing to say because $\mathbb{R} = (-\infty, 1] \cup [1, \infty)$, so we are done by Example 6.41. Thus, suppose that we have some $y \in \mathbb{R} \setminus U$.

Otherwise, given some $x \in U$, we note that there is some $\varepsilon > 0$ such that $B(x, \varepsilon) \subseteq U$. Note that $\varepsilon < |x - y|$ because $\varepsilon > |x - y|$ will force $y \in B(x, \varepsilon) \setminus U$. As such, we may let r_x be the supremum of all such ε, which we see is finite. Note $r_x > 0$ because $\varepsilon > 0$ always.

We now note that $B(x, r_x) \subseteq U$. Indeed, if $x' \in B(x, r_x)$, then

$$|x - x'| < r_x$$

implies that $|x - x'|$ is not an upper-bound for our set of εs, so we can find some $\varepsilon > 0$ such that $B(x, \varepsilon) \subseteq U$ and $|x - x'| < \varepsilon$, so $x' \in B(x, \varepsilon) \subseteq U$.

We now proceed with the proof directly. The rationals are countable, so enumerate the rationals in U as $\{q_n\}_{n=1}^{\infty}$. For each q_n, set $r_n := r_{q_n}$. We now claim that

$$U \stackrel{?}{=} \bigcup_{n=1}^{\infty} B(q_n, r_n).$$

Certainly $B(q_n, r_n) \subseteq U$ for each n, as shown above. Conversely, if $x \in U$, find $r > 0$ such that $B(x, r) \subseteq U$. Because the rationals are dense in $\mathbb{R}$, we may find some rational $q \in B(x, r/3)$. But now we see that

$$B(q, 2r/3) \subseteq B(x, r) \subseteq U,$$

so $r_q \geq 2r/3$. Thus, $x \in B(q, 2r/3) \subseteq B(q, r_q)$, so $x \in \bigcup_{n=1}^{\infty} B(q_n, r_n)$ follows because each rational q is some q_n. ■

Example 6.43. Any closed subset $V \subseteq \mathbb{R}$ has $V = \mathbb{R} \setminus U$ for some open U, so V is a measurable set by Exercise 6.42.

Definition 6.44 (Borel–Stieltjes measure)**.** Let $\mathcal{S}$ be the σ-ring of Borel sets. Given some increasing, left-continuous function $\alpha \colon \mathbb{R} \to \mathbb{R}$, we note that $\mathcal{M}(\mu_\alpha^*)$ contains $\mathcal{P}$ by Theorem 6.24 and is a σ-ring by Theorem 6.21 and thus contains $\mathcal{S}$ by definition of $\mathcal{S}$. Thus, we define $\mu_\alpha^*|_{\mathcal{S}}$ (which is a measure from Remark 5.40) to be the corresponding *Borel–Stieltjes measure*.

We now note that these are actually all the measures.

Proposition 6.45. Fix a Borel measure μ on the Borel sets $\mathcal{B}$ of $\mathbb{R}$ such that $\mu([a, b)) < \infty$ for any $a, b \in \mathbb{R}$. Then there exists an increasing, left-continuous function $\alpha \colon \mathbb{R} \to \mathbb{R}$ such that $\mu = \mu_\alpha$.

Proof. Define the function $\alpha \colon \mathbb{R} \to \mathbb{R}$ by

$$\alpha(t) := \begin{cases} \mu([0, t)) & t \geq 0, \\ -\mu([t, 0)) & t \leq 0. \end{cases}$$

Notably, at $t = 0$, $\mu([0, 0)) = 0 = -\mu([0, 0))$. We now run our checks.

- Suppose $a, b \in \mathbb{R}$ has $a \leq b$. We claim $\mu([a, b)) = \alpha(b) - \alpha(a)$; there's nothing to say if $a = b$. We have the following cases.
  - If $a \geq 0$, then we note $b \geq a \geq 0$, so

    $$\mu([a, b)) = \mu([0, b) \setminus [0, a)) \stackrel{*}{=} \mu([0, b)) - \mu([0, a)) = \alpha(b) - \alpha(a),$$

    where $\stackrel{*}{=}$ is by Remark 5.21.

- If $b \geq 0 \geq a$, then

$$\mu([a,b)) = \mu([a,0) \sqcup [0,b)) = \mu([a,0)) + \mu([0,b)) = -\alpha(a) + \alpha(b).$$

- Lastly, if $0 \geq b \geq a$, then

$$\mu([a,b)) = \mu([a,0) \setminus [b,0)) \stackrel{*}{=} \mu([a,0)) - \mu([b,0)) = -\alpha(a) + \alpha(b),$$

where $\stackrel{*}{=}$ is by Remark 5.21.

- Increasing: given real numbers $a, b \in \mathbb{R}$ such that $a \leq b$, then we note $\alpha(b) - \alpha(a) = \mu([a,b)) \geq 0$, so $\alpha(b) \geq \alpha(a)$ follows.

- Left-continuous: fix some real number $b \in \mathbb{R}$ and some $\varepsilon > 0$ so that we need some $\delta > 0$ such that $b - \delta < a \leq b$ implies $|\alpha(b) - \alpha(a)| < \varepsilon$. To begin, we at least note that $\alpha(a) \leq \alpha(b)$, so $\alpha(b) - \alpha(a) \geq 0 > -\varepsilon$, so it suffices for

$$b - \delta < a \leq b \implies \mu([a,b)) = \alpha(b) - \alpha(a) < \varepsilon.$$

To begin, we note $\mu([b-1,b)) < \infty$ by hypothesis on μ (here is where we use this hypothesis!), so we set $a_n := b - \frac{1}{n}$ and $E_n := [a_n, b)$ so that $\mu(E_1) < \infty$ and $E_n \supseteq E_{n+1}$ for each n. It follows from Corollary 6.37 that

$$\lim_{n\to\infty} \mu(E_n) = \mu\left(\bigcap_{i=1}^{\infty} E_i\right) = \mu(\varnothing) = 0.$$

Indeed, we note that $\bigcap_{i=1}^{\infty} E_i = \varnothing$ because any $x \in E_i$ for each i must have $x < b$, but then any $i > 1/(b-x)$ will force $x \notin E_i$. And also, $\mu(\varnothing) = 0$ by Remark 5.39 because $\mu(\varnothing) \leq \mu([b-1,b)) < \infty$, where we are using the fact that μ is monotone from Lemma 5.51.

In total, we see that there is some N such that $n \geq N$ implies $\mu(E_n) < \varepsilon$. Set $\delta := \frac{1}{N}$ so that $b - \delta < a \leq b$ implies that $a \in E_N$, so $[a,b) \subseteq E_n$, so

$$\mu([a,b)) \leq \mu(E_N) < \varepsilon$$

by Lemma 5.51.

- Lastly, we show $\mu = \mu_\alpha|_{\mathcal{B}}$. Let $\mathcal{P}$ be the prering of right-half-open intervals. Note that μ_α at least makes sense from the above checks, so the fact that $\mathcal{B} \subseteq \mathcal{M}(\mu_\alpha^*)$ as discussed in the definition of the Borel–Stieltjes measure $\mu_\alpha^*|_{\mathcal{B}}$.

Now, we note that $\mu_\alpha([a,b)) = \alpha(b) - \alpha(a) = \mu([a,b))$ as checked above (we have used Lemma 6.16), so $\mu_\alpha^*|_{\mathcal{B}}$ and μ are both extensions of the premeasure μ_α on $\mathcal{P}$, so Theorem 6.35 follows $\mu_\alpha^*|_{\mathcal{B}} = \mu$. This finishes. ■

6.4.3 The Haar Measure

Let's build up to talking about the Haar measure.

Remark 6.46. We'll show on the homework that the Lebesgue measure μ on $\mathbb{R}$ is translation-invariant: if E is measurable, and $t \in \mathbb{R}$, then $E + t = \{r + t : r \in E\}$ is measurable has the same measure as E. In fact, any translation-invariant measure on the Borel sets is a multiple of μ.

There is a different definition of Borel sets is a little different in general.

Definition 6.47 (Borel set)**.** Fix a locally compact Hausdorff space X. Then the σ-ring of *Borel sets* is the σ-ring generated by the compact subsets of X. A *Borel measure* is a measure on the Borel sets of X.

Example 6.48. Certainly any compact subset of $\mathbb{R}$ is closed by Corollary 4.13 and thus Borel by Example 6.43, so the Borel subsets of $\mathbb{R}$ coming from the above definition are indeed Borel subsets of $\mathbb{R}$. Conversely, for any $a, b \in \mathbb{R}$, we note that $[a, b) = [a, b+1] \setminus [b, b+1]$ is a Borel subset from the above definition, so Borel subsets from the above definition are indeed Borel subsets of $\mathbb{R}$.

We quickly note that we have the following uniqueness result.

Theorem 6.49 (Haar). Fix a locally compact Hausdorff topological group G. Then there is a (nonzero) Borel measure, unique up to scaling, which is finite on compact subsets of X and invariant under left-translation.

In some sense, the above result explains Remark 6.46.

Remark 6.50. On the homework, we construct the Haar measure on the circle group S^1.

In fact, we have the following converse to Theorem 6.49.

Theorem 6.51 (Weil). Fix a group G and a σ-ring $\mathcal{S}$ on G equipped with a σ-finite measure μ and some extra separating property. Given that both $\mathcal{S}$ and μ are suitably translation-invariant, there is a topology $\mathcal{T}$ on G making G into a locally compact Hausdorff topological group where μ is a Haar measure for G.

Despite all our work, it's not even obvious which sets are Lebesgue-measurable or even that there are sets which are not Lebesgue-measurable. We will be able to answer at least this second question in the negative next class.

6.5 October 26

Today we explain why we keep marking our sets as being measurable.

6.5.1 A Non-measurable Set

Here is our result.

Exercise 6.52 (Vitali). Let $T = \mathbb{R}/\mathbb{Z}$ be the circle group, and let μ be the translation-invariant measure on $\mathbb{R}/\mathbb{Z}$ with $\mu(T) = 1$. It turns out that μ is complete. We produce a subset of T which is not μ-measurable.

Proof. Let T_{tors} be the torsion subgroup of T. Namely, $r \in T_{\text{tors}}$ if and only if there exists some $n \in \mathbb{Z}_{>0}$ for which $nr = 0$ in T, which means that $nr = k$ in $\mathbb{R}$ for some integer k and so $r = k/n$. Thus, $T_{\text{tors}} = \mathbb{Q}/\mathbb{Z}$; the important point is that T_{tors} is countable.

Now, for each coset in T/T_{tors}, let $V \subseteq T$ be a set of representatives of these cosets.[1] In particular, it follows that

$$T = \bigsqcup_{q \in T_{\text{tors}}} (q + V).$$

Indeed, there are two checks.

- To see the union, for any $r \in T$, we see that $r \in x + T_{\text{tors}}$ for some $x \in V$, so $r = xq$ for some $q \in T_{\text{tors}}$, so $r \in qV$.
- To see that the union is disjoint, suppose $q_1 + V = q_2 + V$. Then we can find $r_1, r_2 \in V$ such that $q_1 + r_1 = q_2 + r_2$. It follows that $r_1 = r_2 + (q_2 - q_1) \in r_2 + T_{\text{tors}}$, so $r_1 + T_{\text{tors}} = r_2 + T_{\text{tors}}$, so $r_1 = r_2$ because V is made of representatives of T/T_{tors}. Thus, $q_1 + r_1 = q_2 + r_2$ has forced $q_1 = q_2$.

[1] Note that we have used the Axiom of Choice here.

We are now ready to complete the proof. Suppose for the sake of contradiction that V is measurable. It follows that

$$1 = \mu(T) = \mu\left(\bigsqcup_{q \in T_{\mathrm{tors}}} (x + V)\right) = \sum_{q \in T_{\mathrm{tors}}} \mu(q + V) \overset{*}{=} \sum_{q \in T_{\mathrm{tors}}} \mu(V).$$

Note that we have used the translation-invariance of μ in $\overset{*}{=}$. However, this is impossible: if $\mu(V) > 0$, then the rightmost sum does not converge, and if $\mu(V) = 0$, then the rightmost sum vanishes, so it is impossible for the sum to actually equal 1. ■

> **Remark 6.53.** The above proof used the Axiom of Choice to construct V. It is a result of Solovay that there are models of the real numbers where all subsets are measurable. Of course, the model does not include the Axiom of Choice.

THEME 7

MEASURABLE FUNCTIONS

7.1 October 26

We now transition to integration. Here is a warning about our exposition.

Warning 7.1. We are going to do integration valued in general Banach spaces instead of just $\mathbb{R} \cup \{\infty\}$.

The above convention is non-standard. See [Lan12] for perhaps another treatment along these lines, but Professor Rieffel doesn't like Lang's exposition.

Example 7.2. The Banach spaces we care about will essentially all be $\mathbb{R}^n$ or $\mathbb{C}^n$ for some integer n.

Example 7.3. We may also use any completion of a normed vector space $(V, \|\cdot\|)$, such as the p-adic rationals $\mathbb{Q}_p$ or $C([0,1])$ using the p-norm $\|\cdot\|_p$.

7.1.1 Simple Measurable Functions

Let's begin with the easiest possible functions we might hope to integrate.

Definition 7.4 (Simple measurable function)**.** Fix a ring $\mathcal{S}$ on a set X and a normed vector space B. Then a *simple $\mathcal{S}$-measurable B-valued function* is a function $f\colon X \to B$ such that $\operatorname{im} f$ is finite and $f^{-1}(\{y\}) \in \mathcal{S}$ for any $y \in B \setminus \{0\}$.

Remark 7.5. It is possible to tell a lot of this story by allowing B to be any metric space with a chosen point $0 \in B$. (In other words, we may allow B to be a "pointed metric space.") Professor Rieffel made some comments about this, but I do not think that keeping track of this is particularly important.

Example 7.6. Given any $y \in B$ and $E \in \mathcal{S}$, the function $y1_E$ is a simple $\mathcal{S}$-measurable function. For one, the image is $\{0, y\}$, which is finite. Further, if $b \in B \setminus \{0\}$, then either $b \neq y$ as well and so $f^{-1}(\{b\}) = \varnothing \in \mathcal{S}$ or $f^{-1}(\{y\}) = E \in \mathcal{S}$.

As from the above example, it turns out that we should think of simple measurable functions as just linear combinations of indicators.

Lemma 7.7. Fix a ring $\mathcal{S}$ on a set X and a normed vector space B. Then any simple $\mathcal{S}$-measurable function $f\colon X \to S$ can be written as

$$f = \sum_{y\in(\operatorname{im} f)\setminus\{0\}} y1_{f^{-1}(\{y\})}.$$

Proof. Fix any $x_0 \in X$, and we want to show that

$$f(x_0) = \sum_{y\in(\operatorname{im} f)\setminus\{0\}} y1_{f^{-1}(\{y\})}(x_0).$$

Well, if $f(x_0)=0$, then note that $x_0 \notin f^{-1}(\{y\})$ for any $y \in (\operatorname{im} f)\setminus\{0\}$, so the right-hand sum vanishes.

Otherwise, say that $f(x_0)=y_0$ where $y_0 \in (\operatorname{im} f)\setminus\{0\}$. Then note that $x_0 \in f^{-1}(\{y_0\})$, and further we see that $x_0 \in f^{-1}(\{y\})$ forces $y = f(x_0) = y_0$, so y_0 is the only y for which $x_0 \in f^{-1}(\{y_0\})$, so

$$\sum_{y\in(\operatorname{im} f)\setminus\{0\}} y1_{f^{-1}(\{y\})}(x_0) = y_0 1_{f^{-1}(\{y_0\})}(x_0) = y_0 = f(x_0),$$

which is what we wanted. ■

Lemma 7.8. Fix a ring $\mathcal{S}$ on a set X and a normed vector space B. Then any simple $\mathcal{S}$-measurable function $f\colon X \to S$ can be written as

$$f = \sum_{i=1}^n y_i 1_{E_i}$$

where the $y_i \in B$ are distinct and nonzero and the $E_i \in \mathcal{S}$ are pairwise disjoint and nonempty. In fact, we must have $\{y_1,\ldots,y_n\} = (\operatorname{im} f)\setminus\{0\}$ and $E_i = f^{-1}(\{y_i\})$.

Proof. We show the claims in sequence.

- Existence: by Lemma 7.7, we can write

  $$f = \sum_{y\in(\operatorname{im} f)\setminus\{0\}} y1_{f^{-1}(\{y\})}.$$

  Here, the elements of $(\operatorname{im} f)\setminus\{0\}$ are surely distinct, and there are finitely many of them, so we enumerate them by $\{y_1,\ldots,y_n\}$. Then we set $E_i := f^{-1}(\{y_i\})$, which is in $\mathcal{S}$ by hypothesis on f.

  Lastly, we note that the $E_\bullet$ are pairwise disjoint: if $x \in E_i$, then $f(x) = y_i$, so if $x \in E_i \cap E_j$, then $f(x) = y_i = y_j$, so $y_i = y_j$, so $i = j$ because the $y_\bullet$ are distinct.

- Uniqueness: suppose we can write

  $$f = \sum_{i=1}^n y_i 1_{E_i}$$

  where the $y_i \in B$ are distinct and nonzero, and the $E_i \in \mathcal{S}$ are pairwise disjoint. We claim that $\{y_1,\ldots,y_n\} = (\operatorname{im} f)\setminus\{0\}$ and $E_i = f^{-1}(\{y_i\})$ for each i.

  Certainly $\{y_1,\ldots,y_n\} \subseteq (\operatorname{im} f)\setminus\{0\}$. Indeed, if $x \in E_i$, then

  $$f(x) = \sum_{i=1}^n y_i 1_{E_i}(x) = y_i 1_{E_i}(x) = y_i$$

  because the $E_\bullet$ are pairwise disjoint, so $y_i \in (\operatorname{im} f)\setminus\{0\}$. In fact, observe that we have also shown that $x \in E_i$ implies $f(x) = y_i$.

Conversely, if $y \in (\operatorname{im} f) \setminus \{0\}$, then find $x \in X$ with $f(x) = y$. Because $f(x) \neq 0$, some term in the sum of

$$f(x) = \sum_{i=1}^{n} y_i 1_{E_i}(x)$$

must be nonzero, so say that $y_i 1_{E_i}(x) \neq 0$, so $x \in E_i$. However, $x \in E_i$ now forces $f(x) = y_i$ as we saw above, so $y \in \{y_1, \ldots, y_n\}$. In fact, observe that we have also shown that $f(x) \neq 0$ implies $x \in E_j$ for some j.

It remains to show that $E_i = f^{-1}(\{y_i\})$. Well, above we showed that $x \in E_i$ implies $f(x) = y_i$. Conversely, we showed that $f(x) = y_i \neq 0$ implies that $x \in E_j$ for some j. But then $f(x) = y_j$ from the above, so $y_j = y_i$, so $i = j$ because the y_i are distinct, so $x \in E_i$. ■

Here's a sanity check.

Lemma 7.9. Fix a ring $\mathcal{S}$ on a set X and a normed k-vector space B. Then the simple $\mathcal{S}$-measurable functions valued in B form a k-vector space.

Proof. We know that the set of all functions $X \to B$ forms a k-vector space under the pointwise operations, so we just need to check that we form a subspace. Here are those checks.

- Scalar multiplication: suppose f is a simple $\mathcal{S}$-measurable function, and let $r \in k$, and we show rf is still a simple $\mathcal{S}$-measurable function. Well, if $r = 0$, then $rf = 0$, so $rf = 0 \cdot 1_\varnothing$ is a simple $\mathcal{S}$-measurable function by Example 7.6.

  Otherwise, take $r \neq 0$. For one, note that

  $$\operatorname{im}(rf) = \{rf(x) : x \in X\} = \{ry : y \in \operatorname{im} f\}$$

  is still finite, with cardinality upper-bounded by $\#(\operatorname{im} f)$.

  Further, we need to show that $y \in B \setminus \{0\}$ will have $f^{-1}(\{y\}) \in \mathcal{S}$. Well, we compute

  $$(rf)^{-1}(\{y\}) = \{x \in X : rf(x) = y\} = \{x \in X : f(x) = 1/r \cdot y\} = f^{-1}(\{1/r \cdot y\}),$$

  where we are using the fact that $r \neq 0$. Because $y \neq 0$, we see $1/r \cdot y \neq 0$, so $f^{-1}(\{1/r \cdot y\}) \in \mathcal{S}$ still.

- Addition: suppose f and g are simple $\mathcal{S}$-measurable so that we want to show $f + g$ is still a simple $\mathcal{S}$-measurable function. Indeed, we claim that

  $$\operatorname{im}(f+g) \subseteq \{b + c : b \in \operatorname{im} f \text{ and } c \in \operatorname{im} g\}.$$

  To see this note that any element of $\operatorname{im}(f+g)$ can be written as $(f+g)(x) = f(x) + g(x)$, which does take the form $b + c$ where $b = f(x) \in \operatorname{im} f$ and $c = g(x) \in \operatorname{im} g$. Thus, we do indeed see that $\operatorname{im}(f+g)$ is finite, with cardinality at most $\#(\operatorname{im} f) \cdot \#(\operatorname{im} g)$.

  Now, suppose that $y \in B \setminus \{0\}$, and we show $(f+g)^{-1}(\{y\}) \in \mathcal{S}$. Indeed, we see that $f(x) + g(x) = y$ is equivalent to $f(x) = y - g(x)$, so

  $$(f+g)^{-1}(\{y\}) = \bigcup_{c \in (\operatorname{im} g)} \left(f^{-1}(\{y - c\}) \cap g^{-1}(\{c\})\right).$$

  Because σ-rings are closed under finite unions, it suffices to show that $f^{-1}(\{y-c\}) \cap g^{-1}(\{c\}) \in \mathcal{S}$ for each $c \in \operatorname{im} g$. We have three cases.

  - If $y \neq c$ and $c \neq 0$, then we see that $f^{-1}(\{y - c\}), g^{-1}(\{c\}) \in \mathcal{S}$, so their intersection remains in $\mathcal{S}$ by Remark 5.24.

- If $y = c$, then note that $c = y \neq 0$. Here, we are showing $f^{-1}(\{0\}) \cap g^{-1}(\{y\}) \in \mathcal{S}$. Well, $f^{-1}(\{0\}) \cap g^{-1}(\{y\})$ is

$$\left(X \Big\backslash \bigcup_{b \in (\operatorname{im} f)\setminus\{0\}} f^{-1}(\{b\})\right) \cap g^{-1}(\{y\}) = g^{-1}(\{y\}) \Big\backslash \bigcup_{b \in (\operatorname{im} f)\setminus\{0\}} f^{-1}(\{b\}).$$

  Well, $f^{-1}(\{b\}) \in \mathcal{S}$ for each of the finitely many $b \in (\operatorname{im} f)\setminus\{0\}$, so the full union lives in $\mathcal{S}$ because $\mathcal{S}$ is a σ-ring. Lastly, the subtraction still lives in $\mathcal{S}$ because $g^{-1}(\{y\}) \in \mathcal{S}$, and $\mathcal{S}$ is still a σ-ring.
- If $c = 0$, then we still have $y \neq 0$. Here, we are showing that $f^{-1}(\{y\}) \cap g^{-1}(\{0\}) \in \mathcal{S}$, so we may just reverse the roles of f and g in the above case to finish.

The above cases finish the proof. ■

Corollary 7.10. Fix a ring $\mathcal{S}$ on a set X and a normed vector space B. For any sets $\{E_i\}_{i=1}^n \subseteq \mathcal{S}$ and outputs $\{y_i\}_{i=1}^n \in B$, the function

$$\sum_{i=1}^n y_i 1_{E_i}$$

is a simple $\mathcal{S}$-measurable function.

Proof. Note that each $y_i 1_{E_i}$ is a simple $\mathcal{S}$-measurable function by Example 7.6, so the finite sum of these remains a simple $\mathcal{S}$-measurable function by Lemma 7.9. ■

Lemma 7.11. Fix a ring $\mathcal{S}$ on a set X and a normed vector space $(B, \|\cdot\|)$. If f is a simple $\mathcal{S}$-measurable function, then the function $x \mapsto \|f(x)\|$ is as well.

Proof. Set $g(x) := \|f(x)\|$ to be a function $g\colon X \to \mathbb{R}$. Because $\operatorname{im} f$ is finite, it follows that $\operatorname{im} g = \{\|y\| : y \in \operatorname{im} f\}$ is still finite, so it remains to check our fibers. Fix some $r \in (\operatorname{im} g) \setminus \{0\}$. Because $\operatorname{im} f$ is finite, we note that

$$B_r := \{y \in \operatorname{im} f : \|y\| = r\}$$

is still finite; notably, each $y \in B_r$ is nonzero because r is nonzero. Now,

$$\begin{aligned} g^{-1}(\{r\}) &= \{x \in X : \|f(x)\| = r\} \\ &= \bigcup_{y \in (\operatorname{im} f)} \{x \in X : f(x) = y \text{ and } \|y\| = r\} \\ &= \bigcup_{y \in B_r} \{x \in X : f(x) = y\} \\ &= \bigcup_{y \in B_r} f^{-1}(\{y\}). \end{aligned}$$

Thus, $g^{-1}(\{y\})$ is the finite union of sets of the form $f^{-1}(\{y\})$ with $y \neq 0$, which are in $\mathcal{S}$ by definition of f. In particular, $g^{-1}(\{y\}) \in \mathcal{S}$ as well. ■

Lemma 7.12. Fix a ring $\mathcal{S}$ on a set X. Given two simple $\mathcal{S}$-measurable functions $f, g\colon X \to \mathbb{R}$, the function fg is simple $\mathcal{S}$-measurable.

Proof. By Lemma 7.8, we may write

$$f = \sum_{i=1}^{m} a_i 1_{E_i} \qquad \text{and} \qquad g = \sum_{j=1}^{n} b_i 1_{F_i}$$

for some $\{a_i\}_{i=1}^m, \{b_j\}_{j=1}^n \subseteq \mathbb{R}$ and $\{E_i\}_{i=1}^m, \{F_j\}_{j=1}^n \subseteq \mathcal{S}$. Thus, we can write

$$(fg) = \sum_{i=1}^{m}\sum_{j=1}^{n}(a_ib_j)1_{E_i}1_{F_j},$$

but $1_{E_i}1_{F_j} = 1_{E_i \cap F_j}$, so

$$(fg) = \sum_{i=1}^{m}\sum_{j=1}^{n}(a_ib_j)1_{E_i\cap F_j}.$$

Now, each $E_i \cap F_j$ lives in $\mathcal{S}$ because $\mathcal{S}$ is a ring, so fg is simple $\mathcal{S}$-measurable by Corollary 7.10. ■

7.1.2 Simple Integrable Functions

We are finally ready to define integrals.

Definition 7.13 (Simple integrable function)**.** Fix a ring $\mathcal{S}$ on a set X and a metric space B. Further, let μ be a finitely additive measure μ on $\mathcal{S}$. Then a function $f\colon X \to B$ is a *simple $\mathcal{S}$-integrable function* if and only if $\operatorname{im} f$ is finite, and $f^{-1}(\{y\}) \in \mathcal{S}$ has finite measure for each $y \in (\operatorname{im} f) \setminus \{0\}$.

Remark 7.14. In fact, if $f\colon X \to B$ is a simple μ-integrable function, for any subset $E \subseteq B \setminus \{0\}$, we see

$$f^{-1}(E) = \bigsqcup_{y \in (E \cap \operatorname{im} f)\setminus\{0\}} f^{-1}(\{y\}),$$

where the union is disjoint because $x \in f^{-1}(\{y\}) \cap f^{-1}(\{y'\})$ implies $f(y) = x = f(y')$. As such, finite additivity of μ implies

$$\mu\left(f^{-1}(E)\right) = \sum_{y \in (E \cap \operatorname{im} f)\setminus\{0\}} \mu\left(f^{-1}(\{y\})\right)$$

is a finite sum of finite numbers and is thus finite.

Definition 7.15 (Integral)**.** Fix a ring $\mathcal{S}$ on a set X and a metric space B. Further, let μ be a finitely additive measure μ on $\mathcal{S}$. Given a simple μ-integrable function f, we define the *integral*

$$\int_X f\,d\mu := \sum_{y \in (\operatorname{im} f)\setminus\{0\}} \mu\left(f^{-1}(\{y\})\right) y.$$

Note this is a finite sum with $\mu\left(f^{-1}(\{y\})\right)$ finite, so $\int_X f\,d\mu$ is finite.

Example 7.16. Given some $E \in \mathcal{S}$ with $\mu(E) < \infty$, we note $\int_X 1_E\,d\mu = \mu(E)$. This function is simple μ-integrable: $\operatorname{im}(1_E) = \{0, 1\}$, and $1_E^{-1}(\{1\}) = E$ has $\mu(E) < \infty$. Thus,

$$\int_X 1_E\,d\mu = 1\mu(E) = \mu(E),$$

as desired.

Note that the sum in the above definition is a finite sum by definition of f, and each $f^{-1}(\{y\})$ is also in $\mathcal{S}$ by definition of f again.

Here are the usual sanity checks once we've defined some functions.

Lemma 7.17. Fix a ring $\mathcal{S}$ on a set X equipped with a finitely additive measure μ and a normed k-vector space B. Then the simple μ-integrable functions valued in B form a k-vector space.

Proof. By definition, note that simple μ-integrable functions are also simple $\mathcal{S}$-measurable functions, so it suffices to show that we form a k-subspace of the space of simple $\mathcal{S}$-measurable functions (see Lemma 7.9). We use ideas from Lemma 7.9 to run our checks.

- Scalar multiplication: fix a simple μ-integrable function f, and let $r \in k$, and we show rf is still a simple μ-integrable function. As usual, $r = 0$ gives $rf = 0$, so checking $f^{-1}(\{y\}) \in \mathcal{S}$ with finite measure for $y \in (\operatorname{im} f) \setminus \{0\}$ is vacuous.

  Otherwise, we have $r \neq 0$. We showed in Lemma 7.9 that any $y \in (\operatorname{im} rf) \setminus \{0\}$ will have
  $$(rf)^{-1}(\{y\}) = f^{-1}(\{1/r \cdot y\}).$$
  Thus, if $y \in (\operatorname{im} rf) \setminus \{0\}$ so that the left-hand side is nonempty, then $1/r \cdot y \in (\operatorname{im} f) \setminus \{0\}$ as well; notably, $1/r \cdot y = 0$ would force $y = 0$, so we must have $1/r \cdot y \neq 0$. Now, $f^{-1}(\{1/r \cdot y\}) \in \mathcal{S}$ has finite measure by hypothesis on f, so $(rf)^{-1}(\{y\})$ has finite measure as well.

- Addition: suppose f and g are simple μ-integrable functions so that we want to show $f + g$ is still a simple μ-integrable function. We showed in Lemma 7.9 that any $y \in B \setminus \{0\}$ will have
  $$(f+g)^{-1}(\{y\}) = \bigcup_{c \in (\operatorname{im} g)} \left(f^{-1}(\{y - c\}) \cap g^{-1}(\{c\})\right).$$
  In particular, $(f + g)^{-1}(\{y\}) \in \mathcal{S}$ as we discussed in Lemma 7.9, and then Lemma 5.55 tells us that
  $$\mu\left((f+g)^{-1}(\{y\})\right) \le \sum_{c \in (\operatorname{im} g)} \mu\left(g^{-1}(\{c\})\right),$$
  which is a finite sum of finite real numbers and therefore finite. It follows that $f + g$ is in fact a simple μ-integrable function. ■

Lemma 7.18. Fix a ring $\mathcal{S}$ on a set X equipped with a finitely additive measure μ and a normed vector space $(B, \|\cdot\|)$. If f is a simple μ-integrable function, then $x \mapsto \|f(x)\|$ is also a simple μ-integrable function.

Proof. As before, we continue from Lemma 7.11. Namely, we set $g(x) := \|f(x)\|$, and we know that g is already a simple $\mathcal{S}$-measurable function.

Thus, we pick up any $r \in (\operatorname{im} g) \setminus \{0\}$, and we need to show that $g^{-1}(\{r\})$ has finite measure. Well, in Lemma 7.11, we defined
$$B_r := \{y \in \operatorname{im} f : \|y\| = r\}$$
and showed that
$$g^{-1}(\{r\}) = \bigcup_{y \in B_r} f^{-1}(\{y\}).$$
Now, each $f^{-1}(\{y\})$ has finite measure by hypothesis on f, so the total finite union $g^{-1}(\{r\})$ will have finite measure by Lemma 5.55. ■

7.2 October 28

We continue our story with integration by defining what we mean by a measurable function.

7.2.1 Measurable Functions

The following definition is non-standard but is how to think about our integrals in practice.

Definition 7.19 (Measurable function)**.** Fix a set X and a σ-ring $\mathcal{S}$ on X. Given a normed vector space B, an $\mathcal{S}$*-measurable function* is a function $f\colon X \to B$ such that there is a sequence of simple $\mathcal{S}$-measurable functions $\{f_n\}_{n\in\mathbb{N}}$ which converge to f pointwise.

Remark 7.20. Later in life, when we take $B = \mathbb{R}$, we will allow the functions f_n to output at ∞, but we will not do so while we allow B to be a normed vector space.

Sometimes we won't converge "on the nose," so we will want a little freedom.

Definition 7.21 (Null set)**.** Fix a set X and a σ-ring $\mathcal{S}$ on X equipped with a measure μ. A *null set* is a subset $N \subseteq X$ such that there is some $E \in \mathcal{S}$ such that $N \subseteq E$ and $\mu(N) = 0$.

Definition 7.22 (Almost everywhere)**.** Fix a set X and a σ-ring $\mathcal{S}$ on X equipped with a measure μ. A property $P(x)$ for points $x \in X$ holds *almost everywhere* if and only if $\{x \in X : \neg P(x)\}$ is a null set.

Definition 7.23 (Converges almost everywhere)**.** Fix a set X and a σ-ring $\mathcal{S}$ on X equipped with a measure μ. Given a metric space B, a sequence of functions $f_n\colon X \to B$ with $n \in \mathbb{N}$ converges to a function $f\colon X \to B$ *almost everywhere* if and only if $f_n(x) \to f(x)$ almost everywhere.

Definition 7.24 (Measurable function)**.** Fix a set X and a σ-ring $\mathcal{S}$ on X equipped with a measure μ. Given a metric space B, a μ*-measurable function* is a function $f\colon X \to B$ such that there is a sequence of simple $\mathcal{S}$-measurable functions $\{f_n\}_{n\in\mathbb{N}}$ which converge to f almost everywhere.

Here is the usual sanity check.

Lemma 7.25. Fix a normed k-vector space B and a set X with a σ-ring $\mathcal{S}$ on X. Then the set of all $\mathcal{S}$-measurable functions forms a k-vector space under pointwise operations.

Proof. We already know that the set of all functions $X \to B$ will form a k-vector space under the pointwise operations, so we just need to show that we have a subspace. Well, pick up $\mathcal{S}$-measurable functions f and g and some scalars $a, b \in k$. We show that $h := af + bg$ is still $\mathcal{S}$-measurable.

Well, f being $\mathcal{S}$-measurable promises simple $\mathcal{S}$-measurable functions $\{f_n\}_{n\in\mathbb{N}}$ with $f_n \to f$ pointwise; similarly, we get simple $\mathcal{S}$-measurable functions $\{g_n\}_{n\in\mathbb{N}}$ with $g_n \to g$ pointwise. Now, we define

$$h_n := af_n + bg_n,$$

which is a simple $\mathcal{S}$-measurable function by Lemma 7.9.

It remains to check that $h_n \to h$ as $n \to \infty$. Let $|\cdot|$ be the norm on k, and let $\|\cdot\|$ be the norm on B, and fix some $x \in X$. Now, for any $\varepsilon > 0$, find $N_f > 0$ such that

$$n > N_f \implies \|f(x) - f_n(x)\| < \frac{\varepsilon}{2(\|a\| + 1)},$$

where we note that $|a| + 1 > 0$ makes this division legal. Similarly, we find $N_g > 0$ such that

$$n > N_f \implies \|f(x) - f_n(x)\| < \frac{\varepsilon}{2(\|b\| + 1)}.$$

Thus, $n > \max\{N_f, N_g\}$ will have

$$\|h(x) - h_n(x)\| \le |a| \cdot \|f(x) - f_n(x)\| + |b| \cdot \|g(x) - g_n(x)\| < |a| \cdot \frac{\varepsilon}{2(\|a\| + 1)} + |b| \cdot \frac{\varepsilon}{2(\|b\| + 1)} < \varepsilon,$$

which finishes. ■

Lemma 7.26. Fix a ring $\mathcal{S}$ on a set X. Given two $\mathcal{S}$-measurable functions $f, g\colon X \to \mathbb{R}$, the function fg is $\mathcal{S}$-measurable.

Proof. We are given sequences of simple $\mathcal{S}$-measurable functions $\{f_n\}_{n\in\mathbb{N}}$ and $\{g_n\}_{n\in\mathbb{N}}$ such that $f_n \to f$ and $g_n \to g$ pointwise. Thus, for each $x \in X$, we see $(f_n g_n)(x) \to (fg)(x)$ by taking products of limits, so we conclude $f_n g_n \to fg$ pointwise. However, $f_n g_n$ is simple $\mathcal{S}$-measurable by Lemma 7.12. ■

7.2.2 Properties of Simple Measurable Functions

Something annoying about our definition is that we can only work simple $\mathcal{S}$-measurable functions "directly." One might wonder, for example, if a looking at limits $f_n \to f$ as $n \to \infty$ where each f_n is $\mathcal{S}$-measurable might give a function f which is not $\mathcal{S}$-measurable. This turns out to not be the case, but it will take some work to prove.

In particular, we will want a better description of $\mathcal{S}$-measurable functions. For today, we will content ourselves with necessary conditions.

Definition 7.27 (Separable)**.** A topological space M is *separable* if and only if there is a countable dense subset of M. As such, a subset $A \subseteq M$ is *separable* if and only if A is separable with the restricted metric; in other words, $A \subseteq M$ is separable if and only if there is a countable subset $B \subseteq A$ such that $A \subseteq \overline{B}$.

Example 7.28. If $A \subseteq M$ is countable, then we can see that $A \subseteq \overline{A}$ by definition of the closure, so A is a countable dense subset with $A \subseteq \overline{A}$.

Here is a quick sanity check.

Lemma 7.29. Fix a metric space (M, d). A subset $A \subseteq M$ is separable if and only if there is a countable subset $B \subseteq M$ such that $A \subseteq \overline{B}$.

Proof. In the forward direction, having a countable subset $B \subseteq A$ with $A \subseteq \overline{B}$ will certainly imply having a countable subset $B \subseteq M$ with $A \subseteq \overline{B}$.

In the reverse direction, we begin with a countable subset $B \subseteq M$ with $A \subseteq \overline{B}$. For now, fix some $\varepsilon > 0$. Then each $a \in A$ has $B(a, \varepsilon/2) \cap B \neq \varnothing$ by Lemma 2.57, so choose some $b_a \in B$ with $d(a, b_{\varepsilon,a}) < \varepsilon/2$. Now, the subset

$$B_\varepsilon := \{b_{\varepsilon,a} : a \in A\} \subseteq B$$

must be countable, so enumerate its elements by $B_\varepsilon = \{b_{\varepsilon,1}, b_{\varepsilon,2}, \ldots\}$, and for each $b_{\varepsilon,k}$, we select some $a_{\varepsilon,k} \in A$ such that $d(b_{\varepsilon,k}, a_{\varepsilon,k}) < \varepsilon/2$, which exists by construction of B_ε.

We now go back to letting $\varepsilon > 0$ vary. As our countable subset, we now set

$$B' := \bigcup_{n=1}^{\infty} \{a_{1/n,k} : k \in \mathbb{Z}_{>0}\}.$$

Indeed, we claim that $A \subseteq \overline{B'}$, which shows density by Lemma 2.57. For this, we pick up any $a \in A$ and $\varepsilon > 0$, and we show that $B(a, \varepsilon) \cap B' \neq \varnothing$. Well, find some N with $N > 1/\varepsilon$. By construction of $B_{1/N}$, we may find some k with $b_{1/N,k} = b_{1/N,a}$, which means that

$$d(a, a_{1/N,k}) \leq d_(a, b_{1/N,k}) + d(b_{1/N,k}, a_{1/N,k}) < \frac{1}{2N} + \frac{1}{2N} < \varepsilon.$$

Thus, $a_{1/N,k} \in B'$ is the element we are looking for. ■

Example 7.30. We give $\mathbb{R}$ the usual metric. Then any subset $A \subseteq \mathbb{R}$ is separable: set $B := \mathbb{Q}$. Then $\overline{B} = \mathbb{R}$ contains A, but B is countable, so we are done by Lemma 7.29.

Remark 7.31. To help our intuition that this should be a smallness condition, we note that if M is a separable space, then any subspace $A \subseteq M$ is still separable. Indeed, there is some countable subset $B \subseteq M$ with $\overline{B} = M$, so $A \subseteq \overline{M}$ follows.

Example 7.32. Given countably many separable subsets $\{A_n\}_{n \in \mathbb{N}}$ of a metric space B, the union $A := \bigcup_{n \in \mathbb{N}} A_n$ is separable. Indeed, each A_n has a countable $B_n \subseteq M$ with $A_n \subseteq \overline{B_n}$.

Now, set $B := \bigcup_{n \in \mathbb{N}} B_n$, which is countable; we claim that $A \subseteq \overline{B}$, which will finish. Because $B_n \subseteq B \subseteq \overline{B}$, we see that $\overline{B}$ is a closed subset containing B_n, so $A_n \subseteq \overline{B_n} \subseteq \overline{B}$ follows. Thus, $A \subseteq \overline{B}$.

Here is why we just defined separable subsets.

Lemma 7.33. Fix a normed vector space B and a set X with σ-ring $\mathcal{S}$ on X. Any simple $\mathcal{S}$-measurable function $f \colon X \to B$ has $\operatorname{im} f \subseteq B$ separable.

Proof. By definition of simple measurable functions, $\operatorname{im} f$ is finite and hence separable by Example 7.28. ■

Here is a last moderately silly checks.

Lemma 7.34. Fix a normed vector space B and a set X with σ-ring $\mathcal{S}$ on X. Any simple $\mathcal{S}$-measurable function $f \colon X \to B$ has $f^{-1}(U \setminus \{0\}) \in \mathcal{S}$ for any open $U \subseteq X$.

Proof. Recall that $\operatorname{im} f$ is finite by definition, so enumerate $\operatorname{im} f$ by $(\operatorname{im} f) \cap (U \setminus \{0\}) = \{y_1, \ldots, y_n\}$. Then we note that

$$f^{-1}(U \setminus \{0\}) = f^{-1}((\operatorname{im} f) \cap (U \setminus \{0\})) = \bigcup_{k=1}^{n} f^{-1}(\{y_k\}).$$

However, $f^{-1}(\{y_k\}) \in \mathcal{S}$ for each k, so the total union lives in $\mathcal{S}$ because $\mathcal{S}$ is a ring. ■

7.2.3 Properties Preserved by Limits

Now, to upgrade from simple $\mathcal{S}$-measurable functions to $\mathcal{S}$-measurable functions, we take limits. Here is the separability check.

Lemma 7.35. Fix a metric space M and a set X. Suppose a sequence of functions $f_n \colon X \to B$ for $n \in \mathbb{N}$ have a pointwise limit $f_n \to f$ as $n \to \infty$. If each $\operatorname{im} f_n \subseteq M$ is separable, then $\operatorname{im} f \subseteq M$ is separable as well.

Proof. For each n, find the countable subset $C_n \subseteq \operatorname{im} f_n$ with $\operatorname{im} f_n \subseteq \overline{C_n}$. Then we set

$$C := \bigcup_{n\in\mathbb{N}} C_n,$$

and we note that C is the countable union of countable subsets and hence countable. We thus claim that $\operatorname{im} f \subseteq \overline{C}$, which will finish by Lemma 7.29.

Well, fix any $y \in \operatorname{im} f$, and find some $x \in X$ with $y = f(x)$. For any $\varepsilon > 0$, we need to show that $B(y,\varepsilon) \cap C \neq \varnothing$; this is enough by Lemma 2.57. For this, we note that there is some N such that $n \geq N$ implies

$$d(f(x), f_n(x)) < \varepsilon/2,$$

where d is the metric of M; set $n := N$. Further, we recall $\operatorname{im} f_n \subseteq \overline{C_n}$, so Lemma 2.57 promises us some $c \in C_n$ such that $d(f_n(x), c) < \varepsilon/2$. In total,

$$d(y,c) \leq d(f(x), f_n(x)) + d(f_n(x), c) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

which is what we wanted. ■

Corollary 7.36. Fix a normed vector space B and a set X with σ-ring $\mathcal{S}$ on X. Any $\mathcal{S}$-measurable function $f\colon X \to B$ has $\operatorname{im} f \subseteq B$ separable.

Proof. By definition of being $\mathcal{S}$-measurable, there is a sequence of simple $\mathcal{S}$-measurable functions $f_n\colon X \to B$ with $f_n \to f$ as $n \to \infty$ pointwise. Each f_n has $\operatorname{im} f_n$ separable by Lemma 7.33, so f does as well by Lemma 7.35. ■

Making Lemma 7.34 work in limits requires a little more care.

Lemma 7.37. Fix a normed vector space B and a set X with σ-ring $\mathcal{S}$ on X. Suppose that a sequence of functions $f_n\colon X \to B$ have $f_n^{-1}(U\setminus\{0\}) \in \mathcal{S}$ for each open $U \subseteq B$. Then satisfy the following conditions for each f_n. If $f_n \to f$ pointwise as $n \to \infty$, then $f^{-1}(U \setminus \{0\}) \in \mathcal{S}$ for each open $U \subseteq B$ as well.

Proof. This is a little tricky. We will replace U with $U \setminus \{0\}$ and simply remember that $0 \notin U$.

The main point is that any $x \in f^{-1}(U)$ will have $f(x) \in U$, and elements of U should have some small positive distance away from $B \setminus U$. Namely, we set

$$U_m := \{x \in U : d(x, B \setminus U) > 1/m\}$$

for any $m \geq 1$; here, d is the metric of B, and $d(x, B\setminus U) = \inf_{y\in B\setminus U} d(x,y)$. Here are a few checks.

- As an intermediate claim, we note that $d(x,y) + d(y, B\setminus U) \geq d(x, B\setminus U)$. Indeed, for any $a \in B \setminus U$, note that
$$d(x,y) + d(y,a) \geq d(b,a) \geq d(x, B\setminus U),$$
so
$$d(y,a) \geq d(x, B\setminus U) - d(x,y).$$
Letting $a \in B\setminus U$ vary in this last inequality tells us that $d(y, B\setminus U) \geq d(x, B\setminus U) - d(x,y)$.
- Note that each U_m is open. Indeed, if $x \in U_m$, then set $\varepsilon := 1/m - d(x, B\setminus U) > 0$. This means that $d(x,y) < \varepsilon$ implies, from the previous check,
$$d(y, B\setminus U) \leq d(x,y) + d(x, B\setminus U) < \varepsilon + d(x, B\setminus U) = 1/m,$$
so $y \in U_m$. Thus, $B(x,\varepsilon) \subseteq U$.

- We claim
$$U \stackrel{?}{=} \bigcup_{m=1}^{\infty} U_m.$$
In one direction, $x \in U_m$ implies $d(x, B \setminus U) \neq 0$, so $x \notin B \setminus U$, so $x \in U$.
In the other direction, note $x \in U$ implies there is some $\varepsilon > 0$ with $B(x,\varepsilon) \subseteq U$, so $B(x,\varepsilon)\cap(B\setminus U) = \varnothing$, so $d(x, B \setminus U) > \varepsilon$. Thus, there is some $m > 1/\varepsilon$ with $d(x, B \setminus U) > 1/m$ and so $x \in U_m$ for this m.
- From the above claim, we note that $0 \notin U_m$ for each m because $0 \notin U$.
- We claim $\overline{U_m} \subseteq U_{m+1}$ for each m; set $\varepsilon := \frac{1}{m} - \frac{1}{m+1} > 0$. Now, if $x \in \overline{U_m}$, then we see $B(x,\varepsilon)\cap U_m \neq \varnothing$ by Lemma 2.57, so we may find $y \in U_m$ with $d(x,y) < \varepsilon$. It follows from the first check that
$$d(x, B \setminus U) \le d(x,y) + d(y, B \setminus U) < \varepsilon + \frac{1}{m} = \frac{1}{m+1},$$
so $x \in U_{m+1}$ follows.

Now, we see from the above that
$$f^{-1}(U) = \bigcup_{m=1}^{\infty} f^{-1}(U_m).$$
Thus, $x \in f^{-1}(U_m)$ implies that there is some $\varepsilon > 0$ with $B(x,\varepsilon) \subseteq U_m$; because $f_n \to f$ as $n \to \infty$ pointwise (!), there is some N for which $f_n(x) \in B(x,\varepsilon) \subseteq U_m$ for each $n \ge N$, so
$$f^{-1}(U) \subseteq \bigcup_{m=1}^{\infty} \bigcup_{N=1}^{\infty} \bigcap_{n \ge N} f_n^{-1}(U_m).$$
Conversely, if x lives in this right-hand set, we have some m and N with $f_n(x) \in U_m \subseteq \overline{U_m}$ for all $n \ge N$. So $f(x) \in \overline{U_m}$ by Lemma 2.50, so $f(x) \in U_{m+1} \subseteq U$ follows. Thus, equality in the above containment follows.

In total, we see that
$$f^{-1}(U) = \bigcup_{m=1}^{\infty} \bigcup_{N=1}^{\infty} \bigcap_{n \ge N} f_n^{-1}(U_m).$$
Notably, $f_n^{-1}(U_m) \cap E \in \mathcal{S}$ for each n and m, by construction of the f_ns, so this full union of unions of intersections is still in $\mathcal{S}$, using the fact that $\mathcal{S}$ is a σ-ring and Remark 5.24. ■

Corollary 7.38. Fix a normed vector space B and a set X with σ-ring $\mathcal{S}$ on X. Any $\mathcal{S}$-measurable function $f\colon X \to B$ has $f^{-1}(U \setminus \{0\}) \in \mathcal{S}$ for each open $U \subseteq B$.

Proof. Any simple $\mathcal{S}$-measurable function satisfies the conclusion by Lemma 7.34. However, because $\mathcal{S}$-measurable functions are limits of simple $\mathcal{S}$-measurable functions, $\mathcal{S}$-measurable functions satisfy the conclusion as well by Lemma 7.37. ■

Remark 7.39. Note that the case of $B = \mathbb{R}$, we see that $f^{-1}(U)$ is measurable for any open $U \subseteq \mathbb{R}$, where f is an $\mathcal{S}$-measurable function. By taking unions and complements appropriately, we in fact see that $f^{-1}(U)$ is measurable for any Borel set $U \subseteq \mathbb{R}$. This is the usual definition of a (Borel) measurable function $X \to \mathbb{R}$, and we will show it is equivalent to the one we gave next class.

7.3 October 31

We continue our discussion of measurable functions by giving an alternate definition.

7.3.1 A Better Measurable

Last class, we saw that measurable functions have some nice properties. Today we show that these properties actually characterize our measurable functions.

Theorem 7.40. Fix a normed vector space B and a set X with a σ-ring $\mathcal{S}$ on X. Then a function $f\colon X \to B$ is $\mathcal{S}$-measurable if and only if

(i) $\operatorname{im} f$ is separable, and

(ii) for any open $U \subseteq B$, we have $f^{-1}(U \setminus \{0\}) \in \mathcal{S}$.

Remark 7.41. Using ideas of Proposition 2.31, it suffices to check (ii) on a sub-base for the topology on B. In particular, it suffices to check (ii) on open balls.

Proof. Last class we provided the forward direction; namely, (i) follows from Corollary 7.36, and (ii) follows from Corollary 7.38. Today we show that (i) and (ii) imply that f is the limit of simple $\mathcal{S}$-measurable functions. There are two steps.

1. We construct our simple $\mathcal{S}$-measurable functions $\{f_n\}_{n\in\mathbb{N}}$. Because $\operatorname{im} f$ is separable by (ii), we may find some countable subset $\{b_i\}_{i=1}^\infty \subseteq \operatorname{im} f$ dense in $\operatorname{im} f$. Now, for each $i, j \in \mathbb{N}$, define

$$C_{ji} \coloneqq f^{-1}(B(b_i, 1/j) \setminus \{0\}),$$

which is always in $\mathcal{S}$ by (ii). Our goal is to carefully make the C_{ji} disjoint in order to define our sequence $\{f_n\}_{n\in\mathbb{N}}$ of simple $\mathcal{S}$-measurable functions, and we prefer C_{ji} with j large because these will give a finer approximation of f. In particular, we order C_{ji} lexicographically by (j,i): namely, $(j,i) < (\ell,k)$ if and only if $j < \ell$ or $j = \ell$ and $i < k$.

We now fix n and define our f_n. To make our C_{ji} appropriately disjoint, we will focus on the (j,i) bounded above by (n,n). Namely, for $(j,i) \in \{1,2,\ldots,n\}^2$, we set

$$E^n_{ji} \coloneqq C_{ji} \Bigg\backslash \bigcup_{\substack{(j,i)<(\ell,k)\\ 1\le \ell,k\le n}} C_{\ell k}.$$

For example, $E^n_{nn} = C_{nn}$ and $E_{n,n-1} = C_{n,n-1} \setminus C_{n,n}$ and $E_{n,n-2} = C_{n,n-2} \setminus (C_{n,n} \cup C_{n,n-2})$ and so on.

Notably, $E^n_{ji} \subseteq C_{ji}$ always, which means that the $E^n_\bullet$ are all disjoint: note $(j,i) \ne (j',i')$ implies that $(j,i) < (j',i')$ or $(j',i') < (j,i)$. Taking $(j,i) < (j',i')$ without loss of generality, we see that $E^n_{ji} \subseteq C_{ji} \setminus C_{j'i'}$ is disjoint from $E^n_{j'i'} \subseteq C_{j'i'}$.

With this in mind, we define

$$f_n \coloneqq \sum_{j=1}^n \sum_{i=1}^n b_i 1_{E^n_{ji}}.$$

Note that $\operatorname{im} f_n = \{0, b_1, \ldots, b_n\}$ because the E^n_{ji} are disjoint, which we see is finite. Further, for any b_i, we can compute

$$f_n^{-1}(\{b_i\}) = \bigcup_{j=1}^n E^n_{ji},$$

which is in $\mathcal{S}$ because $\mathcal{S}$ is a ring. Thus, f_n is in fact a simple $\mathcal{S}$-measurable function.

2. It remains to check that $f_n \to f$ pointwise as $n \to \infty$. If $x \notin f^{-1}(B \setminus \{0\})$, then $f(x) = 0$ while $x \notin C_{ji}$ always and so $x \notin E^n_{ji}$ always and so $f_n(x) = 0$ for all n; so $f_n(x) \to f(x)$ follows with nothing to say in this case. Thus, we may assume $f(x) \ne 0$.

Now, take $\varepsilon > 0$, and we need to find N such that $n > N$ implies $|f(x) - f_n(x)| < \varepsilon$ for $n > N$. This has two steps: first, take some j with $\frac{1}{j} < \varepsilon$, and second, we choose i_0 by density such that $f(x) \in B(b_{i_0}, 1/j)$.[1] As such, set $N := \max\{j, i_0\} + 1$ so that $\frac{1}{N} < \frac{1}{j} < \varepsilon$ and $i_0 < N$. Notably, $f(x) \in B(b_{i_0}, 1/h) \setminus \{0\}$ implies that

$$x \in C_{j_0 i_0}.$$

We now begin our check. If $n > N$, then $x \in E^n_{\ell k}$, where

$$(\ell, k) := \max\{(j, i) : x \in C_{ji} \text{ and } 1 \le j, i \le n\}.$$

Namely, there is certainly some (j, i) with $x \in C_{ji}$ and $1 \le j, i \le n$ because $x \in C_{j_0, i_0}$ while $j_0, i_0 < N < n$, so the maximum certainly exists. And we see $x \in E^n_{\ell k}$ because having $(j, i) > (\ell, k)$ with $1 \le j, i \le n$ will imply that $x \notin C_{ji}$ by maximality of (ℓ, k).

Now, $f_n(x) = b_k$ by construction, and $(j_0, i_0) \le (\ell, k)$ by maximality implies that $j_0 \le \ell$ and so

$$f(x) \in B(b_k, 1/\ell) \subseteq B(b_k, 1/j_0) \subseteq B(b_k, \varepsilon),$$

so $|f(x) - f_n(x)| < \varepsilon$ follows.

The above steps complete the proof. ■

Corollary 7.42. Fix a set X with σ-ring $\mathcal{S}$. A function $f \colon X \to \mathbb{R}$ is $\mathcal{S}$-measurable if and only if $f^{-1}(U \setminus \{0\}) \in \mathcal{S}$ for each open $U \subseteq \mathbb{R}$.

Proof. If f is $\mathcal{S}$-measurable, then this follows from Corollary 7.38. Conversely, if $f^{-1}(U \setminus \{0\}) \in \mathcal{S}$ for each open $U \subseteq \mathbb{R}$, then we note $\operatorname{im} f \subseteq \mathbb{R}$ is separable by Example 7.30, so f is $\mathcal{S}$-measurable by Theorem 7.40. ■

Corollary 7.43. Fix a set X with σ-ring $\mathcal{S}$. If $f \colon X \to \mathbb{R}$ is $\mathcal{S}$-measurable, and $g \colon \mathbb{R} \to \mathbb{R}$ is continuous such that $g(0) = 0$, then $g \circ f$ is still $\mathcal{S}$-measurable.

Proof. For any open $U \subseteq \mathbb{R}$, we note

$$(g \circ f)^{-1}(U \setminus \{0\}) = f^{-1}\left(g^{-1}(U \setminus \{0\})\right) = f^{-1}\left(g^{-1}(U) \setminus g^{-1}(\{0\})\right).$$

Now, $g^{-1}(U) \subseteq \mathbb{R}$ is open because $g \colon \mathbb{R} \to \mathbb{R}$ is continuous, and $0 \in g^{-1}(\{0\})$ because $g(0) = 0$, so $f^{-1}\left(g^{-1}(U) \setminus g^{-1}(\{0\})\right) \in \mathcal{S}$ by Corollary 7.38. Thus, $g \circ f$ is $\mathcal{S}$-measurable by Corollary 7.42. ■

7.4 November 2

We begin class by finishing the proof of Theorem 7.40. I have simply edited that proof for continuity reasons.

7.4.1 Some Measurable Facts

We now use Theorem 7.40 for fun and profit.

[1] Note that $B(f(x), 1/j) \cap \{b_i : i \in \mathbb{N}\}$ is nonempty because $\operatorname{im} f \subseteq \overline{\{b_i : i \in \mathbb{N}\}}$; we are choosing i_0 with $b_{i_0} \in B(f(x), 1/j)$, which means $f(x) \in B(b_{i_0}, 1/j)$

Corollary 7.44. Fix a normed vector space B and a set X with a σ-ring $\mathcal{S}$ on X. If a sequence of functions $\{f_n\}_{n\in\mathbb{N}}$ are $\mathcal{S}$-measurable, and $f_n \to f$ pointwise as $n \to \infty$, then f is also $\mathcal{S}$-measurable.

Proof. By Theorem 7.40, we have two checks.

(i) We show that $\operatorname{im} f$ is separable. Well, each $\operatorname{im} f_n$ is separable by Theorem 7.40, so this follows from Lemma 7.35.

(ii) We show that $f^{-1}(U \setminus \{0\}) \in \mathcal{S}$ for each open $U \subseteq B$. Well, each f_n has $f_n^{-1}(U \setminus \{0\}) \in \mathcal{S}$ for each open $U \subseteq B$, so the same holds for f by Lemma 7.35.

The above checks show that f is $\mathcal{S}$-measurable by Theorem 7.40. ■

Corollary 7.45. Fix a normed vector space $(B, \|\cdot\|)$ and a set X with a σ-ring $\mathcal{S}$ on X. If f is an $\mathcal{S}$-measurable function, then $x \mapsto \|f(x)\|$ is as well.

Proof. For brevity, set $g\colon X \to \mathbb{R}$ by $g(x) := \|f(x)\|$. By Theorem 7.40, there are two checks.

(i) Note that $\operatorname{im} g \subseteq \mathbb{R}$ must be separable by Example 7.30, so there is nothing more to say here.

(ii) For any open $U \subseteq \mathbb{R}$, we see that
$$U' := \{x \in B : \|x\| \in U\}$$
is open in B because $x \mapsto \|x\|$ is continuous by Example 1.38. Thus, $g^{-1}(U) = f^{-1}(U') \in \mathcal{S}$ because f is $\mathcal{S}$-measurable. ■

Example 7.46. If $f\colon X \to \mathbb{R}$ is $\mathcal{S}$-measurable, then Corollary 7.45 tells us that $|f|$ is also $\mathcal{S}$-measurable. As such, if $f, g\colon X \to \mathbb{R}$ are $\mathcal{S}$-measurable, then $(f+g)$ and $(f-g)$ are $\mathcal{S}$-measurable by Lemma 7.25, so $|f-g|$ is $\mathcal{S}$-measurable, so
$$\min\{f, g\} = \frac{(f+g) + |f-g|}{2} \qquad \text{and} \qquad \max\{f, g\} = \frac{(f+g) - |f-g|}{2}$$
are $\mathcal{S}$-measurable by Lemma 7.25 again. Inducting, for any $\mathcal{S}$-measurable functions $\{f_i\}_{i=1}^n$, the minimum function $\min\{f_1, \ldots, f_n\}$ and maximum function $\max\{f_1, \ldots, f_n\}$ are both $\mathcal{S}$-measurable.

We next talk a little about restriction.

Lemma 7.47. Fix a normed vector space B and a measure space $(X, \mathcal{S}, \mu)$ and a set $E \in \mathcal{S}$. If $f\colon X \to B$ is simple $\mathcal{S}$-measurable or $\mathcal{S}$-measurable or simple μ-integrable, then $f1_E$ is as well.

Proof. Before doing anything, we pick up a few facts. Note that
$$\operatorname{im} f1_E = \{f(x)1_E(x) : x \in X\} \subseteq \{0\} \cup \{f(x) : x \in E\} \subseteq \{0\} \cup \operatorname{im} f.$$
Also, if $S \subseteq B \setminus \{0\}$, then we claim
$$(f1_E)^{-1}(S) = E \cap f^{-1}(S).$$
In one direction, note $x \in E \cap f^{-1}(S)$ implies that $(f1_E)(x) = f(x) \in S$. In the other direction, if $x \in (f1_E)^{-1}(S)$, then note $x \in E$ is forced because otherwise $f(x) = 0 \notin S$. Thus, with $x \in E$, we have $(f1_E)(x) = f(x)$, so $(f1_E)(x) \in S$ forces $x \in f^{-1}(S)$ as well.

We now note that we actually have three claims to show, which we show in sequence.

- Suppose that f is a simple $\mathcal{S}$-measurable function. As such, $\operatorname{im} f$ is finite, so $\operatorname{im} f1_E \subseteq \{0\} \cup \operatorname{im} f$ is also finite.

  Further, for each $y \in (\operatorname{im} f1_E) \setminus \{0\}$, we see that $(f1_E)^{-1}(\{y\}) = E \cap f^{-1}(\{y\})$ as discussed above, which lives in $\mathcal{S}$ because $E \in \mathcal{S}$ and $f^{-1}(\{y\}) \in \mathcal{S}$.

- Suppose that f is an $\mathcal{S}$-measurable function. Then $\operatorname{im} f$ is separable, so it follows $\{0\} \cup \operatorname{im} f$ is separable (by Example 7.32), so $\operatorname{im} f1_E \subseteq \{0\} \cup \operatorname{im} f$ is separable (by Remark 7.31).

  Now, for any open subset $U \subseteq B \setminus \{0\}$, we see $(f1_E)^{-1}(U) = E \cap f^{-1}(U)$ as discussed above, which lives in $\mathcal{S}$ because $E \in \mathcal{S}$ and $f^{-1}(\{y\}) \in \mathcal{S}$.

- Suppose that f is a simple μ-integrable function. As before, $\operatorname{im} f$ is finite implies that $\operatorname{im} f1_E \subseteq \{0\} \cup \operatorname{im} f$ is still finite.

  Further, for each $y \in (\operatorname{im} f1_E) \setminus \{0\}$, we see $(f1_E)^{-1}(\{y\}) = E \cap f^{-1}(\{y\})$, which saw in our first point to live in $\mathcal{S}$, but now we note that Lemma 5.51 tells us

  $$\mu\left((f1_E)^{-1}(\{y\})\right) \le \mu\left(f^{-1}(\{y\})\right) < \infty$$

  is finite. ∎

Remark 7.48. On the other hand, if $X \setminus E \in \mathcal{S}$, then we see that $f1_E$ still gets the relevant adjectives. Indeed, each of the classes is a vector space (by Lemma 7.9 and Lemma 7.25 and Lemma 7.17), so it's enough to see $f1_E = f - f1_{X \setminus E}$ and apply Lemma 7.47.

Corollary 7.49. Fix a measure space $(X, \mathcal{S}, \mu)$ and a normed vector space B. Further, fix a μ-measurable function $f\colon X \to B$. Then there is some $N \in \mathcal{S}$ such that $\mu(N) = 0$ while $f1_N$ is $\mathcal{S}$-measurable.

Proof. Because f is μ-measurable, there is a sequence of simple $\mathcal{S}$-measurable functions $\{f_n\}_{n \in \mathbb{N}}$ such that $f_n \to f$ almost everywhere. Thus, there is some $N \in \mathcal{S}$ such that $\mu(N) = 0$ while $f_n(x) \to f(x)$ as $n \to \infty$ for each $x \in X \setminus N$.

We now show that $f1_N$ is $\mathcal{S}$-measurable. Indeed, we claim that $f_n 1_N \to f1_N$ as $n \to \infty$ pointwise, which will finish because each $f_n 1_N$ is simple $\mathcal{S}$-measurable by Lemma 7.47. If $x \notin N$, then we're just asking for $f_n(x) \to f(x)$ as $n \to \infty$, which we know. On the other hand, if $x \notin N$, then we're asking for $0 \to 0$ as $n \to \infty$, for which there's nothing to say. ∎

Theme 8
Integration

Having thus refreshed ourselves in the oasis of a proof, we now turn again into the desert of definitions

—Theodor Bröcker and Klaus Jänich, [BJ82]

8.1 November 2

We now switch gears and begin moving towards integration more directly.

8.1.1 Integrating Simple Functions

We begin by picking up some facts about our integral.

Lemma 8.1. Fix a normed vector space B and a ring $\mathcal{S}$ on a set X equipped with a finitely additive measure μ. Then the mapping

$$f \mapsto \int_X f \, d\mu$$

from simple μ-integrable functions to B is k-linear.

Proof. Unsurprisingly, we use the ideas of Lemma 7.17 to compute our integrals. We have two checks.

- Scalar multiplication: fix a simple μ-integrable function f and a scalar $r \in k$. If $r = 0$, then $rf = 0$, so $\int_X (rf) \, d\mu = 0 = r \int_X f \, d\mu$ vacuously, so there is nothing more to say.

  Otherwise, we have $r \neq 0$, and we remarked in Lemma 7.17 that we have

$$(rf)^{-1}(\{y\}) = f^{-1}(\{1/r \cdot y\}).$$

In other words, $\operatorname{im} rf = \{ry : y \in \operatorname{im} f\}$ with $(rf)^{-1}(\{ry\}) = f^{-1}(\{y\})$ for each $y \in \operatorname{im} f$, so

$$\begin{aligned}\int_X (rf)\,d\mu &= \sum_{y\in(\operatorname{im} rf)\setminus\{0\}} \mu\left((rf)^{-1}(\{y\})\right) y \\ &= \sum_{ry\in(\operatorname{im} rf)\setminus\{0\}} \mu\left((rf)^{-1}(\{ry\})\right)\cdot ry \\ &= r\sum_{y\in(\operatorname{im} f)\setminus\{0\}} \mu\left(f^{-1}(\{y\})\right) y \\ &= r\int_X f\,d\mu.\end{aligned}$$

- Addition: fix simple μ-integrable functions f and g. We remarked in Lemma 7.17 that any $y \in B\setminus\{0\}$ will have

$$(f+g)^{-1}(\{y\}) = \bigcup_{c\in(\operatorname{im} g)} \left(f^{-1}(\{y-c\})\cap g^{-1}(\{y\})\right) = \bigcup_{\substack{b\in(\operatorname{im} f), c\in(\operatorname{im} g)\\ b+c=y}} \left(f^{-1}(\{b\})\cap g^{-1}(\{c\})\right).$$

Now, note that this union is in fact disjoint because the fibers $f^{-1}(\{b\})$ are disjoint. Thus, we may say that

$$\mu\left((f+g)^{-1}(\{y\})\right) = \sum_{\substack{b\in(\operatorname{im} f), c\in(\operatorname{im} g)\\ b+c=y}} \mu\left(f^{-1}(\{b\})\cap g^{-1}(\{c\})\right).$$

Looping through all y, we see

$$\begin{aligned}\int_X (f+g)\,d\mu &= \sum_{y\in\operatorname{im}(f+g)\setminus\{0\}} \mu\left((f+g)^{-1}(\{y\})\right) y \\ &= \sum_{y\in\operatorname{im}(f+g)\setminus\{0\}} \sum_{\substack{b\in(\operatorname{im} f), c\in(\operatorname{im} g)\\ b+c=y}} \mu\left(f^{-1}(\{b\})\cap g^{-1}(\{c\})\right)(b+c) \\ &= \sum_{b\in(\operatorname{im} f)} \sum_{c\in(\operatorname{im} g)} \mu\left(f^{-1}(\{b\})\cap g^{-1}(\{c\})\right)(b+c) \\ &= \sum_{b\in(\operatorname{im} f)} \sum_{c\in(\operatorname{im} g)} \mu\left(f^{-1}(\{b\})\cap g^{-1}(\{c\})\right) b + \sum_{c\in(\operatorname{im} g)} \sum_{b\in(\operatorname{im} f)} \mu\left(f^{-1}(\{b\})\cap g^{-1}(\{c\})\right) c.\end{aligned}$$

Now, we note that

$$\bigsqcup_{b\in(\operatorname{im} f)} f^{-1}(\{b\}) = X \qquad\text{and}\qquad \bigsqcup_{c\in(\operatorname{im} g)} g^{-1}(\{c\}) = X$$

because the fibers should cover the domain and are disjoint. It follows that from the finite additivity of μ that

$$\int_X (f+g)\,d\mu = \sum_{b\in(\operatorname{im} f)} \mu\left(f^{-1}(\{b\})\right) b + \sum_{c\in(\operatorname{im} g)} \mu\left(g^{-1}(\{c\})\right) c,$$

which is $\int_X f\,d\mu + \int_X g\,d\mu$, which is what we wanted. ■

Lemma 8.2. Fix a normed vector space $(B, \|\cdot\|)$ and a ring $\mathcal{S}$ on a set X equipped with a finitely additive measure μ. Given a simple μ-integrable function $f\colon X\to B$, we have

$$\left\|\int_X f\,d\mu\right\| \le \int_X \|f\|\,d\mu.$$

Proof. Note $g := \|f\|$ is a simple μ-integrable function by Lemma 7.18. Now, the statement is essentially the triangle inequality for $\|\cdot\|$. Indeed, we compute

$$\left\|\int_X f\,d\mu\right\| = \left\|\sum_{y\in(\operatorname{im} f)\setminus\{0\}} \mu\left(f^{-1}(\{y\})\right)y\right\| \le \sum_{y\in(\operatorname{im} f)\setminus\{0\}} \mu\left(f^{-1}(\{y\})\right)\|y\|\,.$$

Back in Lemma 7.18, we established that

$$g^{-1}(\{r\}) = \bigcup_{\substack{y\in\operatorname{im} f\\ \|y\|=r}} f^{-1}(\{y\})$$

for each $r \in (\operatorname{im} g)\setminus\{0\}$. Note also that the above is a disjoint union: if $x \in f^{-1}(\{y\}) \cap f^{-1}(\{y'\})$, then $y = f(x) = y'$. As such, the finite additivity of μ tells us

$$\left\|\int_X f\,d\mu\right\| \le \sum_{r\in(\operatorname{im} g)\setminus\{0\}} \sum_{\substack{y\in\operatorname{im} f\\ \|y\|=r}} \mu\left(f^{-1}(\{y\})\right)\|y\| = \sum_{r\in(\operatorname{im} g)\setminus\{0\}} \mu\left(g^{-1}(\{r\})\right)\|r\| = \int_X g\,d\mu,$$

which is what we wanted. ■

Lemma 8.3. Fix a ring $\mathcal{S}$ on a set X equipped with a finitely additive measure μ. If a simple μ-integrable function $f\colon X \to \mathbb{R}$ has $f(x) \ge 0$ for each $x \in X$, then

$$\int_X f\,d\mu \ge 0.$$

Proof. Note that each $y \in (\operatorname{im} f)\setminus\{0\}$ has $y \ge 0$ and so

$$\int_X f\,d\mu = \sum_{y\in(\operatorname{im} f)\setminus\{0\}} \mu\left(f^{-1}(\{y\}\right)y$$

is nonnegative term-by-term, so $\int_X f\,d\mu \ge 0$ follows. ■

Corollary 8.4. Fix a ring $\mathcal{S}$ on a set X equipped with a finitely additive measure μ. Given simple μ-integrable functions $f, g\colon X \to \mathbb{R}$, if $f(x) \ge g(x)$ for each x, then $\int f\,d\mu \ge \int f\,d\mu$.

Proof. Set $h(x) := f(x) - g(x)$, which is a simple μ-integrable function by Lemma 7.17. Note $h(x) \ge 0$ for each x, so Lemma 8.3 tells us that

$$\int_X h(x)\,d\mu \ge 0.$$

However, by Lemma 8.1, we conclude that

$$\int_X h(x)\,d\mu = \int_X f(x)\,d\mu - \int_X g(x)\,d\mu,$$

so $\int_X f(x)\,d\mu \ge \int_X g(x)\,d\mu$ follows. ■

The above positivity result suggests a semi-norm on our space.

Notation 8.5. Fix a normed vector space B and a ring $\mathcal{S}$ on a set X equipped with a finitely additive measure μ. Given a simple integrable function $f\colon X \to B$, we define

$$\|f\|_1 := \int_X \|f\| \, d\mu.$$

Note $\|f\|$ is in fact a simple μ-integrable function by Lemma 7.18.

Lemma 8.6. Fix a normed vector space B and a ring $\mathcal{S}$ on a set X equipped with a finitely additive measure μ. Then the function $f \mapsto \|f\|_1$ on simple μ-integrable functions defines a semi-norm on the space of simple μ-integrable functions.

Proof. Note that simple μ-integrable functions already form a space by Lemma 7.9. Here are our checks.

- Positivity: given a simple μ-integrable function f, note that $\|f(x)\| \geq 0$ for any $x \in X$, so Lemma 8.3 tells us that $\int_X \|f\| \, d\mu \geq 0$.
- Zero: we show $\|z\|_1 = 0$, where $z\colon X \to B$ is the zero function. Well, $\|0\|$ is the zero function $X \to \mathbb{R}$ because $\|0\| = 0$, so the linearity of Lemma 8.1 forces $\int_X \|z\| \, d\mu = 0$.
- Scaling: given a simple μ-integrable function $f\colon X \to B$ and some scalar r, we need $\|rf\|_1 = \|r\| \cdot \|f\|_1$. Well, rf is still a simple integrable function by Lemma 7.17, as is $\|rf\|$ by Lemma 7.18.

  However, the main point is that $\|rf\| = \|r\| \cdot \|f\|$ by checking pointwise: any $x \in X$ has

  $$\|rf\|(x) = \|rf(x)\| = \|r\| \cdot \|f(x)\| = (\|r\| \cdot \|f\|)(x).$$

  Thus, linearity of Lemma 8.1 forces

  $$\int_X \|rf\| \, d\mu = \int_X (\|r\| \cdot \|f\|) \, d\mu = \|r\| \cdot \int_X \|f\| \, d\mu.$$

- Triangle inequality: given simple μ-integrable functions $f, g\colon X \to B$, we note that the triangle inequality gives

  $$\|f+g\|(x) = \|f(x)+g(x)\| \leq \|f(x)\| + \|g(x)\| = (\|f\| + \|g\|)(x)$$

  for any $x \in X$. Thus, noting as usual that $f+g$ and hence $\|f+g\|$ are both simple μ-integrable, we note Corollary 8.4 tells us

  $$\int_X \|f+g\| \, d\mu \leq \int_X (\|f\| + \|g\|) \, d\mu.$$

  As such, linearity of the integral from Lemma 8.1 tells us that $\|f+g\|_1 \leq \|f\|_1 + \|g\|_1$, which is what we wanted. ■

To make this a norm, we need to remove the problematic functions.

Notation 8.7. Fix a normed vector space B and a ring $\mathcal{S}$ on a set X equipped with a finitely additive measure μ. We define

$$\mathcal{SN}(X, \mathcal{S}, \mu, B) = \{f \text{ simple integrable} : \|f\|_1 = 0\}.$$

Thus, Proposition 1.13 tells us that we're going to get a norm on the quotient of all simple integrable functions by $\mathcal{SN}(X, \mathcal{S}, \mu)$. In our story of integration, we are essentially interested in the completion of this normed vector space.

Example 8.8. Give $[0,1]$ the usual Lebesgue measure μ, and let $\{E_i\}_{i=1}^\infty$ be pairwise disjoint Borel subsets of $\mathbb{R}$, where $\mu(E_i) \le 4^{-i}$ for each i. Then we see that

$$\sum_{i=1}^n 1_{E_i} \to \sum_{i=1}^\infty 1_{E_i}$$

as $n \to \infty$, but the function on the right may be potentially quite hard to handle. Namely, we want

$$\int_X \left(\sum_{i=1}^\infty 1_{E_i}\right) d\mu = \lim_{n\to\infty} \int_X \left(\sum_{i=1}^n 1_{E_i}\right) d\mu = \sum_{i=1}^\infty \mu(E_i),$$

but changing the order of this integral and sum is somewhat tricky.

8.1.2 Convergence in Measure

In order to avoid the constant repetition of hypotheses, we pick up the following definition.

Definition 8.9 (Measure space)**.** A *measure space* is a triple $(X, \mathcal{S}, \mu)$, where $\mathcal{S}$ is a σ-ring and μ is a measure on $\mathcal{S}$. We also require $\mu(\varnothing) < \infty$ so that $\mu(\varnothing) = 0$ by Remark 5.39.

Now, let me tell you the bad news.

Warning 8.10. A sequence $\{f_n\}_{n\in\mathbb{N}}$ of simple integrable functions which is Cauchy for $\|\cdot\|_1$ need not converge pointwise, at any point!

Example 8.11. Give $[0,1)$ the usual Lebesgue measure μ, and for $k \ge 1$, define $E_k := \left[\frac{k-2^n}{2^n}, \frac{k+1-2^n}{2^n}\right)$, where n is the integer such that $2^n \le k < 2^{n+1}$. Then the sequence of functions $\{1_{E_k}\}_{k\in\mathbb{Z}_{>0}}$ approaches 0 according to $\|\cdot\|_1$, but it does not converge to 0 pointwise anywhere! We will be brief.

- To see $1_{E_k} \to 0$ as $k \to \infty$ according to $\|\cdot\|_1$, we note $\|1_{E_k}\|_1 = 1/2^n$ by Example 7.16, which goes to 0 as $k \to \infty$. (Namely, $n = \lfloor \log_2 k \rfloor \to \infty$ as $k \to \infty$.)
- However, at particular $x \in [0,1)$, there are infinitely many k for which $x \in E_k$ (so that $1_{E_k}(x) = 1$) and $x \notin E_k$ (so that $1_{E_k}(x) = 0$), meaning $1_{E_k}(x)$ does not converge pointwise.

  Indeed, fix any N, and we find some $k \ge N$ with $x \in E_k$ and some $k \ge N$ with $x \notin E_k$. Well, choose any $n \ge \max\{N, 2\}$, and we see that the sets $E_{2^n}, E_{2^n+1}, \ldots, E_{2^{n+1}-1}$ are disjoint and cover $[0,1)$ by construction, so x will live in exactly one of them.

The main point of the above example is that our functions are allowed to look small according to $\|\cdot\|_1$ but be relatively large for (say) $\|\cdot\|_\infty$.

To fix this bad news, we have the following definition.

Definition 8.12 (Converge in measure)**.** Fix a measure space $(X, \mathcal{S}, \mu)$ and normed vector space $(B, \|\cdot\|)$. Then a sequence $\{f_n\}_{n\in\mathbb{N}}$ of $\mathcal{S}$-measurable functions *converges in measure* to an $\mathcal{S}$-measurable function f if and only if all $\varepsilon > 0$ have

$$\lim_{n\to\infty} \mu(\{x \in X : \|f(x) - f_n(x)\| \ge \varepsilon\}) = 0.$$

Remark 8.13. Notably, f and f_n are $\mathcal{S}$-measurable, so $f - f_n$ is $\mathcal{S}$-measurable by Lemma 7.25, so $g := \|f - f_n\|$ is $\mathcal{S}$-measurable by Corollary 7.45, so

$$\{x : \|f(x) - f_n(x)\| \geq \varepsilon\} = g^{-1}([\varepsilon, \infty)) = g^{-1}((0, \infty)) \setminus g^{-1}((0, \varepsilon))$$

is in fact in $\mathcal{S}$ by Corollary 7.38. In particular, the limit in Definition 8.12 actually makes sense.

Example 8.14. The sequence from Example 8.11 converges in measure to the zero function. Indeed, for any k, we see

$$\mu(\{x \in X : \|0 - 1_{E_k}(x)\| \geq \varepsilon\}) = \mu(E_k) = \frac{1}{2^{\lfloor \log_2 k \rfloor}}$$

by Example 7.16, which goes to 0 as $k \to \infty$.

Of course, with a notion of convergence, we also have a notion of being Cauchy.

Definition 8.15 (Cauchy in measure)**.** Fix a normed vector space $(B, \|\cdot\|)$ and a σ-ring $\mathcal{S}$ on a set X equipped with a measure μ. Then a sequence $\{f_n\}_{n \in \mathbb{N}}$ of $\mathcal{S}$-measurable functions is *Cauchy in measure* if and only if all $\varepsilon > 0$ have

$$\lim_{m,n \to \infty} \mu(\{x \in X : \|f_m(x) - f_n(x)\| \geq \varepsilon\}) = 0.$$

Remark 8.16. Again, we note that $f_n - f_m$ is $\mathcal{S}$-measurable by Lemma 7.25, so $g := \|f_m - f_n\|$ is $\mathcal{S}$-measurable by Corollary 7.45, so

$$\{x : \|f_m(x) - f_n(x)\| \geq \varepsilon\} = g^{-1}([\varepsilon, \infty)) = g^{-1}((0, \infty)) \setminus g^{-1}((0, \varepsilon))$$

is in fact in $\mathcal{S}$ by Corollary 7.38. So we do see the limit in Definition 8.12 actually makes sense.

Remark 8.17. In fact, if f_m and f_n are simple μ-integrable functions, then $f_m - f_n$ is also by Lemma 7.17, as is $g := \|f_m - f_n\|$ by Lemma 7.18. Thus,

$$g^{-1}([\varepsilon, \infty)) = \bigcup_{y \in (\operatorname{im} g) \cap [\varepsilon, \infty)} g^{-1}(\{y\})$$

is a finite union of sets $g^{-1}(\{y\})$ of finite measure, so $\mu\left(g^{-1}([\varepsilon, \infty))\right)$ is finite by Lemma 5.55.

8.2 November 4

We continue our journey towards integrating functions.

8.2.1 Sequences Converging in Measure

We pick up some basic tools on sequences converging in measure.

Lemma 8.18. Fix a normed vector space $(B, \|\cdot\|)$ and a measure space $(X, \mathcal{S}, \mu)$. Now, suppose a sequence $\{f_n\}_{n \in \mathbb{N}}$ of $\mathcal{S}$-measurable functions converges to both f and g in measure, where f and g are both $\mathcal{S}$-measurable. Then $f = g$ almost everywhere; i.e., $\{x \in X : f(x) \neq g(x)\}$ is a null set.

Proof. Before we do anything at all, we note that $f-g$ is $\mathcal{S}$-measurable by Lemma 7.25, so

$$N \coloneqq \{x \in X : f(x) \neq g(x)\} = (f-g)^{-1}(B \setminus \{0\})$$

is $\mathcal{S}$-measurable by Corollary 7.38.

Now, fix any $\varepsilon > 0$; we show $\mu(N) < \varepsilon$. The key observation is that

$$\|f(x) - g(x)\| \leq \|f(x) - f_n(x)\| + \|f_n(x) - g(x)\|,$$

so it follows that $\|f(x) - g(x)\| \geq \varepsilon$ forces $\|f(x) - f_n(x)\| \geq \varepsilon/2$ or $\|g(x) - g_n(x)\| \geq \varepsilon/2$. Thus,

$$\{x : \|f(x) - g(x)\| \geq \varepsilon\} \subseteq \{x : \|f(x) - f_n(x)\| \geq \varepsilon/2\} \cup \{x : \|g(x) - f_n(x)\| \geq \varepsilon/2\},$$

so Lemma 5.55 tells us

$$\mu(\{x \in X : \|f(x) - g(x)\| \geq \varepsilon\}) \leq \mu(\{x \in X : \|f(x) - f_n(x)\| \geq \varepsilon/2\}) + \mu(\{x \in X : \|g(x) - f_n(x)\| \geq \varepsilon/2\}).$$

But now, as $n \to \infty$, we see that the right-hand side goes to $0 + 0 = 0$ because $f_n \to f$ and $f_n \to g$ in measure, so it follows that

$$\mu(\{x \in X : \|f(x) - g(x)\| \geq \varepsilon\}) = 0. \tag{8.1}$$

We now send $\varepsilon \to 0^+$. Namely, we see $f(x) \neq g(x)$ is equivalent to $\|f(x) - g(x)\| > 0$ is equivalent to $\|f(x) - g(x)\| \geq 1/n$ for some $n \in \mathbb{N}$, so

$$N \coloneqq \{x \in X : f(x) \neq g(x)\} = \{x \in X : \|f(x) - g(x)\| > 0\} = \bigcup_{n \in \mathbb{N}} \{x \in X : \|f(x) - g(x)\| \geq 1/n\}.$$

Thus,

$$\mu(N) \leq \sum_{n \in \mathbb{N}} \mu(\{x \in X : \|f(x) - g(x)\| \geq 1/n\}) \stackrel{*}{=} \sum_{x \in X} 0 = 0,$$

so N is in fact a null set. Notably, $\stackrel{*}{=}$ has used (8.1). ■

Lemma 8.19. Fix a normed k-vector space $(B, \|\cdot\|)$ and a measure space $(X, \mathcal{S}, \mu)$. Fix sequences of $\mathcal{S}$-measurable functions $\{f_n\}_{n \in \mathbb{N}}$ and $\{g_n\}_{n \in \mathbb{N}}$ with $f_n \to f$ and $g_n \to g$ in measure as $n \to \infty$.

(a) We have $f_n + g_n \to f + g$ in measure.

(b) Given some scalar $a \in k$, we have $af_n \to af$ in measure.

(c) We have $\|f_n\| \to \|f\|$ as $n \to \infty$.

Proof. We go ahead and let $|\cdot|$ denote the norm on base field k of B.

(a) Note that the $f_n + g_n$ and $f + g$ are all $\mathcal{S}$-measurable by Lemma 7.25.

Now, by the triangle inequality, we see

$$\|(f(x) + g(x)) - (f_n(x) + g_n(x))\| \leq \|f(x) - f_n(x)\| + \|g(x) - g_n(x)\|.$$

We now proceed as in Lemma 8.18. Fix $\varepsilon > 0$. If the left-hand side exceeds ε, then one of the terms on the right-hand side must exceed $\varepsilon/2$, so

$$\begin{aligned}\{x : \|(f(x) + g(x)) - (f_n(x) + g_n(x))\| \geq \varepsilon\} \subseteq {} & \{x : \|f(x) - f_n(x)\| \geq \varepsilon/2\} \\ & \cup \{x : \|g(x) - g_n(x)\| \geq \varepsilon/2\}.\end{aligned}$$

Thus, Lemma 5.55 tells us

$$\begin{aligned}\mu(\{x : \|(f(x) + g(x)) - (f_n(x) + g_n(x))\| \geq \varepsilon\}) \leq {} & \mu(\{x : \|f(x) - f_n(x)\| \geq \varepsilon/2\}) \\ & + \mu(\{x : \|g(x) - g_n(x)\| \geq \varepsilon/2\}).\end{aligned}$$

However, $\varepsilon/2 > 0$, so taking $n \to \infty$ and using our convergence in measure tells us that

$$\lim_{n\to\infty} \mu(\{x : \|(f(x) + g(x)) - (f_n(x) + g_n(x))\| \ge \varepsilon\}) \le 0 + 0 = 0,$$

so we are done after noting that μ will only output nonnegative values, so the limit is at least nonnegative.

(b) Note that the af_n and af are all $\mathcal{S}$-measurable by Lemma 7.25.

Now, fix some $\varepsilon > 0$ so that we want to show that

$$L := \lim_{n\to\infty} \mu(\{x \in X : \|af_n(x) - af(x)\| \ge \varepsilon\}) \stackrel{?}{=} 0.$$

If $a = 0$, then $af_n(x) = af(x) = 0$ for all $x \in X$, so $\{x \in X : \|af_n(x) - af(x)\| \ge \varepsilon\}$ is empty, so the result follows.

Otherwise, take $a \neq 0$ so that $|a| > 0$. Now, note $\|af_n(x) - af(x)\| = |a| \cdot \|f_n(x) - f(x)\|$, so it follows $\|af_n(x) - af(x)\| \ge \varepsilon$ if and only if $\|f_n(x) - f(x)\| \ge \varepsilon/|a|$. Thus,

$$L = \lim_{n\to\infty} \mu(\{x \in X : \|f_n(x) - f(x)\| \ge \varepsilon/|a|\}).$$

However, $\varepsilon/|a| > 0$ because $\varepsilon > 0$, so the above limit vanishes because $f_n \to f$ in measure as $n \to \infty$.

(c) Observe that the g_n and g are all $\mathcal{S}$-measurable by Corollary 7.45.

Now, fix some $\varepsilon > 0$. By the (reverse) triangle inequality,

$$|\,\|f(x)\| - \|f_n(x)\|\,| \le \|f(x) - f_n(x)\|,$$

so any $\varepsilon > 0$ has

$$\{x : |\,\|f(x)\| - \|f_n(x)\|\,| \ge \varepsilon\} \subseteq \{x : \|f(x) - f_n(x)\| \ge \varepsilon\}.$$

Thus, Lemma 5.51 tells us

$$\lim_{n\to\infty} \mu(\{x : |\,\|f(x)\| - \|f_n(x)\|\,| \ge \varepsilon\}) \le \lim_{n\to\infty} \mu(\{x : \|f(x) - f_n(x)\| \ge \varepsilon\}).$$

The right-hand limit vanishes because $f_n \to f$ in measure, so the left-hand limit must vanish as well because the limit's terms are nonnegative. ■

Here is the analogous result for sequences Cauchy in measure.

Lemma 8.20. Fix a normed k-vector space $(B, \|\cdot\|)$ and a measure space $(X, \mathcal{S}, \mu)$. Fix sequences of $\mathcal{S}$-measurable functions $\{f_n\}_{n\in\mathbb{N}}$ and $\{g_n\}_{n\in\mathbb{N}}$ which are Cauchy in measure.

(a) The sequence $\{f_n + g_n\}_{n\in\mathbb{N}}$ is Cauchy in measure.

(b) Given some scalar $a \in k$, the sequence $\{af_n\}_{n\in\mathbb{N}}$ is Cauchy in measure.

(c) The sequence of functions $\{\|f_n\|\}_{n\in\mathbb{N}}$ is Cauchy in measure.

Proof. These proofs are essentially the same as Lemma 8.19 with the appropriate names changed. Again, we let $|\cdot|$ denote the norm on base field k of B.

(a) Note that the $f_n + g_n$ are all $\mathcal{S}$-measurable by Lemma 7.25.

Now, by the triangle inequality, we see

$$\|(f_m(x) + g_m(x)) - (f_n(x) + g_n(x))\| \le \|f_m(x) - f_n(x)\| + \|g_m(x) - g_n(x)\|.$$

Fix $\varepsilon > 0$. As usual

$$\{x : \|(f_m(x) + g_m(x)) - (f_n(x) + g_n(x))\| \ge \varepsilon\} \subseteq \{x : \|f_m(x) - f_n(x)\| \ge \varepsilon/2\} \\ \cup \{x : \|g_m(x) - g_n(x)\| \ge \varepsilon/2\},$$

so Lemma 5.55 tells us

$$\mu(\{x : \|(f_m(x) + g_m(x)) - (f_n(x) + g_n(x))\| \ge \varepsilon\}) \le \mu(\{x : \|f_m(x) - f_n(x)\| \ge \varepsilon/2\}) \\ + \mu(\{x : \|g_m(x) - g_n(x)\| \ge \varepsilon/2\}).$$

However, $\varepsilon/2 > 0$, so taking $m, n \to \infty$ and using our Cauchy in measure conditions tells us that

$$\lim_{m,n\to\infty} \mu(\{x : \|(f_m(x) + g_m(x)) - (f_n(x) + g_n(x))\| \ge \varepsilon\}) \le 0 + 0 = 0,$$

so we are done after noting that μ will only output nonnegative values, so the limit is at least nonnegative.

(b) Note that the af_n are all $\mathcal{S}$-measurable by Lemma 7.25.

Now, fix some $\varepsilon > 0$ so that we want to show that

$$L := \lim_{m,n\to\infty} \mu(\{x \in X : \|af_m(x) - af_n(x)\| \ge \varepsilon\}) \stackrel{?}{=} 0.$$

If $a = 0$, then $af_n(x) = af(x) = 0$ for all $x \in X$, so $\{x \in X : \|af_m(x) - af_n(x)\| \ge \varepsilon\}$ is empty, so the result follows.

Otherwise, take $a \neq 0$ so that $|a| > 0$. Now, note $\|af_m(x) - af_n(x)\| = |a| \cdot \|f_m(x) - f_n(x)\|$, so it follows $\|af_m(x) - af_n(x)\| \ge \varepsilon$ if and only if $\|f_m(x) - f_n(x)\| \ge \varepsilon/|a|$. Thus,

$$L = \lim_{m,n\to\infty} \mu(\{x \in X : \|f_m(x) - f_n(x)\| \ge \varepsilon/|a|\}).$$

However, $\varepsilon/|a| > 0$ because $\varepsilon > 0$, so the above limit vanishes because $\{f_n\}_{n\in\mathbb{N}}$ is Cauchy in measure.

(c) Observe that the g_n are all $\mathcal{S}$-measurable by Corollary 7.45.

Now, fix some $\varepsilon > 0$. By the (reverse) triangle inequality,

$$|\,\|f_m(x)\| - \|f_n(x)\|\,| \le \|f_m(x) - f_n(x)\|\,,$$

so any $\varepsilon > 0$ has

$$\{x : |\,\|f_m(x)\| - \|f_n(x)\|\,| \ge \varepsilon\} \subseteq \{x : \|f_m(x) - f_n(x)\| \ge \varepsilon\}.$$

Thus, Lemma 5.51 tells us

$$\lim_{m,n\to\infty} \mu(\{x : |\,\|f_m(x)\| - \|f_n(x)\|\,| \ge \varepsilon\}) \le \lim_{m,n\to\infty} \mu(\{x : \|f_m(x) - f_n(x)\| \ge \varepsilon\}).$$

The right-hand limit vanishes because $\{f_n\}_{n\in\mathbb{N}}$ is Cauchy in measure, so the left-hand limit must vanish as well because the limit's terms are also nonnegative. ■

We will want a few results on subsequences later on.

Lemma 8.21. Fix a normed vector space $(B, \|\cdot\|)$ and a measure space $(X, \mathcal{S}, \mu)$. Now, fix$\mathcal{S}$-measurable functions $\{f_n\}_{n\in\mathbb{N}}$ and f. If $f_n \to f$ in measure as $n \to \infty$, then $f_{n_i} \to f$ in measure as $i \to \infty$ for any subsequence $\{f_{n_i}\}_{i\in\mathbb{N}}$.

Proof. Fix some $\varepsilon > 0$. Then any $\delta > 0$ has some N for which $n \geq N$ has

$$\mu(\{x \in X : \|f(x) - f_n(x)\| \geq \varepsilon\}) < \delta.$$

As such, for any $i \geq N$, we see $n_i \geq i \geq N$, so

$$\mu(\{x \in X : \|f(x) - f_{n_i}(x)\| \geq \varepsilon\}) < \delta,$$

which finishes. ■

Lemma 8.22. Fix a measure space $(X, \mathcal{S}, \mu)$ and a normed vector space B. Further, fix a sequence of simple $\mathcal{S}$-measurable functions $\{f_n\}_{n\in\mathbb{N}}$ which is Cauchy in measure. If a subsequence $\{f_{n_i}\}_{i\in\mathbb{N}}$ converges to a function f in measure, then the full sequences $\{f_n\}_{n\in\mathbb{N}}$ converges to f in measure.

Proof. Fix any $\varepsilon > 0$ and $\delta > 0$. We need N such that $n \geq N$ implies

$$0 \leq \mu(\{x \in X : \|f(x) - f_n(x)\| \geq \varepsilon\}) \stackrel{?}{<} \delta.$$

Well, we note that any n and i will have

$$\|f(x) - f_n(x)\| \leq \|f(x) - f_{n_i}(x)\| + \|f_{n_i}(x) - f_n(x)\|,$$

so

$$\{x \in X : \|f(x) - f_n(x)\| \geq \varepsilon\} \subseteq \{x \in X : \|f(x) - f_{n_i}(x)\| \geq \varepsilon/2\} \cup \{x \in X : \|f_{n_i}(x) - f_n(x)\| \geq \varepsilon/2\}.$$

Now, $\{f_n\}_{n\in\mathbb{N}}$ being Cauchy in measure allows us to pick N such that $m, n \geq N$ implies

$$\mu(\{x \in X : \|f_m(x) - f_n(x)\| \geq \varepsilon/2\}) < \frac{\delta}{2}.$$

Additionally, $f_{n_i} \to f$ in measure grants N' such that $i \geq N'$ implies

$$\mu(\{x \in X : \|f(x) - f_{n_i}(x)\| \geq \varepsilon/2\}) < \frac{\delta}{2}.$$

Thus, for any $n \geq N$, we select any $i \geq \max\{N, N'\}$. Notably, $n_i \geq i \geq N$ as well, so

$$\mu(\{x \in X : \|f(x) - f_n(x)\| \geq \varepsilon\}) < \frac{\delta}{2} + \frac{\delta}{2} = \delta,$$

where we have used the above inequalities in addition to Lemma 5.55. ■

Lastly, here is the expected uniqueness results.

Lemma 8.23. Fix a normed vector space $(B, \|\cdot\|)$ and a measure space $(X, \mathcal{S}, \mu)$. Further, fix a sequence $\{f_n\}_{n\in\mathbb{N}}$ of $\mathcal{S}$-measurable functions converging in measure to an $\mathcal{S}$-measurable function $f\colon X \to B$. Given an $\mathcal{S}$-measurable function $g\colon X \to B$, we have $f = g$ almost everywhere if and only if $f_n \to g$ in measure.

Proof. In one direction, suppose $f = g$ almost everywhere so that we have some $E \in \mathcal{S}$ with $\mu(E) = 0$ such that $f(x) \neq g(x)$ implies $x \in E$. Now, fix some $\varepsilon > 0$. For any $\delta > 0$, we are promised N such that $n \geq N$ implies

$$\mu(\{x \in X : \|f(x) - f_n(x)\| \geq \varepsilon\}) < \delta.$$

Now, we note that $\|g(x) - f_n(x)\| \geq \varepsilon$ implies that either $f(x) \neq g(x)$ so that $x \in E$ or $\|f(x) - f_n(x)\| \geq \varepsilon$, so it follows

$$\{x \in X : \|g(x) - f_n(x)\| \geq \varepsilon\} = \{x \in X : \|f(x) - f_n(x)\| \geq \varepsilon\} \cup \{x \in E : \|g(x) - f_n(x)\| \geq \varepsilon\}.$$

Notably, $\mu(\{x \in E : \|g(x) - f_n(x)\| \geq \varepsilon\}) = 0$ as a subset of E, so we see that

$$\mu(\{x \in X : \|g(x) - f_n(x)\|) \leq \{x \in X : \|f(x) - f_n(x)\| \geq \varepsilon\} + 0 < \delta,$$

which finishes this direction.

Conversely, suppose $f_n \to g$ in measure. Note that $f(x) \neq g(x)$ if and only if $\|f(x) - g(x)\| > 0$ if and only if $\|f(x) - g(x)\| \geq 1/m$ for some positive integer m. Thus,

$$\{x \in X : f(x) \neq g(x)\} \subseteq \bigcup_{m=1}^{\infty} \{x \in X : \|f(x) - g(x)\| \geq 1/m\}.$$

By Lemma 6.2, to show that $\{x \in X : f(x) \neq g(x)\}$ is a null set, it suffices to show that $\{x \in X : \|f(x) - g(x)\| \geq 1/m\}$ is a null set. Well, note that any positive integer n has

$$\|f(x) - g(x)\| \leq \|f(x) - f_n(x)\| + \|g(x) - f_n(x)\|,$$

implying

$$\{x : \|f(x) - g(x)\| \geq 1/m\} \subseteq \{x : \|f(x) - f_n(x)\| \geq 1/(2m)\} \cup \{x : \|g(x) - f_n(x)\| \geq 1/(2m)\}.$$

Now, for any $\delta > 0$, because $f_n \to f$ and $f_n \to g$ in measure, we are promised some N large enough so that $n \geq N$ has

$$\mu(\{x : \|f(x) - f_n(x)\| \geq 1/(2m)\}), \mu(\{x : \|g(x) - f_n(x)\| \geq 1/(2m)\}) < \frac{\delta}{2}.$$

It follows by Lemma 5.55 that

$$\mu(\{x \in X : \|f(x) - g(x)\| \geq 1/m\}) < \frac{\delta}{2} + \frac{\delta}{2} = \delta$$

for any $\delta > 0$. Thus, $\mu(\{x \in X : \|f(x) - g(x)\| \geq 1/m\}) = 0$ follows. ■

8.2.2 Restricting Measurable Functions

Analogously to Lemma 8.19 and Lemma 8.20, we have the following.

Lemma 8.24. Fix a normed vector space $(B, \|\cdot\|)$ and a measure space $(X, \mathcal{S}, \mu)$, and fix some $E \in \mathcal{S}$. Given a sequence $\{f_n\}_{n\in\mathbb{N}}$ of $\mathcal{S}$-measurable functions with $f_n \to f$ in measure as $n \to \infty$, then $f_n 1_E \to f 1_E$ in measure as $n \to \infty$.

Proof. Note that the $f_n 1_E$ and $f 1_E$ are all $\mathcal{S}$-measurable by Lemma 7.47, so the claim at least makes sense.

For brevity, we set $g_n := \|f 1_E - f_n 1_E\|$ for each n. We would like to show

$$\lim_{n\to\infty} \mu\left(g_n^{-1}([\varepsilon, \infty))\right) \stackrel{?}{=} 0.$$

If $x \notin E$, then note $g_n(x) = 0$; otherwise, $g_n(x) = \|f(x) - f_n(x)\|$ because $1_E(x) = 1$. As such, for $\varepsilon > 0$, we see $x \in g_n^{-1}([\varepsilon, \infty))$ requires $x \in E$ and then $\|f(x) - f_n(x)\| \geq \varepsilon$; conversely, $x \in E$ with $\|f(x) - f_n(x)\| \geq \varepsilon$ does give $g_n(x) \geq \varepsilon$.

Thus, we note that

$$g_n^{-1}([\varepsilon, \infty)) \subseteq \{x : \|f(x) - f_n(x)\| \geq \varepsilon\},$$

so Lemma 5.51 tells us

$$\lim_{n\to\infty} \mu\left(g_n^{-1}([\varepsilon, \infty))\right) \leq \lim_{n\to\infty} \mu\left(\{x : \|f(x) - f_n(x)\| \geq \varepsilon\}\right),$$

where the right-hand limit vanishes because $f_n \to f$ in measure as $n \to \infty$. Thus, the left-hand limit also vanishes because the terms of the limit are nonnegative. ■

Lemma 8.25. Fix a normed vector space $(B, \|\cdot\|)$ and a measure space $(X, \mathcal{S}, \mu)$, and fix some $E \in \mathcal{S}$. Given a sequence $\{f_n\}_{n\in\mathbb{N}}$ of $\mathcal{S}$-measurable functions which is Cauchy in measure, then $\{f_n 1_E\}_{n\in\mathbb{N}}$ is still Cauchy in measure.

Proof. As usual, the proof is exactly the same as before. Note that the $f_n 1_E$ and $f 1_E$ are all $\mathcal{S}$-measurable by Lemma 7.47, so the claim at least makes sense.

For brevity, we set $g_{m,n} := \|f_m 1_E - f_n 1_E\|$ for each m and n. We would like to show

$$\lim_{m,n\to\infty} \mu\left(g_{m,n}^{-1}([\varepsilon, \infty))\right) \stackrel{?}{=} 0.$$

If $x \notin E$, then note $g_{m,n}(x) = 0$; otherwise, $g_{m,n}(x) = \|f_m(x) - f_n(x)\|$ because $1_E(x) = 1$. As such, for $\varepsilon > 0$, we see $x \in g_{m,n}^{-1}([\varepsilon,\infty))$ requires $x \in E$ and then $\|f_m(x) - f_n(x)\| \ge \varepsilon$; conversely, $x \in E$ with $\|f_m(x) - f_n(x)\| \ge \varepsilon$ does give $g_{m,n}(x) \ge \varepsilon$.

Thus, we note that

$$g_{m,n}^{-1}([\varepsilon,\infty)) \subseteq \{x : \|f_m(x) - f_n(x)\| \ge \varepsilon\},$$

so Lemma 5.51 tells us

$$\lim_{m,n\to\infty} \mu\left(g_{m,n}^{-1}([\varepsilon,\infty))\right) \le \lim_{m,n\to\infty} \mu\left(\{x : \|f_m(x) - f_n(x)\| \ge \varepsilon\}\right),$$

where the right-hand limit vanishes because $\{f_n\}_{n\in\mathbb{N}}$ is Cauchy in measure. Thus, the left-hand limit also vanishes because the terms of the limit are nonnegative. ■

The above corollary promises the following notation.

Notation 8.26. Fix a normed vector space $(B, \|\cdot\|)$ and a measure space $(X, \mathcal{S}, \mu)$. Then a simple integrable function f on X and $E \in \mathcal{S}$ will have

$$\int_E f\,d\mu := \int_X f 1_E\,d\mu.$$

Remark 8.27. One can define

$$\mu_f(E) := \int_E f\,d\mu,$$

and it is not too hard to check that this defines a measure on $\mathcal{S}$ which is valued in B. This μ_f will later be called the "indefinite integral for f." We will postpone writing this out until we are ready to talk about what this looks like when f is a general μ-integrable function instead of a simple μ-integrable function.

8.2.3 Almost Uniform Convergence

As we tend to do, we now return to a context which is perhaps too general.

Definition 8.28 (Almost uniformly)**.** Fix a normed vector space B and a measure space $(X, \mathcal{S}, \mu)$. Then a sequence of functions $f_n \colon X \to B$ for $n \in \mathbb{N}$ converges *almost uniformly* to f if and only if every $\varepsilon > 0$ has some $E^\varepsilon \in \mathcal{S}$ such that $\mu(E^\varepsilon) < \varepsilon$ and $f_n|_{X\setminus E} \to f|_{X\setminus E}$ uniformly.

Remark 8.29. The term "almost" above is different from the "almost everywhere" that we've been seeing.

As usual, with a convergence definition, we have a Cauchy definition.

Definition 8.30 (Almost uniformly Cauchy)**.** Fix a measure space $(X, \mathcal{S}, \mu)$ and a normed vector space B. Then a sequence of functions $f_n \colon X \to B$ for $n \in \mathbb{N}$ is *almost uniformly Cauchy* if and only if every $\varepsilon > 0$ has some $E^\varepsilon \in \mathcal{S}$ such that $\mu(E^\varepsilon) < \varepsilon$ and $\{f_n|_{X \setminus E}\}_{n\in\mathbb{N}}$ is uniformly Cauchy.

We take a deep breath and run some of our usual checks.

Lemma 8.31. Fix a normed vector space B and a measure space $(X, \mathcal{S}, \mu)$. Now, suppose a sequence $\{f_n\}_{n\in\mathbb{N}}$ converges almost uniformly to a function f. Then, for a function $g \colon X \to B$, we have $f = g$ almost everywhere if and only if $f_n \to g$ almost uniformly.

Proof. In one direction, suppose $f = g$ almost everywhere so that $\{x \in X : f(x) \neq g(x)\}$ is contained in some $N \in \mathcal{S}$ such that $\mu(N) = 0$. Then for any $\varepsilon > 0$, we note that $f_n \to f$ almost uniformly promises $F \in \mathcal{S}$ such that $\mu(X \setminus F) < \varepsilon$ while $f_n|_F \to f|_F$ uniformly. Now, $f|_{X\setminus N} = g|_{X\setminus N}$, so we note $f_n|_{F\setminus N} \to g|_{F\setminus N}$ uniformly (by restricting $f_n|_F \to f|_F$ uniformly) while

$$\mu(X \setminus (F \setminus N)) = \mu\big((X \setminus F) \cup N\big) = \mu(X \setminus F) + \mu(N \cap F),$$

where $\mu(N \cap F) = 0$ because $\mu(N) = 0$.

The other direction is harder. Define $N := \{x \in X : f(x) \neq g(x)\}$, and we show that N is a null set. Well, for any $d > 0$, we are promised subsets $F_d, G_d \in \mathcal{S}$ such that $\mu(F_d), \mu(G_d) < 1/d$ and $f_n|_{X\setminus F} \to f|_{X\setminus F}$ and $f_n|_{X\setminus G} \to g|_{X\setminus G}$ uniformly as $n \to \infty$.

In particular, if $x \notin (F_d \cup G_d)$, then our uniform convergence will imply pointwise convergence at x, so $f_n(x) \to f(x)$ and $f_n(x) \to g(x)$ as $n \to \infty$. It follows that $f(x) = g(x)$ by properties of convergence. Apply contraposition, we conclude that $N \subseteq F_d \cup G_d$; as such, we use Lemma 5.55 to note that

$$\mu(F_d \cup G_d) \le \mu(F_d) + \mu(G_d) < \frac{2}{d}.$$

We now send $d \to \infty$. Define

$$E := \bigcap_{d\ge 1} (F_d \cup G_d),$$

which lives in $\mathcal{S}$ because $\mathcal{S}$ is a σ-ring. As above, we see that $N \subseteq F_d \cup G_d$ for each d, so $N \subseteq E$. Further, $E \subseteq F_d \cup G_d$ tells us by Lemma 5.51 that

$$\mu(E) \le \mu(F_d \cup G_d) < \frac{2}{d}$$

for any positive integer d. In particular, sending $d \to \infty$ forces $\mu(E) = 0$, which finishes the proof that N is a null set. ■

Lemma 8.32. Fix a normed vector space B and a measure space $(X, \mathcal{S}, \mu)$. Given a sequence of functions $\{f_n\}_{n\in\mathbb{N}}$ converging to $f \colon X \to B$ almost uniformly, we have $f_{n_i} \to f$ almost uniformly for any subsequence $\{f_{n_i}\}_{i\in\mathbb{N}}$.

Proof. For any $\varepsilon > 0$, we are promised $E \in \mathcal{S}$ such that $\mu(X \setminus E) < \varepsilon$ and $f_n \to f$ uniformly on $X \setminus E$. This means that any $\delta > 0$ has some N such that $n \ge N$ implies

$$|f_n(x) - f(x)| < \delta$$

for each $x \in X \setminus E$. However, this implies that $i \ge N$ gives $n_i \ge i \ge N$ and thus

$$|f_n(x) - f(x)| < \delta$$

for each $x \in X \setminus E$, so $f_{n_i} \to f$ uniformly on $X \setminus E$ as well. ■

Lemma 8.33. Fix a normed k-vector space $(B, \|\cdot\|)$ and a measure space $(X, \mathcal{S}, \mu)$. Fix sequences of functions $f_n \colon X \to B$ and $g_n \colon X \to B$ with $f_n \to f$ and $g_n \to g$ almost uniformly as $n \to \infty$.

(a) We have $f_n + g_n \to f + g$ almost uniformly.

(b) Given some scalar $a \in k$, we have $af_n \to af$ almost uniformly.

(c) We have $\|f_n\| \to \|f\|$ almost uniformly as $n \to \infty$.

Proof. We go ahead and let $|\cdot|$ denote the norm on k. For any $\varepsilon > 0$, we will also go ahead and let $F_\varepsilon, G_\varepsilon \in \mathcal{S}$ denote the subsets of X with $\mu(F_\varepsilon), \mu(G_\varepsilon) < \varepsilon$ for which $f_n|_{X\setminus F_\varepsilon} \to f$ and $g_n|_{X\setminus G_\varepsilon} \to g$ uniformly.

(a) For any $\varepsilon > 0$, define $E_\varepsilon := F_{\varepsilon/2} \cup G_{\varepsilon/2}$. Namely, $E_\varepsilon \in \mathcal{S}$, and by Lemma 5.55, we see

$$\mu(E_\varepsilon) \le \mu(F_{\varepsilon/2}) + \mu(G_{\varepsilon/2}) = \varepsilon.$$

As such, we claim that $(f_n + g_n)|_{X\setminus E_\varepsilon} \to (f+g)|_{X\setminus E_\varepsilon}$ uniformly as $n \to \infty$.

Well, for any $\delta > 0$, we are promised N_f such that any $n \ge N_f$ and $x \notin F_{\varepsilon/2}$ will have

$$\|f(x) - f_n(x)\| < \delta/2.$$

We are promised an analogous constant N_g for g_n going to g, so we set $N := \max\{N_f, N_g\}$. Then $n \ge N$ implies $n \ge N_f$ and $n \ge N_g$; as such, if $x \notin E_\varepsilon$, then $x \notin F_{\varepsilon/2}$ and $x \notin G_{\varepsilon/2}$, so

$$\|(f+g)(x) - (f_n+g_n)(x)\| \le \|f(x) - f_n(x)\| + \|g(x) - g_n(x)\| < \frac{\delta}{2} + \frac{\delta}{2} = \delta.$$

(b) If $a = 0$, then we see that $af_n = af = 0$. As such, for any $\varepsilon > 0$, we set $E = \varnothing$ so that $\mu(E) = 0$ while $af_n \to af$ uniformly as $n \to \infty$ because $af_n(x) = 0 = af(x)$ for any $x \in X$.

Otherwise, we have $a \ne 0$ and so $|a| > 0$. Thus, for any $\varepsilon > 0$, we note F_ε will have $\mu(F_\varepsilon) < \varepsilon$, so we claim that $(af_n)|_{X\setminus F_\varepsilon} \to (af)|_{X\setminus F_\varepsilon}$ uniformly as $n \to \infty$.

Well, we already know that $f_n|_{X\setminus F_\varepsilon} \to f|_{X\setminus F_\varepsilon}$ uniformly as $n \to \infty$. Thus, for any $\delta > 0$, there is a constant N so that any $n \ge N$ and $x \notin F_\varepsilon$ will have

$$\|f(x) - f_n(x)\| < \frac{\delta}{|a|}.$$

It follows that $n \ge N$ and $x \notin F_\varepsilon$ gives

$$\|(af)(x) - (af_n)(x)\| = |a| \cdot \|f(x) - f_n(x)\| < |a| \cdot \frac{\delta}{|a|} = \delta.$$

(c) Unsurprisingly, for any $\varepsilon > 0$, we note that F_ε has $\mu(F_\varepsilon) < \varepsilon$, so we claim that $\|f\|_n|_{X\setminus F_\varepsilon} \to \|f\||_{X\setminus F_\varepsilon}$ almost uniformly as $n \to \infty$.

Well, we know that $f_n|_{X\setminus F_\varepsilon} \to f|_{X\setminus F_\varepsilon}$ as $n \to \infty$. Thus, for any $\delta > 0$, we are promised a constant N such that $n \ge N$ and $x \notin F_\varepsilon$ will have

$$\|f(x) - f_n(x)\| < \delta.$$

As such, we note that the (reverse) triangle inequality gives

$$|\,\|f(x)\| - \|f_n(x)\|\,| \le \|f(x) - f_n(x)\| < \delta,$$

which finishes. ■

Lemma 8.34. Fix a normed k-vector space $(B, \|\cdot\|)$ and a measure space $(X, \mathcal{S}, \mu)$. Fix sequences of functions $f_n\colon X \to B$ and $g_n\colon X \to B$ which are almost uniformly Cauchy.

(a) The sequence $\{f_n + g_n\}_{n\in\mathbb{N}}$ is almost uniformly Cauchy.

(b) Given some scalar $a \in k$, the sequence $\{af_n\}_{n\in\mathbb{N}}$ is almost uniformly Cauchy.

(c) The sequence $\{\|f_n\|\}_{n\in\mathbb{N}}$ is almost uniformly Cauchy.

Proof. As usual, these proofs are basically the same.

We go ahead and let $|\cdot|$ denote the norm on k. For any $\varepsilon > 0$, we will also go ahead and let $F_\varepsilon, G_\varepsilon \in \mathcal{S}$ denote the subsets of X with $\mu(F_\varepsilon), \mu(G_\varepsilon) < \varepsilon$ for which $\{f_n|_{X\setminus F_\varepsilon}\}_{n\in\mathbb{N}}$ and $\{g_n|_{X\setminus G_\varepsilon}\}_{n\in\mathbb{N}}$ are uniformly Cauchy.

(a) For any $\varepsilon > 0$, define $E_\varepsilon := F_{\varepsilon/2} \cup G_{\varepsilon/2}$. Namely, $E_\varepsilon \in \mathcal{S}$, and by Lemma 5.55, we see

$$\mu(E_\varepsilon) \le \mu(F_{\varepsilon/2}) + \mu(G_{\varepsilon/2}) = \varepsilon.$$

As such, we claim that $\{(f_n + g_n)|_{X\setminus E_\varepsilon}\}_{n\in\mathbb{N}}$ is uniformly Cauchy.

Well, for any $\delta > 0$, we are promised N_f such that any $m, n \ge N_f$ and $x \notin F_{\varepsilon/2}$ will have

$$\|f_m(x) - f_n(x)\| < \delta/2.$$

We are promised an analogous constant N_g for g_n going to g, so we set $N := \max\{N_f, N_g\}$. Then $m, n \ge N$ implies $m, n \ge N_f$ and $m, n \ge N_g$; as such, if $x \notin E_\varepsilon$, then $x \notin F_{\varepsilon/2}$ and $x \notin G_{\varepsilon/2}$, so

$$\|(f_m + g_m)(x) - (f_n + g_n)(x)\| \le \|f_m(x) - f_n(x)\| + \|g_m(x) - g_n(x)\| < \frac{\delta}{2} + \frac{\delta}{2} = \delta.$$

(b) If $a = 0$, then we see that $af_n = af_m = 0$. As such, for any $\varepsilon > 0$, we set $E = \varnothing$ so that $\mu(E) = 0$ while $\{af_n\}_{n\in\mathbb{N}}$ is uniformly Cauchy because $af_n(x) = 0 = af_m(x)$ for any $x \in X$.

Otherwise, we have $a \neq 0$ and so $|a| > 0$. Thus, for any $\varepsilon > 0$, we note F_ε will have $\mu(F_\varepsilon) < \varepsilon$, so we claim that $\{(af_n)|_{X\setminus F_\varepsilon}\}_{n\in\mathbb{N}}$ is uniformly Cauchy.

Well, we already know that $\{f_n|_{X\setminus F_\varepsilon}\}_{n\in\mathbb{N}}$ is uniformly Cauchy. Thus, for any $\delta > 0$, there is a constant N so that any $m, n \ge N$ and $x \notin F_\varepsilon$ will have

$$\|f_m(x) - f_n(x)\| < \frac{\delta}{|a|}.$$

It follows that $m, n \ge N$ and $x \notin F_\varepsilon$ gives

$$\|(af_m)(x) - (af_n)(x)\| = |a| \cdot \|f_m(x) - f_n(x)\| < |a| \cdot \frac{\delta}{|a|} = \delta.$$

(c) Unsurprisingly, for any $\varepsilon > 0$, we note that F_ε has $\mu(F_\varepsilon) < \varepsilon$, so we claim that $\{\|f\|_n|_{X\setminus F_\varepsilon}\}_{n\in\mathbb{N}}$ is uniformly Cauchy.

Well, we know that $\{f_n|_{X\setminus F_\varepsilon}\}_{n\in\mathbb{N}}$ is uniformly Cauchy. Thus, for any $\delta > 0$, we are promised a constant N such that $m, n \ge N$ and $x \notin F_\varepsilon$ will have

$$\|f_m(x) - f_n(x)\| < \delta.$$

As such, we note that the (reverse) triangle inequality gives

$$|\,\|f_m(x)\| - \|f_n(x)\|\,| \le \|f_m(x) - f_n(x)\| < \delta,$$

which finishes. ■

Now, here is the main result, which we will not prove today.

Theorem 8.35 (Riesz–Weyl)**.** Fix a measure space $(X, \mathcal{S}, \mu)$ and a normed vector space B. Let $\{f_n\}_{n\in\mathbb{N}}$ be a sequence of $\mathcal{S}$-measurable B-valued functions which are Cauchy in measure. Then there is a subsequence $\{f_n\}_{n\in\mathbb{N}}$ which is almost uniformly Cauchy.

In particular, we will be able to define a limit function for the sequence $\{f_n\}_{n\in\mathbb{N}}$ outside some null set, which will finally allow us to take limits of simple integrable functions in a way that makes sense.

8.3 November 7

Today we prove Theorem 8.35.

8.3.1 Rapidly Cauchy Intermission

As an intermission, we introduce the following definition.

Definition 8.36 (Rapidly Cauchy)**.** Fix a metric space (X, d). Then a sequence $\{x_n\}_{n\in\mathbb{N}}$ in X is *rapidly Cauchy* if and only if all $\varepsilon > 0$ have some N for which

$$\sum_{k=1}^{\infty} d(x_k, x_{k+1}) < \infty.$$

We won't use this definition in any meaningful way, but it will be enlightening to note that the main idea to the proof of Theorem 8.35 is similar to the proof that a Cauchy sequence has a rapidly Cauchy subsequence.

As such, let's see our checks on being rapidly Cauchy.

Lemma 8.37. Fix a metric space (X, d). Then any rapidly Cauchy sequence $\{x_n\}_{n\in\mathbb{N}}$ in X is also a Cauchy sequence.

Proof. Fix any $\varepsilon > 0$. We want N for which $n, m \geq N$ give $d(x_n, x_m) < \varepsilon$. Well, set $S := \sum_{k=1}^{\infty} d(x_k, x_{k+1})$, so we note that there is some N for which

$$\left| S - \sum_{k=1}^{n} d(x_k, x_{k+1}) \right| < \varepsilon$$

for each $n \geq N$. It follows that

$$S - \sum_{k=1}^{n} d(x_k, x_{k+1}) < \varepsilon$$

for any $n \geq N$. Thus, for any $m \geq n \geq N + 1$, the triangle inequality yields

$$d(x_n, x_m) \leq \sum_{k=n}^{m-1} d(x_k, x_{k+1}) = \sum_{k=1}^{m-1} d(x_k, x_{k+1}) - \sum_{k=1}^{n-1} d(x_k, x_{k+1}) \overset{*}{\leq} S - \sum_{k=1}^{n-1} d(x_k, x_{k+1}) < \varepsilon.$$

Notably, $\overset{*}{\leq}$ holds because all terms in the series of S are nonnegative, so the sequence of partial sums is increasing, so S is greater than or equal to any individual partial sum. ■

Proposition 8.38. Fix a metric space (X, d). Then any Cauchy sequence $\{x_n\}_{n\in\mathbb{N}}$ has a rapidly Cauchy subsequence.

Proof. We proceed inductively. Set $n_0 = 1$. Next, suppose we already have some n_k. Because $\{x_n\}_{n\in\mathbb{N}}$, we can find a constant $n_{k+1} \geq n_g$ such that $m, n \geq n_{k+1}$ implies $d(x_m, x_n) < 2^{-k}$. In particular, we see that $n_{k+1} \geq n_k$ in this construction tells us that

$$d(x_{n_k}, x_{n_{k+1}}) \leq 2^{-k}$$

for $j \geq 1$. Summing, we see

$$\sum_{k=1}^{\infty} d(x_k, x_{k+1}) \leq \sum_{k=1}^{\infty} 2^{-k} = 1 < \infty,$$

so $\{x_{n_k}\}_{k\in\mathbb{N}}$ is the desired rapidly Cauchy subsequence. ■

8.3.2 The Riesz–Weyl Theorem

And now, our feature presentation.

Theorem 8.35 (Riesz–Weyl)**.** Fix a measure space $(X, \mathcal{S}, \mu)$ and a normed vector space B. Let $\{f_n\}_{n\in\mathbb{N}}$ be a sequence of $\mathcal{S}$-measurable B-valued functions which are Cauchy in measure. Then there is a subsequence $\{f_n\}_{n\in\mathbb{N}}$ which is almost uniformly Cauchy.

Proof. We proceed as in Proposition 8.38. Set $n_0 = 1$. Then we proceed inductively: suppose we already know our n_k for some k, and we construct n_{k+1}. Note that

$$\lim_{m,n\to\infty} \mu\left(\{x \in X : \|f_m(x) - f_n(x)\| \geq 2^{-k}\}\right) = 0,$$

so we can find a constant $n_{k+1} > n_k$ such that $m, n \geq n_{k+1}$ gives

$$\mu\left(\{x \in X : \|f_m(x) - f_n(x)\| > 2^{-k}\}\right) < 2^{-k}.$$

We now claim that the sequence $\{f_{n_k}\}_{k\in\mathbb{N}}$ is almost uniformly Cauchy. This has two steps. Fix any $\varepsilon > 0$.

1. We select our small $E \in \mathcal{S}$ to avoid. Choose N for which

$$\sum_{k=N}^{\infty} 2^{-k} = 2^{-N+1} < \varepsilon.$$

   As such, we set

$$E_k := \{x \in X : \left\|f_{n_k}(x) - f_{n_{k+1}}(x)\right\| \geq 2^{-k}\}$$

   so that $\mu(E_k) < 2^{-k}$ by construction of the sequence $\{n_k\}_{k\in\mathbb{N}}$ because $n_k, n_{k+1} \geq n_k$. Thus, we define our E as

$$E := \bigcup_{k=N}^{\infty} E_k.$$

   Indeed, $E \in \mathcal{S}$ because $E_k \in \mathcal{S}$ for each $k \geq N$, so by Lemma 6.2, we may say

$$\mu(E) \leq \sum_{k=N}^{\infty} \mu(E_k) = \sum_{k=N}^{\infty} 2^{-k} < \varepsilon,$$

   where the last inequality is by construction of N.

2. It remains to check that the subsequence $\{f_{n_k}|_{X\setminus E}\}_{k\in\mathbb{N}}$ is uniformly Cauchy. Well, given $\delta > 0$, we need M so that $i, j > M$ and $x \notin E$ gives

$$\left\|f_{n_i}(x) - f_{n_j}(x)\right\| \stackrel{?}{<} \delta$$

for all $x \notin E$. Well, find some $M > N$ such that

$$\sum_{j \geq M} 2^{-j} = 2^{-M+1} < \delta.$$

As such, it follows from the triangle inequality that any $j > i > M$ will have

$$\left\|f_{n_i}(x) - f_{n_j}(x)\right\| \leq \sum_{k=i}^{j-1} \left\|f_{n_k}(x) - f_{n_{k+1}}(x)\right\| \overset{*}{\leq} \sum_{k=i}^{j-1} 2^{-k} < \sum_{k=M}^{\infty} 2^{-k} < \varepsilon,$$

which is what we wanted. Notably, $\overset{*}{\leq}$ holds by construction of E as a subset of E_k. ■

Example 8.39. Even in Example 8.11, there is a subsequence which is almost uniformly converging to 0. Indeed, consider the subsequence $\{1_{E_{2^n}}\}_{n\in\mathbb{N}}$. Then for any $\varepsilon > 0$, we find some N for which $2^{-N} < \varepsilon$ and set $E := \left[0, 1/2^N\right)$ to have measure less than ε. But now, for $n \geq N$, we see that $1_{E_{2^n}}|_{X\setminus E} = 0$ because $E_{2^n} \subseteq E$. Thus, $1_{E_{2^n}}|_{X\setminus E} \to 0$ uniformly as $n \to \infty$.

We are now ready to use the condition that we are integrating into a Banach space!

Lemma 8.40. Fix a measure space $(X, \mathcal{S}, \mu)$ and a Banach space B. Further, fix an almost uniformly Cauchy sequence $\{f_n\}_{n\in\mathbb{N}}$ of $\mathcal{S}$-measurable functions. Then there is an $\mathcal{S}$-measurable function $f \colon X \to B$ such that $f_n \to f$ almost uniformly as $n \to \infty$.

Proof. The main idea is that the almost uniformity condition allows us to define f outside a null set, which is good enough.

For each $n \in \mathbb{N}$, we get some E_n such that $\mu(E_n) < 1/n$ and such that $\{f_i|_{X\setminus E_n}\}_{i\in\mathbb{N}}$ is uniformly Cauchy. We now set

$$E := \bigcap_{n=1}^{\infty} E_n.$$

Note $E \in \mathcal{S}$ by Remark 5.24, and Lemma 5.51 tells us that $\mu(E) \leq \mu(E_n) < 1/n$ for each n, so it follows $\mu(E) = 0$.

Now, for any $x \in X \setminus E$, we can find k for which $x \notin E_k$. Thus, because $\{f_n|_{X\setminus E_k}\}_{n\in\mathbb{N}}$ is uniformly Cauchy, we see that $\{f_n(x)|_{X\setminus E_k}\}_{n\in\mathbb{N}}$ is Cauchy; we define $f(x)$ as its limit. Note we have used the fact that B is a Banach space here! This defines f outside the null set E.

It doesn't really matter what f does on E, so we just define $f(x) = 0$ for $x \in E$. We will quickly run checks to show that f is $\mathcal{S}$-measurable, but they are not terribly important.

- We show $f_n 1_{X\setminus E}$ is $\mathcal{S}$-measurable for each n. This proof is similar to Lemma 7.47, so we will use the ideas of that proof. For example, $\operatorname{im} f_n$ is separable, and as remarked in Lemma 7.47, we have

$$\operatorname{im} f_n 1_{X\setminus E} \subseteq \{0\} \cup \operatorname{im} f_n.$$

  As such, $\{0\} \cup \operatorname{im} f_n$ is separable by Example 7.32, so $\operatorname{im} f_n 1_{X\setminus E}$ is separable by Remark 7.31.

  Further, for any open $U \subseteq B \setminus \{0\}$, we again note from the proof of Lemma 7.47 that

$$(f_n 1_E)^{-1}(U) = (X \setminus E) \cap f_n^{-1}(U) = f_n^{-1}(U) \setminus E.$$

  In particular, $f_n^{-1}(U) \in \mathcal{S}$ because f_n is $\mathcal{S}$-measurable, so $f_n^{-1}(U) \setminus E \in \mathcal{S}$ follows.

- We show $f_n 1_{X\setminus E} \to f$ pointwise as $n \to \infty$. If $x \notin E$, then we recall that we defined $f(x)$ as the limit of $\{f_n(x)\}_{n\in\mathbb{N}}$, so the result follows by real analysis because $(f_n 1_{X\setminus E})(x) = f_n(x)$ in this case.

  Otherwise, $x \in E$, so $(f_n 1_{X\setminus E})(x) = 0$ for each n, and $f(x) = 0$ by construction of f. So this case is just looking at a constant sequence.

It follows that f is $\mathcal{S}$-measurable by Corollary 7.44.

We now check that $f_n \to f$ almost uniformly as $n \to \infty$. This is done in steps. Fix some $\varepsilon > 0$. We begin by selecting our small subset to avoid. Choose some N with $N > 1/\varepsilon$. Note $\mu(E_N) < 1/N < \varepsilon$, so we will show that $f_n|_{X\setminus E_N} \to f|_{X\setminus E_N}$ uniformly as $n \to \infty$.

For this, we proceed as in Proposition 3.24. Fix any $\delta > 0$. Because $\{f_n|_{X\setminus E_N}\}$ is uniformly, we are promised some M such that $m, n \geq M$ and $x \notin E_N$ gives

$$\|f_m(x) - f_n(x)\| < \delta/2.$$

Now, for any $x \in X$ and $n \geq M$, we see

$$\|f(x) - f_n(x)\| \leq \|f(x) - f_m(x)\| + \|f_m(x) - f_n(x)\| < \|f(x) - f_m(x)\| + \frac{\delta}{2}$$

for any $m \geq N$. However, we see $x \notin E_N$ implies $x \notin E$, so $f(x)$ was constructed to be the limit of $\{f_m(x)\}_{m\in\mathbb{N}}$, so all m sufficiently large have $\|f(x) - f_m(x)\| < \delta/2$. Ensuring that we choose an m with $m \geq M$ as well allows us to conclude

$$\|f(x) - f_n(x)\| < \|f(x) - f_m(x)\| + \frac{\delta}{2} < \delta$$

for any $n \geq N$ and $x \notin E_N$. ■

Remark 8.41. Note that the limit f is unique by Lemma 8.31.

8.4 November 9

Today we define integrable functions. We went through this discussion quickly last class but are now going through it in more detail, so I have just moved the exposition to today.

8.4.1 Convergence in Mean

We are going to want yet another notion of convergence, to align with our desire to integrate.

Definition 8.42 (Converge in mean)**.** Fix a measure space $(X, \mathcal{S}, \mu)$ and a normed vector space B. Then a sequence $\{f_n\}_{n\in\mathbb{N}}$ of simple μ-integrable functions *converges in mean* to a simple μ-integrable function f if and only if $\|f - f_n\|_1 \to 0$ as $n \to \infty$.

Definition 8.43 (Mean Cauchy)**.** Fix a normed vector space B and a measure space $(X, \mathcal{S}, \mu)$. A sequence of simple μ-integrable functions $\{f_n\}_{n\in\mathbb{N}}$ is *mean Cauchy* if and only if it is Cauchy for the semi-norm $\|\cdot\|_1$. In other words, we require

$$\lim_{m,n\to\infty} \|f_m - f_n\|_1 = 0.$$

Remark 8.44. Because simple μ-integrable functions form a vector space by Lemma 7.17, we see that $\|f - f_n\|_1$ and $\|f_m - f_n\|_1$ are legal expressions.

Here are the usual checks.

Lemma 8.45. Fix a normed k-vector space $(B, \|\cdot\|)$ and a measure space $(X, \mathcal{S}, \mu)$. Further, fix a sequence of simple μ-integrable functions $f_n \colon X \to B$ and $g_n \colon X \to B$ with $f_n \to f$ and $g_n \to g$ in mean as $n \to \infty$.

(a) We have $f_n + g_n \to f + g$ in mean.

(b) Given some scalar $a \in k$, we have $af_n \to af$ in mean.

(c) We have $\|f_n\| \to \|f\|$ in mean.

Proof. For (a) and (b), note the relevant functions are simple μ-integrable by Lemma 7.17; for (c), the relevant functions are simple μ-integrable by Lemma 7.18. Now, (a) and (b) follow directly from Lemma 1.59, where we are using the fact that $\|\cdot\|_1$ is a semi-norm by Lemma 8.6.

It remains to show (c). For any $\varepsilon > 0$, we are promised N such that $n \geq N$ implies

$$\|f - f_n\|_1 < \varepsilon.$$

By the reverse triangle inequality, we see

$$|\,\|f(x)\| - \|f_n(x)\|\,| \leq \|f(x) - f_n(x)\|$$

for each $x \in X$, so Corollary 8.4 tells us

$$|\,\|f\| - \|f_n\|\,|_1 = \int_X |\,\|f\| - \|f_n\|\,|\,d\mu \leq \int_X \|f - f_n\|\,d\mu = \|f - f_n\|_1 < \varepsilon$$

for each $n \geq N$. This finishes. ■

Lemma 8.46. Fix a normed k-vector space $(B, \|\cdot\|)$ and a measure space $(X, \mathcal{S}, \mu)$. Further, fix a sequence of simple μ-integrable functions $\{f_n\}_{n\in\mathbb{N}}$ and $\{g_n\}_{n\in\mathbb{N}}$ which are mean Cauchy.

(a) The sequence $\{f_n + g_n\}_{n\in\mathbb{N}}$ is mean Cauchy.

(b) Given some scalar $a \in k$, the sequence $\{af_n\}_{n\in\mathbb{N}}$ is mean Cauchy.

(c) The sequence $\{\|f_n\|\}_{n\in\mathbb{N}}$ is mean Cauchy.

Proof. These proofs are essentially identical. For (a) and (b), note the relevant functions are simple μ-integrable by Lemma 7.17; for (c), the relevant functions are simple μ-integrable by Lemma 7.18. As before, (a) and (b) follow from Lemma 1.65 upon noting $\|\cdot\|_1$ is a semi-norm by Lemma 8.6.

It remains to show (c). For any $\varepsilon > 0$, we are promised N such that $m, n \geq N$ implies

$$\|f_m - f_n\|_1 < \varepsilon.$$

By the reverse triangle inequality, we see

$$|\,\|f_m(x)\| - \|f_n(x)\|\,| \leq \|f(x) - f_n(x)\|$$

for each $x \in X$, so Corollary 8.4 tells us

$$|\,\|f_m\| - \|f_n\|\,|_1 = \int_X |\,\|f_m\| - \|f_n\|\,|\,d\mu \leq \int_X \|f_m - f_n\|\,d\mu = \|f_m - f_n\|_1 < \varepsilon$$

for each $m, n \geq N$. This finishes. ■

Lemma 8.47. Fix a normed k-vector space $(B, \|\cdot\|)$ and a measure space $(X, \mathcal{S}, \mu)$. Given a sequence of simple μ-integrable function $\{f_n\}_{n\in\mathbb{N}}$ which is mean Cauchy and a set $E \in \mathcal{S}$, then the sequence $\{f_n 1_E\}_{n\in\mathbb{N}}$ is still mean Cauchy.

Proof. Fix some $\varepsilon > 0$. We are promised some N such that $m, n \geq N$ implies $\|f_m - f_n\|_1 < \varepsilon$. Now, for any $x \in X$, we see

$$\|f_m 1_E - f_n 1_E\|(x) = \big(\|f_m - f_n\| 1_E\big)(x) \leq \|f_m - f_n\|(x),$$

so Corollary 8.4 tells us

$$\|f_m 1_E - f_n 1_E\|_1 = \int_X \|f_m 1_E - f_n 1_E\| \, d\mu \leq \int_X \|f_m - f_n\| \, d\mu = \|f_m - f_n\|_1,$$

which is less than ε for $m, n \geq N$. This is what we wanted. ■

Lemma 8.48. Fix a normed vector space $(B, \|\cdot\|)$ and a measure space $(X, \mathcal{S}, \mu)$. If $\{f_n\}_{n\in\mathbb{N}}$ is a mean Cauchy sequence of simple μ-integrable functions, then any subsequence $\{f_{n_i}\}_{i\in\mathbb{N}}$ is a mean Cauchy sequence of simple μ-integrable functions.

Proof. For any $\varepsilon > 0$, we are given N such that $m, n \geq N$ implies $\|f_m - f_n\|_1 < \varepsilon$. Because $n_i \geq n$ for each i, we see $i, j \geq N$ has $\|f_{n_i} - f_{n_j}\|_1 < \varepsilon$ as well, which is what we wanted. ■

8.4.2 Comparing Convergences

We are going to want to see the comparative strengths of different convergences. Here is a starting result, which was moved from an earlier lecture for thematic reasons. Note this generalizes Example 8.14.

Lemma 8.49. Fix a normed vector space $(B, \|\cdot\|)$ and a measure space $(X, \mathcal{S}, \mu)$. Then a sequence of simple μ-integrable functions $f_n \colon X \to B$ for $n \in \mathbb{N}$ which is mean Cauchy is also Cauchy in measure.

Proof. Fix $\varepsilon > 0$ and set

$$E_{m,n}^{\varepsilon} \coloneqq \{x \in X : \|f_m(x) - f_n(x)\| \geq \varepsilon\},$$

which has finite measure by Remark 8.17. We need to show that

$$\lim_{m,n\to\infty} \mu(E_{m,n}^{\varepsilon}) \stackrel{?}{=} 0.$$

Notably, for each $x \in X$, we must have

$$1_{E_{m,n}^{\varepsilon}}(x) \leq \frac{\|f_m(x) - f_n(x)\|}{\varepsilon} \tag{8.2}$$

by definition of $E_{m,n}^{\varepsilon}$. Now, both sides of this equation are simple μ-integrable functions: $1_{E_{m,n}^{\varepsilon}}$ is by Example 7.6; and $f_m - f_n$ is simple μ-integrable by Lemma 7.17, as is $\|f_m - f_n\|$ by Lemma 7.18, so $\frac{1}{\varepsilon}\|f_m - f_n\|$ is simple μ-integrable by Lemma 7.17 again.

Thus, we may integrate, for which Corollary 8.4 tells us

$$\mu(E_{m,n}^{\varepsilon}) = \int_X 1_{E_{m,n}^{\varepsilon}} \, d\mu \leq \int_X \frac{\|f_m - f_n\|}{\varepsilon} \, d\mu = \frac{\|f_m - f_n\|_1}{\varepsilon},$$

where the first integral was computed using Example 7.16. But as $m, n \to \infty$, the right-hand value goes to 0 because $\{f_n\}_{n\in\mathbb{N}}$ is Cauchy for $\|\cdot\|_1$, so the left-hand value must also go to 0. ■

Remark 8.50. A similar proof works for when we are Cauchy for $\|\cdot\|_p$ for finite p by taking pth powers of (8.2). For example, in probability theory, the result for $\|\cdot\|_2$ is essentially Chebyshev's inequality.

We now note that converging almost uniformly is stronger than in measure.

Lemma 8.51. Fix a normed vector space $(B, \|\cdot\|)$ and a measure space $(X, \mathcal{S}, \mu)$. Further, fix a sequence $\{f_n\}_{n\in\mathbb{N}}$ of $\mathcal{S}$-measurable functions and an $\mathcal{S}$-measurable function f.

(a) If $f_n \to f$ almost uniformly as $n \to \infty$, then $f_n \to f$ in measure.

(b) If $\{f_n\}_{n\in\mathbb{N}}$ is almost uniformly Cauchy, then $\{f_n\}_{n\in\mathbb{N}}$ is Cauchy in measure.

Proof. Here we go.

(a) For any $\varepsilon > 0$, we need to show that

$$\lim_{n\to\infty} \mu(\{x \in X : \|f(x) - f_n(x)\|\}) \stackrel{?}{=} 0.$$

Well, for any $\delta > 0$, we need N such that $n \geq N$ has

$$\mu(\{x \in X : \|f(x) - f_n(x)\|) \stackrel{?}{<} \delta.$$

Now, by the almost uniform convergence, we are promised $F \in \mathcal{S}$ such that $\mu(X \setminus F) < \delta$ and $f_n \to f$ uniformly as $n \to \infty$ on F. Now using our uniform convergence, we choose N such that $n \geq N$ implies

$$\|f(x) - f_n(x)\| < \varepsilon$$

for each $x \in F$. In particular, for $n \geq N$, we see

$$\{x \in X : \|f(x) - f_n(x)\| \geq \varepsilon\} \subseteq X \setminus F,$$

so Lemma 5.51 tells us

$$\mu(\{x : \|f(x) - f_n(x)\| \geq \varepsilon\}) \leq \mu(X \setminus F) < \delta,$$

which finishes.

(b) This proof is essentially the same. For any $\varepsilon > 0$, we need to show that

$$\lim_{n\to\infty} \mu(\{x \in X : \|f_m(x) - f_n(x)\|\}) \stackrel{?}{=} 0.$$

Well, for any $\delta > 0$, we need N such that $m, n \geq N$ has

$$\mu(\{x \in X : \|f_m(x) - f_n(x)\|) \stackrel{?}{<} \delta.$$

Now, by the almost uniform convergence, we are promised $F \in \mathcal{S}$ such that $\mu(X \setminus F) < \delta$ and $\{f_n\}_{n\in\mathbb{N}}$ is uniformly Cauchy on F. Now using the fact we're uniformly Cauchy, we choose N such that $m, n \geq N$ implies

$$\|f_m(x) - f_n(x)\| < \varepsilon$$

for each $x \in F$. In particular, for $m, n \geq N$, we see

$$\{x \in X : \|f_m(x) - f_n(x)\| \geq \varepsilon\} \subseteq X \setminus F,$$

so Lemma 5.51 tells us

$$\mu(\{x : \|f_m(x) - f_n(x)\| \geq \varepsilon\}) \leq \mu(X \setminus F) < \delta,$$

which finishes. ■

Further, convergence almost uniformly is stronger than convergence almost everywhere.

Lemma 8.52. Fix a normed vector space $(B, \|\cdot\|)$ and a measure space $(X, \mathcal{S}, \mu)$. Further, fix a sequence $\{f_n\}_{n\in\mathbb{N}}$ of functions which converge to f almost uniformly as $n \to \infty$. Then $f_n \to f$ almost everywhere.

Proof. Let E be the set of points such that $\{f_n(x)\}_{n\in\mathbb{N}}$ does not converge to $f(x)$ as $n \to \infty$. Note $x \in E$ is equivalent to having some $\varepsilon > 0$ and N such that $\|f_n(x) - f(x)\| \geq \varepsilon$ for each $n \geq N$, so

$$E = \bigcup_{\varepsilon>0} \bigcup_{N=1}^{\infty} \bigcap_{n\geq N} \{x \in X : \|f(x) - f_n(x)\| \geq \varepsilon\}.$$

Note that $\|f_n(x) - f(x)\| \geq \varepsilon$ for some $\varepsilon > 0$ and all $n \geq N$ is equivalent to $\|f(x) - f_n(x)\| \geq 1/m$ for some m and all $n \geq N$ because $1/m \to 0$ as $m \to \infty$. So in fact

$$E = \bigcup_{m=1}^{\infty} \bigcup_{N=1}^{\infty} \bigcap_{n\geq N} \{x \in X : \|f(x) - f_n(x)\| \geq 1/m\}.$$

By Lemma 6.2, it suffices to show that

$$\mu\left(\bigcap_{n\geq N} \{x \in X : \|f(x) - f_n(x)\| \geq 1/m\} \right) \stackrel{?}{=} 0$$

for each m and N.

Well, $f_n \to f$ almost uniformly, so any $\delta > 0$ has $F \in \mathcal{S}$ such that $\mu(X \setminus F) < \delta$ and $f_n \to f$ uniformly as $n \to \infty$. Thus, there is N' such that $n \geq N'$ implies $\|f(x) - f_n(x)\| < 1/m$ for all $x \in X$ so that

$$\{x \in X : \|f(x) - f_n(x)\| \geq 1/m\} \subseteq X \setminus F.$$

In particular, choosing any n_0 greater than both N and N', we see from Lemma 5.51 that

$$\mu\left(\bigcap_{n\geq N} \{x \in X : \|f(x) - f_n(x)\| \geq 1/m\} \right) \leq \mu(\{x \in X : \|f(x) - f_{n_0}(x)\| \geq 1/m\}) \leq \mu(X \setminus F) < \delta.$$

It follows $\mu\left(\bigcap_{n\geq N}\{x \in X : \|f(x) - f_n(x)\| \geq 1/m\}\right) = 0.$ ■

8.4.3 Integrable Functions

Our payoff to our hard work is a definition of integrable functions. Here it is.

Theorem 8.53. Fix a normed k-vector space $(B, \|\cdot\|)$ and a measure space $(X, \mathcal{S}, \mu)$. Then given an $\mathcal{S}$-measurable function f, the following are equivalent.

(a) There is a mean Cauchy sequence of simple μ-integrable functions that converges to f in measure.

(b) There is a mean Cauchy sequence of simple μ-integrable functions that converges to f almost uniformly.

(c) There is a mean Cauchy sequence of simple μ-integrable functions that converges to f almost everywhere.

Proof. We show our implications in sequence. In all parts, let $\{f_n\}_{n\in\mathbb{N}}$ be the requested mean Cauchy sequence of simple μ-integrable functions.

- We show (a) implies (b). This holds from the Riesz–Weyl theorem. Namely, by Theorem 8.35, $\{f_n\}_{n\in\mathbb{N}}$ will have a subsequence $\{f_{n_i}\}_{i\in\mathbb{N}}$ which is almost uniformly Cauchy; this subsequence remains mean Cauchy by Lemma 8.48.

  It remains to show that $f_{n_i} \to f$ almost uniformly as $i \to \infty$. By Lemma 8.40, we see that $f_{n_i} \to g$ almost uniformly for some $\mathcal{S}$-measurable function $g\colon X \to B$, but then Lemma 8.51 tells us $f_{n_i} \to g$ in measure.

  However, $f_n \to f$ in measure implies that $f_{n_i} \to f$ in measure by Lemma 8.21, so $f = g$ almost everywhere by Lemma 8.23, so $f_{n_i} \to f$ almost uniformly by Lemma 8.31.

- We show (b) implies (c) and (a). Well, converging almost uniformly automatically forces us to converge in measure by Lemma 8.51 and almost everywhere by Lemma 8.52.

- We show (c) implies (a). Well, if $\{f_n\}_{n\in\mathbb{N}}$ is mean Cauchy, then the sequence is Cauchy in measure by Lemma 8.51 and therefore has a subsequence $\{f_{n_i}\}_{i\in\mathbb{N}}$ which is almost uniformly Cauchy by Theorem 8.35. Notably, $f_{n_i} \to f$ almost everywhere because $f_n \to f$ almost everywhere, using the same null set.

  However, this subsequence $\{f_{n_i}\}_{i\in\mathbb{N}}$ will then converge to some $\mathcal{S}$-measurable $g\colon X \to B$ almost uniformly by Lemma 8.40, so $f_{n_i} \to g$ almost everywhere by Lemma 8.52. It follows that $f = g$ almost everywhere,[1] so $f_{n_i} \to f$ almost uniformly by Lemma 8.31. ■

As such, we have the following definition.

Definition 8.54 (Integrable)**.** Fix a measure space $(X, \mathcal{S}, \mu)$ and a normed vector space B. Then an $\mathcal{S}$-measurable function $f\colon X \to B$ is *μ-integrable* if and only if one of the equivalent conditions from Theorem 8.53 is satisfied. This set of integrable functions is often denoted $\mathcal{L}^1(X, \mathcal{S}, \mu, B)$, where some data might be omitted when we want to.

Remark 8.55. Later on, we will define

$$\int_X f\,d\mu := \lim_{n\to\infty} \int_X f_n\,d\mu.$$

However, we have not yet checked that this definition is well-defined.

Remark 8.56. Later on we will also define $\mathcal{L}^\infty(X, \mathcal{S}, \mu, B)$ as the bounded $\mathcal{S}$-measurable functions as well as more general $\mathcal{L}^p(X, \mathcal{S}, \mu, B)$ for finite p where

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

As an example fact, we can see that $\mathcal{L}^1(X, \mathcal{S}, \mu, B)$ is a module over the ring $\mathcal{L}^\infty(X, \mathcal{S}, \mu, k)$, where B is a normed k-vector space. We will not check this here.

Remark 8.57. Morally perhaps, one should define integrable functions to be merely μ-measurable instead of $\mathcal{S}$-measurable. I have not done this for technical reasons because I find it exceedingly annoying to have to keep removing a null set. If this distinction is distressing, then replace $\mathcal{S}$ with the σ-algebra generated by $\mathcal{S}$ and the null sets of μ.

Example 8.58. If f is a simple μ-integrable function, then the sequence $\{f\}_{n\in\mathbb{N}}$ is mean Cauchy and converges to f almost everywhere, so f is also a μ-integrable function.

[1] We know $f_n \to f$ outside some null set F, and $f_n \to g$ outside some null set G, so $f(x) = g(x)$ outside the null set $F \cup G$.

Here are the usual checks.

Lemma 8.59. Fix a measure space $(X, \mathcal{S}, \mu)$ and a normed k-vector space B. Then $\mathcal{L}^1(X, \mathcal{S}, \mu, B)$ forms a k-vector space.

Proof. Here are our checks. For brevity, set $\mathcal{L}^1 := \mathcal{L}^1(X, \mathcal{S}, \mu, B)$.

- Zero: note that the zero function is $1_\varnothing$ and thus a simple μ-integrable function (by Example 7.16) and thus a simple μ-integrable function (by Example 8.58).
- Addition: given $f, g \in \mathcal{L}^1$, we show $f + g \in \mathcal{L}^1$. Well, pick up mean Cauchy sequences $\{f_n\}_{n\in\mathbb{N}}$ and $\{g_n\}_{n\in\mathbb{N}}$ of simple μ-integrable functions which converge in measure to f and g respectively.

  Now, note $\{f_n+g_n\}_{n\in\mathbb{N}}$ is mean Cauchy by Lemma 8.46, and $f_n+g_n \to f+g$ in measure by Lemma 8.20, so $f + g \in \mathcal{L}^1$.
- Scalar multiplication: given a scalar $a \in k$ and $f \in \mathcal{L}^1$, we show $af \in \mathcal{L}^1$. Well, pick up our mean Cauchy sequence $\{f_n\}_{n\in\mathbb{N}}$ of simple μ-integrable functions which converges to f in measure. Then $\{af_n\}_{n\in\mathbb{N}}$ is mean Cauchy by Lemma 8.46 and converges in measure to af by Lemma 8.19. ■

Lemma 8.60. Fix a measure space $(X, \mathcal{S}, \mu)$ and a normed vector space B. Given a μ-integrable function $f: X \to B$ and measurable set $E \in \mathcal{S}$, the function $f1_E$ is still μ-integrable.

Proof. As usual, pick up our mean Cauchy sequence $\{f_n\}_{n\in\mathbb{N}}$ of simple μ-integrable functions converging to f in measure. Then Lemma 7.47 tells us that f_n1_E is still simple μ-integrable. Further, Lemma 8.47 tells us $\{f_n1_E\}_{n\in\mathbb{N}}$ is still mean Cauchy, and Lemma 8.24 tells us $f_n1_E \to f1_E$ in measure. Thus, $f1_E$ is in fact μ-integrable. ■

Remark 8.61. As in Remark 7.48, we note that $E \in \mathcal{S}$ will have $1_X = 1_E + 1_{X\setminus E}$, so $f1_{X\setminus E} = f - f1_E$ is still μ-integrable by Lemma 8.60 and Lemma 8.59.

Lemma 8.62. Fix a measure space $(X, \mathcal{S}, \mu)$ and a normed vector space B. Given a μ-integrable function $f: X \to B$, the function $\|f\|$ is still μ-integrable.

Proof. As usual, pick up our mean Cauchy sequence $\{f_n\}_{n\in\mathbb{N}}$ of simple μ-integrable functions converging to f in measure. Then Lemma 7.18 tells us that each $\|f_n\|$ is still a simple μ-integrable function. As such, we see Lemma 8.46 tells us $\{\|f_n\|\}_{n\in\mathbb{N}}$ is mean Cauchy, and $\|f_n\| \to \|f\|$ in measure by Lemma 8.19. It follows $\|f\|$ is μ-integrable. ■

Example 8.63. We mimic Example 7.46. If $f: X \to \mathbb{R}$ is μ-measurable, then Lemma 8.62 tells us that $|f|$ is also μ-measurable. As such, if $f, g: X \to \mathbb{R}$ are $\mathcal{S}$-measurable, then $(f + g)$ and $(f - g)$ are $\mathcal{S}$-measurable by Lemma 8.59, so $|f - g|$ is $\mathcal{S}$-measurable, so

$$\min\{f, g\} = \frac{(f+g)+|f-g|}{2} \qquad \text{and} \qquad \max\{f, g\} = \frac{(f+g)-|f-g|}{2}$$

are μ-measurable by Lemma 8.59 again. Inducting, for any μ-measurable functions $\{f_i\}_{i=1}^n$, the minimum function $\min\{f_1, \ldots, f_n\}$ and maximum function $\max\{f_1, \ldots, f_n\}$ are both μ-measurable.

8.4.4 Towards Defining Integrals

We now move towards defining integration.

Lemma 8.64. Fix a measure space $(X, \mathcal{S}, \mu)$ and a normed k-vector space B. Further, fix mean Cauchy sequences of simple μ-integrable functions $\{f_n\}_{n\in\mathbb{N}}$ and $\{g_n\}_{n\in\mathbb{N}}$ which converges to f and g in measure, respectively. If $\|f_n - g_n\|_1 \to 0$ as $n \to \infty$, then $f = g$ almost everywhere.

Proof. The key trick is to consider the sequence $f_1, g_1, f_2, g_2, \ldots$. To be explicit, define $\{h_n\}_{n\in\mathbb{N}}$ by $h_{2n} = f_n$ and $h_{2n-1} = g_n$. Here are our checks on $\{h_n\}_{n\in\mathbb{N}}$.

- Note that each $n \in \mathbb{N}$ has h_n is either an f_i or g_i and is therefore a simple μ-integrable functions.
- We claim that $\{h_n\}_{n\in\mathbb{N}}$ is Cauchy in measure; it suffices to show that $\{h_n\}_{n\in\mathbb{N}}$ is mean Cauchy by Lemma 8.49.

  Well, fix any $\varepsilon > 0$. Because $\{f_n\}_{n\in\mathbb{N}}$ and $\{g_n\}_{n\in\mathbb{N}}$ are mean Cauchy, we get N_f and N_g such that

  $$m, n \geq N_f \implies \|f_m - f_n\|_1 < \varepsilon \qquad \text{and} \qquad m, n \geq N_g \implies \|f_m - f_n\|_1 < \varepsilon.$$

  Further, $\|f_n - g_n\|_1 \to 0$ as $n \to \infty$, so we get N' such that $n \geq N'$ implies $\|f_n - g_n\|_1 < \varepsilon$.

  Combining, set $N := \max\{2N_f, 2N_g + 1, 2N'\}$. Then, for $m, n \geq N$, we have three cases according to parity.

  - If $m = 2k$ and $n = 2\ell$, then $k, \ell \geq N_f$, so $\|h_m - h_n\|_1 = \|f_k - f_\ell\|_1 < \varepsilon$.
  - If $m = 2k+1$ and $n = 2\ell + 1$

  The case with m odd and n even is analogous to the last one, by symmetry. This finishes our check.
- Note $h_{2n} = f_n$ for each n, so the subsequence $\{h_{2n}\}_{n\in\mathbb{N}}$ converges to f in measure. Thus, Lemma 8.22 tells us $h_n \to f$ in measure.
- Analogously, note $h_{2n-1} = g_n$ for each n, so the subsequence $\{h_{2n-1}\}_{n\in\mathbb{N}}$ converges to g in measure. Thus, Lemma 8.22 tells us $h_n \to g$ in measure.

From the above checks, we see from Lemma 8.23 that $f = g$ almost everywhere. ■

The point here is that we can take equivalence classes in $\mathcal{L}^1(X, \mathcal{S}, \mu, B)$ to get a bona fide norm from our semi-norm $\|\cdot\|_1$.

To finish our discussion of completeness, we will need the following result, which we will state but not prove today.

Proposition 8.65. Fix a measure space $(X, \mathcal{S}, \mu)$ and a normed vector space $(B, \|\cdot\|)$. Suppose $\{f_n\}_{n\in\mathbb{N}}$ and $\{g_n\}_{n\in\mathbb{N}}$ are mean Cauchy sequences of simple μ-integrable functions which both converge to some $\mathcal{S}$-measurable function f in measure. Then $\|f_n - g_n\|_1 \to 0$ as $n \to \infty$.

Roughly speaking, this will imply that the integral $\int_X f\,d\mu$ is well-defined.

8.5 November 14

Today we show that the space L^1 is complete. Here is a challenge problem.

Remark 8.66. Here is a challenge problem. Fix a sequence of continuous functions $f_n \colon [0,1] \to [0,1]$. Show that if $f_n \to f$ pointwise, then $\|f_n\|_1 \to 0$ as $n \to \infty$, where $\|f_n\|_1$ is defined using the Riemann integral. There are proofs which do not use any measure theory!

8.5.1 Equivalent Mean Cauchy Sequences

Last class we were about to prove the following result.

Proposition 8.65. Fix a measure space $(X, \mathcal{S}, \mu)$ and a normed vector space $(B, \|\cdot\|)$. Suppose $\{f_n\}_{n\in\mathbb{N}}$ and $\{g_n\}_{n\in\mathbb{N}}$ are mean Cauchy sequences of simple μ-integrable functions which both converge to some $\mathcal{S}$-measurable function f in measure. Then $\|f_n - g_n\|_1 \to 0$ as $n \to \infty$.

For this proof, we will want the following lemma.

Lemma 8.67. Fix a measure space $(X, \mathcal{S}, \mu)$ and a mean Cauchy sequence of nonnegative simple μ-integrable functions $f_n \colon X \to \mathbb{R}$. If $f_n \to 0$ in measure, then $\|f_n\|_1 \to 0$.

Proof. By Lemma 8.49, we see $\{f_n\}_{n\in\mathbb{N}}$ is Cauchy in measure, so we may use Theorem 8.35 to extract an almost uniformly Cauchy subsequence $\{f_{n_i}\}_{i\in\mathbb{N}}$, which we then see almost uniformly converges to 0 by Lemma 8.32.

For another reduction, we note that $\{\|f_n\|_1\}_{n\in\mathbb{N}}$ is a Cauchy sequence: for any $\varepsilon > 0$, find N such that $\|f_m - f_n\|_1 < \varepsilon$ for $m, n \geq N$. Then

$$|\,\|f_m\|_1 - \|f_n\|_1\,| \leq \|f_m - f_n\|_1 < \varepsilon$$

for $m, n \geq N$ by the (reverse) triangle inequality from Lemma 8.6. As such, $\{\|f_n\|_1\}_{n\in\mathbb{N}}$ does converge to some real number r, and the subsequence $\{f_{n_i}\}_{i\in\mathbb{N}}$ will also converge to the same real number r. So we will show that $\|f_{n_i}\|_1 \to 0$ as $i \to \infty$.

Now, to simplify notation, set $g_i := f_{n_i}$ so that $g_i \to 0$ almost uniformly, and we want to show $\|g_i\|_1 \to 0$ as $i \to \infty$. Fix $\varepsilon > 0$; we want N such that $n \geq N$ has

$$0 \leq \|g_n\|_1 \overset{?}{<} \varepsilon.$$

This means we have to bound an integral, which we do in many pieces. To begin, our sequence $\{g_n\}_{n\in\mathbb{N}}$ is mean Cauchy, so we start with some N_1 such that $m, n \geq N_1$ implies $\|g_m - g_n\|_1 < \varepsilon/4$. Now here are the pieces of our integral.

1. Set $F := g_{N_1}^{-1}(B \setminus \{0\})$, which is in $\mathcal{S}$ by Lemma 7.34 and has finite measure by definition of a simple μ-integrable function. Now, for $n \geq N_1$, we see any $x \in X \setminus F$ has

$$g_n(x) = |g_n(x) - g_{N_1}(x)|,$$

so

$$\int_X g_n 1_{X\setminus F}\,d\mu = \int_X |g_n(x) - g_N(x)| 1_{X\setminus F}\,d\mu \overset{*}{\leq} \int_X |g_n(x) - g_N(x)|\,d\mu = \|g_n - g_N\|_1 < \frac{\varepsilon}{4},$$

where we have used Corollary 8.4 in $\overset{*}{\leq}$. (Note $g_n 1_{X\setminus F}$ is simple μ-integrable by Remark 7.48.)

2. To continue, we recall $g_n \to 0$ almost uniformly, so we use $\delta := \frac{\varepsilon}{4(1+\|g_N\|_\infty)} > 0$ to find $G \in \mathcal{S}$ with $\mu(G) < \delta$ and $g_n \to 0$ uniformly on $X \setminus G$.

As such, we can choose N_2 for which $n \geq N_2$ has

$$g_n(x) \leq \frac{\varepsilon}{4(1+\mu(F))}$$

for each $x \in X \setminus G$. Integrating, we see $n \geq N_2$ gives

$$\int_X g_n 1_{F\setminus G}\,d\mu \leq \int_X \left(\frac{\varepsilon}{4(1+\mu(F))} \cdot 1_{F\setminus G}\right)\,d\mu$$

by Corollary 8.4. (Note $g_n1_{F\setminus G} = g_n1_F - g_n1_{F\cap G}$ is a simple μ-integrable function by Lemma 7.47 and Lemma 7.17.) Using Lemma 8.1 and then Example 7.16 to compute the integral, we see

$$\int_X g_n 1_{F\setminus G}\,d\mu \le \frac{\varepsilon}{4(1+\mu(F))}\cdot\mu(F\setminus G) < \frac{\varepsilon}{4(1+\mu(F))}\cdot(1+\mu(F)) \le \frac{\varepsilon}{4},$$

where $\mu(F\setminus G)\le\mu(F)$ by Lemma 5.51.

3. It remains to handle what's happening on $F\cap G$. Well, $\mu(F\cap G)\le\mu(G)<\delta$ by Lemma 5.51, so whatever happens here is pretty small. Indeed, note any n gives

$$g_n1_{F\cap G}\le|g_n-g_N|1_{F\cap G}+|g_N|1_{F\cap G}\le|g_n-g_N|+\|g_N\|_\infty 1_G,$$

so Corollary 8.4 and Lemma 8.1 tells us

$$\int_X g_n1_{F\cap G}\,d\mu \le \underbrace{\int_X |g_n-g_N|\,d\mu}_{\|g_n-g_N\|_1}+\|g_N\|_\infty\int_X 1_G\,d\mu.$$

(As usual, the relevant restricted functions are simple μ-integrable by Lemma 7.47.) By Example 7.16, we see

$$\int_X 1_G\,d\mu=\mu(G)<\delta=\frac{\varepsilon}{4(1+\|g_N\|_\infty)},$$

so we have the bound

$$\int_X g_n1_{F\cap G}\,d\mu\le\|g_n-g_N\|_1+\|g_N\|_\infty\cdot\frac{\varepsilon}{4(1+\|g_N\|_\infty)}<\|g_n-g_N\|_1+\frac{\varepsilon}{4}.$$

As such, $n\ge N_1$ will give

$$\int_X g_n1_{F\cap G}\,d\mu<\frac{\varepsilon}{2}.$$

In total, we set $N:=\max\{N_1,N_2\}$ so that $n\ge N$ implies

$$\int_X g_n\,d\mu=\int_X g_n1_{X\setminus F}\,d\mu+\int_X g_n1_{F\setminus G}\,d\mu+\int_X g_n1_{F\cap G}\,d\mu<\frac{\varepsilon}{4}+\frac{\varepsilon}{4}+\frac{\varepsilon}{2}=\varepsilon,$$

where we have used Lemma 8.1 in the first equality. This finishes. ■

We are now ready to prove Proposition 8.65.

Proof of Proposition 8.65. Unsurprisingly, set $h_n:=f_n-g_n$, which is a simple μ-integrable function by Lemma 7.17. Note $\{h_n\}_{n\in\mathbb{N}}$ is a mean Cauchy sequence by Lemma 8.46 and converges to $f-f=0$ in measure by Lemma 8.19.

We want to show that $\|h_n\|_1\to0$ as $n\to\infty$, so we define $j_n:=\|h_n\|$, which is mean Cauchy by Lemma 8.46. Furthermore, $j_n\to0$ in measure by Lemma 8.19 (note $\|0\|=0$), so Lemma 8.67 tells us that $\|j_n\|_1\to0$ as $n\to\infty$. However, for any n,

$$\|j_n\|_1=\int_X|j_n|\,d\mu=\int_X\|h_n\|\;d\mu=\|h_n\|_1\,,$$

so it follows $\|h_n\|_1\to0$ as $n\to\infty$ as well. ■

The above result grants us a natural bijection between equivalence classes of mean Cauchy sequences of simple μ-integrable functions and "almost everywhere" equivalence classes of μ-integrable functions. So we have constructed our completion of simple μ-integrable functions.

Remark 8.68. As an aside, we note that the $\|\cdot\|_1$ norm is pretty poorly behaved at points. For example, the function $C([0,1])\to[0,1]$ by $f\mapsto f(1)$ is not continuous for $\|\cdot\|_1$. Namely, define $f_n(x)=x^n$ so that $f_n\to0$ in mean (we will be able to check this eventually) as $n\to\infty$, but $f_n(1)\to1$ as $n\to\infty$.

8.5.2 Defining Integrals

The main use of Proposition 8.65 is the following corollary.

Corollary 8.69. Fix a measure space $(X, \mathcal{S}, \mu)$ and a normed vector space $(B, \|\cdot\|)$. Given mean Cauchy sequences $\{f_n\}_{n\in\mathbb{N}}$ and $\{g_n\}_{n\in\mathbb{N}}$ of simple μ-integrable functions both converging to an $\mathcal{S}$-measurable function f in measure, we have

$$\lim_{n\to\infty} \int_X f_n \, d\mu = \lim_{n\to\infty} \int_X g_n \, d\mu.$$

Namely, if the limits exist, then they are equal. If B is a Banach space, then the limits exist.

Proof. There are two claims here, which we will show in sequence.

- Suppose the limits exist. We show they are equal. Using Lemma 8.1, it suffices to show that

  $$\lim_{n\to\infty} \int_X (f_n - g_n) \, d\mu$$

  vanishes. Well, by Proposition 8.65, we see $\|f_n - g_n\|_1 \to 0$ as $n \to \infty$. Thus, for any $\varepsilon > 0$, we are promised some N for which $n \ge N$ has $\|f_n - g_n\|_1 < \varepsilon$. But then Lemma 8.2 implies

  $$\left\| \int_X (f_n - g_n) \, d\mu \right\| \le \int_X \|f_n - g_n\| \, d\mu = \|f_n - g_n\|_1 < \varepsilon$$

  for $n \ge N$, which is what we wanted.

- Now suppose that B is a Banach space, and we must show the limits exist. By symmetry, it suffices to show that $\lim_{n\to\infty} \int_X f_n \, d\mu$ exists. Because B is complete, it suffices to show that the sequence $\int_X f_n \, d\mu$ of elements in B is Cauchy.

  Well, fix some $\varepsilon > 0$. We see $\{f_n\}_{n\in\mathbb{N}}$ is mean Cauchy, so there is some N such that $m, n \ge N$ implies $\|f_m - f_n\|_1 < \varepsilon$. We now bound. Using Lemma 8.1 and Lemma 8.2, we see

  $$\left\| \int_X f_m \, d\mu - \int_X f_n \, d\mu \right\| = \left\| \int_X (f_m - f_n) \, d\mu \right\| \le \int_X \|f_m - f_n\| \, d\mu = \|f_m - f_n\|_1 ,$$

  which is less than ε for $m, n \ge N$. This finishes. ■

Remark 8.70. It is not too hard to extend the above proof to show that if just one of the limits exist, then both of them exist. We will not need this.

As such, we are prepared to finally define integrals.

Definition 8.71 (Integral)**.** Fix a measure space $(X, \mathcal{S}, \mu)$ and a Banach space B. Given an integrable function $f \colon X \to B$, find the corresponding sequence mean Cauchy sequence $\{f_n\}_{n\in\mathbb{N}}$ of simple μ-integrable functions with $f_n \to f$ in measure. Then we define the *integral* by

$$\int_X f \, d\mu := \lim_{n\to\infty} \int_X f_n \, d\mu.$$

Example 8.72. If $f \colon X \to B$ is already a simple μ-integrable function, then $\{f\}_{n\in\mathbb{N}}$ is mean Cauchy with $f \to f$ in measure, so our new integral $\int_X f \, d\mu$ takes the intended value.

Note that this limit exists and is well-defined by Corollary 8.69. We now pick up some facts about our integral. The main theme here is to just reduce these facts to the corresponding one about simple μ-integrable functions.

Proposition 8.73. Fix a measure space $(X, \mathcal{S}, \mu)$ and a Banach k-space B. Further, fix μ-integrable functions f and g and scalars $a, b \in k$. Then

$$\int_X (af + bg)\, d\mu = a \int_X f\, d\mu + b \int_X g\, d\mu.$$

Proof. Because f and g are μ-integrable, we are promised mean Cauchy sequences $\{f_n\}_{n\in\mathbb{N}}$ and $\{g_n\}_{n\in\mathbb{N}}$ of simple μ-integrable functions such that $f_n \to f$ and $g_n \to g$ in measure.

Now, it follows from the proof of Lemma 8.59 that $\{af_n\}_{n\in\mathbb{N}}$ and $\{bg_n\}_{n\in\mathbb{N}}$ are mean Cauchy sequences of simple μ-integrable functions converging to af and bg in measure, so $\{af_n + bg_n\}_{n\in\mathbb{N}}$ is a mean Cauchy sequence of simple μ-integrable functions converging to $af + bg$ in measure. As such, we begin by using Lemma 8.1 to compute

$$\int_X (af_n + bg_n)\, d\mu = a \int_X f_n\, d\mu + b \int_X g_n\, d\mu,$$

so

$$\int_X (af + bg)\, d\mu = \lim_{n\to\infty} \int_X (af_n + bg_n)\, d\mu = a \lim_{n\to\infty} \int_X f_n\, d\mu + b \lim_{n\to\infty} \int_X g_n\, d\mu = a \int_X f\, d\mu + b \int_X g\, d\mu,$$

which is what we wanted. ■

Here are the usual bounding results.

Lemma 8.74. Fix a measure space $(X, \mathcal{S}, \mu)$. Given a μ-integrable function $f\colon X \to \mathbb{R}$, if $f(x) \ge 0$ almost everywhere, we have

$$\int_X f\, d\mu \ge 0.$$

Proof. The main point is that $f = |f|$ almost everywhere. Indeed, we are promised some $E \in \mathcal{S}$ such that $\mu(E) = 0$ and $f(x) \ge 0$ for $x \in X \setminus E$. Now, pick up our mean Cauchy sequence $\{f_n\}_{n\in\mathbb{N}}$ of simple μ-integrable functions such that $f_n \to f$ in measure. It follows from the proof of Lemma 8.62 that $\{|f_n|\}_{n\in\mathbb{N}}$ is also a mean Cauchy sequence of simple μ-integrable functions but with $|f_n| \to |f|$ in measure. However,

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

for each $x \in X \setminus E$, so $|f| = f$ almost everywhere, so $|f_n| \to f$ in measure by Lemma 8.23. Thus,

$$\int_X f\, d\mu = \lim_{n\to\infty} \int_X |f_n|\, d\mu.$$

However, $|f_n|(x) \ge 0$ for each $x \in X$, so the integrals on the right-hand side are nonnegative by Lemma 8.3. It follows $\int_X f\, d\mu \ge 0$. ■

Lemma 8.75. Fix a measure space $(X, \mathcal{S}, \mu)$. Given μ-integrable functions $f, g\colon X \to \mathbb{R}$ such that $f(x) \ge g(x)$ almost everywhere, we have

$$\int_X f\, d\mu \ge \int_X g\, d\mu.$$

Proof. Quickly, note $f - g$ is μ-integrable by Lemma 8.59. By Proposition 8.73, it suffices to show that

$$\int_X (f - g)\, d\mu \ge 0.$$

However, $(f - g)(x) = f(x) - g(x) \ge 0$ almost everywhere, so this follows directly from Lemma 8.74. ■

Lemma 8.76. Fix a measure space $(X, \mathcal{S}, \mu)$. Given a μ-integrable function $f\colon X \to B$, we have

$$\left\|\int_X f\,d\mu\right\| \le \int_X \|f\|\,d\mu.$$

Proof. Quickly, note $\|f\|$ is μ-integrable by Lemma 8.62. Now, as usual, pick up our mean Cauchy sequence $\{f_n\}_{n\in\mathbb{N}}$ of simple μ-integrable functions such that $f_n \to f$ in measure. It follows from the proof of Lemma 8.62 that $\{\|f_n\|\}_{n\in\mathbb{N}}$ is a mean Cauchy sequence of simple μ-integrable functions with $\|f_n\| \to \|f\|$ in measure. It follows

$$\int_X \|f\|\,d\mu = \lim_{n\to\infty}\int_X \|f_n\|\,d\mu.$$

Now, using Lemma 8.2, we see

$$\int_X \|f\|\,d\mu \ge \lim_{n\to\infty}\left\|\int_X f_n\,d\mu\right\|.$$

To finish, we note that $\|\cdot\| : B \to \mathbb{R}$ is continuous (Example 1.38), so Lemma 1.57 grants

$$\int_X \|f\|\,d\mu \ge \left\|\lim_{n\to\infty}\int_X f_n\,d\mu\right\| = \left\|\int_X f\,d\mu\right\|,$$

which is what we wanted. ■

8.5.3 A Semi-Norm for $\mathcal{L}^1$

Here is our semi-norm.

Notation 8.77. Fix a measure space $(X, \mathcal{S}, \mu)$ and a Banach space $(B, \|\cdot\|)$. Given a μ-integrable function, we define

$$\|f\|_1 := \int_X \|f\|\,d\mu.$$

Note $\|f\|$ is in fact μ-integrable by Lemma 8.62.

Remark 8.78. As before, we see $\|f\|_1$ extends our definition from simple μ-integrable functions because our definition of integral also extended our definition from simple μ-integrable functions.

And here is the check.

Corollary 8.79. Fix a measure space $(X, \mathcal{S}, \mu)$ and a Banach space $(B, \|\cdot\|)$. Then $\|\cdot\|_1$ defines a semi-norm on $\mathcal{L}^1(X, \mathcal{S}, \mu, B)$.

Proof. Here are our checks.

- Zero: suppose $f = 0$. Then f is a simple μ-integrable function, so this follows from Lemma 8.6.
- Nonnegative: for some μ-integrable $f\colon X \to B$, note that $\|f\|(x) \ge 0$ for each $x \in X$, so Lemma 8.75 tells us
$$\|f\|_1 = \int_X \|f\|\,d\mu \ge \int_X 0\,d\mu = 0,$$
where the relevant functions are μ-integrable by Lemma 8.62.
- Homogeneous: fix a scalar c and a μ-integrable function $f\colon X \to B$. Then Proposition 8.73 tells us
$$\|cf\|_1 = \int_X \|cf\|\,d\mu = \int_X c\cdot\|f\|\,d\mu = c\int_X \|f\|\,d\mu = c\,\|f\|_1,$$
where the relevant functions are μ-integrable by Lemma 8.59 and Lemma 8.62.

- Triangle inequality: given μ-integrable functions $f, g\colon X \to B$, we note that

$$\|f\|(x) + \|g\|(x) = \|f(x)\| + \|g(x)\| \geq \|f(x) + g(x)\| = \|f + g\|(x)$$

for each $x \in X$, so Lemma 8.75 tells us

$$\int_X (\|f\| + \|g\|)\, d\mu \geq \int_X \|f + g\|\, d\mu = \|f + g\|_1.$$

Thus, Proposition 8.73 tells us $\|f\|_1 + \|g\|_1 \geq \|f + g\|_1$, which is what we wanted. ■

We will show that $\mathcal{L}^1$ is complete in some sense next lecture.

8.6 November 16

Here we go.

8.6.1 Integration Facts

We continue our fact-collection.

Lemma 8.80. Fix a measure space $(X, \mathcal{S}, \mu)$ and a Banach space B. Further, fix a μ-integrable function $f\colon X \to B$ with corresponding sequence mean Cauchy sequence $\{f_n\}_{n\in\mathbb{N}}$ of simple μ-integrable functions such that $f_n \to f$ in measure. Then $\|f - f_n\|_1 \to 0$ as $n \to \infty$.

Proof. This essentially follows directly from the definition of integration. Indeed, fix some $\varepsilon > 0$. Our sequence is mean Cauchy, so choose some N for which $m, n \geq N$ implies $\|f_m - f_n\|_1 < \varepsilon/2$.

Now, for some fixed m, define $g_n := f_m - f_n$ for each $n \in \mathbb{N}$, which is a simple μ-integrable function by Lemma 7.17, and we see $\{g_n\}_{n\in\mathbb{N}}$ is mean Cauchy by Lemma 8.46 with $g_n \to f_m - f$ in measure by Lemma 8.19. Now, it follows from the proof that Lemma 8.62 that $\{\|g_n\|\}_{n\in\mathbb{N}}$ is still a mean Cauchy sequence of simple μ-integrable functions such that $\|g_n\| \to \|f_m - f\|$ in measure, so

$$\|f_m - f\|_1 = \int_X \|f_m - f\|\, d\mu = \lim_{n\to\infty} \int_X \|g_n\|\, d\mu = \lim_{n\to\infty} \|f_m - f_n\|_1.$$

(All the relevant functions are μ-integrable by Lemma 8.59 and Lemma 8.62.) Thus, taking $m \geq N$, we see $\|f_m - f_n\|_1 < \varepsilon/2$ for $n \geq N$, so

$$\|f_m - f\|_1 = \lim_{n\to\infty} \|f_m - f_n\|_1 \leq \varepsilon/2 < \varepsilon.$$

This completes the proof. ■

The point of the above lemma is the following density result.

Corollary 8.81. Fix a measure space $(X, \mathcal{S}, \mu)$ and a Banach space B. For any μ-integrable function $f\colon X \to B$ and error $\varepsilon > 0$, there is a simple μ-integrable function $g\colon X \to B$ such that $\|f - g\|_1 < \varepsilon$.

Proof. Because $f\colon X \to B$ is integrable, there is a mean Cauchy sequence $\{f_n\}_{n\in\mathbb{N}}$ of simple μ-integrable functions such that $f_n \to f$ in measure. But then Lemma 8.80 tells us that

$$\lim_{n\to\infty} \|f - f_n\|_1 = 0,$$

so there is some N such that $n \geq N$ implies $\|f - f_n\|_1 < \varepsilon$. Choosing any $n \geq N$ and setting $g := f_n$ thus finishes. ■

Lemma 8.82. Fix a measure space $(X, \mathcal{S}, \mu)$ and a normed vector space $(B, \|\cdot\|)$. Given a μ-integrable function $f\colon X \to B$ and bound $\varepsilon > 0$, there is some $F \subseteq X$ with $F \in \mathcal{S}$ and $\mu(F) < \infty$ such that

$$\int_X \|f\| \, 1_{X \setminus F} \, d\mu < \varepsilon.$$

Proof. Because $f\colon X \to B$ is μ-integrable, we may choose some simple μ-integrable function $g\colon X \to B$ such that $\|f - g\|_1 < \varepsilon$, where we are using Corollary 8.81. Now, choose $F := g^{-1}(B \setminus \{0\})$, which is in $\mathcal{S}$ again using Lemma 7.34, and we note Lemma 5.51 implies

$$\mu(F) \le \mu\left(g^{-1}(B \setminus \{0\})\right),$$

where $\mu\left(g^{-1}(B \setminus \{0\})\right)$ is finite by Remark 7.14. Thus, $\mu(F) < \infty$.

It remains to compute $\int_X \|f\| \, 1_{X \setminus F} \, d\mu$. Well, we see $g(x) = 0$ for $x \notin F$, so

$$(\|f\| \, 1_{X \setminus F})(x) = (\|f - g\| \, 1_{X \setminus F})(x) \le \|f - g\| \, (x)$$

for each $x \in X$, so Lemma 8.75 tells us

$$\int_X \|f\| \, 1_{X \setminus F} \, d\mu \le \int_X \|f - g\| \, d\mu = \|f - g\|_1 < \varepsilon,$$

which is what we wanted. ■

Remark 8.83. The above result basically says that f is almost supported on a set of finite measure.

Lemma 8.84. Fix a measure space $(X, \mathcal{S}, \mu)$. Given a μ-integrable function $f\colon X \to \mathbb{R}$, given $E \in \mathcal{S}$ with $f(x) \ge 1_E(x)$ almost everywhere, then

$$\mu(E) \le \int_X f \, d\mu.$$

Proof. The main difficulty here is that we don't actually know if 1_E is an integrable function at the outset.

For convenience, we set $F := f^{-1}(B \setminus \{0\})$. We claim that F is contained in the countable union of sets of finite measure; this is annoying, so we will brief. Well, because $f\colon X \to \mathbb{R}$ is μ-integrable, we can find a mean Cauchy sequence $\{f_n\}_{n \in \mathbb{N}}$ of simple μ-integrable functions such that $f_n \to f$ in measure. Now, $g_n \to f$ almost everywhere (because $g_n \to f$ in measure), so there is some $N \in \mathcal{S}$ such that $\mu(N) = 0$ while $g_n 1_{X \setminus N} \to f 1_{X \setminus N}$. We now define

$$G_n := g_n^{-1}(B \setminus \{0\}),$$

which is in $\mathcal{S}$ has finite measure by Remark 7.14. In particular, $f(x) \ne 0$ implies that either $x \in N$ or $g_n(x) \to f(x)$ as $n \to \infty$, which requires $g_n(x) \ne 0$ for some n and thus $x \in G_n$ for some n. As such, we see

$$F \subseteq N \cup \bigcup_{n=1}^{\infty} G_n,$$

which competes the proof of the claim.

Now, $f(x) \ge 1_E(x)$ almost everywhere, so select some $N' \in \mathcal{S}$ such that $\mu(N') = 0$ and $x \in X \setminus N'$ implies $f(x) \ge 1_E(x)$. With this in mind, we define

$$E_n := (E \setminus N') \cap \left(N \cup \bigcup_{i=1}^{n} G_i \right).$$

In particular, we see that Lemma 5.51 and Lemma 5.55 imply

$$\mu(E_n) \le \mu(N') + \mu(N) + \sum_{i=1}^{n} \mu(G_i)$$

is a finite sum of finite real numbers and is therefore finite.

As such, we note $x \in E \setminus N'$ implies $f(x) \neq 0$ and thus $x \in F$, so $E \setminus N' \subseteq F$, so $E \setminus N' = \bigcup_{n=1}^{\infty} E_n$. Further, we see $E_n \subseteq E_n \cup G_{n+1} = E_{n+1}$ straight from the definition, so Proposition 6.36 tells us

$$\mu(E \setminus N') = \lim_{n\to\infty} \mu(E_n).$$

However, $E_n \subseteq E \setminus N'$ implies $1_{E_n}(x) \le 1_{E\setminus N'}(x)$ for each $x \in X$, so $1_{E_n}(x) \le f(x)$ for $x \in X \setminus N'$, so $1_{E_n}(x) \le f(x)$ almost everywhere, so Lemma 8.75 tells us

$$\int_X 1_{E_n}\,d\mu \le \int_X f\,d\mu.$$

Noting $\mu(E_n) = \int_X 1_{E_n}\,d\mu$ by Example 7.16, we see $\mu(E_n) \le \int_X f\,d\mu$ for each n. It follows that

$$\mu(E \setminus N') \le \int_X f\,d\mu.$$

However, $\mu(N') = 0$, so $\mu(E \cap N') = 0$ by Lemma 5.51, so $\mu(E \setminus N') = \mu(E) - \mu(E \cap N') = \mu(E)$. This finishes. ■

Corollary 8.85. Fix a measure space $(X, \mathcal{S}, \mu)$ and a normed vector space B. Further, fix a simple $\mathcal{S}$-measurable function $f\colon X \to B$ and a μ-integrable function $g\colon X \to \mathbb{R}$. If $\|f(x)\| \le g(x)$ almost everywhere, then f is simple μ-integrable.

Proof. Fixing any $y \in (\operatorname{im} f) \setminus \{0\}$, we have to show that $f^{-1}(\{y\})$ has finite measure. Well, by Lemma 5.51, we can just show $E := f^{-1}(B \setminus \{0\})$ has finite measure, where $E \in \mathcal{S}$ already.

For this, we note that $\operatorname{im} f$ is finite, so $\{\|y\| : y \in (\operatorname{im} f) \setminus \{0\}\}$ is finite and therefore has a minimum value r. Note $r > 0$ because $\|y\| = 0$ implies $y = 0$. As such, we note that

$$r1_E(x) \le \|f\|(x)$$

for all $x \in X$ because either $x \notin E$ and thus $f(x) = 0$ or $x \in E$ and thus $r \le \|f(x)\|$. It follows $1_E(x) \le \frac{1}{r}\|f\|(x) \le \frac{1}{r} g(x)$ almost everywhere, so Lemma 8.84 tells us that E has finite measure. In particular, $\frac{1}{r} g$ is μ-integrable by Lemma 8.59. ■

8.6.2 Convergence in Mean, Again

We now move towards showing that $\mathcal{L}^1$ is complete. To state the result, we need to (re)define converging in mean for our μ-integrable functions.

Definition 8.86 (Converge in mean)**.** Fix a measure space $(X, \mathcal{S}, \mu)$ and a Banach space B. Then a sequence $\{f_n\}_{n\in\mathbb{N}}$ of μ-integrable functions *converges in mean* to a μ-integrable function $f\colon X \to B$ if and only if $\|f - f_n\|_1 \to 0$ as $n \to \infty$.

Definition 8.87 (Mean Cauchy)**.** Fix a measure space $(X, \mathcal{S}, \mu)$ and a Banach space B. Then a sequence $\{f_n\}_{n\in\mathbb{N}}$ of μ-integrable functions is *mean Cauchy* if and only if $\|f_m - f_n\|_1 \to 0$ as $m, n \to \infty$.

Remark 8.88. If everything is simple μ-integrable, then we note the fact that $\|f\|_1$ is the same for μ-integrable functions as for simple μ-integrable functions means that our definitions above also do not change.

Remark 8.89. Roughly speaking, convergence in mean lets us compute integrals. Namely, if $f_n \to f$ in mean, then we claim $\int_X f_n \, d\mu \to \int_X f \, d\mu$. Indeed, note

$$\left\| \int_X f \, d\mu - \int_X f_n \, d\mu \right\| = \left\| \int_X (f - f_n) \, d\mu \right\| \leq \int_X \|f - f_n\| \, d\mu = \|f - f_n\|_1$$

by Proposition 8.73 and Lemma 8.76. Thus, for any $\varepsilon > 0$, we use $f_n \to f$ in mean to find N such that $n \geq N$ implies $\|f - f_n\|_1 < \varepsilon$, which implies $\left\| \int_X f \, d\mu - \int_X f_n \, d\mu \right\| < \varepsilon$ as well. This finishes.

We now take a deep breath and run a few checks. Here is a comparison result.

Lemma 8.90. Fix a measure space $(X, \mathcal{S}, \mu)$ and a Banach space $(B, \|\cdot\|)$. Further, fix a sequence of μ-integrable functions $\{f_n\}_{n\in\mathbb{N}}$ and another μ-integrable function $f \colon X \to B$. If $f_n \to f$ in mean, then $f_n \to f$ in measure.

Proof. We imitate Lemma 8.49; note the statement makes sense because the f_n and f are $\mathcal{S}$-measurable. Now, fix some $\varepsilon > 0$. Then, for any n, we define

$$E_n := \{x \in X : \|f(x) - f_n(x)\| \geq \varepsilon\}$$

so that we want to show $\mu(E_n) \to 0$ as $n \to \infty$; note E_n is $\mathcal{S}$-measurable because all the relevant functions are $\mathcal{S}$-measurable. Well, we see that each $x \in X$ has

$$1_{E_n}(x) \leq \frac{\|f(x) - f_n(x)\|}{\varepsilon},$$

so Lemma 8.84 tells us that

$$\mu(E_n) \leq \int_X \frac{\|f(x) - f_n(x)\|}{\varepsilon} \, d\mu = \frac{1}{\varepsilon} \int_X \|f(x) - f_n(x)\| \, d\mu = \frac{\|f - f_n\|_1}{\varepsilon},$$

where we have used Proposition 8.73. Thus,

$$\lim_{n\to\infty} \mu(E_n) \leq \lim_{n\to\infty} \frac{\|f - f_n\|_1}{\varepsilon} = \frac{1}{\varepsilon} \lim_{n\to\infty} \|f - f_n\|_1,$$

which is 0 because $f_n \to f$ in mean. The fact that $\mu(E_n) \geq 0$ for each n tells us $\lim_{n\to\infty} \mu(E_n) \geq 0$, so $\mu(E_n) \to 0$ as $n \to \infty$ follows. ■

Here is a nice consequence.

Lemma 8.91. Fix a measure space $(X, \mathcal{S}, \mu)$ and a normed vector space B. Given a μ-integrable function $f \colon X \to B$, if $\|f\|_1 = 0$, then $f(x) = 0$ almost everywhere.

Proof. Let z denote the zero function so that we want to show $f(x) = 0 = z(x)$ almost everywhere. Note that the sequence $\{f_n\}_{n\in\mathbb{N}}$ of functions defined by $f_n := z$ are all simple μ-integrable functions (vacuously). As such, we see that $f_n \to z$ in measure.

On the other hand, we see $f_n \to f$ in mean because any $\varepsilon > 0$ can set $N = 0$ so that $n \geq N$ has

$$\|f - f_n\|_1 = \|f - 0\|_1 = \|f\|_1 = 0 < \varepsilon.$$

However, $f_n \to f$ in mean implies that $f_n \to f$ in measure by Lemma 8.90. It follows from Lemma 8.23 that $f(x) = z(x)$ almost everywhere. ■

And here is the converse.

Lemma 8.92. Fix a measure space $(X, \mathcal{S}, \mu)$ and a Banach space B. Given some μ-integrable function $f\colon X \to B$, if $f(x) = 0$ almost everywhere, then $\|f\|_1 = 0$.

Proof. Define $f_n\colon X \to B$ to be the zero function for each n. The main claim is that $f_n \to f$ in measure. Indeed, because $f(x) = 0$ almost everywhere, we can find $E \in \mathcal{S}$ such that $\mu(E) = 0$ while $f(x) = 0$ for $x \in X \setminus E$. As such, for any $\varepsilon > 0$, we note that any $n \geq 1$ has

$$\{x : \|f(x) - f_n(x)\| \geq \varepsilon\} = \{x : \|f(x)\| \geq \varepsilon\} \subseteq \{x : f(x) \neq 0\} \subseteq E.$$

Thus, Lemma 5.51 tells us $\mu(\{x : \|f(x) - f_n(x)\| \geq \varepsilon\}) \leq \mu(E) = 0$, finishing.

We now note that each $\|f_n\|$ is the zero function and hence (vacuously) a simple $\mathcal{S}$-integrable function, and $\|f_n\| \to \|f\|$ by Lemma 8.19. Thus, by definition of our integral,

$$\|f\|_1 = \int_X \|f\|\, d\mu = \lim_{n\to\infty} \int_X \|f_n\|\, d\mu = \lim_{n\to\infty} \|f_n\|_1 = \lim_{n\to\infty} 0 = 0.$$

This is what we wanted. ■

And here is the total result.

Lemma 8.93. Fix a measure space $(X, \mathcal{S}, \mu)$ and a normed vector space B. Then two μ-integrable functions $f, g\colon X \to B$ have $\|f - g\|_1 = 0$ if and only if $f(x) = g(x)$ almost everywhere.

Proof. We let $[h] \in L^1(X, \mathcal{S}, \mu, B)$ denote the equivalence class of a μ-integrable function $h\colon X \to B$.

In one direction, if $[f] = [g]$, then $f - g \in \mathcal{N}(X, \mathcal{S}, \mu, B)$, so $\|f - g\|_1 = 0$. It follows that $f(x) - g(x) = 0$ almost everywhere by Lemma 8.91, so we can select $E \in \mathcal{S}$ such that $\mu(E) = 0$ while $f(x) - g(x) = 0$ for $x \in X \setminus E$. As such, $f(x) = g(x)$ for $x \in X \setminus E$, so $f(x) = g(x)$ almost everywhere.

In the other direction, suppose $f(x) = g(x)$ almost everywhere. Then we can select $E \in \mathcal{S}$ such that $\mu(E) = 0$ while $f(x) = g(x)$ for $x \in X \setminus E$. It follows $f(x) - g(x) = 0$ for $x \in X \setminus E$, so $(f - g)(x) = 0$ almost everywhere. Thus, $\|f - g\|_1 = 0$ by Lemma 8.92. ■

Remark 8.94. Fix a measure space $(X, \mathcal{S}, \mu)$ and a normed vector space B. Lemma 8.93 tells us that μ-integrable functions $f, g\colon X \to B$ equal almost everywhere have $\|f - g\|_1 = 0$. As an application, we note Proposition 8.73 and Lemma 8.76 imply

$$\left\| \int_X f\, d\mu - \int_X g\, d\mu \right\| = \left\| \int_X (f - g)\, d\mu \right\| \leq \int_X \|f - g\|\, d\mu = \|f - g\|_1 = 0,$$

so $\int_X f\, d\mu = \int_X g\, d\mu$ follows.

8.6.3 Completeness of L^1

And now for our feature presentation.

Proposition 8.95. Fix a measure space $(X, \mathcal{S}, \mu)$ and a Banach space B. Then a mean Cauchy sequence $\{f_n\}_{n\in\mathbb{N}}$ of μ-integrable functions converges in mean to some μ-integrable function $f\colon X \to B$.

Proof. For each n, Corollary 8.81 grants some simple μ-integrable function $g_n\colon X \to B$ with $\|f_n - g_n\|_1 < 1/n$. We now proceed in steps.

1. We claim that $\{g_n\}_{n\in\mathbb{N}}$ is mean Cauchy. Well, fix any $\varepsilon > 0$. Then we can find some N_f such that $m,n \geq N_f$ implies
$$\|f_m - f_n\|_1 < \frac{\varepsilon}{3}.$$
Thus, we define $N := \max\{N_f, 3/\varepsilon\}$ so that $m,n \geq N$ implies (by Corollary 8.79) that
$$\|g_m - g_n\|_1 \leq \|f_m - g_m\|_1 + \|f_m - f_n\|_1 + \|f_n - g_n\|_1 < \frac{1}{m} + \frac{\varepsilon}{3} + \frac{1}{n} \leq \frac{2}{N} + \frac{\varepsilon}{3} \leq \frac{2\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon,$$
which finishes.

2. Next, we construct the limit function. Because $\{g_n\}_{n\in\mathbb{N}}$ is mean Cauchy, it is Cauchy in measure by Lemma 8.49, so there is a uniformly Cauchy subsequence $\{g_{n_i}\}_{i\in\mathbb{N}}$ by Theorem 8.35. However, this subsequence $\{g_{n_i}\}_{i\in\mathbb{N}}$ will then converge to some $\mathcal{S}$-measurable $g\colon X \to B$ almost uniformly by Lemma 8.40. Note that g is then μ-integrable by definition.

3. It remains to show that $f_n \to g$ in mean. Well, note $g_{n_i} \to g$ almost uniformly, so $g_{n_i} \to g$ in measure by Lemma 8.51, so $g_{n_i} \to g$ in mean by Lemma 8.80. (This step is why it is important for the $g_\bullet$ to be simple μ-integrable!) Finishing up, we fix any $\varepsilon > 0$ and note that there is N_g such that $i \geq N_g$ implies
$$\|g_{n_i} - g\|_1 < \varepsilon/3.$$
Further, $\{f_n\}_{n\in\mathbb{N}}$ is mean Cauchy, so there is N_f such that $m,n \geq N_f$ implies
$$\|f_m - f_n\|_1 < \varepsilon/3.$$
In total, we set $N := \max\{N_f, N_g, 3/\varepsilon\}$. Then $n \geq N$ implies $n \geq N_f$ and $n_n \geq n \geq N \geq N_g$, so (using Corollary 8.79 some more)
$$\|g - f_n\|_1 \leq \|g - g_{n_n}\|_1 + \|g_{n_n} - f_{n_n}\|_1 + \|f_{n_n} - f_n\|_1 < \frac{\varepsilon}{3} + \frac{1}{n_n} + \frac{\varepsilon}{3} \leq \frac{\varepsilon}{3} + \frac{1}{3/\varepsilon} + \frac{\varepsilon}{3} = \varepsilon,$$
which is what we wanted. ■

Corollary 8.96. Fix a measure space $(X, \mathcal{S}, \mu)$ and a Banach space $(B, \|\cdot\|)$. Given a mean Cauchy sequence $\{f_n\}_{n\in\mathbb{N}}$ of μ-integrable functions converging to some $\mathcal{S}$-measurable function $f\colon X \to B$ almost everywhere, we know f is μ-integrable, and $f_n \to f$ in mean.

Proof. The main annoyance here is proving that f is actually μ-integrable. As such, we divide the proof into two steps.

1. Note Proposition 8.95 promises some μ-integrable function $f'\colon X \to B$ such that $f_n \to f'$ in mean. However, $f_n \to f'$ in mean implies that $f_n \to f'$ in measure by Lemma 8.90, so $f_n \to f'$ almost everywhere, so $f = f'$ almost everywhere.

   However, f' is already μ-integrable, so there exists some mean Cauchy sequence $\{g_n\}_{n\in\mathbb{N}}$ of simple μ-integrable functions such that $g_n \to f'$ almost everywhere. It follows that $g_n \to f$ almost everywhere as well, so f is in fact μ-integrable.

2. Now, $f = f'$ almost everywhere implies that $\|f - f'\|_1 = 0$ by Lemma 8.93. Thus, for any $\varepsilon > 0$, we use $f_n \to f'$ in mean to find N such that $n \geq N$ implies
$$\|f' - f_n\| < \varepsilon.$$
However, we now see $\|f - f_n\|_1 \leq \|f - f'\|_1 + \|f' - f_n\| < \varepsilon$ by Corollary 8.79, finishing. ■

In order to actually state this as a completeness result, we need to turn the semi-norm $\|\cdot\|_1$ into an actual norm.

Notation 8.97. Fix a measure space $(X, \mathcal{S}, \mu)$ and a Banach space B. We set $\mathcal{N}(X, \mathcal{S}, \mu, B) := \{f \in \mathcal{L}^1(X, \mathcal{S}, \mu, B) : \|f\|_1 = 0\}$ and

$$L^1(X, \mathcal{S}, \mu, B) := \mathcal{L}^1(X, \mathcal{S}, \mu, B)/\mathcal{N}(X, \mathcal{S}, \mu, B).$$

Remark 8.98. Given μ-integrable functions $f, g\colon X \to B$, we claim that the equivalence classes $[f], [g] \in L^1(X, \mathcal{S}, \mu, B)$ are equal if and only if $f(x) = g(x)$ almost everywhere. Indeed, $[f] = [g]$ if and only if $f - g \in \mathcal{N}(X, \mathcal{S}, \mu, B)$, which is equivalent to $\|f - g\|_1 = 0$. However, by Lemma 8.93, $\|f - g\|_1 = 0$ is equivalent to $f(x) = g(x)$ almost everywhere.

Lemma 8.99. Fix a measure space $(X, \mathcal{S}, \mu)$ and a Banach space B. The function $\|\cdot\|_1$ descends to a norm on $L^1(X, \mathcal{S}, \mu, B)$.

Proof. This is a direct consequence Proposition 1.13, applied to $\|\cdot\|_1$ on $\mathcal{L}^1(X, \mathcal{S}, \mu, B)$. ■

Corollary 8.100. Fix a measure space $(X, \mathcal{S}, \mu)$ and a Banach space B. Then $L^1(X, \mathcal{S}, \mu, B)$ is the completion of the vector space of simple μ-integrable functions.

Proof. The normed vector space $L^1(X, \mathcal{S}, \mu, B)$ is complete by Proposition 8.95. Further, the space of simple μ-integrable functions (modded out by the functions of norm zero) are dense in $L^1(X, \mathcal{S}, \mu, B)$ by Corollary 8.81. ■

Next class we will begin trying to compute integrals.

Theme 9
Integration Applications

What we didn't do is make the construction at all usable in practice!
This time we will remedy this.

—Kiran S. Kedlaya, [Ked21]

9.1 November 18

There will be at most two more homework assignments.

9.1.1 Measures from Integrals

Now that we have a reasonable notion of what functions to integrate, given a measure, we would like to take these integrable functions to build measures. It will be convenient to have the following notation.

Notation 9.1. Fix a measure space $(X, \mathcal{S}, \mu)$ and a Banach space B. Given a μ-integrable function f and some $E \subseteq X$ such that either $E \in \mathcal{S}$ or $X \setminus E \in \mathcal{S}$, we define

$$\mu_f(E) := \int_E f \, d\mu := \int_X f 1_E \, d\mu.$$

Note that $f1_E$ is μ-integrable by Lemma 8.60 when $E \in \mathcal{S}$ and by Remark 8.61 when $X \setminus E \in \mathcal{S}$.

Remark 9.2. We note that μ_f has good additivity properties. Namely, given scalars $a, b \in k$, where k is the base field of B, and two μ-integrable functions $f, g \colon X \to B$, we have

$$\int_X (af + bg)1_E \, d\mu = \int_X \big(a(f1_E) + b(g1_E)\big) \, d\mu \stackrel{*}{=} a \int_X f1_E \, d\mu + \int_X b1_E \, d\mu,$$

where $\stackrel{*}{=}$ is by Proposition 8.73. Thus, $\mu_{af+bg}(E) = a\mu_f(E) + b\mu_g(E)$.

Here are a few quick inequalities.

Lemma 9.3. Fix a measure space $(X, \mathcal{S}, \mu)$ and a Banach space $(B, \|\cdot\|)$. Further, fix a μ-integrable function $f\colon X \to B$ and some $E \in \mathcal{S}$.

(a) We have $\|\mu_f(E)\| \le \|f\|_1$.

(b) Given a bound $M \ge 0$ such that $\|f(x)\| \le M$ almost everywhere for $x \in E$, then

$$\|\mu_f(E)\| \le M\mu(E).$$

Proof. By Lemma 8.76, we see

$$\|\mu_f(E)\| = \left\| \int_X f1_E \, d\mu \right\| \le \int_X \|f1_E\| \, d\mu.$$

We now approach the two parts separately.

(a) For each $x \in X$, we note that $\|f1_E\|(x)$ is either 0 or $\|f(x)\|$, so $\|f1_E\|(x) \le \|f\|(x)$ for each $x \in X$. Thus, Lemma 8.75 tells us

$$\int_X \|f1_E\| \, d\mu \le \int_X \|f\| \, d\mu = \|f\|_1 ,$$

which finishes.

(b) We claim that $\|f1_E\|(x) \le M1_E(x)$ almost everywhere for $x \in X$. Indeed, $\|f(x)\| \le M$ almost everywhere for $x \in E$, so there is some $N \in \mathcal{S}$ such that $\mu(N) = 0$ and $x \in E \setminus N$ implies $\|f(x)\| \le M$. Thus, $x \in X \setminus N$ either has $x \in X \setminus E$ so that $\|f1_E\|(x) = 0 \le 0 = M1_E(x)$ or $x \in E \setminus N$ so that

$$\|f1_E\|(x) = \|f(x)\| \le M = M1_E(x).$$

Finishing up, Lemma 8.75 kicks in to tell us that

$$\int_X \|f1_E\| \, d\mu \le \int_X M1_E \, d\mu.$$

The right-hand side is $M \int_X 1_E \, d\mu$ by Lemma 8.1, which is $M\mu(E)$ by Example 7.16. This finishes. ■

Now, the notation μ_f is intended to be suggestive that we're going to have a measure. Finite additivity is relatively quick.

Remark 9.4. Suppose $f\colon X \to B$ is μ-integrable. It's pretty fast to see that $\mu_f\colon \mathcal{S} \to B$ is finitely additive: if $E, F \in \mathcal{S}$ are disjoint, we need to show that $\mu_f(E \sqcup F) = \mu_f(E) + \mu_f(F)$. (By induction, this extends to any finite collection.) Well, $1_E + 1_F = 1_{E \sqcup F}$ because $x \in E \sqcup F$ if and only if $x \in E$ or $x \in F$, but only one of $x \in E$ or $x \in F$ is possible. Thus, Proposition 8.73 tells us

$$\mu_f(E \sqcup F) = \int_X f1_{E\sqcup F} \, d\mu = \int_X f(1_E + 1_F) \, d\mu = \int_X f1_E \, d\mu + \int_X f1_F \, d\mu = \mu_f(E) + \mu_f(F).$$

In fact, we can extend Remark 9.4 to make μ_f countably additive.

Proposition 9.5. Fix a measure space $(X, \mathcal{S}, \mu)$ and a Banach space $(B, \|\cdot\|)$. Given some μ-integrable function $f\colon X \to B$, the function $\mu_f\colon \mathcal{S} \to B$ is countably additive.

Proof. Suppose we have a pairwise disjoint collection $\{E_i\}_{i\in\mathbb{N}} \subseteq \mathcal{S}$. Set $E := \bigsqcup_{i=1}^\infty E_i$ (which is in $\mathcal{S}$) so that we want to show

$$\mu_f(E) \stackrel{?}{=} \sum_{i=1}^\infty \mu_f(E_i). \tag{9.1}$$

We have two steps.

1. Suppose $f\colon X \to B$ is a simple μ-integrable function; we show (9.1). Well, by Lemma 7.8, we can write
$$f = \sum_{j=1}^{n} y_j 1_{F_j}$$
for some nonzero distinct points $y_j \in B$ and pairwise disjoint $F_j \coloneqq f^{-1}(\{y_j\}) \in \mathcal{S}$.
Now, we note $1_{F\cap F_j} = 1_F 1_{F_j}$ for any $F \in \mathcal{S}$ because $(1_F 1_{F_j})(x) = 1$ if and only if $1_F(x) = 1$ and $1_{F_j}(x) = 1$, which is equivalent to $x \in F$ and $x \in F_j$. Applying this multiple times, we compute
$$\begin{aligned}\mu_f(F) &= \int_X f 1_F \, d\mu \\ &= \int_X \left(\sum_{j=1}^{n} y_j 1_{F_j} 1_F\right) d\mu \\ &= \int_X \left(\sum_{j=1}^{n} y_j 1_{F_j\cap F}\right) d\mu \\ &= \sum_{j=1}^{n} \left(y_j \int_X 1_{F_j\cap F}\, d\mu\right) \\ &= \sum_{j=1}^{n} y_j \mu(F_j \cap F),\end{aligned}$$
where the last two equalities follow from Proposition 8.73 and then Example 7.16. As such, we can use the fact that μ is countably additive already: for each j, note that $\{F_j \cap E_i\}_{i\in\mathbb{N}}$ is a pairwise disjoint collection because $x \in (F_j \cap E_i) \cap (F_j \cap E_{i'})$ implies $x \in E_i \cap E_{i'}$ implies $i = i'$. Thus,
$$F_j \cap E = F_j \cap \bigsqcup_{i=1}^{\infty} E_i = \bigsqcup_{i=1}^{\infty} (F_j \cap E_i)$$
implies
$$\mu(F_j \cap E) = \sum_{i=1}^{\infty} \mu(F_j \cap E_i).$$
Summing over all j, we can write
$$\begin{aligned}\mu_f(E) &= \sum_{j=1}^{n} y_j \mu(F_j \cap E) \\ &= \sum_{j=1}^{n} \left(y_j \sum_{i=1}^{\infty} \mu(F_j \cap E_i)\right) \\ &\overset{*}{=} \sum_{i=1}^{\infty} \left(\sum_{j=1}^{n} y_j \mu(F_j \cap E_i)\right) \\ &= \sum_{i=1}^{\infty} \mu_f(F_j \cap E_i),\end{aligned}$$
which is what we wanted. Note that we are allowed to switch the order of summation in $\overset{*}{=}$ because the outer sum is finite. (This is effectively just the linearity of limits.)

2. We now let $f\colon X \to B$ be an arbitrary μ-integrable function. Fix any $\varepsilon > 0$, and we need some N such that $n \ge N$ implies
$$\left\| \mu_f(E) - \sum_{i=1}^{n} \mu_f(E_i) \right\| < \varepsilon.$$

The idea is to relate f to a simple μ-integrable function: Corollary 8.81 grants us some simple μ-integrable function $g\colon X \to B$ such that $\|f-g\|_1 < \varepsilon/3$. Now, for any finite n, we can compute

$$\begin{aligned}\left\|\mu_f(E) - \sum_{i=1}^n \mu_f(E_i)\right\| &\le \|\mu_f(E)-\mu_g(E)\| + \left\|\mu_g(E) - \sum_{i=1}^n \mu_g(E_i)\right\| + \left\|\sum_{i=1}^n \left(\mu_g(E_i)-\mu_f(E_i)\right)\right\| \\ &= \|\mu_f(E)-\mu_g(E)\| + \left\|\mu_g(E) - \sum_{i=1}^n \mu_g(E_i)\right\| + \left\|\mu_g\left(\bigsqcup_{i=1}^n E_i\right) - \mu_f\left(\bigsqcup_{i=1}^n E_i\right)\right\|,\end{aligned}$$

where the last equality is because μ_f and μ_g are already finitely additive by Remark 9.4. Now, for any $F \in \mathcal{S}$, we note Remark 9.2 tells us

$$\|\mu_f(F)-\mu_g(F)\| = \|\mu_{f-g}(F)\|,$$

which is upper-bounded by $\|f-g\|_1$ by Lemma 9.3. Thus,

$$\left\|\mu_f(E) - \sum_{i=1}^n \mu_f(E_i)\right\| \le 2\|f-g\|_1 + \left\|\mu_g(E) - \sum_{i=1}^n \mu_g(E_i)\right\|.$$

To finish, we use the previous step to note that there is some N such that $n \ge N$ implies

$$\left\|\mu_g(E) - \sum_{i=1}^n \mu_g(E_i)\right\| < \frac{\varepsilon}{3}.$$

In total,

$$\left\|\mu_f(E) - \sum_{i=1}^n \mu_f(E_i)\right\| < 2\cdot\frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon,$$

which is what we wanted. ■

Having a notion of countably additive functions encourages us to extend our definition of measure.

Definition 9.6 (Measure)**.** Fix a measure space $(X, \mathcal{S}, \mu)$ and a Banach space $(B, \|\cdot\|)$. Then a B*-valued measure* μ *on* $\mathcal{S}$ is a countably additive function $\mu\colon \mathcal{S} \to B$.

Example 9.7. By Proposition 9.5, we see that a μ-integrable function $f\colon X \to B$ gives a B-valued measure μ_f.

We now note that μ_f cannot be terribly large.

Lemma 9.8. Fix a measure space $(X, \mathcal{S}, \mu)$ and a normed vector space B. Further, fix a μ-integrable function $f\colon X \to B$. For any $\varepsilon > 0$, there is some $\delta > 0$ such that $E \in \mathcal{S}$ with $\mu(E) < \delta$ implies $\|\mu_f(E)\| < \varepsilon$.

Proof. As usual, use Corollary 8.81 to find some simple μ-integrable function $g\colon X \to B$ with $\|f-g\|_1 < \varepsilon/2$. Then any $E \in \mathcal{S}$ grants

$$\|\mu_f(E)\| \le \|\mu_f(E)-\mu_g(E)\| + \|\mu_g(E)\|.$$

We now bound the terms individually.

- By Remark 9.2, we see $\mu_f(E)-\mu_g(E) = \mu_{f-g}(E)$, so

$$\|\mu_f(E)-\mu_g(E)\| = \|\mu_{f-g}(E)\| \le \|f-g\|_1,$$

where the inequality is by Lemma 9.3.

- Note that g has finite image, so we may set $M := \max\{\|g(x)\| : x \in X\}$ so that $\|g(x)\| \le M$ for each $x \in X$. Thus, Lemma 9.3 tells us
$$\|\mu_g(E)\| \le M\mu(E).$$

In total, we see
$$\|\mu_f(E)\| \le \frac{\varepsilon}{2} + M\mu(E)$$
for any $E \in \mathcal{S}$. Thus, we set $\delta := \frac{\varepsilon}{2(M+1)}$ so that $\mu(E) < \delta$ implies
$$\|\mu_f(E)\| < \frac{\varepsilon}{2} + M \cdot \frac{\varepsilon}{2(M+1)} < \varepsilon,$$
which is what we wanted. ■

The above lemma motivates the following definition.

Definition 9.9 (Strongly absolutely continuous)**.** Fix a measure space $(X, \mathcal{S}, \mu)$ and some Banach space $(B, \|\cdot\|)$. Then a B-valued measure $\nu\colon \mathcal{S} \to B$ is *strongly absolutely continuous* if and only if each $\varepsilon > 0$ have some $\delta > 0$ such that $\mu(E) < \delta$ implies $\|\nu(E)\| < \varepsilon$.

Example 9.10. By Lemma 9.8, each μ_f coming from a μ-integrable function f is strongly absolutely continuous.

Remark 9.11. If ν is strongly absolutely continuous, then note that any $E \in \mathcal{S}$ with $\mu(E) = 0$ will have $\nu(E) = 0$. Indeed, for any $\varepsilon > 0$, there is $\delta > 0$ such that $\mu(E') < \delta$ implies $\|\nu(E')\| < \varepsilon$. But we will always have $\mu(E) = 0 < \delta$, so $\|\nu(E)\| < \varepsilon$ for all $\varepsilon > 0$, so $\|\nu(E)\| = 0$, so $\nu(E) = 0$. (This condition is that ν is "absolutely continuous." We will not need it later.)

Remark 9.12. The Radon–Nikodym theorem says that sufficiently nice B-valued measures ν which are absolutely continuous will have $\nu = \mu_f$ for some μ-integrable function f.

9.1.2 Egorov's Theorem

To help us later, we pick up the following result on $\mathcal{S}$-measurable functions.

Theorem 9.13 (Egorov's)**.** Fix a measure space $(X, \mathcal{S}, \mu)$ and a Banach space B. Further, fix some sequence $\{f_n\}_{n\in\mathbb{N}}$ of $\mathcal{S}$-measurable functions. Suppose $E \in \mathcal{S}$ has $\mu(E) < \infty$ such that the $\{f_n\}_{n\in\mathbb{N}}$ converge almost everywhere on E to a function $f\colon X \to B$. Then $f_n|_E \to f$ almost uniformly on E.

Proof. This is a little tricky. We'll take this in steps.

1. We begin by removing a few null sets, for psychological reasons. Note we are given some $N \in \mathcal{S}$ such that $\mu(N) = 0$ while $f_k(x) \to f(x)$ as $n \to \infty$ for each $x \in E \setminus N$. As such,
$$f_n 1_{E\setminus N} \to f 1_{E\setminus N}$$
on E because if $x \notin N$, then $f_n 1_{E\setminus N}(x) = 0 = f 1_{E\setminus N}(x)$ for each n.
We thus claim that $f_n 1_{E\setminus N} \to f 1_{E\setminus N}$ almost uniformly on E. To see that this is enough, note that any $\varepsilon > 0$ has some $F \subseteq E$ with $\mu(E \setminus F) < \varepsilon$ while $f_n 1_{E\setminus N}|_F \to f 1_{E\setminus N}|_F$ uniformly. But then we set $F' := F \setminus N$ so that Lemma 5.51 and Lemma 5.55 tells us
$$\mu(E \setminus F') \le \mu((E \setminus F) \cup N) \le \mu(E \setminus F) + \mu(N) < \delta + 0 = \delta.$$

But now $f_n 1_{E\setminus N}|_{F'} = f_n|_{F'}$ and $f 1_{E\setminus N}|_{F'} = f|_{F'}$ because each $x \in F'$ has $x \notin N$ already. Thus, $f_n|_{F'} \to f|_{F'}$ uniformly, which is what we needed.

In total, we are given $f_n 1_{E\setminus N} \to f 1_{E\setminus N}$ everywhere on E and would like to show this convergence is almost uniform. As such, we replace each $f_n 1_{E\setminus N}$ with f_n and $f 1_{E\setminus N}$ with f to no detriment, except now we know $f_n \to f$ everywhere. In particular, f is $\mathcal{S}$-measuralbe by Corollary 7.44.

2. Now, for each m and n, set

$$E_n^m := \bigcup_{k\ge n}\{x \in E : \|(f-f_k)(x)\| \ge 1/m\} = \{x \in E : \|(f-f_k)(x)\| \ge 1/m \text{ for some } k \ge n\}.$$

Note each $\|f - f_k\|$ is $\mathcal{S}$-measurable by Lemma 7.25 and Corollary 7.45, so the union E_n^m is in $\mathcal{S}$ by Corollary 7.38. Now, for fixed m, we note that $f_k(x) \to f(x)$ for $x \in E$ forces $\bigcap_{n=1}^\infty E_n^m = \varnothing$. However, $\mu(E_1) \le \mu(E) < \infty$ by Lemma 5.51, so Corollary 6.37 tell us that

$$\lim_{n\to\infty} \mu\left(E_n^m\right) = \mu\left(\bigcap_{n=1}^\infty E_n^m\right) = \mu(\varnothing) = 0.$$

3. We now attack the proof directly. Set $\varepsilon > 0$. For each m, we may choose n_m so that $\mu\left(E_n^m\right) < \varepsilon/2^m$ for $n \ge n_m$. As such, we set

$$F := E \Big\backslash \bigcup_{m=1}^\infty E_{n_m}^m$$

so that Lemma 6.2 tells us

$$\mu(E\setminus F) = \mu\left(\bigcup_{m=1}^\infty E_{n_m}^m\right) \le \sum_{m=1}^\infty \mu\left(E_{n_m}^m\right) < \sum_{m=1}^\infty \frac{\varepsilon}{2^m} = \varepsilon.$$

It remains to show $f_n|_F \to f|_F$ uniformly. Fix any $\delta > 0$. To set N, find m with $m > 1/\delta$, and we set $N := n_m$.

To see that this construction works, fix some $n \ge N$ and $x \in F$. Well, $x \in F$ implies that $x \notin E_{n_m}^m$ for our m, so

$$\|f(x) - f_k(x)\| < 1/m < \delta$$

for each $k \ge n_m$. In particular, $n \ge n_m$, so $\|f(x) - f_n(x)\| < \delta$ follows. ■

The point of picking up Theorem 9.13 is so that we can prove the Dominated convergence theorem.

Theorem 9.14 (Dominated convergence)**.** Fix a measure space $(X, \mathcal{S}, \mu)$ and a Banach space $(B, \|\cdot\|)$. Further, fix some sequence $\{f_n\}_{n\in\mathbb{N}}$ of μ-integrable functions converging almost everywhere to a function f. If there is a μ-integrable function $g\colon X \to \mathbb{R}$ such that $\|f_n(x)\| \le g(x)$ almost everywhere for each n, then $\{f_n\}_{n\in\mathbb{N}}$ is in fact mean Cauchy.

We will prove this next class.

Remark 9.15. It will follow from the conclusion that $f_n \to f$ in mean and so

$$\int_X f\,d\mu = \lim_{n\to\infty}\int_X f_n\,d\mu.$$

9.2 November 20

Today we prove Theorem 9.14.

9.2.1 Dominated Convergence

Here is the statement.

Theorem 9.14 (Dominated convergence)**.** Fix a measure space $(X, \mathcal{S}, \mu)$ and a Banach space $(B, \|\cdot\|)$. Further, fix some sequence $\{f_n\}_{n\in\mathbb{N}}$ of μ-integrable functions converging almost everywhere to a function f. If there is a μ-integrable function $g\colon X \to \mathbb{R}$ such that $\|f_n(x)\| \le g(x)$ almost everywhere for each n, then $\{f_n\}_{n\in\mathbb{N}}$ is in fact mean Cauchy.

Proof. Note that $g(x) \le |g|(x)$, so $\|f_n(x)\| \le |g|(x)$ almost everywhere for each n. Further, $|g|$ is μ-integrable by Lemma 8.62. Thus, we may replace g with $|g|$ so that $g = |g|$. Also, before doing any heavy lifting, for each n, we select our $E_n \in \mathcal{S}$ with $\mu(E_n) = 0$ while $\|f_n(x)\| \le g(x)$ for each $x \in X \setminus E_n$.

Fix any $\varepsilon > 0$. Observe that we are interested in bounding the integral

$$\|f_m - f_n\|_1 = \int_X \|f_m - f_n\| \, d\mu$$

for large m and n. We do this in three steps.

1. Because g is μ-integrable, we use Lemma 8.82 to find $E \in \mathcal{S}$ such that $\mu(E) < \infty$ and

   $$\int_X g 1_{X\setminus E} \, d\mu = \int_X |g| 1_{X\setminus E} \, d\mu < \frac{\varepsilon}{6}.$$

   In particular, note $g1_{X\setminus E}$ is μ-integrable by Remark 8.61. Now, for any $m, n \in \mathbb{N}$, we note

   $$\|f_m(x) - f_n(x)\| \le \|f_m(x)\| + \|f_n(x)\| \le 2g(x)$$

   almost everywhere: if $x \notin (E_m \cup E_n)$, then $\|f_m(x)\|, \|f_n(x)\| \le 2g(x)$. However, $E_m \cup E_n$ is a null set because $\mu(E_m \cup E_n) \le \mu(E_m) + \mu(E_n) = 0 + 0 = 0$ by Lemma 5.55. Thus, for any $E' \subseteq X$, we see

   $$\|f_m(x) - f_n(x)\| \, 1_{X\setminus E} \le 2g(x) 1_{X\setminus E} \tag{9.2}$$

   almost everywhere as well because $x \notin (E_m \cup E_n)$ has either $x \notin E'$ so that both sides are zero or $x \in E'$ so that we reduce to the inequality.

   As such, we use $E' = E$ and integrate with Lemma 8.75 to get

   $$\int_X \|f_m - f_n\| \, 1_{X\setminus E} \, d\mu \le \int_X 2g 1_{X\setminus E} \, d\mu \stackrel{*}{=} 2 \int_X g 1_{X\setminus E} \, d\mu < \frac{\varepsilon}{3}.$$

   Note we have used Proposition 8.73 at $\stackrel{*}{=}$.

2. It remains to bound what's happening on E. Note $f_n \to f$ almost everywhere on E,[1] so Theorem 9.13 tells us $f_n|_E \to f|_E$ converges almost uniformly. In particular, for any $\delta > 0$, we can find $F \subseteq E$ with $\mu(F) < \delta$ such that $f_n|_{E\setminus F} \to f|_{E\setminus F}$ uniformly.

   We get some choice in this δ, so we use the fact that the measure μ_g is strongly absolutely continuous (by Lemma 9.8) to find $\delta > 0$ such that $\mu(F) < \delta$ implies $\mu_g(F) < \varepsilon/6$. As such, using $E' = F$ in (9.2), Lemma 8.75 lets us bound

   $$\int_X \|f_m(x) - f_n(x)\| \, 1_F \, d\mu \le \int_X 2g 1_F \, d\mu \stackrel{*}{=} 2 \int_X g 1_F \, d\mu = 2\mu_g(F) < \frac{\varepsilon}{3}.$$

   Again, we have used Proposition 8.73 at $\stackrel{*}{=}$.

[1] Whatever null set witnessed $f_n(x) \to f(x)$ almost everywhere on X will work for E.

3. Thus, it now remains to bound what's happening on $E \setminus F$. Well, $f_n|_{E\setminus F} \to f|_{E\setminus F}$ uniformly, so $\{f_n|_{E\setminus F}\}_{n\in\mathbb{N}}$ is uniformly Cauchy, so we may find N such that $m, n \geq N$ has

$$\|f_m(x) - f_n(x)\| < \frac{\varepsilon}{3(1+\mu(E\setminus F))}$$

for $x \in E \setminus F$. (Note $\mu(E\setminus F) < \infty$ because $\mu(E\setminus F) \leq \mu(E) < \infty$ by Lemma 5.51.) Thus, Lemma 9.3 grants

$$\int_X \|f_m(x) - f_n(x)\| \, 1_{E\setminus F} \, d\mu = \mu_{\|f_m - f_n\|}(E\setminus F) \leq \frac{\varepsilon}{3(1+\mu(E\setminus F))} \cdot \mu(E\setminus F) < \frac{\varepsilon}{3}.$$

We now add our integrals together. Note $X = (X\setminus E) \sqcup E = (X \setminus E) \sqcup F \sqcup (E \setminus F)$ because $F \subseteq E \subseteq X$. Thus, Remark 9.4 promises

$$\|f_m - f_n\|_1 = \mu_{\|f_m-f_n\|}(X) = \mu_{\|f_m-f_n\|}(X\setminus E) + \mu_{\|f_m-f_n\|}(F) + \mu_{\|f_m-f_n\|}(E\setminus F) < \varepsilon$$

for each $m, n \geq N$, where N was chosen in item 3. ■

Remark 9.16. We manifest Remark 9.15; we continue in the context of Theorem 9.14 but now assume that f is $\mathcal{S}$-measurable. In this case, we see that f is μ-integrable and that $f_n \to f$ in mean by Corollary 8.96. Lastly, Remark 8.89 implies

$$\int_X f \, d\mu = \lim_{n\to\infty} \int_X f_n \, d\mu.$$

As an application of Theorem 9.14, we upgrade Corollary 8.85.

Corollary 9.17. Fix a measure space $(X, \mathcal{S}, \mu)$ and a Banach space $(B, \|\cdot\|)$. Further, fix an $\mathcal{S}$-measurable function $f \colon X \to B$. If there is a μ-integrable function $g \colon X \to \mathbb{R}$ such that $\|f(x)\| \leq g(x)$ almost everywhere, then f is μ-integrable.

Proof. Because f is $\mathcal{S}$-measurable, there is a sequence of simple $\mathcal{S}$-measurable functions $\{f_n\}_{n\in\mathbb{N}}$ such that $f_n \to f$ almost everywhere. The main idea is to coerce the f_n into being a mean Cauchy sequence of simple μ-integrable functions, which will finish.

To begin, set $C := g^{-1}(B \setminus \{0\})$ (which is in $\mathcal{S}$ by Corollary 7.38), and define $g_n := f_n 1_C$. Each g_n is still simple $\mathcal{S}$-measurable by Lemma 7.47, and we see $g_n \to f$ almost everywhere still: there is some $E \in \mathcal{S}$ with $\mu(E) = 0$ while $f_n(x) \to f(x)$ for $x \in X \setminus E$. But then $x \in X \setminus E$ implies $g_n(x) \to f(x)$ as well: if $x \in C$, then $g_n(x) = f_n(x)$ for all n; otherwise if $x \notin C$, then $g_n(x) = 0$ for all n while $g(x) = 0$ and thus $f(x) = 0$.

Now, the key restriction is to define

$$E_n := \{x \in X : \|f_n(x)\| \leq 2|g(x)|\}$$

and $h_n := g_n 1_{E_n}$. Notably, $2|g| - \|f_n\|$ is $\mathcal{S}$-measurable by Lemma 7.11 and Lemma 7.9, so $E_n \in \mathcal{S}$ by Corollary 7.38, so h_n is simple $\mathcal{S}$-measurable by Lemma 7.47. But now we see

$$\|h_n(x)\| \leq 2g(x)$$

for each $x \in X$ because $x \in E_n$ grants this inequality for free by definition of E_n, and $x \notin E_n$ gives $\|h_n(x)\| = 0 \leq 2|g(x)|$.

Further, we claim $h_n \to f$ almost everywhere. Fix some $x \in X \setminus E$ so that $f_n(x) \to f(x)$. There are two cases.

- If $x \notin C$, then $h_n(x) = 0$ for all n while $g(x) = 0$ and thus $f(x) = 0$.

- Otherwise, $x \in C$ so that $|g(x)| > 0$. Now, $f_n(x) \to f(x)$ for each x, so $\|f_n(x)\| \to \|f(x)\|$ by Example 1.38, so $\|f(x)\| < 2|g(x)|$ tells us there is some N_g with
$$\|f_n(x)\| < 2|g(x)|$$
for $n \geq N_g$. (Namely, use the error bound $|g(x)| > 0$ so that $n \geq N$ implies $\|f_n(x)\| - \|f(x)\| < |g(x)|$.) Thus, for $n \geq N$, we see $x \in E_n$, so $h_n(x) = g_n(x) = f_n(x)$. So $f_n(x) \to f(x)$ implies that $h_n(x) \to f(x)$ because the sequences match on large terms.[2]

Finishing up, Corollary 8.85 tells us that each h_n is simple μ-integrable (and thus μ-integrable), so Theorem 9.14 tells us $\{h_n\}_{n\in\mathbb{N}}$ is mean Cauchy. Thus, $h_n \to f$ almost everywhere implies f is μ-integrable. ■

9.2.2 Monotone Convergence

We finish class by picking up another convergence theorem, for real-valued functions.

Theorem 9.18 (Monotone convergence)**.** Fix a measure space $(X, \mathcal{S}, \mu)$. Given μ-integrable functions $f_n \colon X \to \mathbb{R}$ such that $f_m(x) \geq f_n(x) \geq 0$ almost everywhere for each $m \geq n$. If we can find some $C \in \mathbb{R}$ such that
$$\int_X f_n \, d\mu \leq C$$
for each n, then $\{f_n\}_{n\in\mathbb{N}}$ is a mean Cauchy sequence.

Proof. There are two steps. For brevity, we set $I_n \coloneqq \int_X f_n \, d\mu$.

1. We compute $\|f_m - f_n\|_1$ when $m \geq n$. The main point is that $m \geq n$ implies $|f_m - f_n| = f_m - f_n$ almost everywhere. Indeed, there exists $E \in \mathcal{S}$ such that $\mu(E) = 0$ while $f_m(x) \geq f_n(x) \geq 0$ for each $x \in X \setminus E$, so
$$|f_m - f_n|(x) = |f_m(x) - f_n(x)| = f_m(x) - f_n(x) = (f_m - f_n)(x)$$
for each $x \in X \setminus E$. Thus, Remark 8.94 tells us that
$$\|f_m - f_n\|_1 = \int_X |f_m - f_n| \, d\mu = \int_X (f_m - f_n) \, d\mu.$$
As usual, the linearity of integration from Proposition 8.73 gives
$$\|f_m - f_n\|_1 \leq \int_X f_m \, d\mu - \int_X f_n \, d\mu = I_m - I_n.$$

2. We complete the proof. We know that $m \geq n$ implies $I_m - I_n = \|f_m - f_n\|_1 \geq 0$ (say, using Corollary 8.79), so $I_m \geq I_n$. Thus, $\{I_n\}_{n\in\mathbb{N}}$ is an increasing sequence, but we are given that $I_n \leq C$ for each n. It follows that $\{I_n\}_{n\in\mathbb{N}}$ is a Cauchy sequence!

   Finishing up, for any $\varepsilon > 0$, we are promised N such that $m, n \geq N$ implies $|I_m - I_n| < \varepsilon$. Thus, $m \geq n \geq N$ implies
$$\|f_m - f_n\|_1 \leq |I_m - I_n| < \varepsilon,$$
which finishes. ■

Remark 9.19. We work in the context of Theorem 9.18. Notably, by Proposition 8.95, we are granted some μ-integrable function $f \colon X \to B$ such that $f_n \to f$ in mean. Thus, Remark 8.89 tells us
$$\int_X f \, d\mu = \lim_{n\to\infty} \int_X f_n \, d\mu.$$

[2] Explicitly, for any $\varepsilon > 0$, find N_f such that $\|f(x) - f_n(x)\| < \varepsilon$, and define our N as $N \coloneqq \max\{N_f, N_g\}$.

9.3 November 28

It's the last week of class, so it's time to go off the rails.

9.3.1 Infinite Integrals

As an application of Theorem 9.18, we get the following convention.

Definition 9.20. Fix a measure space $(X, \mathcal{S}, \mu)$. Given an $\mathcal{S}$-measurable function $f\colon X \to \mathbb{R}$ such that $f(x) \geq 0$ always, we say

$$\int_X f\,d\mu := +\infty$$

if and only if there is a sequence of μ-integrable functions $f_n\colon X \to \mathbb{R}$ such that $f_{n+1}(x) \geq f_n(x) \geq 0$ for each x and n such that $f_n \to f$ pointwise and $\int_X f_n\,d\mu \to +\infty$.

For this definition to make sense, we need a few lemmas.

Lemma 9.21. Fix a measure space $(X, \mathcal{S}, \mu)$. Given an $\mathcal{S}$-measurable function $f\colon X \to \mathbb{R}$ such that $f(x) \geq 0$, there exists some sequence of μ-integrable functions $f_n\colon X \to \mathbb{R}$ such that $f_{n+1}(x) \geq f_n(x) \geq 0$ for each x and n such that $f_n \to f$ pointwise.

Proof. We have two steps.

1. Because f is $\mathcal{S}$-measurable, we may find a sequence of simple $\mathcal{S}$-measurable functions $g_n\colon X \to \mathbb{R}$ such that $g_n \to f$ pointwise. Thus, by Lemma 7.47, the functions $g_n 1_{[-n,n]}$ are still simple $\mathcal{S}$-measurable, but now we claim they are simple μ-integrable. Indeed, for any $y \in \mathbb{R} \setminus \{0\}$, we see that
$$\left(g_n 1_{[-n,n]}\right)^{-1}(\{y\}) \subseteq [-n, n+1)$$
because $x \notin [-n, n+1)$ gives $1_{[-n,n]}(x) = 0$; thus, $\mu\left((g_n 1_{[-n,n]})^{-1}(\{y\})\right) \leq (n+1) - -n < \infty$ by Lemma 5.51, which is what we wanted.
Further, we claim that $g_n 1_{[-n,n]} \to f$ pointwise as $n \to \infty$. Indeed, for any $x \in X$, fix some $\varepsilon > 0$. We may find some N_1 such that $n \geq N$ implies
$$|f(x) - g_n(x)| < \varepsilon.$$
As such, we set $N := \max\{N_1, |x|\}$ so that $n \geq N \geq |x|$ gives $g_n 1_{[-n,n]}(x) = g_n(x)$, and so $n \geq N \geq N_1$ gives $|f(x) - g_n(x)| < \varepsilon$.
2. Relabeling, the previous step constructed a sequence of simple μ-integrable functions $g_n\colon X \to \mathbb{R}$ converging to f pointwise. It remains to deal with our bounding. For this, we delete our sequence of functions f_n recursively. Define $f_1 = 0$, which is μ-integrable by Lemma 8.59.
Now, given f_n, we define f_{n+1} by
$$f_{n+1} := \max\{f_n, \min\{g_{n+1}, f\}\}.$$
Note $\min\{g_{n+1}, f\}$ is $\mathcal{S}$-measurable by Example 7.46 and thus μ-integrable by Corollary 9.17 because $\min\{g_{n+1}(x), f(x)\} \leq g_{n+1}(x)$ for each $x \in X$. Thus, we see f_{n+1} is μ-integrable (inductively) by Example 8.63. We also note $f_n(x) \leq f_{n+1}(x)$ for any n and x by construction, so we get $f_n(x) \geq f_1(x) = 0$.
It remains to check $f_n \to f$ pointwise; fix any $x \in X$. To begin, note $f_n(x) \leq f(x)$ for each n. For $n = 1$, this is by hypothesis on f, and in general we note that $f_n(x) \leq f(x)$ and $\min\{g_{n+1}(x), f(x)\} \leq f(x)$ forces $f_{n+1}(x) \leq f(x)$.
However, for all $\varepsilon > 0$, we can find some $N > 0$ such that $n \geq N$ implies $|g_n(x) - f(x)| < \varepsilon$. We claim that $|f_n(x) - f(x)| < \varepsilon$ for each $n \geq N > 0$ as well. There are two cases.

- If $g_n(x) \le f(x)$, then we note
$$g_n(x) = \min\{g_n(x), f(x)\} \le f_n(x) \le f(x),$$
so $|f(x) - f_n(x)| = f(x) - f_n(x) \le f(x) - g_n(x) < \varepsilon$.
- If $g_n(x) \ge f(x)$, then we note $f_n(x) \le f(x)$ while $f_n(x) \ge \min\{g_n(x), f(x)\} = f(x)$, so $f_n(x) = f(x)$.

The above checks complete the proof. ■

Proposition 9.22. Fix a measure space $(X, \mathcal{S}, \mu)$ and an $\mathcal{S}$-measurable function $f\colon X \to \mathbb{R}$ with $f(x) \ge 0$ for all $x \in X$. Then exactly one of the following is true.

- f is μ-integrable.
- $\int_X f\, d\mu = +\infty$.

Proof. By Lemma 9.21, there certainly exists some sequence of μ-integrable functions $f_n\colon X \to \mathbb{R}$ such that $f_{n+1}(x) \ge f_n(x) \ge 0$ for each x and n such that $f_n \to f$ pointwise. As such, note that the sequence of integrals
$$I_n := \int_X f_n\, d\mu$$
are increasing by Lemma 8.75. Thus, if the sequence is bounded above, we note f is μ-integrable by Theorem 9.18. Otherwise, $I_n \to +\infty$ as $n \to \infty$, so $\int_X f\, d\mu = +\infty$.

Thus, we have so far shown that at least one of the conclusions is true. It remains to show that they cannot both be true. Well, suppose $f\colon X \to \mathbb{R}$ is μ-integrable, and we show $\int_X f\, d\mu \ne +\infty$. If we have any increasing sequence $\{g_n\}_{n\in\mathbb{N}}$ of μ-integrable functions such that $g_n \to f$ pointwise, then we see $g_n(x) \le f(x)$ for each x, so we may use Lemma 8.75 to upper-bound
$$\int_X f_n\, d\mu \le \int_X f\, d\mu.$$
Thus, the sequence $\int_X f_n\, d\mu$ does not go to $+\infty$. ■

While we're here, we pick up a few of our standard bounds.

Lemma 9.23. Fix a measure space $(X, \mathcal{S}, \mu)$. Given $\mathcal{S}$-measurable functions $f, g\colon X \to \mathbb{R}$ such that $f(x), g(x) \ge 0$ for all $x \in X$, we have
$$\int_X (f+g)\, d\mu = \int_X f\, d\mu + \int_X g\, d\mu,$$
where we permit values to be $+\infty$.

Proof. Note that $f + g$ is $\mathcal{S}$-measurable by Lemma 7.25. We have two cases.

- If $\int_X (f+g)\, d\mu \ne \infty$, then $(f+g)$ is μ-integrable by Proposition 9.22. However, we note $f(x), g(x) \le (f+g)(x)$ for each $x \in X$ because $f(x), g(x) \ge 0$, so Lemma 8.75 tells us that f and g are both μ-integrable. Thus, the result follows from Proposition 8.73.
- Suppose $\int_X (f+g)\, d\mu = +\infty$. If f and g are both μ-integrable, $(f+g)$ is μ-integrable by Lemma 8.59, which violates the hypothesis of this case by Proposition 9.22. Thus, one of f or g is not μ-integrable. Without loss of generality, we say f is not μ-integrable, so
$$\int_X f\, d\mu = +\infty$$
follows from Proposition 9.22. Because $\int_X g\, d\mu \ge 0$ either when g is μ-integrable (by Lemma 8.75 and Proposition 8.73) or when g is not μ-integrable (by Proposition 9.22), the result follows. ■

Lemma 9.24. Fix a measure space $(X, \mathcal{S}, \mu)$. Given $\mathcal{S}$-measurable functions $f, g\colon X \to \mathbb{R}$ such that $f(x) \geq g(x) \geq 0$ for all $x \in X$, we have

$$\int_X f(x)\, d\mu \geq \int_X g(x)\, d\mu,$$

where we permit values to be $+\infty$.

Proof. We have two cases.

- If f is μ-integrable, then g is μ-integrable by Corollary 9.17, so the result follows from Lemma 8.75.
- If f is not μ-integrable, then $\int_X f\, d\mu = +\infty$ by Proposition 9.22, so the result follows. ■

9.3.2 Defining $\mathcal{L}^p$

Here is our definition.

Definition 9.25 ($\mathcal{L}^p$-space)**.** Fix a measure space $(X, \mathcal{S}, \mu)$ and a Banach space $(B, \|\cdot\|)$. Given some $p \in (0, \infty)$, we define

$$\mathcal{L}^p(X, \mathcal{S}, \mu, B) := \{\mathcal{S}\text{-measurable} f : \|f\|^p \text{ is } \mu\text{-integrable}\}.$$

Lemma 9.26. Fix a measure space $(X, \mathcal{S}, \mu)$ and a k-Banach space $(B, \|\cdot\|)$. Then $\mathcal{L}^p(X, \mathcal{S}, \mu, B)$ is a k-vector space.

Proof. We have the following checks. As usual, let $|\cdot|$ denote the norm on k.

- Zero: note the zero function $z\colon X \to B$ has $\|z\|^p(x) = 0$ for each $x \in X$, which is μ-integrable by Lemma 8.59.
- Scalar multiplication: if $f \in \mathcal{L}^p(X, \mathcal{S}, \mu, B)$, then $\|f\|^p$ is μ-integrable. However, for $r \in k$, we note that rf is $\mathcal{S}$-maesurable by Lemma 7.25, and

  $$\|rf\|^p(x) = (|r|^p \cdot \|f\|^p)(x)$$

  for each $x \in X$, so the fact that $\|f\|^p$ is μ-integrable implies that $\|rf\|^p$ is also μ-integrable. This finishes.
- Additive: suppose $f, g \in \mathcal{L}^p(X, \mathcal{S}, \mu, B)$ so that we want to show $f + g \in \mathcal{L}^p(X, \mathcal{S}, \mu, B)$, which means that we want $\|f+g\|^p \in \mathcal{L}^1(X, \mathcal{S}, \mu, B)$. Well, $\|f+g\|^p$ is $\mathcal{S}$-measurable by applying Lemma 7.25 and Corollary 7.45 and Corollary 7.43 (with the continuous function $x \mapsto |x|^p$), so we merely need to upper-bound $\|f+g\|^p$ and use Corollary 9.17.

  Indeed, the triangle inequality implies that

  $$\|f(x) + g(x)\| \leq \|f(x)\| + \|g(x)\| \leq 2\max\{\|f(x)\|, \|g(x)\|\}$$

  for each $x \in X$, so

  $$\begin{aligned} \|f(x) + g(x)\|^p &\leq 2^p \big(\max\{\|f(x)\|, \|g(x)\|\}\big)^p \\ &= 2^p \big(\max\{\|f(x)\|^p, \|g(x)\|^p\}\big) \\ &\leq 2^p \left(\|f(x)\|^p + \|g(x)\|^p\right). \end{aligned}$$

  However, each $\|f\|^p$ and $\|g\|^p$ are μ-integrable by hypothesis, so $2^p\|f\|^p + 2^p\|g\|^p$ is μ-integrable by Lemma 8.59. Thus, Corollary 9.17 finishes. ■

Here is a reason to care about $\mathcal{L}^p$: just like $\mathcal{L}^1$, they have a well-behaved semi-norm (in good cases).

Notation 9.27. Fix a measure space $(X, \mathcal{S}, \mu)$ and a Banach space $(B, \|\cdot\|)$. For $p \in (0, \infty)$ and $f \in \mathcal{L}^p(X, \mathcal{S}, \mu, B)$, we define

$$\|f\|_p := \left(\int_X \|f\|^p \, d\mu\right)^{1/p}.$$

Note that this integral is well-defined by definition of $L^p(X, \mathcal{S}, \mu, B)$.

Here is the analogue for Lemma 8.93.

Lemma 9.28. Fix a measure space $(X, \mathcal{S}, \mu)$ and a Banach space $(B, \|\cdot\|)$, and fix $p \in (0, \infty)$. Given some $f \in \mathcal{L}^p(X, \mathcal{S}, \mu, B)$, we see $\|f\|_p = 0$ if and only if $f(x) = 0$ almost everywhere.

Proof. By Lemma 8.93, we see that

$$\|f\|_p^p = \int_X \|f\|^p \, d\mu$$

equals zero if and only if $\|f\|^p(x) = 0$ almost everywhere. However, $\|f\|^p(x) = 0$ is equivalent to saying $\|f(x)\|^p = 0$, which is equivalent to $\|f(x)\| = 0$, which is equivalent to $f(x) = 0$. Thus, one side of

$$\{x \in X : \|f\|^p(x) \neq 0\} = \{x \in X : f(x) \neq 0\}$$

is a null set if and only if the other is, which finishes. ■

Corollary 9.29. Fix a measure space $(X, \mathcal{S}, \mu)$ and a Banach space $(B, \|\cdot\|)$, and fix $p \in (0, \infty)$. Given some $f, g \in \mathcal{L}^p(X, \mathcal{S}, \mu, B)$, we see $\|f - g\|_p = 0$ if and only if $f(x) = g(x)$ almost everywhere.

Proof. Note that $\|f - g\|_p = 0$ is equivalent to $(f - g)(x) = 0$ almost everywhere by Lemma 9.28, which is equivalent to $\{x \in X : f(x) \neq g(x)\}$ being a null set, which is what we wanted. ■

Continuing, we now show that $\|\cdot\|_2$ is a semi-norm.

Proposition 9.30. Fix a measure space $(X, \mathcal{S}, \mu)$ and a Banach space $(B, \|\cdot\|)$. Then the function $\|\cdot\|_2$ defines a semi-norm on $\mathcal{L}^2(X, \mathcal{S}, \mu, B)$.

Proof. We quickly run our easier checks. Let $|\cdot|$ denote the norm on our base field k, and fix some $f \in \mathcal{L}^2(X, \mathcal{S}, \mu, B)$.

- Zero: the zero function $z \colon X \to B$ has $\|z\|^2 : X \to \mathbb{R}$ equal to zero everywhere, so $\|z\|^2 = 1_\varnothing$, so

  $$\int_X \|z\|^2 \, d\mu = \mu(\varnothing) = 0$$

  by Example 7.16.
- Nonnegative: we note that $\|f\|(x) \geq 0$ for each $x \in X$, so $\|f\|^2(x) \geq 0$ for each $x \in X$, so

  $$\int_X \|f\|^2 \, d\mu \geq 0$$

  by Lemma 8.74, so $\|f\|_2 \geq 0$ follows.
- Homogeneous: if $r \in k$, then we note $\|rf\|(x) = (|r| \cdot \|f\|)(x)$ for each $x \in X$. Thus, by Proposition 8.73 tells us

  $$\|rf\|_2 = \left(\int_X \|rf\|^2 \, d\mu\right)^{1/2} = \left(\int_X |r| \cdot \|f\|^2 \, d\mu\right)^{1/2} = |r| \cdot \left(\int_X \|f\|^2 \, d\mu\right)^{1/2} = |r| \cdot \|f\|_2 .$$

It remains to check the triangle inequality, which is harder. We have the following lemma.

Lemma 9.31 (Cauchy–Schwarz)**.** Fix a measure space $(X, \mathcal{S}, \mu)$ and a Banach space $(B, \|\cdot\|)$. Given $\mathcal{S}$-measurable functions $f, g \in \mathcal{L}^2(X, \mathcal{S}, \mu, B)$, then $\|f\| \cdot \|g\|$ is μ-integrable, and

$$\int_X (\|f\| \cdot \|g\|)\, d\mu \le \frac{\|f\|_2^2 + \|g\|_2^2}{2}.$$

Proof. Quickly, we see $\|f\| \cdot \|g\|$ is $\mathcal{S}$-measurable by applying Lemma 7.26 to Corollary 7.45. It remains to upper-bound $\|f\| \cdot \|g\|$.

Now, the main point is the arithmetic mean-geometric mean inequality: for $r, s \in \mathbb{R}_{\ge 0}$, we see

$$0 \le (r-s)^2 = r^2 + s^2 - 2rs,$$

so $r^2 + s^2 \ge 2rs$. Applying this to our situation, we see

$$2\|f(x)\| \cdot \|g(x)\| \le \|f(x)\|^2 + \|g(x)\|^2$$

for each $x \in X$, so

$$(\|f\| \cdot \|g\|)(x) \le \left(\frac{\|f\|^2 + \|g\|^2}{2}\right)(x)$$

for each $x \in X$. However, $\frac{\|f\|^2+\|g\|^2}{2}$ is μ-integrable by Lemma 8.59, so we conclude that $\|f\| \cdot \|g\|$ is μ-integrable by Corollary 9.17.

Continuing, Lemma 8.75 tells us

$$\int_X (\|f\| \cdot \|g\|)\, d\mu \le \int_X \frac{\|f\|^2 + \|g\|^2}{2}\, d\mu,$$

and now the right-hand side simplifies to $\frac{1}{2}\left(\|f\|_2^2 + \|g\|_2^2\right)$ by Proposition 8.73. This finishes. ■

We now proceed with the proof of the triangle inequality. Fix $f, g \in \mathcal{L}^2(X, \mathcal{S}, \mu, B)$. We quickly deal with the case of $\|f\|_2 = 0$. Here, $\|f\|_2 = 0$ implies that $f(x) = 0$ almost everywhere by Lemma 9.28, so $(f+g)(x) = g(x)$ almost everywhere, so $\|f+g\|^2(x) \le \|g\|^2(x)$ almost everywhere, so Lemma 8.75 implies

$$\int_X \|f+g\|^2\, d\mu \le \int_X \|g\|^2\, d\mu \le \int_X \|f+g\|^2\, d\mu.$$

Thus, $\|f+g\|_2 \le \|g\|_2 = \|f\|_2 + \|g\|_2$ follows. Note that a similar argument works for $\|g\|_2 = 0$.

Thus, we may assume that $\|f\|_2, \|g\|_2 \ne 0$, which allows us to set $h := \frac{f}{\|f\|_2}$ and $k := \frac{g}{\|g\|_2}$. Notably, $\|h\|_2 = \|k\|_2 = 1$ by the homogeneity check above. As such, Lemma 9.31 grants

$$\int_X (\|h\| \cdot \|k\|)\, d\mu \le \frac{1+1}{2} = 1.$$

However, $\|h\| = \|f\| / \|f\|_2$ and $\|k\| = \|g\| / \|g\|_2$, so Proposition 8.73 implies

$$\int_X (\|f\| \cdot \|g\|)\, d\mu \le 2\|f\|_2 \cdot \|g\|_2.$$

This now rearranges to the desired inequality: given $f, g \in \mathcal{L}^2(X, \mathcal{S}, \mu, B)$, we see

$$\begin{aligned}
\|f+g\|_2^2 &= \int_X (\|f+g\|)^2\, d\mu \\
&\overset{*}{\le} \int_X (\|f\| + \|g\|)^2\, d\mu \\
&= \int_X \|f\|^2\, d\mu + \int_X \|g\|^2\, d\mu + 2\int_X (\|f\| \cdot \|g\|)\, d\mu \\
&\le \|f\|_2^2 + \|g\|_2^2 + 2\|f\|_2 \cdot \|g\|_2 \\
&= (\|f\|_2 + \|g\|_2)^2,
\end{aligned}$$

and taking the square root finishes. Notably, $\overset{*}{\leq}$ has used the triangle inequality and Lemma 8.75 (and the following equality used Proposition 8.73). ■

Remark 9.32. In fact, $\|f\|_p$ is a norm in general, but it is somewhat harder to show. Roughly speaking, the difficulty lies in establishing an analogue for Lemma 9.31.

Remark 9.33. If $p \in (0,1)$, then $\|f\|_p$ is not a norm. In particular, it does not satisfy the triangle inequality. For that matter, $p \in (0,1)$ do very strange things. For example, if we define

$$U_r := \left\{ f \in \mathcal{L}^p(X, \mathcal{S}, \mu, B) : \int_X \|f\|^p \, d\mu < r \right\},$$

then the convex hull of U_r recovers all of $\mathcal{L}^p(X, \mathcal{S}, \mu, B)$. In particular, one can show that there are thus no nonzero continuous linear functionals on $\mathcal{L}^p(X, \mathcal{S}, \mu, B)$.

9.3.3 Defining L^2

Now that we have a semi-norm, we can mod out to get our norm.

Notation 9.34. Fix a measure space $(X, \mathcal{S}, \mu)$ and a Banach space B. We set $\mathcal{N}(X, \mathcal{S}, \mu, B) := \{f \in \mathcal{L}^2(X, \mathcal{S}, \mu, B) : \|f\|_2 = 0\}$ and

$$L^2(X, \mathcal{S}, \mu, B) := \mathcal{L}^2(X, \mathcal{S}, \mu, B)/\mathcal{N}(X, \mathcal{S}, \mu, B).$$

Remark 9.35. Given $f, g \in \mathcal{L}^p(X, \mathcal{S}, \mu, B)$, note that $[f] = [g]$ in $L^p(X, \mathcal{S}, \mu, B)$ if and only if $\|f - g\|_p = 0$ by definition, which is equivalent to $f(x) = g(x)$ almost everywhere by Corollary 9.29.

Proposition 9.36. Fix a measure space $(X, \mathcal{S}, \mu)$ and a Banach space B. The function $\|\cdot\|_2$ descends to a norm on $L^2(X, \mathcal{S}, \mu, B)$.

Proof. We have a semi-norm by Proposition 9.30, which descends to a norm on $L^2(X, \mathcal{S}, \mu, B)$ from Proposition 1.13. ■

We would like to show that L^2 is complete, but this requires some work. Namely, we will require Fatou's lemma, a result we will state and prove next class.

9.4 November 30

The final is in about two weeks. Material covered this week may appear on the exam. Material covered in the topology section of the course may also appear on the exam.

9.4.1 Completeness of L^2

We continue moving towards proving the completeness of L^2. We pick up the following result.

Lemma 9.37 (Fatou)**.** Fix a measure space $(X, \mathcal{S}, \mu)$. Further, fix a sequence $\{f_n\}_{n\in\mathbb{N}}$ of $\mathcal{S}$-measurable functions $f_n \colon X \to \mathbb{R}^+$. Then

$$\int_X \left(\liminf_{n\to\infty} f_n\right) d\mu \le \liminf_{n\to\infty} \int_X f_n \, d\mu.$$

Proof. For $m \ge n$, define

$$h_{n,m} := \min\{f_n, f_{n+1}, \ldots, f_m\},$$

which is also $\mathcal{S}$-measurable by Example 7.46. Notably, for fixed n, the functions $h_{n,n}, h_{n+1,n}, h_{n+2,n}, \ldots$ are decreasing as $m \to \infty$ (adding more terms to this minimum requires values to decrease), so there is a limit function

$$g_n(x) := \inf\{h_{n,m}(x) : m \ge n\} = \lim_{m\to\infty} h_{n,m}(x),$$

which is $\mathcal{S}$-measurable as the pointwise limit of $\mathcal{S}$-measurable functions. Note that $g_n(x)$ is always a real number because the set $\{h_{n,m}(x) : n \ge m\}$ is bounded below by 0. However, we can see that the $g_n(x)$ are monotonically increasing (taking fewer terms in our infimum requires values to increase), so we define

$$\left(\liminf_{n\to\infty} f_n\right)(x) := \liminf_{n\to\infty} f_n(x) = \lim_{n\to\infty} g_n(x),$$

where $+\infty$ is a permitted value.

Quickly, fixing some n, any $m \ge n$ has $h_{n,m}(x) \le f_n(x)$ for all x, so the limit function $g_n(x) \le f_n(x)$ for all x. Thus, by Theorem 9.18, we see

$$\int_X \left(\liminf_{n\to\infty} f_n\right) d\mu = \lim_{n\to\infty} \int_X g_n \, d\mu,$$

where we are still permitting $+\infty$. However, $g_n(x) \le f_n(x)$ for each x by construction, so

$$\int_X (\liminf_{n\to\infty} f_n) \, d\mu \le \liminf_{n\to\infty} \int_X f_n \, d\mu,$$

which is what we wanted. ■

And now here is our result.

Theorem 9.38. Fix a measure space $(X, \mathcal{S}, \mu)$ and a Banach space $(B, \|\cdot\|)$. Then $L^2(X, \mathcal{S}, \mu, B)$ is complete.

Proof. Fix a sequence of functions $\{f_n\}_{n\in\mathbb{N}}$ of functions in $\mathcal{L}^2(X, \mathcal{S}, \mu, B)$ which are Cauchy for $\|\cdot\|_2$. We claim that $\{f_n\}_{n\in\mathbb{N}}$ is Cauchy in measure. For this, we pick up the following inequality.

Lemma 9.39 (Chebychev)**.** Fix a measure space $(X, \mathcal{S}, \mu, B)$ and a Banach space $(B, \|\cdot\|)$. Given some $h \in \mathcal{L}^2(X, \mathcal{S}, \mu, B)$ and $\varepsilon > 0$, the set $E := \{x \in X : \|h(x)\| \ge \varepsilon\}$ has finite measure with

$$\mu(E) \le \frac{\|h\|_2^2}{\varepsilon^2}.$$

Proof. Note that E is $\mathcal{S}$-measurable. Now, the indicator function has

$$1_E(x) \le \frac{\|h(x)\|}{\varepsilon} \le \left(\frac{\|h(x)\|}{\varepsilon}\right)^2,$$

so it follows from Lemma 8.75 that 1_E is μ-integrable, and

$$\mu(E) \leq \int_X 1_E \, d\mu \leq \frac{1}{\varepsilon^2} \int_X \|h\|^2 \, d\mu = \frac{\|h\|_2^2}{\varepsilon^2},$$

which is what we wanted. ■

Now, for any m and n, the above inequality gives

$$\mu(\{x \in X : \|f_n(x) - f_m(x)\| \geq \varepsilon\}) \leq \frac{\|f_m - f_n\|_2^2}{\varepsilon^2},$$

but $\|f_m - f_n\|_2 \to 0$ as $m, n \to \infty$. Thus, $\{f_n\}_{n\in\mathbb{N}}$ is Cauchy in measure, so it follows that some subsequence $\{f_{n_k}\}_{k\in\mathbb{N}}$ converges to some $\mathcal{S}$-measurable function $f \colon X \to B$, almost uniformly using arguments similar to Proposition 8.95. It remains to show that $f \in \mathcal{L}^2(X, \mathcal{S}, \mu, B)$ and $f_{n_k} \to f$ in $L^2(X, \mathcal{S}, \mu, B)$. We go ahead and re-index our sequence so that $f_n \to f$ almost uniformly, which is legal because Cauchy sequences with a convergent subsequence will in total converge to the same limit as the subsequence.

We will actually show that $\|f - f_n\|_2 \to 0$ as $n \to \infty$; this will finish because it forces $\|f\|_2$ to be close to $\|f_n\|_2$ for n large enough and in particular finite. For this, we consider the integral

$$\int_X \|f - f_n\|^2 \, d\mu,$$

which is a legal expression provided that we permit $+\infty$. Now, $f_m \to f$ almost uniformly as $m \to \infty$, so $\|f_m - f_n\| \to \|f - f_n\|$ almost uniformly as $m \to \infty$, so

$$\liminf_{m\to\infty} \|f_m - f_n\|^2 (x) = \|f - f_n\|^2 (x)$$

for each $x \in X$. Thus, by Lemma 9.37, we see

$$\int_X \|f - f_n\|^2 \, d\mu \leq \liminf_{m\to\infty} \int_X \|f_m - f_n\|^2 \, d\mu = \liminf_{m\to\infty} \|f_m - f_n\|_2^2 .$$

As such, for any $\varepsilon > 0$, we select N such that $m, n \geq N$ has $\|f_m - f_n\|^2 < \varepsilon/2$, so it follows that $\|f - f_n\|_2 < \varepsilon$ from the above bounding. This completes the proof. ■

Remark 9.40. The above proof will work for any $p \in [1, \infty)$.

We close class by noting we have made a Hilbert space.

Definition 9.41 (Hilbert space)**.** A *Hilbert space* is a vector space V over $\mathbb{R}$ or $\mathbb{C}$ equipped with an inner product $\langle \cdot, \cdot \rangle$ such that V is complete for the norm defined by $\|v\| \coloneqq \langle v, v\rangle^{1/2}$.

Example 9.42. In the usual set-up, we can make $L^2(X, \mathcal{S}, \mu, \mathbb{R})$ into a Hilbert space by

$$\langle f, g\rangle \coloneqq \int_X fg \, d\mu.$$

Notably, $\langle f, f\rangle^{1/2} = \|f\|_2$. A similar definition works for $L^2(X, \mathcal{S}, \mu, \mathbb{C})$ by conjugating g in the integral.

Remark 9.43. One can show that $L^2(X, \mathcal{S}, \mu, \mathbb{R})$ is "self-dual" in that every linear functional arises in the form $\langle f, \cdot\rangle$. (More generally, the dual of L^p is L^q, where $\frac{1}{p} + \frac{1}{q} = 1$.) This is one reason why L^2 is better than other L^ps.

Next class we will discuss L^∞.

BIBLIOGRAPHY

[Jac90] Joseph Jacobs. *English Fairy Tales*. English Fairy Tales. David Nutt, 1890.

[Wei59] Andre Weil. "Correspondence". In: *Annals of Mathematics* 69.1 (1959), pp. 247–251.

[Gei60] Theodor Geisel. *One Fish, Two Fish, Red Fish, Blue Fish*. Beginner Books. Beginner Books, 1960.

[BJ82] Theodor Bröcker and Klaus Jänich. *Introduction to Differential Topology*. Cambridge University Press, 1982.

[Tao09] Terrence Tao. *245B, notes 3: L^p spaces*. 2009. URL: https://terrytao.wordpress.com/2009/01/09/245b-notes-3-lp-spaces/.

[Lan12] Serge Lang. *Real and Functional Analysis*. Graduate Texts in Mathematics. Springer New York, 2012. URL: https://books.google.com/books?id=KOMlBQAAQBAJ.

[Doe14] Anthony Doerr. *All the Light We Cannot See*. Scribner, 2014.

[Ked21] Kiran S. Kedlaya. *Notes on Class Field Theory*. 2021. URL: https://kskedlaya.org/papers/cft-ptx.pdf.

[Che22] Evan Chen. *An Infinitely Large Napkin*. 2022. URL: https://venhance.github.io/napkin/Napkin.pdf.

List of Definitions