Lecture 1

Why measure theory?

We want to answer questions like the following: what is the "total length" of an arbitrary $E \subseteq \mathbb{R}$? What about the "total area" of an arbitrary $E \subseteq \mathbb{R}^2$?
In other words, can we define a function $\mu : 2^{\mathbb{R}^d} \to [0, +\infty)$ so that $\mu(E)$ is sufficiently "nice?"
What properties would we like a function μ (called a measure) to have? Let's stick to $\mathbb{R}$ for now.

1. For $E = [a, b]$ (or (a, b)), we'd like $\mu(E) = b - a$.

2. For a sequence of disjoint intervals $I_i \subseteq \mathbb{R}$,

$$\mu\left(\bigcup_{i=1}^{n} I_i\right) = \sum_{i=1}^{n} \mu(I_i)$$

What about $\mathbb{Q} \cap [0, 1]$? What about the area under a curve?
What if E is any arbitrary set????

Pre measure theory

In the mid 1800s, Riemann first defined the Riemann integral in terms of upper and lower sums.
Fortunately, it's good enough for most "ordinary" functions a student might encounter when doing calculus for the first time.
Unfortunately, it's not good for taking limits.
For example, given $f_1, f_2, f_3, \cdots : [a, b] \to \mathbb{R}$ such that $\lim_{n\to\infty} f_n(x) \stackrel{\text{def}}{=} f(x)$ exists for all x, when can we conclude that

$$\lim_{n\to\infty} \int_a^b f_n(x)\,dx = \int_a^b \lim_{n\to\infty} f_n(x)\,dx = \int_a^b f(x)\,dx$$

We learn in undergraduate real analysis that we may only conclude the above if the f_n converge uniformly to f.

Measure theory

Measure theory allows us to define a much more powerful theory of integration, giving us

- More integrable functions
- An integral which behaves better with limits
- An integral ideally suited for probability theory.

Our first goal will be to define a function $\mu : 2^{\mathbb{R}} \to [0, \infty)$ satisfying the following:

1. If $E_1, E_2, \dots$ is a countable sequence of disjoint sets, then

$$\mu\left(\bigcup E_i\right) = \sum \mu(E_i)$$

If μ satisfies this, we say it is "countably additive."

2. $\mu([a, b]) = b - a$ for all such intervals.

3. μ is translation invariant, i.e. for any $t \in \mathbb{R}$,

$$\mu(E + t) = \mu(E)$$

Where "$E + t$" $\stackrel{\text{def}}{=} \{x + t \mid x \in E\}$

Theorem 0.1. *(Vitali) There is no such μ.*

Proof. Suppose that such a μ exists.

Claim. *If $A \subseteq B$, then $\mu(A) \leq \mu(B)$.*

Proof. Note $B = A \coprod (B \backslash A)$, so

$$\mu(B) = \mu(A) + \mu(B \backslash A)$$

And because μ is always nonnegative, we may conclude that $\mu(B) \geq \mu(A)$.

■

Now, define an equivalence relation on $\mathbb{R}$ as follows:

$$x \sim y <=> x - y \in \mathbb{Q}$$

$$[x] \stackrel{\text{def}}{=} \{y \in \mathbb{R} \mid x \sim y\}$$

Claim. *Every equivalence class contains a point in* $[0, 1]$.

Proof. Homework exercise.

■

Now, for each equivalence class, choose an element in $[0, 1]$ belonging to that class. For this step, we are using choice and, it turns out, there is no way not to in this proof.

Call the resulting set A. So $A \subseteq [0,1]$, and for any x, $[x] \cap A$ is a singleton.
Let $B = \bigcup_{q \in \mathbb{Q} \cap [-1,1]} A + q$
Note that this is a disjoint union: indeed, if $A+q$ intersects nontrivially with $A+q'$ for $q \neq q'$, then there are x, x' in A such that $x = x' + q$, and so $x \sim x'$, which by construction is impossible.

Claim. $[0,1] \subseteq B \subseteq [-1,2]$.

Proof. First, if $x \in [0,1]$, then $x = a + q$ for some $a \in A \subseteq [0,1], q = x - a \in [-1,1]$. Thus, $x \in B$.
Next, if $b \in B$, then $b = a + q$, for $q \in A = [0,1]$ and $q \in [-1,1]$, so $b \in [-1,2]$.
So we must conclude by the lemma that

$$1 = \mu([1,0]) \leq \mu(B) \leq \mu([-1,2]) = 3$$

But by the properties of μ, we also have

$$\mu(B) = \sum_{q \in \mathbb{Q} \cap [0,1]} \mu(A+q) = \sum_{q \in \mathbb{Q} \cap [0,1]} \mu(A)$$

The sum on the right hand side is either 0 or ∞. But we just showed that it is between 1 and 3, a contradiction. Therefore, $\mu(A)$ cannot be defined.

■

So, if this is impossible, which criterion should we weaken to make it possible?
If we weaken the first to get finite additivity, we run into problems for $d \geq 3$, for example the Banach-Tarski paradox.
If we weaken the other two, then μ is no longer compactible with the usual notion of "length."

Two good choices

- Given a measure on a family of sets, it extends to an outer measure on all sets.
- Similarly, given an outer measure, you can single out "nice sets" on which it is a measure.

What kind of family of subsets should we restrict to?
Let X be a set.

Definition 0.1. $\mathcal{A}$ is an <u>algebra of subsets of X</u> if $\mathcal{A} \neq \varnothing$, and

- $E_1, \ldots, E_n \in \mathcal{U} \implies \bigcup_{i=1}^{n} E_i \in \mathcal{U}$. In other words, it is "closed under finite unions."
- $E \in \mathcal{U} \implies X \backslash E \in \mathcal{U}$. In other words it is "closed under compliments."

- $\varnothing, X \in \mathcal{A}$.

Lemma 1. *If $\mathcal{A}$ is an algebra of subsets, then $\mathcal{A}$ is closed under finite intersections.*

Proof. Homework 2

Example 0.1. (i) $\mathcal{A} = 2^X$

(ii) $\mathcal{A} = \{\varnothing, X\}$

(iii) $\mathcal{A}$ = all finite OR cofinite subsets of X (cofinite means the complement is finite).

Definition 0.2. A σ-algebra $\mathcal{A}$ is an algebra that is closed under countable unions.

Remark. σ-algebras are closed under countable intersections.

Example 0.2. Above, (i) and (ii) are σ-algebras, but (iii) is not.

Proposition 1. *Given any family $\mathcal{E}$ of subsets of X, there is a smallest σ-algebra $\mu(\mathcal{E})$ containing $\mathcal{E}$, meaning that if $\mathcal{F}$ is a σ-algebra containing $\mathcal{E}$, then $\mu(\mathcal{E}) \subseteq \mathcal{F}$.*

Lecture 2

Definition 0.3. Given a nonempty set X and $\mathcal{M}$ a σ-algebra of subsets of X, we call $(X, \mathcal{M})$ a measurable space.
Recall:

Proposition 2. *Given any family $\mathcal{E}$ of subsets of X, there is a smallest σ-algebra $\mathcal{M}(\mathcal{E})$ containing $\mathcal{E}$, meaning that if $\mathcal{F}$ is a σ-algebra containing $\mathcal{E}$, then $\mathcal{M}(\mathcal{E}) \subseteq \mathcal{F}$.*

Proof. We begin with a claim.

Claim. Given any nonempty collection $\mathcal{C}$ of σ-algebras on X, then

$$\cap\mathcal{C} \overset{\text{def}}{=} \{E \subseteq X \mid E \in \mathcal{A} \forall \mathcal{A} \in \mathcal{C}\}$$

is a σ-algebra.

Proof. Homework 2

■

Let $\mathcal{C} = \{\mathcal{A} \mid \mathcal{A} \text{ is a } \sigma\text{-algebra on } X \text{ and } \mathcal{E} \subseteq \mathcal{A}\}$. $\mathcal{C}$ is nonempty, because $2^X \in \mathcal{C}$. By the claim, $\cap\mathcal{C}$ is a σ-algebra. By the definition of $\mathcal{C}$, $\mathcal{E} \subseteq \mathcal{C}$ and for any σ-algebra $\mathcal{A}$ such that $\mathcal{E} \subseteq \mathcal{A}$, $\cap\mathcal{C} \subseteq \mathcal{A}$.
Thus $\mathcal{M}(\mathcal{E}) = \cap\mathcal{C}$ is the smallest σ-algebra containing $\mathcal{E}$.

■

Remark. Intuitively, $\mathcal{M}(\mathcal{E})$ is a σ-algebra containing the sets in $\mathcal{E}$ by "going from the outside in," starting with σ-algebras that are "too big" and taking intersections.
Recall: a topology τ is a collection of subsets of a set X (called open sets), which is closed under arbitrary unions and finite intersections, and $X, \varnothing \in \tau$
Let (X, τ) be a topoological space.

Definition 0.4. The Borel σ-algebra of X, denoted $\mathcal{B}X$, is the σ-algebra generated by the open subsets of X. Its members are known as Borel sets.
What do the Borel sets look like? Let's go from the "inside out."
Let $\mathcal{F}$ =open sets in X, $\mathcal{F}^\sigma$ all countable unions of sets in $\mathcal{F}$, $\mathcal{F}^\delta$ all countable intersections, and $\overline{\mathcal{F}}$ complements of sets in $\mathcal{F}$.
To build Borel sets:

$$\mathcal{F} \to \mathcal{F}^\delta \cup \overline{\mathcal{F}^\delta} \to \cdots \to \mathcal{B}X$$

To learn more, look up the "Borel hierarchy."

Proposition 3. *The Borel σ-algebra on $\mathbb{R}$, which we denote $\mathcal{B}_\mathbb{R}$, is generated by each of the following.*

(i) Open intervals $\mathcal{E}_1 = \{(a, b) \mid a < b, a, b \in \mathbb{R}\}$

(ii) Closed intervals $\mathcal{E}_2 = \{[a, b] \mid a \le b, a, b, \in \mathbb{R}\}$

(iii) Half-open intervals $\mathcal{E}_3 = \{[a, b) \mid a < b, a, b, \in \mathbb{R}\}$

(iv) Open rays $\mathcal{E}_4 = \{(a, \infty) \mid a \in \mathbb{R}\}$

(v) Closed rays $\mathcal{E}_5 = \{[a, \infty) \mid a \in \mathbb{R}\}$

That is, $\mathcal{M}(\mathcal{E}_i) = \mathcal{B}_\mathbb{R}$ *for any* $i \in \{1, \ldots, 5\}$.

Proof. Homework 2 ■

Let

- $\{(X_1, \mathcal{M}_i)\}_{i=1}^\infty$ be a collection of measurable spaces.

$$X \overset{\text{def}}{=} \prod_{i=1}^{\infty} X_i$$

- π_i be the projection $X \to X_i$

Example 0.3. If $(X_i, \mathcal{M}_i) = (\mathbb{R}, \mathcal{B}_\mathbb{R})$, for $i \in \{1, \ldots, n\}$. Then $X = \mathbb{R}^n$.

Definition 0.5. The product σ-algebra

$$\bigotimes_{i\in\mathbb{N}}\mathcal{M}_i \stackrel{\text{def}}{=} \mathcal{M}\left(\{\prod_{i\in\mathbb{N}}E_i \mid E_i\in\mathcal{M}_i\}\right)$$

Our goal is to show that $\mathcal{B}_{\mathbb{R}^n} = \otimes_{i=1}^n\mathcal{B}_{\mathbb{R}}$.

Proposition 4. *Given $\mathcal{E}_i \subseteq 2^{X_i}$ such that $X_i\in\mathcal{E}_i$, let $\mathcal{M}_i = \mathcal{M}(\mathcal{E}_i)$. Then*

$$\bigotimes_{i\in\mathbb{N}}\mathcal{M}_i = \mathcal{M}\left(\{\prod_{i\in\mathbb{N}}E_i \mid E_i\in\mathcal{E}_i\}\right)$$

Note: If $\mathcal{E}\subseteq\mathcal{F}$, then $\mathcal{M}(\mathcal{E})\subseteq\mathcal{M}(\mathcal{F})$. If $\mathcal{E}\subseteq\mathcal{M}(\mathcal{F})$, then $\mathcal{M}(\mathcal{E})\subseteq\mathcal{M}\mathcal{F}$.
Recall: Given a function $f: X\to Y$ between arbitrary nonempty sets, then

(i) $f^{-1}\left(\cup_{i\in\mathbb{N}}E_i\right) = \cup_{i\in\mathbb{N}}f^{-1}(E_i)$ *for all* $E_i\subseteq X$.

(ii) $f^{-1}(E^c) = (f^{-1}(C))^c$, *for all* $E\subseteq X$.

Proof. By the first statement of the preceding note,

$$\mathcal{M}\left(\{\prod_{i\in\mathbb{N}}E_i \mid E_i\in\mathcal{E}_i\}\right) \subseteq \mathcal{M}\left(\{\prod_{i\in\mathbb{N}}E_i \mid E_i\in\mathcal{M}_i\}\right) = \bigotimes_{i\in\mathbb{N}}\mathcal{M}_i$$

For equality, it suffices to show

$$\mathcal{M}\left(\{\prod_{i\in\mathbb{N}}E_i \mid E_i\in\mathcal{M}_i\}\right) \subseteq \mathcal{M}\left(\{\prod_{i\in\mathbb{N}}E_i \mid E_i\in\mathcal{E}_i\}\right)$$

Let $\mathcal{M}\left(\{\prod_{i\in\mathbb{N}}E_i \mid E_i\in\mathcal{E}_i\}\right) = \mathcal{A}$.
Note that

$$\begin{aligned}\prod_{i\in\mathbb{N}}E_i &= \{x\in X \mid \pi_i(x)\in E_i\forall i\}\\ &= \bigcap_{i\in\mathbb{N}}\{x\in X : \pi_i(x)\in E_i\}\\ &= \bigcap_{i\in\mathbb{N}}\pi_i^{-1}(E_i)\end{aligned}$$

Because $\mathcal{A}$ is a σ-algebra, it suffices to show that $\pi_i^{-1}(E_i)\in\mathcal{A}$ for all $E_i\in\mathcal{M}_i$.

Claim. *Let* $\mathcal{F}_i \overset{\text{def}}{=} \{E_i \subseteq X_i \mid \pi_i^{-1}(E_i) \in \mathcal{A}\}$.
This is a σ*-algebra.*

Proof.

$$\bigcup_{i=1}^{\infty} \pi_i^{-1}(E_i) = \pi_i^{-1}(\bigcup_{i=1}^{\infty} E_i)$$

So $\{E_i\}_{i=1}^{\infty} \subseteq \mathcal{F}_i$.
Similarly,

$$\pi_i^{-1}(E^c) = (\pi_i^{-1}(E))^c$$

so $E \in \mathcal{F}_i$ implies $E^c \in \mathcal{F}_i$.

Lecture 3

Because $X_i \in \mathcal{E}_i$and $\pi_i^{-1}(E_i) = X_1 \times X_2 \times \cdots \times E_i \times \cdots$, we know $\pi_i^{-1}(E_i) \in \mathcal{A}$ for all $E_i \in \mathcal{E}_i$.
In other words, $\mathcal{E}_i \subseteq \mathcal{F}_i$. Since $\mathcal{F}_i$ is a σ-algebra, $\mathcal{M}(\mathcal{E}_i) = \mathcal{M}_i \subseteq \mathcal{F}_i$.
Thus, $\pi_i^{-1}(\mathcal{E}_i) \in \mathcal{A}$ for all $E_i \in \mathcal{M}_i$ and all i.

■

In order to characterize the Borel product σ-algebra, it will be convenient to assume our undderlying spaces have a metric that induces the topology.
Let $(X_i, d_i), i = 1, \ldots, n$ be metric spaces. Let

$$X = \prod_{i=1}^{n} X_i$$

Endow the product space with the metric

$$d_{\max}((x_1, x_2, \ldots, x_n), (y_1, y_2, \ldots, y_n)) = \max_{i=1,\ldots,n} (d_1(x_1, y_1), d_2(x_2, y_2), \ldots, d_n(x_n, y_n))$$

Theorem 0.2. *Given metric spaces* $X_1, X_2, \ldots, X_d$ *and their product*

$$X = \prod_{i=1}^{d} X_i$$

endowed with the metric $d_{\max}$*, then* $\otimes_{i=1}^{n} \mathcal{B}_{X_i} \subseteq \mathcal{B}_X$
If the X_i *are all seperable, then* $\otimes_{i=1}^{n} \mathcal{B}_{X_i} = \mathcal{B}_X$

Remark. Since the definition of $\mathcal{B}_X$ only depends on the topology of X, then this statement holds even if $d_{\max}$ is replaced by an equivalent metric, where "equivalent" means "generates the same topology."

Remark. $d_{\max}$ is convenient because:

$$\begin{aligned}
B_r(x_1, \ldots, x_m) &= \{(y_1, \ldots, y_n) | d_{\max}(\vec{x}, \vec{y}) < r\} \\
&= \{(y_2, \ldots, y_n) \mid d_i(x_i, y_i) < r \forall i\} \\
&= \prod_{i=1}^{n} B_r(X_i)
\end{aligned}$$

Recall:
Fact 1: If $X_1, \ldots, X_m$ are seperable, so is $\prod_{i=1}^m X_i$.
Fact 2: In a seperable metric space, every open set can be written as a countable union of balls, $\mathcal{U} = \cup_{i=1}^{\infty} B_i$
Fact 3: $\{\prod_{i=1}^n E_i \mid E_i \subseteq X_i, \text{open}\} \subseteq \{\text{open subsets of } X\}$

Proof. : By the previous proposition, $\otimes_{i=1}^n \mathcal{B}_{X_i}$ is generated by

$$\{\prod_{i=1}^{n} E_i \mid E_i \subseteq X_{\text{open}}\} \subseteq \{\text{open subsets of } X\}$$

Thus $\otimes_{i=1}^n \mathcal{B}_{X_i} \subseteq \mathcal{B}_X$.
Now, suppose $X_1, \ldots, X_n$ are seperable. By facts 1 and 2, every open subset of X can be written as a countable union of balls.
To prove $\otimes_{i=1}^n \mathcal{B}_{X_i} = \mathcal{B}_X$, it suffices to show that

$$\{\text{open subsets of } X\} \subseteq \otimes_{i=1}^n \mathcal{B}_{X_i}$$

The left hand side is equal to $\{\cup_{j=1}^{\infty} B_j \mid B_j \subseteq X_{\text{open}} \text{ball}\}$, and the right hand side is equal to $\mathcal{M}\left(\{\prod_{i=1}^n E_i \mid E_i \text{ open}\}\right)$
This will hold, as long as we can show $B_j \in \mathcal{M}\left(\{\prod_{i=1}^n E_i \mid E_i \text{ open}\}\right)$
Since X is endowed with $d_{\max}$, we know that any open ball in X can be expressed as $B = \prod_{i=1}^n B_i$, where $B_i \subseteq X_i$ is a ball. This gives the result. ■

Now, it is finally time to talk about measures.

Measures

Call $(X, \mathcal{M})$ a measurable space when X is a set and $\mathcal{M}$ is a σ-algebra on X.

Definition 0.6. A measure on a measurable space $(X, \mathcal{M})$ is a function $\mu : \mathcal{M} \to [0, +\infty]$ such that

(i) $\mu(\varnothing) = 0$

(ii) If $\{E_i\}$ is a countable disjoint collection of sets, then

$$\mu(\bigcup E_i) = \sum \mu(E_i)$$

This is called "countable (disjoint) additivity"

Example 0.4. (Dirac mass/Dirac measure)
Let $(X, \mathcal{M}) = (X, 2^X)$.
Fix $x_0 \in X$ and define

$$\mu(A) = \begin{cases} 1 & \text{if } x_0 \in A \\ 0 & \text{otherwise} \end{cases}$$

Example 0.5. (Counting measure)
Let $(X, \mathcal{M}) = (X, 2^X)$. Define

$$\mu(A) = |A| = \text{ the number of elements in } A$$

Given a measurable space $(X, \mathcal{M})$ and a measure μ, we call $(X, \mathcal{M}, \mu)$ a measure space and $E \in \mathcal{M}$ a measurable set

Theorem 0.3. *For any measure space $(X, \mathcal{M}, \mu)$ and measurable sets $A, B, A_1, A_2, \dots \in \mathcal{M}$,*

(i) $A \subseteq B \implies \mu(A) \leq \mu(B)$. This is called "monotonicity"

(ii) $\mu(\bigcup_i A_i) \leq \sum_i \mu(A_i)$. This is called "(countable) sub additivity)."

(iii) If $A_i \subseteq A_{i+1}$, then $\mu(\bigcup_i A_i) = \lim_{n\to\infty} \mu(A_i)$. This is called "continuity from below."

(iv) If $A_{i+1} \subseteq A_i$ for all i, and $\mu(A_i) < \infty$, then $\mu(\bigcap A_i) = \lim_{i\to\infty} \mu(A_i)$. This is called "continuity from above."

Remark. For (iv), why do we need the additional hypothesis $\mu(A_1) < \infty$?.
Consider the counting measure on $(\mathbb{N}, 2^{\mathbb{N}})$, and $A_i = \{n \in \mathbb{N} \mid n \geq i\}$, which satisfies $A_{i+1} \subseteq A_i$, but it fails $\mu(A_1) < \infty$:

$$0 = \mu(\varnothing) = \mu(\cap_{i=1}^{\infty} A_i) \neq \lim_{i\to\infty} \mu(A_i) = +\infty$$

Proof. (i) Since $A \subseteq B$, $B = A \cup (B\backslash A)$, so $\mu(B) = \mu(A) + \mu(B\backslash A)$ by countable additivity. $\mu(B\backslash A) \geq 0$, so (i) follows.

(ii) Define $B_1 = A_1$, $B_2 = A_2 \backslash A_1$, $B_3 = A_3 \backslash (A_1 \cup A_2), \ldots, B_n = A_n \backslash (\cup_{i=1}^{n-1} A_i)$. Then $\cup_i A_i = \cup_i B_i$, so by countable disjoint additivity,

$$\mu(\cup_i A_i) = \mu(\cup_i B_i) = \sum_i \mu(B_i) \leq \sum_i \mu(A_i)$$

(iii) Define $B_1 = A_1$, and $B_i = A_i \backslash A_{i-1}$. Then $A_n = \cup_{i=1}^n B_i$, so $\cup_{i=1}^\infty A_i = \cup_{i=1}^\infty B_i$. Thus $\mu(A_n) = \mu(\cup_{i=1}^n B_i) = \sum_{i=1}^n \mu(B_i)$. Consequently,

$$\mu(\cup_{n=1}^\infty A_n) = \mu(\cup_{i=1}^\infty B_i) = \sum_{i=1}^\infty \mu(B_i) = \lim_{n\to\infty} \sum_{i=1}^n \mu(B_i) = \lim_{n\to\infty} \mu(A_n)$$

(iv) Next time!

Lecture 4

Recall
Let

- (X_i, d_i), $i = 1, \ldots, n$ metric spaces
- $\{(X_i, \mathscr{M}_i)\}_{i=1}^n$ a collection of measurable spaces.
- $\mathcal{X} = \prod_{i=1}^n X_i$ product space.
- $d_{\max}((x_1, \ldots, x_n), (y_1, \ldots, y_n)) = \max\{d_i(x_i, y_i)\}$.

Definition 0.7.

$$\bigotimes_{\alpha \in A} \mathscr{M}_\alpha = \mathcal{M}\left(\{\prod_{\alpha \in A} E_\alpha \mid E_\alpha \in \mathscr{M}_\alpha\}\right)$$

We have the following theorem

Theorem 0.4.

$$\mathcal{B}_{\mathcal{X}} = \otimes_{i=1}^n \mathcal{B}_{X_i}$$

That is, the Borel σ-algebra generated by the products of the X_i is equal to the products of the Borel σ-algebras generated by the X_i.
Now, back to measure spaces.

Remark. The definition of Borel sets only depends on the notion of open sets, do $d_{\max}$ could be replaced with any equivalent metric.
We will now prove that a measure satisfies continuity from above.

Proof. Let $\{A_i\}_{i\in\mathbb{N}}$ be a descending sequence of measurable sets.
Define $B_i = A_1 \backslash A_i$
We have $B_1 \subseteq B_2 \subseteq B_3 \subseteq \cdots$
Note $\mu(A_1) = \mu(B_i \cup A_i) = \mu(B_i) + \mu(A_i)$ by disjoint additivity.

$$\bigcup_{i=1}^{\infty} B_i = \bigcup_{i=1}^{\infty}(A_1\backslash A_i) = \bigcup_{i=1}^{\infty} A_1 \cap A_i^c = A_1\backslash \left(\bigcap_{i=1}^{\infty} A_i\right)$$

So

$$\begin{aligned}
\mu(A_1) &= \mu\left(\left(A_1\backslash \left(\bigcap_{i=1}^{\infty} A_i\right)\right) \cup \bigcap_{i=1}^{\infty} A_i\right) \\
&= \mu\left(A_1\backslash \left(\bigcap_{i=1}^{\infty} A_i\right)\right) + \mu\left(\bigcap_{i=1}^{\infty} A_i\right) \\
&= \mu\left(\bigcap_{i=1}^{\infty} B_i\right) + \mu\left(\bigcap_{i=1}^{\infty} A_i\right) \\
&= \lim_{i\to\infty} \mu(B_i) + \mu\left(\bigcap_{i=1}^{\infty} A_i\right)
\end{aligned}$$

Since $\mu(A_1) < \infty$, by monotonicity, $\mu(B_i), \mu(A_i)$ are also finite, and, recalling from before,

$$\mu(B_i) = \mu(A_1) - \mu(A_i)$$

So,

$$\mu(A_1) = \lim_{i\to\infty}(\mu(A_1) - \mu(A_i)) + \mu(\cap_{i=1}^{\infty} A_i)$$

So $\lim_{i\to\infty}\mu(A_i) = \mu(\cap_{i=1}^{\infty} A_i)$.

Measure Terminology

- μ is a finite measure if $\mu(\mathcal{X}) < +\infty$.
- μ is a σ-finite measure if there exists $\{E_i\}_{i=1}^{\infty} \in \mathcal{M}^{\mathbb{N}}$ such that $\cup_{i=1}^{\infty} E_i = \mathcal{X}$ and $\mu(E_i) < +\infty$. In other words, we can chop $\mathcal{X}$ into countably many measurable pieces of finite size.

- E is a null set of μ if $E \in \mathcal{M}$ and $\mu(E) = 0$.
- We say that a property holds for μ-almost every $x \in \mathcal{X}$ if the set of points where it doesn't hold is a null set.

Recall our ultimate goal: a measure μ on $(\mathbb{R}, \mathcal{B}_\mathbb{R})$ where $\mu((a,b)) = b - a$, and it is translation invariant.

Outer Measures

Definition 0.8. An outer measure on a set $\mathcal{X}$ is a function $\mu^* : 2^\mathcal{X} \to [0, +\infty]$ satisfying

(i) $\mu^*(\varnothing) = 0$

(ii) $A \subseteq B \implies \mu^*(A) \leq \mu^*(B)$

(iii) $\mu^*(\cup_{i=1}^\infty A_i) \leq \sum_{i=1}^\infty \mu^*(A_i)$

Remark. *(ii)* + *(iii)* is equivalent to the statement that if $E \subseteq \cup_{i=1}^\infty A_i$, then $\mu^*(E) \leq \sum_{i=1}^\infty \mu^*(A_i)$.

Example 0.6. Let $\mathcal{X} = \mathbb{R}$. The Lebesgue Outer Measure is defined by

$$\mu^*(A) = \inf\{\sum_{i=1}^\infty |b_i - a_i| : A \subseteq \bigcup_{i=1}^\infty (a_i, b_i]\}$$

We will prove that μ^* is an outer measure. We will also show $\mu^*((a,b]) = b - a$, and μ^* is translation-invariant.
Is μ^* countably additive? No, by Vitali's theorem.
While we will be able to show that μ^* is an outer measure, it is not a measure on $2^\mathbb{R}$.

Definition 0.9. Let $\mathcal{X}$ be a nonempty set, and μ^* an outer measure on $\mathcal{X}$. We say $A \subseteq \mathcal{X}$ is μ^*-measurable if, for all $E \subseteq \mathcal{X}$,

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

Remark. We know that if, in the above expression, "=" is replaced by "$\leq$", it holds for any $E \subseteq \mathcal{X}$ by countable subadditivity

Proposition 5. *If $\mu^*(B) = 0$ for $B \subseteq \mathcal{X}$, then B is μ^*-measurable.*

Proof. Fix an arbitrary $E \subseteq \mathcal{X}$. Then, by monotonicity, $\mu^*(E) \geq \mu^*(E \cap B^c) = \mu^*(B) + \mu^*(E \cap B^c)$, so $\mu^*(E) = E \cap B) + \mu^*(E \cap B^c)$.

Theorem 0.5. *(Caratheodory): Given an outer measure μ^* on $\mathcal{X}$, let*

$$\mathscr{M} \stackrel{\text{def}}{=} \{A \subseteq X : A \textit{ is } \mu^* - \textit{measurable}\}$$

Then

(i) $\mathscr{M}$ is a σ-algebra

(ii) μ^* is a measure on $\mathscr{M}$.

Question: Is this the "largest" σ-algebra on which μ^* can be defined as a measure? In general, the answer is no - see hw3.

Proof. $\mathscr{M}$ is nonempty, because by the proposition, $\varnothing$ is μ^*-measurable.
Now we want to see that $\mathscr{M}$ is closed under complements. This clearly holds by the definition of μ^*.
We will now show $\mathscr{M}$ is closed under finite unions. It will suffice to show that if $A, B \in \mathscr{M}$, then $A \cup B \in \mathscr{M}$.
Fix an arbitrary $E \subseteq \mathcal{X}$. We have

$$\begin{aligned}
\mu^*(E) &= \mu^*(E \cap A) + \mu^*(E \cap A^c) \\
&= \mu^*(E \cap A) + \mu^*(E \cap A^c \cap B) + \mu^*(E \cap A^c \cap B^c) \\
&\geq \mu^*((E \cap A) \cup (E \cap A^c \cap B^c)) + \mu^*(E \cap A^c \cap B^c) \\
&= \mu^*(E \cap (A \cup B)) + \mu^*(E \cap (A \cup B)^c)
\end{aligned}$$

So $A \cup B$ is μ^*-measurable.
Remark. "$\leq$" always holds by countable subadditivity.
Now, we will show that $\mu^*|_{\mathscr{M}}$ is finitely additive.

Claim. given $\{B_i\}_{i=1}^n \subseteq \mathscr{M}$ disjoint, then for all $A \subseteq \mathcal{X}$,

$$\mu^*(E \cap (\cup_{i=1}^n)) = \sum_{i=1}^n \mu^*(E \cap B_i)$$

Proof. We will proceed by induction. The base case is obvious.
Now, assume the result holds for $n - 1$. We will show it holds for n. We have

$$\begin{aligned}
\mu^*(E \cap (\cup_{i=1}^n B_i)) &= \mu^*(E \cap (\cup_{i=1}^n) \cap B_n) + \mu^*(E \cap (\cup_{i=1}^n B_i) \cap B_n^c \\
&= \mu^*(E \cap B_n) + \mu^*(E \cap (\cup_{i=1}^{n-1} B_i)) \\
&= \mu^*(E \cap B_n) + \sum_{i=1}^{n-1} \mu^*(E \cap B_i)
\end{aligned}$$

We will finish next time!

Lecture 5

Now for the exciting conclusion.
Taking $E = \mathcal{X}$ in the above claim, we see $\mu^*|_{\mathscr{M}}$ is finitely additive.

Claim. Given $\{B_i\}_{i=1}^\infty \subseteq \mathscr{M}$ *disjoint, for all* $E \subseteq X$*,*

$$\mu^*(E) = \sum_{i=1}^{\infty}(\mu^*(E \cap B_i)) + \mu^*(E \cap (\cup_{i=1}^\infty B_i)^c)$$

Proof. The left hand side is immediately seen to be less than the right hand side due to the countable subadditivity of μ^*, since

$$\begin{aligned} E &= (E \cap (\cup_{i=1}^\infty B_i)) \cup (E \cap (\cup_{i=1}^\infty B_i)^c) \\ &= (\cup_{i=1}^\infty (E \cap B_i)) \cup (E \cap (\cup_{i=1}^\infty B_i)^c) \end{aligned}$$

It remains to show that the left hand side is greater than or equal to the right hand side.
Since $\mathscr{M}$ is closed under finite unions, $\cup_{i=1}^n B_i \in \mathscr{M}$, so by the definition of μ^*-measurable,

$$\begin{aligned} \mu^*(E) &= \mu^*(E \cap (\cup_{i=1}^n B_i)) + \mu^*(E \cap (\cup_{i=1}^n B_i)^c) \\ &= \sum_{i=1}^{n} \mu^*(E \cap B_i) + \mu^*(E \cap (\cup_{i=1}^\infty B_i)^c) \end{aligned}$$

Taking the limit as $n \to \infty$ gives the result.
Now, to show $\mathscr{M}$ is closed under countable unions. Fix $\{C_i\}_{i=1}^\infty \subseteq \mathscr{M}$. We want to show $\cup_{i=1}^\infty C_i \in \mathscr{M}$.
Define $B_1 = C_1$, and in general $B_n = C_n \backslash (\cup_{i=1}^{n-1} C_i)$
Then $B_n \in \mathscr{M}$ for each n, and $\cup_{i=1}^\infty B_i = \cup_{i=1}^\infty C_i$.
Fix $E \subseteq \mathcal{X}$. Then, by a previous claim, we know

$$\begin{aligned} \mu^*(E) &= \sum_{i=1}^{n} \mu^*(E \cap B_i) + \mu^*(E \cap (\cup_{i=1}^\infty B_i)^c) \\ &\geq \mu^*(E \cap (\cup_{i=1}^\infty B_i)) + \mu^*(E \cap (\cup_{i=1}^\infty B_i)^c) \end{aligned}$$

Since we already have the inequality going in the other direction, we have shown $\cup_{i=1}^\infty B_i = \cup_{i=1}^\infty C_i \in \mathscr{M}$.

Taking $E = \cup_{i=1}^{\infty} B_i$, for B_i disjoint, nonempty, then

$$\begin{aligned}
\mu^*(\cup_{i=1}^{\infty} B_i) &= \mu^*(E) \\
&= \sum_{i=1}^{\infty} \mu^*(E \cap B_i) + \mu^*(E \cap (\cup_{i=1}^{\infty} B_i)^c) \\
&= \sum_{i=1}^{\infty} \mu^*(B_i) + \mu^*(\varnothing)
\end{aligned}$$

Thus $\mu^*(\cup_{i=1}^{\infty} B_i) = \sum_{i=1}^{\infty} B_i$.

■

Back to Lebesgue outer measure. Let $\mathcal{X} = \mathbb{R}$. Recall we define

$$\mu^*(A) = \inf\{\sum_{i=1}^{\infty} |b_i - a_i| : A \subseteq \cup_{i=1}^{\infty} (a_i, b_i]\}$$

We want to show that

1. μ^* is an outer measure

2. It gives the correct lengths to $(c, d]$,

3. It is translation invariant

4. $\mathscr{B}_{\mathbb{R}}$ is contained in the collection of μ^*-measurable sets.

In fact, we will study a generalization of Lebesgue outer measure that will give rise to Lebesgue-Stieljes measures.
Recall: $F : \mathbb{R} \to \mathbb{R} \cup \{\infty\}$ is <u>right continuous</u> if for all $x \in \mathbb{R}$,

$$\lim_{y \to x^+} F(y) = F(x)$$

Definition 0.10. Given $F : \mathbb{R} \to \mathbb{R}$ non-decreasing and right-continuous, define

$$\mu_F^*(A) = \inf\{\sum_{i=1}^{\infty} (F(b_i) - F(a_i)) : A \subseteq \cup_{i=1}^{\infty} (a_i, b_i]\}$$

Note: Katy HATES the term "non-decreasing," and will use it interchangeably with the term "increasing." To denote something which is not constant anywhere, she will say "strictly" increasing.
Why do we require F to be nondecreasing?

Spoiler: We will show any finite measure μ on $\mathscr{B}_{\mathbb{R}}$ satisfies $\mu = \mu_F^*|_{\mathscr{B}_{\mathbb{R}}}$, for

$$F(x) = \mu((-\infty, x])$$

We call F the Cumulative Distribution Function, or CDF.
Note that if μ is a finite measure on $\mathscr{B}_{\mathbb{R}}$ and $F(x)$ is it's CDF, then F is

- Nondecreasing: If $x \leq y$, then $(-\infty, x] \subseteq (-\infty, y]$, which implies $F(x) \leq F(y)$.
- Right-continuous: For any decreasing sequence x_n whose limit is x, $\lim_{n\to\infty} F(x_n) = \lim_{n\to\infty} \mu((-\infty, x_n]) = \mu((-\infty, x]) = F(x)$. The penultimate equality is due to μ being continuous from above, as it is a finite measure.

Theorem 0.6. For any nondecreasing right-continuous F, μ_F^* is an outer measure.

Proof.

- First, $\mu_F^*(\varnothing) = \inf\{\sum_{i=1}^\infty F(b_i) - F(a_i) : \varnothing \subseteq \cup_{i=1}^\infty (a_i, b_i]\} = 0$, as every interval contains $\varnothing$ as a subset so we may set $a_i = b_i \equiv 1$ for all i. Since $\mu_F^* \geq 0$ by definition, $\mu_F^*(\varnothing) = 0$
- Now, we want to show that if $A \subseteq \cup_{i=1}^\infty B_i$, then $\mu_F^*(A) \leq \sum_{i=1}^\infty \mu_F^*(B_i)$.

  If $\mu_F^*(B_i) = \infty$ for some i, we are done. Without loss of generality, suppose $\mu_F^*(B_i) < \infty$ for each i.

  By the definition of inf, for all $\varepsilon > 0$ and each i there exists $\{I_i^\varepsilon\}_{i=1}^\infty$ of intervals depending on ε such that

$$B_j \subseteq \cup_{i=1}^\infty I_i^{j,\varepsilon}$$
$$\mu_F^*(B_j) \leq \sum_{i=1}^\infty |I_i^{j,\varepsilon}| \leq \mu_F^*(B_j) + \frac{\varepsilon}{2j}$$

  Thus

$$\begin{aligned} A &\subseteq \cup_{j=1}^\infty \cup_{i=1}^\infty I_i^{j,\varepsilon} \\ \mu_F^*(A) &\leq \sum_{i,j=1}^\infty |I|_F^{j,\varepsilon} \\ &\leq \sum_{j=1}^\infty \mu_F^*(B_j) + \frac{\varepsilon}{2j} \\ &= \sum_{j=1}^\infty \mu_F^*(B_j) + \varepsilon \end{aligned}$$

Sending $\varepsilon \to 0$ completes the proof.

∎

Theorem 0.7. *For all* $a, b \in \mathbb{R}$, $a \leq b$,

$$\mu^*((a,b]) = F(b) - F(a)$$

Proof. $\leq$ follows quickly, since we know $(a,b] \subseteq (a,b] \cup \varnothing \cup \varnothing \cup \cdots$, so the definition of μ_F^* ensures

$$\mu_F^*((a,b]) \leq \sum_{i=1}^{\infty} F(b_i) - F(a_i) = F(b) - F(a)$$

Now we turn to $\geq$. Note that if $a = b$, we already showed that $\mu_F^*((a,b]) = \mu_F^*(\varnothing) = 0 = F(b) - F(a)$, so without loss of generality $a < b$.
It suffices to show that if $(a,b] \subseteq \cup_{i=1}^{\infty}(a_i, b_i]$, then $F(b) - F(a) \leq \sum_{i=1}^{\infty} F(b_i) - F(a_i)$
Since F is right continuous, for all $\varepsilon > 0$ we can find $\delta_i > 0$ such that $F(b_i + \delta_i) < F(b_i) + \frac{\varepsilon}{2^i}$.
Note that

$$[a+\varepsilon, b] \subseteq (a,b] \subseteq \bigcup_{i=1}^{\infty}(a_i, b_i] \subseteq \bigcup_{i=1}^{\infty}(a_i, b_i + \delta_i)$$

Since $[a+\varepsilon, b]$ is compact and $\{(a_i, b_i + \delta_i)\}_{i=1}^{\infty}$ is an open cover, there exists a finite subcover

$$[s + \varepsilon] \subseteq \bigcup_{i=1}^{N}(a_i, b_i + \delta_i)$$

Without loss of generality we may throw away any unnecessary elements of the cover. The "first" element of the cover must overlap with exactly one other element of the cover, the "second" interval. Thus we may assume that

$$b_i + \delta_i \in (a_{i+1}, b_{i+1} + \delta_{i+1}) \forall i = 1, \ldots, N-1$$

Tune in next time for the continuation!

Lecture 6

And now, for the exciting conclusion...

Since F is nondecreasing,

$$\begin{aligned}
F(b) - F(a+\varepsilon) &\leq f(b_N + \delta_N) - F(a_1) \\
&= F(b_N + \delta_N) - F(a_N) + \sum_{i=1}^{N-1}(F(a_{i+1}) - F(a_i)) \\
&= F(b_N + \delta_N) - F(a_N) + \sum_{i=1}^{N-1}(F(b_i + \delta_i) - F(a_i)) \\
&= \sum_{i=1}^{N}(F(b_i + \delta_i) - F(a_i)) \\
&\leq \sum_{i=1}^{N}(F(b_i) - F(a_i) + \frac{\varepsilon}{2^i}) \\
&\leq \left|\sum_{i=1}^{\infty} F(b_i) - F(a_i)\right| + \varepsilon
\end{aligned}$$

Since $\varepsilon > 0$ was arbitrary, and F is right-continuous, sending $\varepsilon \to 0$ gives the result. ■

Definition 0.11. By Carathéodory's theorem, we know μ_F^* is a measure when restricted to $\mathscr{M}_{\mu_F^*}$, the collection of μ_F^*-measurable sets. We will denote this measure by μ_F, and call it the Legesgue-Stieljes measure associated to F.
How does this help our goals?
Is μ_F a Borel measure (that is, a measure when restricted to the Borel σ-algebra)?
Yes

Theorem 0.8. $\mathscr{B}_{\mathbb{R}} \subseteq \mathscr{M}_{\mu_F^*}$

Proof. It suffices to show that, for all $b \in \mathbb{R}$, $(-\infty, b] \in \mathscr{M}_{\mu_F^*}$.
That is, we must show for all $E \subseteq \mathbb{R}$

$$\mu_F^*(E) \geq \mu_F^*(E \cap (-\infty, b]) + \mu_F^*(E \cap (-\infty, b]^c)$$

We already have $\leq$ by countable additivity.
Fix a $\varepsilon > 0$. By definition of μ_F^*, there exists a cover $\{(a_i, b_i]\}_{i=1}^{\infty}$ such that $E \subseteq \bigcup_{i=1}^{\infty}(a_i, b_i]$ and

$$\sum_{i=1}^{\infty}(F(b_i) - F(a_i)) \leq \mu_F^*(E) + \varepsilon$$

Note that

$$
\begin{aligned}
(a_i, b_i] \cap (-\infty, b] &\subseteq (a_i, b] \\
(a_i, b_i] \cap [b, \infty) &\subseteq (b, b_i]
\end{aligned}
$$

so

$$
\begin{aligned}
E \cap (-\infty, b] &\subseteq \bigcup_{i=1}^{\infty} (a_i, b] \\
E \cap [b, +\infty) &\subseteq \bigcup_{i=1}^{\infty} (b, b_i]
\end{aligned}
$$

$$
\begin{aligned}
\mu_F^*(E \cap (-\infty, b]) + \mu_F^*(E \cap [b, +\infty)) &\leq \sum_{i=1}^{\infty} (F(b) - F(a_i)) + \sum_{j=1}^{\infty} (F(b_j) - F(b)) \\
&= \sum_{i=1}^{\infty} (F(b_i) - F(a_i)) \\
&\leq \mu_F^*(E) + \varepsilon
\end{aligned}
$$

Sending $\varepsilon \to 0$ gives us the result. ∎

Definition 0.12. When $F(x) = x$, we write $\lambda^* \stackrel{\text{def}}{=} \mu_F^*$,
and we call it the Lebesgue outer measure. Similarly, we write $\lambda \stackrel{\text{def}}{=} \mu_F$, and call it the Lebesgue measure. Finally, $\mathscr{M}_{\lambda^*} \stackrel{\text{def}}{=} \mathscr{M}_{\mu_F^*}$, and we call this collection the Lebesgue measurable sets.

Thus, we know all Borel sets are Lebesgue measurable.

In this way, we have found a Borel measure that gives the "right" length to intervals $(a, b]$.

The last "intuitive" property of λ that we seek to show is translation invariance.

Theorem 0.9. *λ^* is translation invariant on $2^{\mathbb{R}}$, and λ is translation invariant on $\mathscr{M}_{\lambda^*}$.*

Proof. For any $a \in \mathbb{R}$, $A \subseteq \mathbb{R}$, $A \subseteq \bigcup_{i=1}^{\infty} (a_i, b_i]$ is equivalent to $A + a \subseteq \bigcup_{i=1}^{\infty} (a_i + a, b_i + a]$.

Therefore $\lambda^*(A) = \lambda^*(A + a)$. The only thing left to show is $\mathscr{M}_\lambda$ is translation invariant.

Claim. Let $A \in \mathscr{M}_{\lambda^*}$. Then $A + a \in \mathscr{M}_{\lambda^*}$.

Proof. Fix $E \subseteq \mathbb{R}$. We want to show

$$\lambda^*(E) = \lambda^*(E \cap (A + a)) + \lambda^*(E \cap (A + a)^c)$$

We know λ^* is translation invariant, so

$$\lambda^*(E) = \lambda^*((E - a) \cap A) + \lambda^*((E - a) \cap A^c)$$

For any $S \subseteq \mathbb{R}$,

$$\begin{aligned}(E - a) \cap S &= \{x - a \mid x \in E, x - a \in S\} \\ &= \{x \mid x \in E, x \in S + a\} \\ &= E \cap (S + a) - a\end{aligned}$$

So

$$\begin{aligned}(S + a)^c &= \{y \mid y \notin S + a\} \\ &= \{y \mid y - a \notin S\} \\ &= \{y - a \mid y - a \notin S\} \\ &= S^c + a\end{aligned}$$

Therefore

$$\begin{aligned}\lambda^*(E) &= \lambda^*(E \cap (A + a)) + a) + \lambda^*((E \cap (A^c + a)) - a) \\ &= \lambda^*((E \cap (A + a)) - a) + \lambda^*((E \cap (A + a)^c) - a) \\ &= \lambda^*(E \cap (A + a)) + \lambda^*(E \cap (A + a)^c)\end{aligned}$$

Thus, for any $A \in \mathscr{M}_{\lambda^*}$, we have $\lambda(A) \stackrel{\text{def}}{=} \lambda^*(A) = \lambda^*(A + a) \stackrel{\text{def}}{=} \lambda(A + a)$

■

In fact, <u>all</u> finite Borel measures are of this form.

Theorem 0.10. *Suppose μ is a finite Borel measure. Then $\mu = \mu_F$, where F is the cumulative distribution function, $F(x) = \mu((-\infty, x))$.*

Proof. Recall, we already showed that for any finite measure μ on $\mathscr{B}_{\mathbb{R}}$, $F(x) = \mu((-\infty, x])$ is nondecreasing and right-continuous.

We seek to show $\mu(E) = \mu_F(E)$ for all $E \in \mathscr{B}_\mathbb{R}$.
First, consider the half-open interval $(a, b]$, $a \le b$. μ is a measure by hypothesis, so in particular is finitely additive:

$$\mu((a,b]) + \underbrace{\mu((-\infty,a])}_{=F(a)} = \underbrace{\mu((-\infty,b])}_{=F(b)}$$

So, $\mu((a,b]) = F(b) - F(a) = \mu_F((a,b])$.
Now, fix $E \in \mathscr{B}_\mathbb{R}$. Consider $\{(a_i, b_i]\}_{i=1}^\infty$ such that $E \subseteq \bigcup_{i=1}^\infty (a_i, b_i]$.
By countable subadditivity,

$$\mu(E) \le \sum_{i=1}^\infty \mu((a_i, b_i]) = \sum_{i=1}^\infty (F(b_i) - F(a_i))$$

Taking the infimum over all such covers, $\mu(E) \le \mu_F(E)$. It remains to show the opposite inequality.
Since $E \in \mathscr{B}_\mathbb{R}$ was arbitrary,

$$\mu(E^c) \le \mu_F(E^c)$$

Thus $\mu(E) + \mu(E^c) = \mu(\mathbb{R})$. Then $\mu(E) = \mu(\mathbb{R}) - \mu(E^c)$. So

$$\begin{aligned}\mu(E) &\ge \mu(\mathbb{R}) - \mu_F(E^c) \\ &= \mu_F(R) - \mu_F(E^c)) \\ &= \mu_F(E)\end{aligned}$$

But the above has a gap: in particular, we don't know that $\mu(\mathbb{R}) = \mu_F(\mathbb{R})$. If we prove that, we are done.

Claim. $\mu(\mathbb{R}) = \mu_F(\mathbb{R})$.

Proof.

$$\begin{aligned}\mu(\mathbb{R}) = \mu\left(\bigcup_{i=1}^\infty [-i,i]\right) &\overbrace{=}^{\text{by upper continuity}} \lim_{i\to\infty} \mu((-i,i]) \\ = \lim_{i\to\infty} \mu_F((-i,i]) &\underbrace{=}_{\text{by lower continuity}} \mu_F\left(\bigcup_{i=1}^\infty [-i,i]\right) = \mu_F(\mathbb{R})\end{aligned}$$

■

We conclude our study of Borel measures on the real line with some regularity properties of Lebesgue-Stieljes measures.

Lemma 2. Given $F : \mathbb{R} \to \mathbb{R}$ nondecreasing, right-continuous, for all $E \in \mathscr{M}_{\mu_F^*}$,

$$\mu_F(E) = \inf\{\sum_{i=1}^{\infty} \mu_F((a_i, b_i)) \mid E \subseteq \bigcup_{i=1}^{\infty}(a_i, b_i), a_i \leq b_i\}$$

Proof. By HW3Q2,

$$\mu_F(E) = \inf\{\sum_{i=1}^{\infty} \mu_F(A_i) \mid E \subseteq \bigcup_{i=1}^{\infty} A_i, \{A_i\}_{i=1}^{\infty} \in \mathscr{M}_{\mu_F^*}^{\mathbb{N}}\}$$

Thus, "$\leq$" must hold. It remains to show the opposite inequality.
By definition, for all $E \in \mathscr{M}_{\mu_F^*}$,

$$\mu_F^*(E) = \inf\{\sum_{i=1}^{\infty}(F(b_i) - F(a_i)) \mid E \subseteq \bigcup_{i=1}^{\infty}(a_i, b_i], a_i \leq b_i\}$$

Fix $\varepsilon > 0$. Then there exists a sequence of intervals $\{(a_i, b_i]\}_{i=1}^{\infty}$ such that $E \subseteq \bigcup_{i=1}^{\infty}(a_i, b_i]$ and

$$\mu_F(E) + E \geq \sum_{i=1}^{\infty}(F(b_i) - F(a_i))$$

Furthermore, for any $(a_i, b_i]$, we may define $B_n \stackrel{\text{def}}{=} (a_i, b_i + \frac{1}{n})$, and since $\mu_F(B_1) < +\infty$, continuity from above ensures

$$\lim_{n\to\infty} \mu(B_n) = \mu_F\left(\bigcap_{i=1}^{\infty} B_n\right) = \mu_F((a_i, b_i])$$

Thus, for all i, there exists a $\delta_i > 0$ such that

$$\mu_F((a_i, b_i + \delta_i]) \leq \mu_F((a_i, b_i]) + \frac{\varepsilon}{2^i}$$

Thus,

$$\begin{aligned}
\mu_F(E) &\leq \sum_{i=1}^{\infty} \mu_F((a_i, b_i + \delta_i]) \\
&\leq \sum_{i=1}^{\infty} \mu_F((a_i, b_i]) + \frac{\varepsilon}{2^i} \\
&\leq \mu_F(E) + 2\varepsilon
\end{aligned}$$

Letting $\varepsilon \to 0$, this shows the other direction.

■

Lecture 7

Recall: Given $F : \mathbb{R} \to \mathbb{R}$ nondecreasing and right continuous,

$$\mu_F^*(A) \overset{\text{def}}{=} \inf\{\sum_i((F(b_i) - F(a_i)) \mid A \subseteq \bigcup_{i=1}^{\infty}(a_i, b_i], a_i \leq b_i\}$$

Theorem 0.11.

$$\mathcal{B}_{\mathbb{R}} \subseteq \mathcal{M}_{\mu_F^*}$$

We will show that in general, this is a strict containment.

Theorem 0.12. *For any* $E \in \mathcal{M}_{\mu_F^*}$,

$$\begin{aligned} \mu_F(E) &= \inf\{\mu_F(U) \mid E \subseteq U, U \, open\} \\ &= \sup\{\mu_F(K) \mid K \subseteq E, K \, compact\} \end{aligned}$$

Proof. Fix $E \in \mathcal{M}_{\mu_F^*}$.

Step 1

Fix $\varepsilon > 0$. The lemma proven in the previous lecture ensures that there exists $\{(a_i, b_i)\}_{i=1}^{\infty}$ such that $E \subseteq \cup_{i=1}^{\infty}(a_i, b_i)$ and

$$\mu_F \underbrace{\left(\bigcup_{i=1}^{\infty}(a_i, b_i)\right)}_{\overset{\text{def}}{=} U} \leq \sum_{i=1}^{\infty} \mu_F((a_i, b_i)) \leq \mu_F(E) + \varepsilon$$

The first inequality is by countable additivity of μ, and the second by construction.

Step 2

There are a few cases:

1. In the first case, assume that E is bounded. If E is closed, then by Heine-Borel E is compact, and taking $K = E$ gives the result. Next, suppose that E is not closed. Fix $\varepsilon > 0$. By step 1, there exists an open $U \supseteq \overline{E} \setminus E$ such that

$$\mu_F(U) \leq \mu_F(\overline{E} \setminus E) + \varepsilon$$

Define $K = \overline{E} \setminus U$. Then K is compact. By definition,

$$\begin{aligned} K = \overline{E} \cap U^c &\subseteq \overline{E} \cap (\overline{E} \cap E^c)^c \\ &= \overline{E} \cap (\overline{E}^c \cup E) = E \end{aligned}$$

So

$$\mu_F(E \cap U) + \mu_F(K) \geq \mu_F(E \cap U) + \mu_F(E \setminus U) = \mu_F(E)$$

Since E is bounded, $\mu_F(E \cap U) < +\infty$, $\mu_F(U) < +\infty$, so

$$\begin{aligned}
\mu_F(K) &\geq \mu_F(E) - \mu_F(E \cap U) \\
&= \mu_F(E) - (\mu_F(U) - \mu_F(E \setminus U)) \\
&\geq \mu_F(E) - \mu_F(U) + \mu_F(\overline{E} \setminus E) \\
&\geq \mu_F(E) - \varepsilon
\end{aligned}$$

Since $\varepsilon > 0$ was arbitrary, this gives the result.

2. In the second case, assume that E is unbounded. Define $E_j = E \cap (j, j+1], j \in \mathbb{Z}$. E_j is clearly bounded, so by case 1, we know that for all $\varepsilon > 0$, there exists a compact $K_j \subseteq E_j$, such that

$$\mu_F(K_j) \geq \mu_F(E_j) - \frac{\varepsilon}{2^{|j|}}$$

Then $H_n = \cup_{j=-n}^{n} E_j$ is compact, with $H_n \subseteq E$. By additivity,

$$\begin{aligned}
\mu_F(H_n) = \sum_{j=-n}^{n} \mu_F(K_j) &\geq \sum_{j=-n}^{n} \mu_F(E_j) - 2\varepsilon \\
&\geq \mu_F\left(\bigcup_{j=-n}^{n} E_j\right) - 2\varepsilon
\end{aligned}$$

By continuity from below, we may pick $N \in \mathbb{N}$ sufficiently large so that

$$\mu_F\left(\bigcup_{j=-N}^{N} E_j\right) \geq \mu_F(E) - \varepsilon$$

Thus,

$$\mu_F(H_N) \geq \mu_F(E) - \varepsilon$$

■

Time for an important example.

Example 0.7. The Cantor Set
Warmup:
$\lambda(\{a\}) = 0$, $\lambda(\mathbb{Q}) = \lambda(\cup_{i=1}^{\infty}\{r_i\})$, where r_i is some enumeration of $\mathbb{Q}$. Then the above is equal to $\sum_{i=1}^{\infty} \lambda(\{r_i\}) = 0$.

On the other hand, fix $\varepsilon > 0$ and define

$$U = (0,1) \cap \left(\bigcup_{j=1}^{\infty} (r_j - \frac{\varepsilon}{2^{j+1}}, r_j + \frac{\varepsilon}{2^{j+1}}) \right)$$

Then U is open and dense in $(0,1)$. From a topological perspective, this means that U is "large" (comeagre).
However, in a measurable sense, U is "small:"

$$\lambda(U) \leq \sum_{j=1}^{\infty} \frac{\varepsilon}{2^j} = \varepsilon$$

Note: U depends on ε, so we have not shown that $\mu(U) = 0$.
Now we construct the Cantor set.
Start with $U_1 = (0,1)$. Let U_2 be U_1 with the middle third removed, so is two disjoint intervals. Inductively, let U_i be U_{i-1}, with the middle thirds of all intervals removed. Then the Cantor set is the intersection of all U_i.
Alternatively, the Cantor set is every real in $(0,1)$ whose base 3 expansion does not contain a 2.

Theorem 0.13. *Let C be the Cantor set. Then*

(i) C is compact, nowhere dense, and totally disconnected (meaning the only connected subsets are singletons). Further, C has no isolated points.

(ii) $\lambda(C) = 0$.

(iii) C has cardinality of the continuum.

Measurable Functions

Definition 0.13. Given $f : X \to Y$, $f^{-1}(E) = \{x \in X \mid f(x) \in E\}$ is called the preimage of f
These are basic set theory facts from HW 1:

$$f^{-1}(\cup_\alpha E_\alpha) = \cup_\alpha f^{-1}(E_\alpha)$$
$$f^{-1}(E^c) = (f^{-1}(E))^c$$
$$f^{-1}(\cap_\alpha E_\alpha) = \cap_\alpha f^{-1}(E\alpha)$$

Definition 0.14. Suppose $(X, \mathcal{M}), (Y, \mathcal{N})$ are measurable spaces, and $f : X \to Y$. Then $\{f^{-1}(E) \mid E \in \mathcal{N}\}$ is the pullback of $\mathcal{N}$, and $\{E \mid f^{-1}(E) \in \mathcal{M}\}$ is the pushforward of $\mathcal{M}$.

Definition 0.15. A function $f : X \to Y$ is $(\mathcal{M}, \mathcal{N})$-measurable if for all $E \in \mathcal{N}$, $f^{-1}(E) \in \mathcal{M}$.
Equivalently, the pullback of $\mathcal{M}$ is a subset of N. Equivalently, $\mathcal{N}$ is a subset of the pushforward of $\mathcal{M}$.
Informally, "the inverse image of every measurable set is measurable" (Katy says the reason this isn't formal is because "measurable set" already means something specific in the context of an outer measure).
If $f : X \to \mathbb{R}(\overline{\mathbb{R}})$, we will suppose that the range is endowed with $\mathscr{B}_{\mathbb{R}}(\mathscr{B}_{\overline{\mathbb{R}}})$.

Definition 0.16. (a) $f : \mathbb{R} \to \overline{\mathbb{R}}$ is Lebesgue Measurable if it is $(\mathcal{M}_{\lambda^*}, \mathscr{B}_{\overline{\mathbb{R}}})$-measurable.

(b) Given topological spaces X, Y, $f : X \to Y$ is Borel Measurable if it is $(\mathscr{B}_X, \mathscr{B}_Y)$-measurable.

Remark. Given $f : \mathbb{R} \to \overline{\mathbb{R}}$, which is a stronger criteria: being Borel measurable, or being Lebesgue measurable? Borel measurable implies Lebesgue measurable since $\mathscr{B}_{\mathbb{R}} \subseteq \mathcal{M}_{\lambda^*}$.

Proposition 6. *Given measureable spaces $(X, \mathcal{M}), (Y, \mathcal{N})$, where $\mathcal{N}$ is generated by $\mathcal{E}$. Then $f : X \to Y$ is $(\mathcal{M}, \mathcal{N})$-measurable is equivalent to $f^{-1}(E) \in \mathcal{M}$ for all $E \in \mathcal{N}$.*

Proof. One direction is immediate. For the other direction, since $\{E : f^{-1}(E) \in \mathcal{M}\}$ (the pushforward of $\mathcal{M}$) is a σ-algebra containing $\mathcal{E}$. By assumption, $\mathcal{N}$ is generated by $\mathcal{E}$, meaning $\mathcal{N} \subseteq \{E \mid f^{-1}(E) \in \mathcal{M}\}$.

Corollary 0.14. *If X and Y are topological spaces, then every continuous function $f : X \to Y$ is Borel measurable.*

Proof. Since the open subsets of Y generate the σ-algebra $\mathcal{N} = \mathcal{B}_Y$, the previous proposition ensures that it suffices to check $f^{-1}(U) \in \mathcal{B}_X$ for all U open, and this is true, since $f^{-1}(U)$ is open.

Corollary 0.15. *If $(X, \mathcal{M})$ is a measurable space and $f : X \to \mathbb{R}$ the following are equivalent:*

(i) f is $(\mathcal{M}, \mathscr{B}_{\mathbb{R}})$-measurable.

(ii) $f^{-1}((a, +\infty)) \in \mathcal{M}$ for all $a \in \mathbb{R}$.

(iii) $f^{-1}((-\infty, a)) \in \mathcal{M}$ for all $a \in \mathbb{R}$

If $(X, \mathcal{M})$ is a measurable space and $f : X \to \overline{\mathbb{R}}$ the following are equivalent:

(i) f is $(\mathcal{M}, \mathscr{B}_{\mathbb{R}})$-measurable.

(ii) $f^{-1}((a, +\infty]) \in \mathcal{M}$ *for all* $a \in \mathbb{R}$.

(iii) $f^{-1}([-\infty, a)) \in \mathcal{M}$ *for all* $a \in \mathbb{R}$

Lecture 8

Proof. Since open rays generate $\mathscr{B}_{\mathbb{R}}$, and half-open rays generate $\mathscr{B}_{\overline{\mathbb{R}}}$, this follows immediately from the proposition.
For the rest of this lecture, let $(X, \mathcal{M})$ be a measurable space. We say a function $f : X \to \overline{\mathbb{R}}$ is measurable if it is $(\Updownarrow, \mathscr{B}_{\overline{\mathbb{R}}})$-measurable.

Theorem 0.16. *Given* $f_1, f_2, \cdots : X \to \overline{\mathbb{R}}$ *measurable, then the fullowing are also measurable:*

(i) $f_i + f_j$

(ii) $f_i f_j$ *(our convention is* $0 \times x = 0$ *even if* $x = \pm\infty$.

(iii) $f_i \vee f_j$, *where* $\vee$ *means pointwise maximum, i.e.* $f_i \vee f_j(x) = \max(f_i(x), f_j(x))$

(iv) $f_i \wedge f_j$, *where* $\wedge$ *means pointwise minimum, i.e.* $f_i \wedge f_j(x) = \min(f_i(x), f_j(x))$.

(v) $\sup_n f_n = \vee_{n=1}^{\infty} f_n$.

(vi) $\inf_n f_n = \wedge_{n=1}^{\infty} f_n$.

(vii) $\limsup_{n\to\infty} f_n$

(viii) $\liminf_{n\to\infty} f_n$

(ix) $\lim_{n\to\infty} f_n$, *if this limit exists everywhere.*

Remark. Question: suppose $\sum_{i=1}^{\infty} f_i(x)$ exists for all $x \in X$. Does it follow that $\sum_{i=1}^{\infty} f_i(x)$ measurable? Yes! By part (i), any finite sum is measurable, and by part (ix) the limit of these partial sums (which is what an infinite sum is) is measurable.

Proof. (i) HW5

(ii) See above

(iii) Fix $a \in \mathbb{R}$. Then

$$\begin{aligned}(f_i \vee f_j)^{-1}((a, +\infty]) &= \{x \in \mathcal{X} \mid f_i \vee f_j(x) > a\} \\ &= \{x \in \mathcal{X} \mid f_i(x) > a\} \cup \{x \in \mathcal{X} \mid f_i \vee f_j(x) > a\} \\ &= f_i^{-1}((a, +\infty]) \cup f_2^{-1}((a, +\infty])\end{aligned}$$

By the previous corollary, we are done.

(iv) Similar to above

(v) Fix $a \in \mathbb{R}$. Then

$$\begin{aligned}(\sup_n f_n)^{-1}((a, +\infty]) &= \{x \in \mathcal{X} \mid \sup_n f_n(x) > a\} \\ &= \cup_{n=1}^{\infty} \{x \in \mathcal{X} \mid f_n(x) > a\}\end{aligned}$$

By assumption, $\{x \in \mathcal{X} \mid f_i > a\}$ is measurable for all a, i.

(vi) Similar to above

(vii)
$$\limsup_{n \to \infty} f_n = \inf_n \sup_{k \, \mathrm{eq} \, n} f_k$$

The right hand side is measurable by parts (v) and (vi)

(viii) Similar to above

(ix) Since $\lim_{n \to \infty} f_n(x)$ exists for all $x \in X$, the lim sup and lim inf of the sequence both exist and are equal to the limit. So by (vii) this is measurable.

■

Remark. What about the composition of measurable functions? For example, if $f : \mathbb{R} \to \mathbb{R}$, $g : \mathbb{R} \to \mathbb{R}$ are both Borel measurable, then $f \circ g$ is Borel measurable.
What if f, g are Lebesgue measurable? No, because a function being "Lebesgue measurable" means it is measurable when the domain has the Lebesgue σ-algebra, and the codomain has the Borel σ-algebra. In other words, it is "Lebesgue Measurable" if the preimage of any Borel set is a Lebesgue set.

Only The Things Above Here Are On Midterm 1

Simple functions

Definition 0.17. For any $A \subseteq \mathcal{X}$, the indicator function of A is the function

$$1_A(x) \overset{\text{def}}{=} \begin{cases} 1 & x \in A \\ 0 & \text{otherwise} \end{cases}$$

In the book, the notation χ_A is also used. Katy will not do this because, in her opinion, the notation "χ_A" is overloaded.

Definition 0.18. A $(\mathcal{M}, \mathscr{B}_{\mathbb{R}})$-measurable function $f : \mathcal{X} \to \mathbb{R}$ is a simple function if its image is a finite subset of $\mathbb{R}$. The standard representation of a simple function is

$$f(x) = \sum_{i=1}^{n} c_i 1_{E_i}(x)$$

where $f(\mathcal{X}) = \{c_1, \ldots, c_n\}$, and $E_i = f^{-1}(c_i)$.

Remark. $\{E_i\}_{i=1}^{n} \subseteq \mathcal{M}$ is a disjoint partition of $\mathcal{X}$.

Example 0.8. There are many ways a simple function can be expressed as a linear combination of indicator functions:

$$\begin{aligned} f(x) &\equiv 2 \\ &= 2 \cdot 1_{\mathbb{R}} \\ &= 2 \cdot 1_{[0,+\infty)} + 2 \cdot 1_{(-\infty,0)} \\ &= etc. \end{aligned}$$

The last two are not in standard representation.
It is "easy" to define the integral of a simple function.

Definition 0.19. For any measure space $(\mathcal{X}, \mathcal{M}, \mu)$, we can define the integral of a simple function to be

$$\int f \, d\mu \stackrel{\text{def}}{=} \sum_{i=1}^{n} c_i \mu(E_i)$$

Again, we use the convention $0 \times \pm\infty = 0$.
For $A \in \mathcal{M}$, define

$$\int_A f \, d\mu \stackrel{\text{def}}{=} \int f \cdot 1_A \, d\mu$$

$f \cdot 1_A$ is the simple function $\sum_{i=1}^{n} c_i 1_{E_i \cap A}$. Note that $E_i \cap A$ could have zero measure, or even be empty. Recall that $\mu(\varnothing) = 0$.

Remark. Sometimes we will write

$$\int_A f \, d\mu = \int_A f$$

and suppress the $d\mu$ in the notation.
Next, we will show that we can approximate any nonnegative measurable functions.

Theorem 0.17. *Given $f : \mathcal{X} \to [0, \infty]$ measurable, there exists a sequence f_n of simple functions so that f_n converges up to f pointwise. To be clear: for each $x \in \mathcal{X}$, $f_n(x)$ is an increasing sequence which converges to $f(x)$.*

Proof. For $n \in \{0\} \cup \mathbb{N}$, $0 \leq k \leq 2^{2^n} - 1$, define

$$E_n^k \overset{\text{def}}{=} f^{-1}((k2^{-n}, (k+1)2^{-n}])$$

and

$$F_n \overset{\text{def}}{=} f^{-1}((2^n, +\infty])$$

Now, let

$$f_n \overset{\text{def}}{=} \sum_{k=0}^{2^{2^n}-1} k2^{-n} 1_{E_n^k} + 2^n 1_{F_n}$$

Key properties of this construction:

- $f_1 \leq f_2 \leq f_3 \leq \cdots$
- $0 \leq f - f_n \leq 2^{-n}$ on F_n^c.

Thus, f_n increases to f pointwise. ∎

In order to apply this to integrate general nonnegative functions, we will use the following properties of integrating simple functions.

Proposition 7. On simple functions, the integral is linear, and preserves order. That is, if $f \leq g$, then $\int f \leq \int g$.
Further, $\nu(A) \overset{\text{def}}{=} \int_A f \, d\mu$ is a measure on $\mathcal{M}$.

Lecture 9

Proof. Let $f = \sum_{i=1}^n a_i 1_{E_i}, g = \sum_{j=1}^n b_j 1_{F_j}$ be the standard representations of simple functions.

(a) Suppose $c \neq 0$. Then,

$$c \int f = c \sum_{i=1}^n a_i \mu(E_i) = \sum_{i=1}^n c a_i \mu(E_i) = \int cf$$

since $\sum_{i=1}^n c a_i 1_{E_i} = cf$ is the standard representation.

(b) $\{E_i\}_{i=1}^n, \{F_j\}_{j=1}^n$ are partitions of X. So

$$E_i = \coprod_{j=1}^m E_i \cap F_j, \; F_j = \coprod_{i=1}^n F_j \cap E_i$$

By definition,

$$\begin{aligned}\int f + \int g &= \sum_i a_i \mu(E_i) + \sum_j b_i \mu(F_j) \\ &= \sum_{i,j} \big(a_i \mu(E_i \cap F_j) + b_j \mu(E_i \cap F_j)\big) \\ &= \sum_{i,j} (a_i + b_j) \mu(E_i \cap F_j)\end{aligned}$$

Let $h = \sum_{i,j}(a_i + b_j)1_{E_i \cap F_j} = f + g$. But this is not necessarily the standard representation of h.

Let $\{c_\ell\}_{\ell=1}^k$ be the distinct values of $\{a_i + b_j\}_{i,j}$.

Likewise, let $G_\ell = h^{-1}(c_\ell) = \cup_{i,j,a_i+b_j=c}(E_i \cap F_j)$.

Then

$$\begin{aligned}\sum_{i,j}(a_i + b_j)\mu(E_i \cap F_j) &= \sum_{\ell=1}^k \sum_{i,j,a_i+b_j=c_\ell} (a_i + b_j)\mu(E_i \cap F_j) \\ &= \sum_{\ell=1}^k c_\ell \mu(G_\ell) \\ &= \int (f+g)\end{aligned}$$

(c) If $f \le g$, then $a_i \le b_j$ whenever $E_i \cap F_j \neq \varnothing$. So

$$\int f = \sum_{i,j} a_i \mu(E_i \cap F_j) \le \sum_{i,j} b_j \mu(E_i \cap F_j) = \int g$$

(d) Let $\nu(A) \stackrel{\text{def}}{=} \int_A f$. This is a nonnegative function on $\mathcal{M}$. Now,

$$\nu(\varnothing) = \int_\varnothing f = \int f 1_\varnothing = \int f0 = 0 \int f = 0$$

Finally, given a disjoint sequence of sets $\{A_k\}_{k=1}^\infty$, $A = \cup_{k=1}^\infty A_k$, then

$$\begin{aligned}
\nu(A) = \int_A f = \int f 1_A &= \sum_{i, E_i \cap A \neq \varnothing, c_i \neq 0} c_i \mu(E_i \cap A) \\
&= \sum_{i=1}^n a_i \mu(E_i \cap A) = \sum_{i,k} a_i \mu(E_i \cap A_k) \\
&= \sum_{k=1}^\infty \sum_{i, E_i \cap A_k \neq 0, a_i \neq 0} \mu(E_i \cap A_k) \\
&= \sum_{k=1}^\infty \int f 1_{A_k} = \sum_{k=1}^\infty \nu(A_k)
\end{aligned}$$

■

Remark. Parts (a) and (b) ensure we no longer have to worry about standard representations.
Suppose $f = \sum_{i=1}^n c_i 1_{E_i} = \sum_{j=1}^n d_j 1_{F_j}$. Then

$$\sum_j d_j \mu(F_j) = \sum_j d_j \int 1_{F_j} = \int \sum_j d_j 1_{F_j} = \int f$$

Integration of nonnegative measurable functions

Let $(X, \mathcal{M}, \mu)$ be a measure space.

Definition 0.20. Given $f : X \to [0, \infty]$ measurable, define the integral of f by

$$\int f d\mu \stackrel{\text{def}}{=} \sup\{\int \phi \, d\mu \mid 0 \leq \phi \leq f, \phi \text{ simple}\}$$

Remark. (i) If f is simple, this agrees with our previous definition.

(ii) For $c \geq 0$,

$$\begin{aligned}
\int cf\,d\mu &\stackrel{\text{def}}{=} \sup\{\int \phi\,d\mu \mid 0 \leq \phi \leq cf, \phi \text{ simple}\} \\
&= \sup\{\int \phi\,d\mu \mid 0 \leq \frac{\phi}{c} \leq f, \phi \text{ simple}\} \\
&= \sup\{\int c\psi\,d\mu \mid 0 \leq \psi \leq f, \psi \text{ simple}\} \\
&= c\sup\{\int \psi\,d\mu \mid 0 \leq \psi \leq f, \psi \text{ simple}\} \\
&= c\int f\,d\mu
\end{aligned}$$

Likewise, if $c = 0$, we see $\int cf\,d\mu = 0$.

(iii) If $f \leq g$, then $\int f \leq \int g$. This follows immediately from the definition.

Recall: A major deficiency of the Riemann integral is that it was difficult to develope minimal criteria to ensure

$$\lim_{n\to\infty} \int f_n\,d\mu = \int \lim_{n\to\infty} f_n\,d\mu$$

We are now ready for our second major theorem (the first being the characterization of all finite measures on $\mathbb{R}$ as Lebesgue-Stieljes measures)

Theorem 0.18. *(Monotonce Convergence Theorem)*
Given $\{f_n\}_{n=1}^{\infty}$ *nonnegative measurable functions such that* $f_n \leq f_{n+1}$, *then*

$$\lim_{n\to\infty} \int f_n\,d\mu = \int \lim_{n\to\infty} f_n\,d\mu$$

Proof. "$\leq$" is easy - we know by monotonicity of f_n, $f_n \leq \lim_{n\to\infty} f_n$, so

$$\int f_n \leq \int \lim_{n\to\infty} f_n$$

Thus

$$\lim_{n\to\infty} \int f_n = \limsup_{n\to\infty} \int f_n \leq \int \lim_{n\to\infty} f_n$$

Now for "$\geq$".
Let ϕ be a simple function such that $0 \leq \phi \leq \lim_{n\to\infty} f_n$.
Then, for any $a \in (0, 1)$, if $\phi(x) \neq 0$,

$$a\phi(x) < \lim_{n\to\infty} f_n(x)$$

Definte $E_n = \{x \mid f_n(x) \geq a\phi(x)\} \in \mathcal{M}$.
Since f_n is increasing, $E_1 \subseteq E_2 \subseteq E_3 \subseteq \cdots$.
Furthermore,$\cup_{n=1}^{\infty} E_n = X$, since if $\phi(x) = 0$, then $x \in E_n$ for all x, and if $\phi(x) \neq 0$, the above ensures that for all $x \in X$, there is some N such that $n \geq N$ implies $f_n(x) > a\phi(x)$.
We have

$$\int f_n \geq \int_{E_n} f_n \geq \int_{E_n} a\phi$$

Since $\nu(A) = \int_A \phi$ is a measure, by continuity from below, $\lim_{n\to\infty} \int_{E_n} \phi = \int_{\cup_n E_n} = \int \phi$.
Thus, taking limits in the above expression,

$$\lim_{n\to\infty} \int f_n \geq a \int \phi$$

Since ϕ was arbitrary, taking the supremum over ϕ,

$$\lim_{n\to\infty} \int f_n \geq a \int \lim_{n\to\infty} f_n$$

Sending $a \to 1$ gives the result.

■

Theorem 0.19. *(Beppo-Levi)*
Given $\{f_n\}_{n=1}^{\infty}$ *a sequence of nonnegative measurable functions,*

$$\sum_{n=1}^{\infty} \int f_n d\mu = \int \sum_{n=1}^{\infty} f_n \, d\mu$$

Proof. First, fix f, g nonnegative, measurable. There exist sequences $\{\phi_i\}_{i=1}^{\infty}, \{\psi_j\}_{j=1}^{\infty}$ simple, with $\phi_i \uparrow f, \psi_j \uparrow g$ pointwise.
In particular, $\phi_i + \psi_i \uparrow f + g$.

$$\begin{aligned}
\int f + g &= \int \lim_{i\to\infty} \phi_i + \psi_i = \lim_{i\to\infty} \int \phi_i + \psi_i \\
&= \lim_{i\to\infty} \int \phi_i + \lim_{i\to\infty} \int \psi_i \\
&= \int \lim_{i\to\infty} \phi_i + \int \lim_{i\to\infty} \psi_i \\
&= \int f + \int g
\end{aligned}$$

By induction, for all $N \in \mathbb{N}$,

$$\int \sum_{n=1}^{N} f_n = \sum_{n=1}^{N} \int f_n$$

Lecture 10

So, for all $N \in N$

$$\int \sum_{n=1}^{N} f_n = \sum_{n=1}^{N} \int f_n$$

Thus, by the monotone convergence theorem,

$$\sum_{n=1}^{\infty} \int f_n = \lim_{n\to\infty} \sum_{n=1}^{N} \int f_n = \int \lim_{n\to\infty} \sum_{n=1}^{N} f_n = \int \sum_{n=1}^{\infty} f_n$$

Without monootnicity of the sequence f_n, you can still get an inequality for liminfs.

Lemma 3. *(Fatou's Lemma)*
Given $\{f_n\}_{n=1}^{\infty}$*, nonnegative, measurable,*

$$\liminf_{n\to\infty} \int f_n \, d\mu \geq \int \liminf_{n\to\infty} f_n \, d\mu$$

Proof. By definition, $\liminf_{n\to\infty} f_n = \lim_{n\to\infty} \underbrace{\inf_{k\geq n}(f_k)}_{=g(n)}$. Further, $g_n \leq g_{n+1}$ for all $n \in \mathbb{N}$.
Thus, by the monotone convergence theorem,

$$\lim_{n\to\infty} \int g_n = \int \lim_{n\to\infty} = \int \liminf_{n\to\infty} f_n$$

By definition, $g_n \leq f_n$ for all $n \in \mathbb{N}$, so $\int g_n \leq \int f_n$. Taking the liminf of both sides,

$$\liminf_{n\to\infty} \int f_n \geq \liminf_{n\to\infty} \int g_n = \lim_{n\to\infty} \int g_n = \int \liminf_{n\to\infty} f_n$$

■

Example 0.9. Here are two important examples where strict inequality occurs. We sill use $(X, \mathcal{M}, \mu) = (\mathbb{R}, \mathcal{M}_{\lambda^*}, \lambda)$.

1. The "runs away to infinity" example. Consider $f_n = 1_{[n,n+1]}$. It is clear that $\lim_{n\to\infty} f_n = 0$. However, $\int f_n = 1$ for all n. So

$$\liminf_{n\to\infty} \int f_n = 1 > 0 = \int \liminf_{n\to\infty} f_n$$

2. The "goes up the spout" example. Let $f_n = n 1_{[0,\frac{1}{n}]}$. We see $\lim_{n\to\infty} f_n = \begin{cases} 0 & x \neq 0 \\ +\infty & x = 0 \end{cases}$

But, for each f_n, $\int f_n = 1$. So

$$\liminf_{n\to\infty} \int f_n = 1 > 0 = \int \liminf_{n\to\infty} f_n$$

In the second example, we accept on faith that the integral of a function which is zero almost everywhere is zero. If we don't want to do that, we can prove it using the simple functions $g_n(x) = \begin{cases} 0 & x \neq 0 \\ n & x = 0 \end{cases}$.

Proposition 8. *Given $f : X \to [0, +\infty]$ measurable, then if $\int f\, d\mu < +\infty$, we know*

(i) $\{x \mid f(x) = +\infty\}$ is a null set.

(ii) $\{x \mid f(x) > 0\}$ is σ-finite (i.e. a countable union of measurable sets of finite measure).

Proof. Homework 6

Proposition 9. *Given $f : X \to [0, +\infty]$ measurable, then $\int f\, d\mu = 0$ if and only if $f = 0$ μ-almost everywhere.*

Proof. First, suppose f is a simple function, so

$$f = \sum_{i=1}^{n} a_i 1_{E_i}$$

Then $\int f.\, d\mu = 0$ is equivalent to $\sum_{i=1}^{n} a_i \mu(E_i)$. By hypothesis, f is nonnegative, so we are forced to conclude that, for every i, either $a_i = 0$ or $\mu(E_i) = 0$. If $a_i = 0$, then $f|_{E_i} \equiv 0$, and if $\mu(E_i) = 0$, f achieves the value a_i only on a measure zero set.
Now for more general f.
Suppose $f = 0$ μ-almost everywhere. Then,

$$\int f\, d\mu = \sup\{\int \phi\, d\mu \mid 0 \leq \phi \leq f, \phi \text{ simple}\}$$

But for every such ϕ, $\int \phi = 0$. So the supremum of f over this set is 0.
Conversely, if $\int f\, d\mu = 0$, then by definition, we must have that $\int \phi\, d\mu = 0$ for all $0 \leq \phi \leq f$. Assume for the sake of contradiction that $f = 0$ μ-almost everywhere fails. That is, there exists a set A with $\mu(A) = 0$ so that $f|_A > 0$.
Note that $\{x \mid f(x) > 0\} = \cup_{i=1}^{\infty} \underbrace{\{x \mid f(x) > \frac{1}{n}\}}_{E_n}$.

Then $\mu(A) \leq \mu(\{x \mid f(x) > 0\}) \leq \sum_{i=1}^{\infty} \mu(E_n)$ by subadditivity. Thus $\mu(E_n) > 0$ for some $n \in \mathbb{N}$. Let $\phi = \frac{1}{n} 1_{E_n}$. Then $0 \leq \phi \leq f$, but $\int \phi \, d\mu = \frac{1}{n}\mu(E_n) > 0$, a contradiction.

■

Integration of Real Functions

Let $(X, \mathcal{M}, \mu)$ be a measure space. Given $f : X \to \overline{\mathbb{R}}$,

$$\begin{aligned} \overbrace{f_+}^{\text{“positive part”}} &= f \vee 0 \\ \underbrace{f_-}_{\text{“negative part”}} &= (-f) \vee 0 \end{aligned}$$

Note that the negative part is a positive function. Note $f = f_+ - f_-, |f| = f_+ + f_-$.

Definition 0.21. Given $f : \mathbb{R} \to \overline{\mathbb{R}}$ measurable, if one of $\int f_+, \int f_-$ is finite, then we define

$$\int f \, d\mu \overset{\text{def}}{=} \int f_+ \, d\mu - \int f_- \, d\mu$$

If both $\int f_+ \, d\mu, \int f_- \, d\mu$ are finite, then $\int |f| \, d\mu$ is finite.
In the case that $\int f$ is finite, we say that f is integrable, and we write $f \in L^1(\mu)$.

Proposition 10. *$L^1(\mu)$, the space of integrable functions, is a real vector space, and $\int \, d\mu$ is a linear functional on it.*

Proof. Fix $f, g \in L^1(\mu), a, b \in \mathbb{R}$. Note $af + bg \in L^1(\mu)$ and

$$|af + bg| \leq |a||f| + |b||g|$$

So

$$\begin{aligned} \int |af + bg| \, d\mu &\leq \int |a||f| \, d\mu + \int |b||g| \, d\mu \\ &= |a| \int |f| \, d\mu + |b| \int |g| \, d\mu \\ &< \infty \end{aligned}$$

Thus, $af + bg \in L^1(\mu)$, which proves that it is a real vector space.
Now, let $f \in L^1(\mu)$, $a \geq 0$,

$$\int af = \int af_+ - \int af_- = a \int f_+ - a \int f_- = a \int f$$

For $a \leq 0$, the result follows by replacing f with $-f$. Thus, $\forall f \in L^1\mu, a, \in \mathbb{R}$, $\int af = a \int f$.
We just now have to check that it behaves well under sums.
We'll come back to this proof. It boils down to a bunch of boring cases, and Katy will fill this in later. In broad strokes,

$$\begin{aligned}\int f+g &= \text{stuff}\\ &= \text{stuff}\\ &= \text{stuff}\\ &= \int f_+ + \int g_+ - \left(\int f_- + \int g_-\right)\\ &= \int f + \int g\end{aligned}$$

■

Proposition 11. *If $f \in L^1(\mu)$, then*

$$\left|\int f\,d\mu\right| \leq \int |f|\,d\mu$$

Proof.

$$\begin{aligned}\left|\int f\,d\mu\right| &= \left|\int f_+ - \int f_-\right|\\ &\leq \left|\int f_+\right| + \left|\int f_-\right|\\ &= \int f_+ + \int f_-\\ &= \int f_+ + f_- = \int |f|\end{aligned}$$

■

Proposition 12. *If $f, g \in L^1(\mu)$, then $\int |f-g|\,d\mu = 0$ if and only if $f = g$ μ-almost everywhere.*

Proof. This is an immediate corollary of previous proposition about nonnegative measurable functions, taking $h = |f-g|$.

■

The moral is that if you modify a function on a null set, it does not change the integral.
Consequently, even if a function f is only defined almost everywhere, $\int f$ is still well-defined, since we may take f to be equal to any particular element of $\overline{\mathbb{R}}$ where it is not defined, and it won't affect $\int f$.
We have already shown $L^1(\mu)$ is a vector space... is it a metric space?
What if we define $d_{L^1}(f,g) = \int |f-g|\, d\mu$? Well, if f and g agree almost everywhere, but not everywhere, then $\int |f-g|\, d\mu = 0$, but $f \neq g$.
We see that, when d_{L^1} is defined on all integrable functions, it fails to be nondegenerate. The solution is to slightly modify our definition of L^1.

Definition 0.22. $L^1(\mu) \overset{\text{def}}{=} \{f : X \to \overline{\mathbb{R}} \text{ measurable}, \int |f|\, d\mu < +\infty\}/\sim$, where $f \sim g$ if and only if $f = g$ μ-almost everywhere.
Remark. By abuse of notation, let $f \in L^1(\mu)$ denote

1. The equivalence class $[f]$ under $\sim$

2. A representative of this equivalence class

3. A representative which is only defined μ-almost everywhere.

Proposition 13. *We can define a norm, $\|f\|_{L^1(\mu)} = \int |f|\, d\mu$ on $L^1(\mu)$.*

Proof. The triangle inequality follows directly from the fact that

$$\left|\int f\right| \leq \int |f|$$

Absolute homogeneity is a consequence of the linearity of the integral, and nondegeneracy is a consequence of the fact that a $\int |f|\, d\mu = 0$ if and only if $f = 0$ μ-almost everywhere.

■

Recall that a hypothesis of Fatou's lemma is that the functions f_n are nonnegative. Is this necessary? It turns out yes, and here is an example of why.

Example 0.10. This is the "goes down the spout" example. Let $(X, \mathcal{M}, \mu) = (\mathbb{R}, \mathcal{M}_{\lambda^*}, \lambda)$.
Let $f_n = -n 1_{[0,\frac{1}{n}]}$. Then $\lim_{n\to\infty} f_n = \begin{cases} 0 & x \neq 0 \\ -\infty & x = 0 \end{cases}$.
So

$$\liminf_{n\to\infty} \int f_n = -1 \not\geq 0 = \int \liminf_{n\to\infty} f_n$$

Is there a theorem that will allow us to interchange the limit and the integral of a wider class of sequences of functions? Tune in next time...

Lecture 11

The following is the 4th major theorem of the course.

Theorem 0.20. *(Dominated Convergence Theorem)*
Given a $\{f_n\}_{n=1}^{\infty} \in (L^1(\mu))^{\mathbb{N}}$ *such that* $\lim_{n\to\infty} f_n$ *exists* μ*-almost everywhere, then if there exists* $g \in L^1(\mu)$ *such that* $|f_n| \leq g$ μ*-almost everywhere for all* $n \in \mathbb{N}$*, then*

$$\lim_{n\to\infty} \int f_n \, d\mu = \int \lim_{n\to\infty} f_n \, d\mu$$

Proof.
Remark. g does <u>not</u> need to be bounded. For example, if we take $g = \frac{1}{\sqrt{x}} 1_{[-1,1]}(x)$, then $\int g(x) d\lambda(x) = \int_{-1}^{1} \frac{dx}{\sqrt{|x|}} = 4 < \infty$.
Remark. Recall Fatou's lemma for nonnegative, measurable functions:

$$\liminf_{n\to\infty} \int f_n \, d\mu \geq \int \liminf_{n\to\infty} f_n \, d\mu$$

No dominating function can exist for the "run off to infinity" example (i.e. $f_n = 1_{[n,n+1)}$).

Proof. Since $\lim_{n\to\infty} f_n$ exists μ-almost everywhere and $|f_n| \leq g$ μ-almost everywhere for all $n \in \mathbb{N}$, we have

$$\left| \lim_{n\to\infty} f_n \right| \leq g, \ \mu - \text{almost everywhere}$$

Thus $\lim_{n\to\infty} f_n \in L^1(\mu)$. Since $g - f_n \geq 0$ and $g + f_n \geq 0$ μ-almost everywhere for all $n \in \mathbb{N}$, Fatou's lemma ensures

$$\begin{aligned} \liminf_{n\to\infty} \int g + f_n &\geq \int \liminf_{n\to\infty} g + f_n \\ &= \int g + \lim_{n\to\infty} f_n \\ &= \int g + \int \lim_{n\to\infty} f_n \end{aligned}$$

and

$$\limsup_{n\to\infty}\int g - f_n \geq \int \limsup_{n\to\infty} g - f_n$$
$$= \int g - \lim_{n\to\infty} f_n$$
$$= \int g - \int \lim_{n\to\infty} f_n$$

Because $\int g < \infty$, we may subtract $\int g$ on both sides, so

$$\liminf_{n\to\infty}\int f_n \geq \int \lim_{n\to\infty} f_n$$
$$\geq \limsup_{n\to\infty}\int f_n$$

Thus, equality holds throughout, which gives the result.

■

Using the dominated convergence theorem, we can identify the useful subsets of functions that are dense in $L^1(\mu)$.
The following can be considered the 5th major theorem of the course.

Theorem 0.21. *For any measure space* $(X, \mathcal{M}, \mu)$,

- *Simple functions are dense in* $L^1(\mu)$

If μ *is a Lebesgue-Stieltjes measure on* $\mathbb{R}$,

- *Simple functions of the form*

$$\xi = \sum_{j=1}^{n} a_j 1_{I_j},\ I_j = \cup_{i=1}^{m} U_{ij}$$

  For any U_{ij} *open, are dense in* $L^1(\mu)$.

- $C_c(\mathbb{R})$ *(the continuous, compactly supported real functions) is dense in* $L^1(\mu)$. *We define the support of a function as* $\operatorname{supp} f = \overline{\{x \in \mathbb{R} \mid f(x) \neq 0\}}$

Our proof of this theorem relies on the following lemma, whose proof will be assigned on HW6

Lemma 4. If μ is a Lebesgue-Stieltjes measure and $E \in \mathcal{M}_{\mu^*}$ with $\mu(E) < \infty$, then for all $\varepsilon > 0$, there exist open intervals $\{I_i\}_{i=1}^n$ so that

$$\mu\left(E \triangle \left(\bigcup_{i=1}^n I_i\right)\right) < \varepsilon$$

Proof. Fix $f \in L^1(\mu)$. Since f_+, f_- are nonnegative, measurable functions, there exist simple functions $\psi_n \uparrow f_+$ and $\zeta_n \uparrow f_-$ pointwise.
Thus, $\psi_n - \zeta_n \to f_+ - f_- = f$ pointwise. Furthermore, $\left|(\psi_n - \zeta_n)\right| \leq \psi_n + \zeta_n + |f| \leq f_+ + f_- + |f| = 2|f|$. Because $f \in L^1(\mu)$, $2|f| \in L^1(\mu)$. So we can use the Dominated Convergence Theorem, which ensures

$$\lim_{n\to\infty} \int |(\psi_n - \zeta_n) - f| = \int \lim_{n\to\infty} |(\psi_n - \zeta_n) - f| = 0$$

Thus, for all $\varepsilon > 0$, there exists a simple function ϕ such that $\|\phi - f\|_{L^1(\mu)} < \varepsilon$.
This shows simple functions are dense in $L^1(\mu)$.
Now, suppose μ is a Lebesgue-Stieltjes measure on $\mathbb{R}$. Fix $\varepsilon > 0$, and let ϕ be a simple function as above such that

$$\|\phi - f\|_{L^1(\mu)} < \frac{\varepsilon}{2},\ |\phi| \leq |f|$$

Let $\phi = \sum_{j=1}^n a_j 1_{E_j}$. Without loss of generality, we may suppose $a_j \neq 0$ for all j and $\{E_j\}_{j=1}^n$ are disjoint.
By construction, for any j,

$$|a_j|\mu(E_j) \leq \int |\phi|\, d\mu \leq \int |f|\, d\mu < +\infty$$

Thus, $\mu(E_j) < +\infty$. So, by the lemma, there exist open intervals

$$\{I_i^j\}_{i=1}^m$$

so that $\mu(E_j \triangle (\cup_{i=1}^m I_i^j)) < \frac{\varepsilon}{2\max_j |a_j| n}$. Thus,

$$\begin{aligned}
\|\phi - \sum_{j=1}^n a_j 1_{\cup_{i=1}^m I_i^j}\|_{L^1(\mu)} &= \|\sum_{j=1}^n a_j 1_{E_j} - \sum_{j=1}^n a_j 1_{\cup_{i=1}^m I_i^j}\|_{L^1(\mu)} \\
&\leq \sum_{j=1}^n |a_j| \|1_{E_j} - 1_{\cup_{i=1}^n I_i^j}\|_{L^1(\mu))} \\
&= \sum_{j=1}^n |a_j| \mu(E_j \triangle (\cup_{i=1}^n I_i^j)) \\
&< \frac{\varepsilon}{2}
\end{aligned}$$

Since $\varepsilon > 0$ was arbitrary, this gives the result. To show $C_c(\mathbb{R})$ is dense in $L^1(\mu)$, it suffices to show that for any

$$\psi = \sum_{j=1}^{n} a_j 1_{\cup_{i=1}^n I_i^j}$$

and $\varepsilon > 0$, there is an $f \in C_c(\mathbb{R})$ such that

$$\|f - \psi\|_{L^1(\mu)} < \varepsilon$$

Note that, for any open interval I_i^j, there exists $f_{ij} \in C_c(\mathbb{R})$ so that

$$\|1_{I_i^j} - f_{ij}\|_{L^1(\mu)} < \varepsilon$$

We can accomplish this via something like continuous bump functions.
We will finish the proof next time.

Lecture 12

For any interval I_i^j, there exists a $f_{ij} \in C_c(\mathbb{R})$ such that

$$\begin{aligned}\|1_{I_i^j} - f_{ij}\|_{L^1(\mu)} &< \int_{(a_j - \frac{1}{k}, a_{ij}) \cup (b_{ij}, b_{ij} + \frac{1}{k})} f_{ij}\, d\mu \\ &\leq \mu((a_{ij} - \frac{1}{k}, a_{ij})) + \mu((b_{ij}, b_{ij} + \frac{1}{k})) \to 0\end{aligned}$$

As $k \to \infty$, the above goes to 0 by continuity from above. So, in particular, for each I_i^j, there is some $f_{ij} \in C_c(\mathbb{R})$ so that $\|1_{I_i^j} - f_{ij}\|_{L^1(\mu)} < \frac{\varepsilon}{(\max_j a_j)nm}$, where n, m are the upper bounds of the indexing sets that i, j come from.
Without loss of generality, we may assume that $I_i^j \cap I_{i'}^{j'} = \varnothing$ unless $\delta_{ii'}\delta_{jj'} = 1$.
Thus,

$$\begin{aligned}\|\sum_{j=1}^{n} a_j 1_{\cup_{i=1}^m I_j^i} - \overbrace{\sum_{i,j} a_j f_{ij}}^{\in C_c(\mathbb{R})}\|_{L^1(\mu)} &\leq \sum_{j=1}^{n} |a_j| \|1_{\cup_{i=1}^m I_j^i} - \sum_{i=1}^{m} f_{ij}\|_{L^1(\mu)} \\ &\leq \sum_{i,j} |a_j| \|1_{I_j^i} - f_{ij}\|_{L^1(\mu)} \\ &< \varepsilon\end{aligned}$$

■

Remark. We showed it suffices to consider simple functions of the form $\xi = \sum_{j=1}^{n} a_j 1_{\cup_{i=1}^m I_j^i}$ for I_i^j open intervals.

Theorem 0.22. *(Differentiation under the integral)*
Consider $f : X \times [a,b] \to \mathbb{R}$ *so that* $(x,t) \mapsto f(x,t)$*. Suppose that* f *satisfies the following properties:*

(i) For all $t \in [a,b]$, $f(\cdot,t) \in L^1(\mu)$.

(ii) For all $(x,t) \in X \times [a,b]$, $\frac{\partial f}{\partial t}(x,t)$ *exists.*

(iii) There exists a $g \in L^1(\mu)$ *such that* $\left|\frac{\partial f}{\partial t}(x,t)\right| \leq g(x)$ *for all* $(x,t) \in X \times [a,b]$.

Then $t \mapsto \int_X f(x,t)\,d\mu$ *is differentiable, and*

$$\frac{d}{dt}\int_X f(x,t)\,d\mu(x) = \int_X \frac{\partial f}{\partial t}(x,t)\,d\mu(x)$$

Proof. Fix $t_0 \in [a,b]$ and let $\{t_n\}_{n=1}^{\infty}$ be any sequence converging to t_0. By (ii),

$$\lim_{n\to\infty} \underbrace{\frac{f(x,t_n) - f(x,t_0)}{t_n - t_0}}_{\overset{\text{def}}{=} h_n(x)} = \frac{\partial f}{\partial t}(x,t_0)$$

Thus, $x \mapsto \frac{\partial f}{\partial t}(x,t_0)$ is measurable. By the Mean Value Theorem and (iii),

$$|h_n(x)| \leq \sup_{t\in[a,b]}\left|\frac{\partial f}{\partial t}(x,t)\right| \leq g(x)$$

Thus, by the Dominated Convergence Theorem,

$$\begin{aligned}\lim_{n\to\infty}\frac{\int f(x,t)\,d\mu(x) - \int f(x,t_0)\,d\mu(x)}{t_n - t_0} &= \lim_{n\to\infty}\int h_n(x)\,d\mu(x) \\ &= \int \frac{\partial f}{\partial t}(x,t_0)\,d\mu(x)\end{aligned}$$

Modes of Convergence

Let $(X,\mathcal{M},\mu)$ be some measure space, and $f_n, f : X \to \overline{\mathbb{R}}$ measurable. We know the following senses in which f_n "converges" to f:

1. Uniform convergence: $\sup_{x\in X}|f_n(x) = f(x)| \to 0$ as $n \to \infty$

2. Pointwise convergence: $f_n(x) \to f(x)$ for all $x \in X$.

3. Pointwise μ-almost everywhere convergence: $f_n(x) \to f(x)$ for μ-almost every $x \in X$.

4. L^1-convergence: $\|f_n - f\|_{L^1(\mu)} \to 0$

Clearly, $1 \implies 2 \implies 3$
The main goal of this section of the course is to figure out how 3 and 4 are related.

Example 0.11. Of the above, 1 does not necessarily imply 4. An example of something that converges uniformly but not in L^1 is the "splat" example: $f_n = \frac{1}{n}1_[0, n]$, and $f = 0$. Clearly $f_n \to f$ uniformly, but $\|f_n - f\|_{L^1(\mu)} = 1$, which does not go to zero.

Example 0.12. 4 does not necessarily imply 3. This is the "refining wave" example. Define $f_1 = 1_{[0,1]}$, $f_2 = 1_{[0,\frac{1}{2}]}$, $f_3 = 1_{[\frac{1}{2},1]}$, $f_4 = 1_{[0,\frac{1}{4}]}$, $f_5 = 1_{[\frac{1}{4},\frac{1}{2}]}$, and so on. This is also sometimes called the "typewriter sequence". We can see that it converges in L^1 to $f = 0$, because the integral will get smaller than any negative power of 2, but $f_n(x)$ does not converge to $f(x)$ for any $x \in [0, 1]$.
From these it would seem to be pretty hopeless that we could actually say anything about the relationship between these notions. But with additional assumptions/lower expectations, something can still be salvaged.

Example 0.13. If we have 3 and a dominating function, then this implies 4. Given $f_n \to f$ μ-almost everywhere and $|f_n| \leq g$ μ-almost everywhere, then $|f_n - f| \to 0$ μ-almost everywhere and $|f_n - f| \leq 2g$, so by the Dominated Convergence Theorem,

$$\|f_n - f\|_{L^1(\mu)} = \int |f_n - f| \, d\mu \to 0$$

Lots of time in the next few lectures will be spent wondering when we can go from 4 to 3.

Definition 0.23. A sequence of measurable functions $f_n : X \to \mathbb{R}$ converges in measure to a meaasurable function f if for all $\varepsilon > 0$,

$$\lim_{n\to\infty} \mu(\{x : |f_n(x) - f(x)| \geq \varepsilon\}) = 0$$

Similarly, a sequence f_n is Cauchy in measure if, for all $\varepsilon > 0$,

$$\lim_{m,n\to\infty} \mu(\{x : |f_n(x) - f_m(x)| \geq \varepsilon\}) = 0$$

Remark. We will show in homework that this topology is metrizable (i.e. a set is closed if it contains all its limit points, and this uniquely specifies a topology).

Example 0.14. Recall the "splat" example, $f_n = \frac{1}{n}1_{[0,n]}$. This converges uniformly, but not in L^1. Does it converge in measure?
For any $\varepsilon > 0$, $\frac{1}{n} < \varepsilon$ for n large enough. So for large enough n, $\{x : |\frac{1}{n}1_{[0,n]}(x)| \geq \varepsilon\} = \varnothing$, so this converges in measure.
The "refining wave" example converges to 0 in measure. However, the "run away to infinity" examples does not converge to zero in measure.

Lecture 13

The best way to improve your score on midterm 2 is to solve homework without help, and re-work old homework without help.
We will work up to showing that if a sequence f_n converges in L^1 to f, then f_n admits a subsequence which converges to f almost everywhere.

Remark. If we have a sequence f_n which converges to f in measure, then the sequence f_n is Cauchy in measure.

Proof. Fix $\varepsilon > 0$. If $|f_n(x) - f(x)| < \frac{\varepsilon}{2}$ and $|f_m(x) - f(x)| < \frac{\varepsilon}{2}$, then by the triangle inequality, $|f_n(x) - f_m(x)| < \varepsilon$.
This implies the following containment of sets:

$$\{x : |f_n(x) - f(x)| < \frac{\varepsilon}{2}\} \cap \{x : |f_m(x) - f(x)| < \frac{\varepsilon}{2}\} \subseteq \{x : |f_n(x) - f_m(x)| < \varepsilon\}$$

Since $A \subseteq B \iff B^c \subseteq A^c$,

$$\{x : |f_n(x) - f_m(x)| \geq \frac{\varepsilon}{2}\} \subseteq \underbrace{\{x : |f_n(x) - f(x)| \geq \frac{\varepsilon}{2}\}}_{\mu(\text{this})\to 0} \cup \overbrace{\{x : |f_m(x) - f(x)| \geq \frac{\varepsilon}{2}\}}^{\mu(\text{this})\to 0}$$

Thus, by subaddivity,

$$\mu(\{x : |f_n(x) - f_m(x)| \geq \frac{\varepsilon}{2}\}) \leq \underbrace{\mu(\{x : |f_n(x) - f(x)| \geq \frac{\varepsilon}{2}\})}_{\to 0} + \overbrace{\mu(\{x : |f_m(x) - f(x)| \geq \frac{\varepsilon}{2}\})}^{\to 0}$$

so the measure of the left hand side goes to zero.

Theorem 0.23. *Consider $f_n : X \to \mathbb{R}$ measurable.*

(i) If f_n is Cauchy in measure,

  (a) There exists $f : X \to \mathbb{R}$ measurable such that $f_n \to f$ in measure

  (b) There is a subsequence f_{n_k} such that $f_{n_k} \to f$ μ-almost everywhere.

(ii) If, in addition, $f_n \to g$ in measure, then $f = g$ μ-almost everywhere.

Proof. We will start with $(i)(b)$. We begin by finding our guess for f.

By assumption, f_n is Cauchy in measure. So, there exists a subsequence f_{n_k} such that

$$\mu(\overbrace{\{x : |\underbrace{f_{n_k}}_{g_k}(x) - \underbrace{f_{n_{k+1}}}_{g_{k+1}}(x)| \geq \frac{1}{2^k}\}}^{E_k}) \leq \frac{1}{2^k}$$

E_k will be our "bad sets" that we want to avoid.
By countable subadditivity,

$$mu \underbrace{\left(\bigcup_{k=\ell}^{\infty} E_k\right)}_{F_\ell} \leq \sum_{k=\ell}^{\infty} \mu(E_k) \leq \frac{1}{2^{\ell-1}}$$

Note that, if $x \notin F_\ell$, then $i \geq j \geq \ell$, so

$$|g_i(x) - g_j(x) \leq \sum_{m=j}^{i-1} |g_m(x) - g_{m+1}(x)|$$

Since $x \notin F_\ell$, it ensures $x \notin E_m$, so the right hand side of the above is less than or equal to $\sum_{m=j}^{i-1} \frac{1}{2^m} \leq \frac{1}{2^{j-1}}$.
Thus, $\{g_i(x)\}_{i=1}^{\infty}$ is a Cauchy sequence of real numbers. Note that $\ell \in \mathbb{N}$ was arbitrary.
Define $F = \cap_{\ell=1}^{\infty} F_\ell$, and note $\mu(F) \leq \mu(F_\ell)$ by montonicity, and $\mu(F_\ell) \leq \frac{1}{2^{\ell-1}}$.
Taking $\ell \to \infty$, we are forced to conclude $\mu(F) = 0$. Note $x \notin F$ ensures there is some ℓ such that $x \notin F_\ell$.
Define

$$f(x) = \begin{cases} \lim_{k\to\infty} g_k(x) & x \notin F_\ell \\ 0 & x \in F \end{cases}$$

By the way $f(x)$ is defined, $g_k \to f$ μ-almost everywhere. Now, we show $(i)(a)$. By our earlier estimation $|g_i(x) - g_j(x)| \leq \frac{1}{2^{j-1}}$, we know that if $x \notin F_\ell$, and $j \geq \ell$, then

$$|f(x) - g_j(x)| = \lim_{i\to\infty} |g_i(x) - g_j(x)| \leq \frac{1}{2^{j-1}}$$

Thus for all $\varepsilon > 0$ and $\ell \in \mathbb{N}$, we may choose j sufficiently large so that

$$\begin{aligned} \mu(\{x : |f(x) - g_j(x)| \geq \varepsilon\} &\leq \mu(\{x : |f(x) - g_j(x)| \geq \frac{1}{2^{j-1}}\}) \\ &\leq \mu(F_\ell) \\ &\leq \frac{1}{2^{\ell-1}} \end{aligned}$$

In particular, this tells us $g_j \to f$ in measure.
Finally, for all $\varepsilon > 0$, $j \in \mathbb{N}$,

$$\{x : |f_n(x) - f(x)| > \varepsilon\} \subseteq \{x : |f_n(x) - g_n(x)| \geq \frac{\varepsilon}{2}\} \cup \{x : |g_j(x) - f(x)| \geq \frac{\varepsilon}{2}\}$$

Since we may choose n, j sufficiently large to make the right have arbitrarily small measure, this shows

$$\lim_{n\to\infty} \mu(\{x : |f_n(x) - f(x)| > \varepsilon\}) = 0$$

That is, $f_n \to f$ in measure.
It remains to show (ii). Fix $\varepsilon > 0$. Then,

$$\{x : |g(x) - f(x)| > \varepsilon\} \subseteq \{x : |f_n(x) - g(x)| > \frac{\varepsilon}{2}\} \cup \{x : |f_n(x) - f(x)| > \frac{\varepsilon}{2}\}$$

Since we may choose n sufficiently large so that the right hand side has arbitrarily small measure,

$$\mu(\{x : |g(x) - f(x)| > \varepsilon\})$$

Taking $\varepsilon = \frac{1}{k}$ and letting

$$E_k = \{x : |g(x) - f(x)| > \frac{1}{k}\},$$

we have

$$\begin{aligned} \mu(\{x : |f(x) - g(x)| > 0\}) &\leq \mu\left(\bigcup_{k=1}^{\infty} E_k\right) \\ &\leq \sum_{k=1}^{\infty} \mu(E_k) = 0 \end{aligned}$$

Thus $f = g$ μ-almost everywhere.

■

We now apply this to study convergence in $L^1(\mu)$.

Proposition 14. *(a) If f_n is Cauchy in $L^1(\mu)$, then it is Cauchy in measure.*

(b) If $f_n \to f$ in $L^1(\mu)$, then $f_n \to f$ in measure.

Proof. Fix $\varepsilon > 0$ and define

$$E_{n,m,\varepsilon} = \{x : |f_n(x) - f_m(x)| \geq \varepsilon\}$$

Then

$$\begin{aligned}
\varepsilon\mu(E_{n,m,\varepsilon}) &= \int_{E_{n,m,\varepsilon}} \varepsilon \, d\mu \\
&\leq \int_{E_{n,m,\varepsilon}} |f_n(x) - f_m(x)| \, d\mu \\
&\leq \int_X |f_n(x) - f_m(x)| \, d\mu = \|f_n - f_m\|_{L^1(\mu)}
\end{aligned}$$

Thus $\lim_{m,n\to\infty} \mu(E_{n,m,\varepsilon}) = 0$, so f_n is Cauchy in measure. The proof of (b) is analagous. ■

Lecture 14

Remark. If f_n is Cauchy in $L^1(\mu)$, then there exists some $M \in \mathbb{N}$ such that if $n, m \geq M$, then

$$\|f_n\|_{L^1(\mu)} - \|f_m\|_{L^1(\mu)}\|f_n - f_m\|_{L^1(\mu)} < 1$$

In particular,

$$\sup_n \|f_n\|_{L^1(\mu)} \leq \max_{k\leq M} \|f_k\|_{L^1(\mu)}\|f_m\|_{L^1(\mu)} + 1 < \infty$$

The following corollary could be considered our sixth major theorem of the course.

Corollary 0.24. *(i) If f_n is Cauchy in $L^1(\mu)$, then there exists $f \in L^1(\mu)$ and a subsequence f_{n_k} such that $f_{n_k} \to f$ μ-almost everywhere.*

(ii) If, in addition, $f_n \to g$ in $L^1(\mu)$, then $f = g$ μ-almost everywhere.

Proof. We will prove part (i) first.
Since f_n is Cauchy in L^1, it is Cauchy in measure by the previous proposition. By the most recent theorem, there exists a measurable $f : X \to \mathbb{R}$ and a subsequence f_{n_k} such that $f_{n_k} \to f$ μ-almost everywhere.
It remains to show $f \in L^1(\mu)$. By the remark, for k sufficiently large,

$$\|f_{n_k}\|_{L^1} < 1 + \|f_M\|_{L^1}$$

By Fatou's Lemma,

$$\int |f| \, d\mu = \int \liminf_{n\to\infty} |f_{n_k}| \, d\mu \leq \liminf_{n\to\infty} \int |f_{n_k}| \, d\mu < \infty,$$

where the final inequality follows from our bound on $\|f_{n_k}\|_{L^1}$ above. So $\int |f| \, d\mu < \infty$, so $f \in L^1(\mu)$.

It remains to prove part (*ii*), which follows easily from the previous proposition and the theorem.

■

Definition 0.24. A normed vector space which is complete with respect to the metric induced by this norm is called a Banach Space
The following can be considered the seventh major theorem.

Corollary 0.25. *$L^1(\mu)$ is a Banach Space.*

Proof. Let f_n be Cauchy in $L^1(\mu)$. By the previous corollary, we know there exists $f \in L^1(\mu)$ and a subsequence f_{n_k} such that $f_{n_k} \to f$ μ-almost everywhere.
By Fatou's Lemma,

$$\begin{aligned} \int |f_{n_k} - f| \, d\mu &= \int \lim_{j\to\infty} |f_{n_k} - f_j| \, d\mu \\ &\leq \liminf_{j\to\infty} \int |f_{n_k} - f_j| \, d\mu \end{aligned}$$

Thus $\lim_{k\to\infty} \int |f_{n_k} - f| \, d\mu = 0$.
Finally, since f_n is Cauchy, $\|f_m - f\|_{L^1(\mu)} \leq \|f_m - f_{n_k}\|_{L^1(\mu)} + \|f_{n_k} - f\|_{L^1(\mu)} \to 0$, so $lim_{n\to\infty} \int |f_n - | \, d\mu = 0$, so $f_n \to f$ in L^1.

■

Summary of Different Modes of Convergence

Let $f_n \to f$ in some sense.

- If it converges in L^1, it converges in measure. The "splat" example (i.e. $f_n = \frac{1}{n}1_{[0,n]}$) is a counterexample to the converse.
- If it converges in measure, it converges up to a subsequence μ-almost everywhere. The converse hold if $\mu(X) < \infty$.

The following can be considered the eighth major theorem.

Theorem 0.26. *(Egoroff)*
Suppose $\mu(X) < \infty$, and $f_n, f : X \to \mathbb{R}$ are measurable functions such that $f_n \to f$ μ-almost everywhere. Then for all $\varepsilon > 0$, there exists $E \in \mathcal{M}$ such that $\mu(E) < \varepsilon$, and $f_n \to f$ uniformly on E^c.

Remark. The "run away to infinity" example (i.e. $f_n = 1_{[n,n+1)}$) shows the necessity of the hypothesis that $\mu(X) < \infty$.
Note also that we can NOT necessarily get a null-set E such that $f_n \to f$ uniformly on E^c. A witness to this is the sequence $f_n(x) = x^n|_{[0,1]}$. $f_n \not\to 0$ uniformly, and will

not converge uniformly on any set which has 1 as a limit point. But the complement of any null subset of $[0,1]$ is dense, so has 1 as a limit point.

Proof. We will prove it by cases.

Case 1

In case 1, suppose $f_n \to f$ pointwise. Define $E_{n,k} = \cup_{m=n}^{\infty}\{x : |f_m(x) - f(x)| \geq \frac{1}{k}\}$. We can see that $E_{n,k} \supseteq E_{n+1,k}$. Because $\mu(X) < \infty$, we can use continuity from above.
$\cap_{n=1}^{\infty} E_{n,k} = \varnothing$, because $f_n \to f$ pointwise, so $\lim_{n\to\infty} \mu(E_{n,k}) = 0$.
Fix $\varepsilon > 0$. For all $k \in \mathbb{N}$, there exists a subsequence n_k such that $\mu(E_{n_k,k}) < \frac{\varepsilon}{2^k}$.
Define $E = \cup_{k=1}^{\infty} E_{n_k,k}$. It is clear that E is measurable. By subadditivity,

$$\mu(E) \leq \sum_{k=1}^{\infty} \mu(E_{n,k}) = \sum_{k=1}^{\infty} \frac{\varepsilon}{2^k} = \varepsilon$$

So $\mu(E) < \varepsilon$.
If $x \in E^c$, then $x \notin E_{n_k,k}$ for any $k \in \mathbb{N}$.
Thus, $x \notin E_{n_k,k}$ for all $k \in \mathbb{N}, n \geq n_k$. That is,

$$x \notin \{x : |f_n(x) - f(x)| \geq \frac{1}{k}\}$$

Thus, $|f_n(x) - f(x)| < \frac{1}{k}$. This shows $f_n \to f$ uniformly on E^c.

Case 2

Suppose $f_n \to f$ pointwise μ-almost everywhere. We will reduce to case 1 by defining

$$N = \{x : f_n(x) \not\to f(x)\}$$

By assumption, this is a null set. Define

$$g_n = f_n 1_{N^c},\ g = f 1_{N^c}$$

Since $g_n \to g$ pointwise, there exists $E \in \mathcal{M}$ with arbitrarily small measure such that $g_n \to g$ uniformly on E^c.
Let $F = E \cup N$. Then $\mu(F) < \varepsilon$ and $f_n \to f$ uniformly on F^c.

■

Corollary 0.27. *Suppose $\mu(X) < \infty$ and $f_n, f : X \to \mathbb{R}$ are measurable such that $f_n \to f$ μ-almost everywhere. Then $f_n \to f$ in measure.*

Proof. Fix $\varepsilon > 0$. By Egeroff's theorem, for all $\delta > 0$, there exists $E \in \mathcal{M}$ such that $\mu(E) < \delta$, and $f_n \to f$ uniformly on E^c.
Thus,

$$\mu(\{x : |f_n(x) - f(x)| \geq \varepsilon\}) \leq \underbrace{\mu(E)}_{\delta} + \mu(\{x \in E^c : |f_n(x) - f(x)| \geq \varepsilon\})$$

This shows that, for n sufficiently large,

$$\mu(\{x : |f_n(x) - f(x)| \geq \varepsilon\}) \leq \delta$$

This shows $f_n \to f$ in measure.

Lecture 15

Recall our 6th major theorem:

Corollary 0.28.
- *If f_n is Cauchy in $L^1(\mu)$, then $\exists f \in L^1(\mu)$ and a subsequence $f_{n_k} \to f$ μ-almost everywhere.*
- *If, in addition, $f_n \to g$ in $L^1(\mu)$, then $f = g$ μ-almost everywhere.*

The following is our 7th major theorem:

Corollary 0.29. *$L^1(\mu)$ is a Banach Space.*
Our 8th major theorem is Egoroff's:

Theorem 0.30. *(Egoroff)*
Suppose $\mu(X) < \infty$ and $f_n, f : X \to \mathbb{R}$ measurable such that $f_n \to f$ μ-almost everywhere. Then for all $\varepsilon > 0$, $\exists E \in \mathcal{M}$ such that $\mu(E) < \infty$ and $f_n \to f$ uniformly on E^c.

Corollary 0.31. *Suppose $\mu(X) < \infty$ and $f_n, f : X \to \mathbb{R}$ are measurable such that $f_n \to f$ μ-almost everywhere. Then $f_n \to f$ in measure.*

Summary of Different Modes of Convergence

$$\begin{tikzcd}
f_n \to f \text{ in } L^1(\mu) \arrow[d, Rightarrow, bend right] \\
f_n \to f \text{ in measure} \arrow[u, Rightarrow, bend right, "f_n = \frac{1}{n}1_{[0,n]}"'] \arrow[d, Rightarrow, bend right, "\text{up to a subsequence}"'] \\
f_n \to f \ \mu - \text{almost everywhere} \arrow[u, Rightarrow, bend right, "\text{if } \mu(X) < \infty"']
\end{tikzcd}$$

Product Measures

Let $(X, \mathcal{M}, \mu), (Y, \mathcal{N}, \nu)$ be measure spaces. We call any set of the form $A \times B$, for $A \in \mathcal{M}$ and $B \in \mathcal{B}$ a rectangle.
Recall:

Definition 0.25. The product σ-algebra is defined by

$$\mathcal{M} \otimes \mathcal{N} \stackrel{\text{def}}{=} \mathscr{M}(\{A \times B : A \in \mathcal{M}, B \in \mathcal{N}\})$$

Goal: We want to prove the existence of a unique measure $\mu \otimes \nu$ on the measurable space $(X \times Y, \mathcal{M} \otimes \mathcal{N})$ so that

$$\mu \otimes \nu(A \times B) = \mu(A)\nu(B)$$

for all rectangles.
Our construction of the prodcut measure will rely on the Monotone Class Theorem
Recall:

Definition 0.26. $\mathcal{A}$ is an algebra of subsets of X if it is a nonempty collection of subsets of X such that

(i) $E_1, \ldots, E_n \in \mathcal{A} \implies \cup_{i=1}^n E_i \in \mathcal{A}$

(ii) $E \in \mathcal{A} \implies E^c \in \mathcal{A}$

Remark. We usually want $\varnothing, X \in \mathcal{A}$.

Definition 0.27. $\mathcal{C}$ is a monotone class of subsets of X if it is a nonempty collection of subsets of X such that

(i) Closed under countable increasing unions. That is, if $\{E_i\}_{i=1}^\infty \subseteq \mathcal{C}$, with $E_1 \subseteq E_2 \subseteq \cdots$, then $\cup_{i=1}^\infty E_i \in \mathcal{C}$.

(ii) Closed under countable decreasing intersections. That is, if $\{E_i\}_{i=1}^\infty \subseteq \mathcal{C}$ with $E_1 \supseteq E_2 \supseteq \cdots$, then $\cap_{i=1}^\infty E_i \in \mathcal{C}$.

The slogan is that "A monotone class is closed under countable monotone unions and intersections"

Example 0.15. Any σ-algebra is a monotone class.

Example 0.16. Let $X = \{1, 2, 3\}$ (i.e. $X = 4$), and $\mathcal{C} = \{\{1\}, \{1, 2\}, \{1, 2, 3\}\}$. This is a monotone class on X but not a σ-algebra.
Recall that given any family $\mathcal{E} \subseteq 2^X$ of subsets of X, there is a smallest σ-algebra containing $\mathcal{E}$, denoted $\mathscr{E}$. Likewise,

Proposition 15. *Given $\mathcal{E} \subseteq 2^X$, there is a smallest montone class $\mathcal{C}(\mathcal{E})$ containing $\mathcal{E}$, known as the montone class generated by $\mathcal{E}$.*

Proof.

Claim. Given any nonempty collection $\mathscr{F}$ of monotone classes on X, $\cap\mathscr{F} \overset{\text{def}}{=} \{E \subseteq X : E \in \mathcal{C} \forall \mathcal{C} \in \mathscr{F}\}$

Proof. Homework 7
Now, let $\mathscr{F} = \{\mathcal{C} : \mathcal{C} \text{ is a monotone class}, \mathcal{E} \subseteq \mathcal{C}\}$. $\mathscr{F}$ is nonempty because $2^X \in \mathscr{F}$.
By the previous claim, $\cap\mathscr{F}$ is a monotone class.
By construction $\mathcal{E} \subseteq \cap\mathscr{F}$. Further, for any monotone class $\mathscr{D}$ such that $\mathcal{E} \subseteq \mathscr{D}$, $\cap\mathscr{F} \subseteq \mathscr{D}$. Thus, $\mathcal{C}(\mathcal{E}) = \cap\mathscr{F}$ is the smallest monotone class containing $\mathcal{E}$.

■

Theorem 0.32. *(Monotone Class Theorem)*
Given an algebra $\mathcal{A} \subseteq 2^X$,

$$\mathcal{C}(\mathcal{A}) = \mathscr{M}(\mathcal{A})$$

Remark. It is clear that $\mathcal{C}(\mathcal{A}) \subseteq \mathscr{M}(\mathcal{A})$, since $\mathscr{M}(\mathcal{A})$ is a monotone class.

Proof. It suffices to show that $\mathcal{C} = \mathcal{C}(\mathcal{A})$ is a σ-algebra.
For any $E \in \mathcal{C}$, define

$$\mathcal{E}_E \overset{\text{def}}{=} \overbrace{\{F \in \mathcal{C} : E \setminus F, F \setminus E, E \cap F \in \mathcal{C}\}}^{\text{“sets that play nicely with } E\text{”}}$$

Note that $\varnothing, E \in \mathcal{E}_E$, and

$$F \in \mathcal{E}_E \iff E \in \mathcal{E}_F$$

Next, note that if $E \in \mathcal{A}$, then the fact that $\mathcal{A}$ is an algebra ensures

$$F \in \mathcal{E}_E, \forall F \in \mathcal{A}$$

Furthermore, for any $E \in \mathcal{C}, \mathcal{E}_E$ is a monotone class.
If $\{E_i\}_{i=1}^\infty \subseteq \mathcal{E}_E$, with $E_1 \subseteq E_2 \subseteq \cdots$, then

(i) $E \setminus (\cup_{i=1}^\infty E_i) = E \cap (\cap_{i=1}^\infty E_i^c) = \cap_{i=1}^\infty (E \cap E_i^c)$

belongs to $\mathcal{C}$, since it's a countable decreasing intersection and the fact that $E_i \in \mathcal{E}_E$ ensures $E \setminus E_i \in \mathcal{C}$.

(ii) $(\cup_{i=1}^\infty E_i) \setminus E = \cup_{i=1}^\infty (E_i \cap E^c)$

belongs to $\mathcal{C}$, since it's a countable increasing union, and the fact that $E_i \in \mathcal{E}_E$ ensures $E_i \setminus E \in \mathcal{C}$.

(iii) $E \cap (\cup_{i=1}^\infty E_i) = \cup_{i=1}^\infty (E \cap E_i) \in \mathcal{C}$

Thus $\mathcal{E}_E$ is closed under increasing unions. Similarly, it is closed under decreasing intersections. Thus, it is a monotone class.
For any $E \in \mathcal{A}, \mathcal{A} \in \mathcal{E}_E$ and $\mathcal{E}_E$ is a monotone class, so $\mathcal{C} \subset \mathcal{E}_E$. That is, for all $E \in \mathcal{A}, F \in \mathcal{C}$, we have

$$F \in \mathcal{E}_E \iff E \in \mathcal{E}_F$$

Hencce, for all $F \in \mathcal{C}, \mathcal{A} \subseteq \mathcal{E}_F$. Thus $\mathcal{C} \subseteq \mathcal{E}_F$ for all $F \in \mathcal{C}$. Therefore for all $E, F \in \mathcal{C}$,

$$(\star) \quad E \setminus F, F \setminus E, E \cap F \in \mathcal{C}$$

Since $X \in \mathcal{A} \subseteq \mathcal{C}$, $(\star)$ ensures that $\mathcal{C}$ is closed under complements and finite unions. Finally, note that, for any $\{E_i\}_{i=1}^{\infty} \subseteq \mathcal{C}$, since $\cup_{i=1}^{n} E_i \in \mathcal{C}$ for all $n \in \mathbb{N}$ and $\mathcal{C}$ is closed under countable increasing unions, we have $\cup_{i=1}^{\infty} E_i = \cup_{i=1}^{\infty}(\cup_{i=1}^{n} E_i) \in \mathscr{D}$.
Thus $\mathcal{C}$ is a σ-algebra.

■

Using the monotne class theorem, we can now prove the following result on uniqueness of measures.

Theorem 0.33. *Suppose that*

- $\mathcal{A}$ *is an algebra on a nonempty set* X
- μ *and* ν *are measures on* $\mathscr{M}(\mathcal{A})$
- $\exists\{A_i\}_{i=1}^{\infty} \subseteq \mathcal{A}$, $X = \cup_{i=1}^{\infty} A_i, \mu(A_i) < \infty$ *for all* i *(i.e.* μ *is* σ*-finite).*

Then if $\mu(A) = \nu(A)$ *for all* $A \in \mathcal{A}$*, we have* $\mu(E) = \nu(E)$ *for all* $E \in \mathscr{M}(\mathcal{A})$*.*

Proof. **Case 1:**

Suppose $\mu(X) < \infty$ (so $\nu(X) < \infty$.
Consider $\mathcal{E} = \{\mathcal{A} \in \mathscr{A}(\mathcal{A}) : \mu(A) = \nu(A)\}$. Since $\mathcal{A} \subseteq \mathcal{E}$, it suffices to show taht $\mathcal{E}$ is a monotone class to conclude that $\mathscr{M}(\mathcal{A}) \subseteq \mathcal{E}$.
If $B_1 \subseteq B_2 \subseteq \cdots$ is a countable increasing sequence in $\mathcal{C}$, by continuity from below, we have

$$\mu\left(\bigcup_{n=1}^{\infty} B_n\right) = \lim_{n\to\infty} \mu(B_n) = \lim_{n\to\infty} \nu(B_n) = \nu\left(\bigcup_{n=1}^{\infty} B_n\right)$$

Thus $\cup_{n=1}^{\infty} B_n \in \mathcal{E}$. If $B_1 \supseteq B_2 \supseteq \cdots$ is a countable decreasing sequence in $\mathcal{E}$, by continuity from above, since $\mu(X) < \infty$ (and hence $\nu(X) < \infty$),

$$\mu\left(\bigcap_{n=1}^{\infty} B_n\right) = \lim_{n\to\infty} \mu(B_n) = \lim_{n\to\infty} \nu(B_n) = \nu\left(\bigcap_{n=1}^{\infty} B_n\right)$$

Thus $\cap_{n=1}^{\infty} B_n \in \mathcal{E}$, so $\mathcal{E}$ is a monotone class.

Case 2:

$\mu(X) = \infty$.
We may assume without loss of generality that $\mathcal{A}\rangle_{i=1}^{\infty}$ are disjoint (other than $B_1 = A_1, B_2 = A_2 \setminus A_1$), define

$$\mu_i(E) = \mu(A_i \cap E)$$
$$\nu_i(E) = \nu(A_i \cap E)$$

Then $\mu_i(X) < \infty$ and $\nu_i(X) < \infty$ and μ_i, ν_i are finite measures on $\mathscr{M}(\mathcal{A})$.
Fact: Given a measure space $(X, \mathcal{M}, \mu)$, then for any $A \in \mathcal{M}$, the funciton

$$\mu_A(E) = \mu(E \cap A)$$

is a measure on $(X, \mathcal{M})$.
Note that, for all $A \in \mathcal{A}$,

$$\mu_i(A) = \mu(A_i \cap A) = \nu(A_i \cap A) = \nu_i(A)$$

By Case 1, $\mu_i(E) = \nu_i(E)$ for all $E \in \mathscr{M}(\mathcal{A})$. Thus, for all $E \in \mathscr{M}(\mathcal{A})$,

$$\mu(E) = \mu\left(\bigcup_{i=1}^{\infty}(E \cap A_i)\right) = \sum_i \mu(E \cap A_i)$$
$$= \sum_i \mu_i(E) = \sum_i \nu_i(E) = \nu(E)$$

■

Remark. (John's note) Recall at the beginning, one of our assumptions was that there exists a countable cover $\{A_i\}$ of X, such that $\mu(A_i) < \infty$ for each i. Measures satisfying this are called σ-finite. We can weaken this hypothesis:

Definition 0.28. We say that a measure μ on a σ-algebra $\mathcal{M}$ is s-finite if there is a sequence $\{\mu_i\}_{i=1}^{\infty}$ of finite measures on $\mathcal{M}$ such that

$$\mu = \sum_i \mu_i$$

Case 2 of our proof ends with showing that being σ-finite implies being s-finite. But, the last line of the proof still goes through if μ, ν are s-finite, and $\{\nu_i\}, \{\mu_i\}$ are finite measures such that $\mu = \sum_i \mu_i$ and $\nu = \sum_i \nu_i$.
Remark. A measure being s-finite is strictly weaker than being σ-finite.

Example 0.17. We can cook up examples of measures which are s-finite but not σ-finite by doing shenanigans like this (thank you Wikipedia for this example): Let $X = \{a\}$, and let $\mathcal{M} = \{\mathcal{A}, \varnothing\}$. Let ν_n be the counting measure on $\mathcal{M}$ for all n, and let

$$\mu = \sum_n \nu_n$$

Then μ is s-finite, because it is the sum of countably many finite measures. However, μ is not σ-finite, because $\mu(\{a\}) = \infty$.

Lecture 16

We can now return to our main goal, which was, given $(X, \mathcal{M}, \mu), (Y, \mathcal{N}, \nu)$ measure spaces, to prove the existence of a unique measure $\mu \times \nu$ on $(X \times Y, \mathcal{M} \otimes \mathcal{N})$ so that

$$\mu \otimes \nu(A \times B) = \mu(A)\nu(B)$$

for all rectangles $A \times B$.

Definition 0.29. For any $E \in \mathcal{M} \otimes \mathcal{N}$, for a fixed $x \in X$ the <u>x-section</u> is $E_x \overset{\text{def}}{=} \{y : (x, y) \in E\}$, and for a fixed $y \in Y$ the <u>y-section</u> is $E^y \overset{\text{def}}{=} \{x : (x, y) \in E\}$. Note that:

(a) If $E = A \times B$, $A \in \mathcal{M}, B \in \mathcal{N}$,

$$E_x = \begin{cases} B & \text{if } x \in A \\ \varnothing & x \notin A \end{cases}, \; E^y = \begin{cases} A & \text{if } y \in B \\ \varnothing & y \notin B \end{cases}$$

(b)

$$\begin{aligned} \left(\bigcup_{i=1}^{\infty} E_i\right)_x &= \{y : (x, y) \in \bigcup_{i=1}^{\infty} E_i\} \\ &= \bigcup_{i=1}^{\infty} \{y : (x, y) \in E_i\} \\ &= \bigcup_{i=1}^{\infty} (E_i)_x \end{aligned}$$

(c)

$$(E^c)_x = \{y : (x, y) \in E^c\} = \{y : (x, y) \in E\}^c = (E_x)^c$$

(d)

$$\nu(E_x) = \int_Y 1_{\underbrace{E_x}_{E_x \subseteq Y}}(y)\, d\nu(y) = \int_Y 1_{\underbrace{E}_{E \subseteq X \times Y}}\, d\nu(y)$$

Proposition 16. *If $E \in \mathcal{M} \otimes \mathcal{N}$, then $E_x \in \mathcal{N}$ and $E^y \in \mathcal{M}$ for all $x \in X, y \in Y$.*

Proof. Let $\mathcal{R} = \{E \in \mathcal{M} \otimes \mathcal{N} : E_x \in \mathcal{N}, E^c \in \mathcal{M} \forall x, y\}$.
Our goal is to show $\mathcal{M} \otimes \mathcal{N} \subseteq \mathcal{R}$.
Note that by (a), all rectangles belong to $\mathcal{R}$, so it suffices to show $\mathcal{R}$ is a σ-algebra.
If $\{E_i\}_{i=1}^{\infty} \subseteq \mathcal{R}$, then (b) ensures

$$\left(\bigcup_{i=1}^{\infty} E_i\right)_x = \bigcup_{i=1}^{\infty} (E_i)_x \in \mathcal{N}, \; \left(\bigcup_{i=1}^{\infty} E_i\right)^y \in \mathcal{M},$$

so $\cup_{i=1}^{\infty} E_i \in \mathcal{R}$, and $\mathcal{R}$ is closed under countable unions.
Finally, $E \in \mathcal{R}$, then

$$(E^c)_x = (E_x)^c \in \mathcal{N} \text{ and } (E^c)^c \in \mathcal{N}$$

So $E^c \in \mathcal{R}$ and $\mathcal{R}$ is closed under complements. ∎

Theorem 0.34. *Consider σ-finite measurable spaces $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$.*
For any $E \in \mathcal{M} \otimes \mathcal{N}$,

(i) The functions $x \mapsto \nu(E_x)$ and $y \mapsto \mu(E^y)$ are $(\mathcal{M}, \mathcal{B}_{\mathbb{R}})$-measurable and $(\mathcal{N}, \mathcal{B}_{\mathbb{R}})$-measurable.

(ii)

$$\int_X \nu(E_x)\, d\mu(x) = \int_Y \mu(E^y)\, d\nu(y)$$

Spoiler: This is equivalent to showing that

$$(\star) = \int_X \left(\int_Y 1_E(x,y)\, d\nu(y)\right) d\mu(x) = \int_Y \left(\int_X 1_E(x,y)\, d\mu(x)\right) d\nu(y)$$

This will ultimately allow us to define the product measure by

$$\mu \otimes \nu(E) \overset{\text{def}}{=} \iint 1_E d\nu \otimes \nu = (\star)$$

Proof. **Case 1:**

Suppose $\mu(X) < \infty, \nu(Y) < \infty$.
Let

$$\mathcal{A} = \underbrace{\{\coprod_{i=1}^{n} E_i : \{E_i\}_{i=1}^{n} \text{ are disjoint rectangles for all } n \in \mathbb{N}\}}_{\text{“all finite disjoint unions of rectangles”}}$$

Fact from homework 8: $\mathcal{A}$ is an algebra.
Note that

$$\mathcal{M} \otimes \mathcal{N} = \mathscr{M}(\{E : E \text{ is a rectangle }\}) \stackrel{?}{=} \mathscr{M}(\mathcal{A})$$

- Since all rectangles belong to $\mathcal{A}$, we have "$\subseteq$"
- Since $\mathcal{A} \subseteq \mathcal{M} \otimes \mathcal{N}$ and $\mathcal{M} \otimes \mathcal{N}$ is a σ-algebra, $\mathscr{M}(\mathcal{A} \subseteq \mathcal{M} \otimes \mathcal{N}$.
- Thus, "$=$" holds.

Let $\mathcal{C}$ be the set of $E \in \mathcal{M} \otimes \mathcal{N}$ which satisfy the hypotheses of this theorem. Our goal is to show $\mathcal{M} \otimes \mathcal{N} \subseteq \mathcal{C}$. We will accomplish this by showing $\mathcal{A} \subseteq \mathcal{C}$ and $\mathcal{C}$ is a monotone class. Then, the monotone class theorem will imply

$$\mathcal{M} \otimes \mathcal{N} = \mathscr{M}(\mathcal{A}) = \mathcal{C}(\mathcal{A}) \subseteq \mathcal{C}$$

First, we will show $\mathcal{A} \subseteq \mathcal{C}$, beginning by showing all rectangles belong to $\mathcal{C}$.
If $E = A \times B$ is a rectangle, then

$$\nu(E_x) = 1_A(x)\nu(B)$$
$$\mu(E^y) = 1_B(y)\mu(A)$$

These are clearly measurable, so the hypotheses of the theorem hold. Part (ii) follows because

$$\int_X \nu(E_x)\, d\mu(x) = \nu(B)\mu(A) = \int_Y \mu(E^y)\, d\nu(y)$$

Thus, $E \in \mathcal{C}$.
Now, suppose $E \in \mathcal{A}$, such that $E = \coprod_{i=1}^{n} E_i$, where $E_i = A_i \times B_i$. Then by part (b) of our earlier remark,

$$E_x = \coprod_{i=1}^{n} (E_i)_x \implies \nu(E_x) = \sum_{i=1}^{n} \nu((E_i)_x) = \sum_{i=1}^{n} 1_{A_i(x)}\nu(B_i)$$

Similarly, $\mu(E^y) = \sum_{i=1}^n 1_{B_i(y)}\mu(A_i)$. Thus, E clearly satisfies the hypotheses of this theorem. Part (ii) follows because

$$\int_X \nu(E_x)\, d\mu(x) = \sum_{i=1}^n \mu(A_i)\nu(B_i) = \int_Y \mu(E^y)\, d\nu(y)$$

That is, $E \in \mathcal{C}$.
Now, we will show that $\mathcal{C}$ is a monotone class.
Let $\{E_n\}_{n=1}^\infty \subseteq \mathcal{C}$, with $E_1 \subseteq E_2 \subseteq \cdots$, and define $E = \cup_{n=1}^\infty E_n$. Then $E_1^y \subseteq E_2^y \subseteq \cdots$ and

$$f_n(y) \stackrel{\text{def}}{=} \mu(E_n^y)$$
$$g_n(x) = \nu((E_n)_x)$$

are sequences of measurable functions that (by continuity from below) satisfy

$$f_n(y) \uparrow \mu(E^y)$$
$$g_n(x) \uparrow \nu(E_x)$$

pointwise for all $x \in X, y \in Y$. So, E satisfies hypothesis (i) of the theorem. Furthermore, by the Monotone Convergence Theorem,

$$\begin{aligned}\int_Y \mu(E^y)\, d\nu(y) &= \lim_{n\to\infty} \int_Y \mu(E_n^y)\, d\nu(y) \\ &= \lim_{n\to\infty} \int_X \nu((E_n)_x)\, d\mu(x) \\ &= \int_X \nu(E_x)\, d\mu(x),\end{aligned}$$

so E satisfies part (ii). So, $E \in \mathcal{C}$.
Now, let $\{E_n\}_{n=1}^\infty \subseteq \mathcal{C}$, with $E_1 \supseteq E_2 \supseteq \cdots$, and define $E = \cap_{n=1}^\infty E_n$. Then, if we define $f_n(y), g_n(x)$ as before, we see that these satisfy $f_n(y) \downarrow \mu(E^y)$ and $g_n(x) \downarrow \nu(E_x)$, where we are able to use continuity from above, since μ and ν are finite measures. So E satisfies (i). Furthermore,

$$f_1(y) = \mu(E_1^y) \leq \mu(X) < \infty$$
$$g_1(x) = \nu((E_1)_x) \leq \nu(Y) < \infty$$

Thus, the constant functions $\varphi(y) \equiv \mu(X$ and $\psi(x) \equiv \nu(Y)$ are dominating functions for our sequences.
Thus, we may use the dominated convergence theorem to finish the above calculation and see that E satisfies (ii), so indeed $E \in \mathcal{C}$.
Therefore, $\mathcal{C}$ is a monotone class.
Next time: case 2 of σ-finite.

■

Remark. (John's note)
I believe that in the proof to come, it will once again be the case that we can replace σ-finite with s-finite. Maybe I'm wrong tho idk.

Lecture 17

Case 2:

Suppose μ, ν are σ-finite, that is there exist sequences $\{A_i\}_{i=1}^{\infty} \in \mathcal{M}^{\mathbb{N}}, \{B_i\}_{i=1}^{\infty} \in \mathcal{N}^{\mathbb{N}}$, such that $\mu(A_i), \mu(B_i) < \infty$ for all i, and $\cup_i A_i = X, \cup_i B_i = Y$.
Without loss of generality, we may assume A_i, B_i increasing (let $A_i' = \cup_{j=1}^{i} A_i$. Then $\{A_i'\}_{i=1}^{\infty}$ is an increasing sequence of finite sets whose union is X, and similarly for B_i).
Because they are increasing, $\cup_{i=1}^{\infty}(A_i \times B_i) = X \times Y$.
For all i, define

$$\begin{aligned} \mu_i(A) &\stackrel{\text{def}}{=} \mu(A \cap A_i) \\ \nu_i(B) &\stackrel{\text{def}}{=} \nu(B \cap B_i) \end{aligned}$$

For all i, μ_i, ν_i are finite measures on $\mathcal{M}, \mathcal{N}$, respectively.
Note that

$$\begin{aligned} &\int 1_A \, d\mu_i = \int 1_A 1_{A_i} \, d\mu && \forall A \in \mathcal{M} \\ &\quad\Downarrow \\ &\int \varphi \, d\mu_i = \int \varphi 1_{A_i} \, d\mu && \forall \varphi \text{ simple} \\ &\quad\Downarrow{\scriptstyle MCT} \\ &\int f \, d\mu_i = \int f 1_{A_i} \, d\mu && \forall f \text{ nonnegative, measurable} \end{aligned}$$

Case 1 ensures that for any $E \in \mathcal{M} \otimes \mathcal{N}$,

(i) The functions $x \mapsto \nu(E_x)$ and $y \mapsto \mu(E^y)$ are $(\mathcal{M}, \mathcal{B}_{\mathbb{R}})$-measurable and $(\mathcal{N}, \mathcal{B}_{\mathbb{R}})$-measurable.

(ii) $\int_X \nu_i(E_x)\, d\mu_i(x) = \int_Y \mu_i(E^y) d\,\nu_i(y)$

By definition of μ_i, ν_i and continuity from below,

$$\nu_i(E_x) \uparrow \nu(E_x),\ \mu_i(E^y) \uparrow \mu(E^y)$$

Finally,

$$\begin{aligned}
\int_X \nu(E_x)\, d\mu(x) &= \lim_{i\to\infty} \int_X \nu_i(E_x) 1_{A_i}(x)\, d\mu(x) \\
&= \lim_{i\to\infty} \int_X \nu_i(E_x)\, d\mu_i(x) \\
&= \lim_{i\to\infty} \int_Y \mu_i(E^y)\, d\nu_i(y) \\
&= \lim_{i\to\infty} \int_Y \mu_i(E^y) 1_{B_i}(y)\, d\nu(y) \\
&= \int_Y \mu(E^y)\, d\nu(y)
\end{aligned}$$

where the first and last equalities are by MCT. So, we are done. ■

Remark. (John's note) As before, note that we may weaken σ-finiteness to s-finiteness: if we have that $\mu = \sum_{i=1}^{\infty} \mu_i$ for a sequence $\{\mu_i\}_{i=1}^{\infty}$ of finite measures, then the proof above goes thru without having to use σ-finiteness, because the only thing we use σ-finiteness for is constructing the μ_i.

Theorem 0.35. *Consider σ-finite (John: or s-finite!) measure spaces $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$.*
Define $\mu \otimes \nu : \mathcal{M} \otimes \mathcal{N} \to [0, +\infty]$ by

$$\mu \otimes \nu(E) = \int_X \nu(E_x)\, d\mu(x) = \int_Y \mu(E^y)\, d\nu(y)$$

Then

(i) $\mu \otimes \nu$ is a σ-finite measfure on $\mathcal{M} \otimes \mathcal{N}$

(ii) $\mu \otimes \nu$ is the unique measure on $\mathcal{M} \otimes \mathcal{N}$ satisfying

$$\mu \otimes \nu(A \times B) = \mu(A)\nu(B)\, \forall A \in \mathcal{M}, B \in \mathcal{N}$$

Proof. We will first show (i). It is clear that $\mu \otimes \nu(\varnothing) = 0$. Furthermore, if $\{E_i\}_{i=1}^\infty \in (\mathcal{M} \otimes \mathcal{N})^\mathbb{N}$ is a sequence of disjoint sets, then $\{(E_i)_x\}_{i=1}^\infty \in \mathcal{N}^\mathbb{N}$ are disjoint and

$$\begin{aligned}\mu \otimes \nu(\cup_{i=1}^\infty E_i) &= \int_X \nu(\cup_{i=1}^\infty (E_i)_x)\, d\mu(x) = \sum_{i=1}^\infty \int_X \nu((E_i)_x)\, d\mu(x) \\ &= \sum_{i=1}^\infty \mu \otimes \nu((E_i)_x)\end{aligned}$$

Thus $\mu \otimes \nu$ is a measure. To see σ-finiteness, let $\{A_i\}_{i=1}^\infty \in \mathcal{M}^\mathbb{N}, \{B_i\}_{i=1}^\infty \in \mathcal{N}^\mathbb{N}$, with $\cup_i A_i = X, \cup_i B_i = Y, \mu(A_i), \mu(B_i) < \infty$ for all i. As before, we may assume without loss of generality that A_i, B_i are increasing sequences, so $\cup_{i=1}^\infty (A_i \times B_i) = X \times Y$.
By definition, $\mu \otimes \nu(A_i \times B_i) = \mu(A_i)\nu(B_i) < \infty$. Thus $\mu \otimes \nu$ is σ-finite.
In order to prove uniqueness, let $\mathcal{A}$ be the algebra of finite disjoint unions of rectangles. Last time, we showed $\mathscr{M}(\mathcal{A}) = \mathcal{M} \otimes \mathcal{N}$.
Suppose σ is a measure on $\mathcal{M} \otimes \mathcal{N}$ satisfying

$$\sigma(A \times B) = \mu(A)\nu(B)\ \forall A \in \mathcal{M}, B \in \mathcal{N}$$

Note that, for any $E \in \mathcal{A}$, $E = \coprod_{i=1}^n (A_i \times B_i)$,

$$\sigma(E) = \sum_{i=1}^n \sigma(A_i \times B_i) = \sum_{i=1}^n \mu \otimes \nu(A_i \times B_i) = \mu \otimes \nu(E)$$

Since $\mu \otimes \nu$ satisfies the stronger σ-finite condition, we conclude

$$\mu \otimes \nu(E) = \sigma(E)\ \forall E \in \mathcal{M} \otimes \mathcal{N}$$

■

The following can be considered the 9th major theorem of the course:

Theorem 0.36. *(Fubini/Tonelli) Consider σ-finite measure spaces $(X, \mathcal{M}, \mu), (Y, \mathcal{N}, \nu)$.*

(i) Tonelli:

Given $f : X \times Y \to [0, +\infty]$ which is $(\mathcal{M} \otimes \mathcal{N}, \mathcal{B}_{\overline{\mathbb{R}}}$ measurable, then $x \mapsto \int f(x,y)\, d\nu(y)$ and $y \mapsto \int f(x,y)\, d\mu(x)$ are $(\mathcal{M}, \mathcal{B}_{\overline{\mathbb{R}}})$ and $(\mathcal{N}, \mathcal{B}_{\overline{\mathbb{R}}})$ measurable, respectively

$$\begin{aligned}\int_{X \times Y} f(x,y)\, d\mu \otimes \nu(x,y) &= \int_X \left(\int_Y f(x,y)\, d\nu(y) \right) d\mu(x) \\ &= \int_Y \left(\int_X f(x,y)\, d\mu(y) \right) d\nu(x)\end{aligned}$$

(ii) *Fubini: Given $f \in L^1(\mu \otimes \nu)$, then $y \mapsto f(x,y) \in L^1(\nu)$ for μ-almost all x, and $x \mapsto f(x,y) \in L^1(\mu)$ for ν-almost all y, and the above holds.*

Proof. We will start with Tonelli. For any $E \in \mathcal{M} \otimes \mathcal{N}$, we have already shown that $x \mapsto \nu(E_x)$ and $y \mapsto \mu(E^y)$ are $(\mathcal{M}, \mathcal{B}_{\mathbb{R}})$-measurable and $(\mathcal{N}, \mathcal{B}_{\mathbb{R}})$-measurable. Now,

$$\begin{aligned}\int_{X\times Y} 1_E \, d\mu \otimes \nu &= \int_X \left(\int_Y 1_E(x,y)\, d\nu(y) \right) d\mu(x) \\ &= \int_X \nu(E_x)\, d\mu(x) \\ &= \int_Y \mu(E^y)\, d\nu(y) \\ &= \int_Y \left(\int_X 1_E(x,y)\, d\mu(x) \right) d\nu(y)\end{aligned}$$

Thus, the conclusion of the theorem holds for all $f = 1_E$ indicator functions.
By linearity of the integral, the theorem holds for simple functions. For general non-negative measurable f, take a sequence of simple functions $\varphi_n \uparrow f$ pointwise. Then, by the monotone convergence theorem, for all $x \in X$,

$$\int \varphi_n(x,y), d\nu(y) \uparrow \int f(x,y)\, d\nu(y)$$

so $x \mapsto \int f(x,y)\, d\nu(y)$ is measurable and

$$\begin{aligned}\int f \, d\mu \otimes \nu = \lim_{n\to\infty} \int \varphi_n \, d\mu \otimes \nu &= \lim_{n\to\infty} \iint \varphi_n \, d\nu \, d\mu \\ &= \int \lim_{n\to\infty} \int \varphi_n \, d\nu \, d\mu = \iint f \, d\nu \, d\mu\end{aligned}$$

and similarly with the roles of μ, ν interchanged. It remains to show Fubini. Since $f \in L^1(\mu \otimes \nu)$, by Tonelli

$$+\infty > \int |f| \, d\mu \otimes \nu = \int \left(\int |f| \, d\nu \right) d\mu = \iint |f| \, d\mu \, d\nu$$

so $x \mapsto \int |f(x,y)|\, d\nu(y) < \infty$ μ-almost everywhere, and $y \mapsto \int |f(x,y)|\, d\mu(x) < \infty$ ν-almost everywhere.

Furthermore,

$$-\int |f(x,y)|\,d\nu(y) \le \int f(x,y)\,d\nu(y) \le \int |f(x,y)|\,d\nu(y)$$

$$\iff$$

$$\left|\int f(x,y)\,d\nu(y)\right| \le \int |f(x,y)|\,d\nu(y)$$

Thus

$$\int\left|\int f(x,y)\,d\nu(y)\right|d\mu(x) \le \iint |f(x,y)\,d\nu(y)\,d\mu(x)$$

so $x \mapsto \int f(x,y)\,d\nu(y) \in L^1(\mu)$ and similarly with μ,ν interchanged. Finally, if we apply Tonelli to the positive and negative parts of f, we are done.

■

Application of Tonelli: "Layer-cake representation of integral"

Proposition 17. *Consider a σ-finite measure $(X,\mathcal{M},\mu)$ and $f : X \to [0,+\infty]$ measurable. Let ν be a measure on $\mathcal{B}_{\mathbb{R}}$ such that*

$$\varnothing(t) \overset{\text{def}}{=} \nu([0,t)) < \infty\ \forall t > 0$$

Then

(i) $\varnothing \circ f$ is measurable.

(ii) $t \mapsto \mu(\{x : f(x) > t\})$ is measurable.

(iii) $\int_X \varnothing(f(x))\,d\mu(x) = \int_{[0,+\infty)} \mu(\{x : f(x) > t\})\,d\nu$

Remark. (a) Choose $\nu(A) = p\int_{A\cap[0,+\infty)} s^{p-1}\,ds$ for $p > 0$,

$$\int_X (f(x))^p\,d\mu(x) = p\int_{[0,+\infty)} t^{p-1}\mu(\{x : f(x) > t\})\,dt$$

(b) In particular, for $p = 1$,

$$\int f(x)\,d\mu(x) = \int_{[0,+\infty)} \mu(\{x : f(x) > t\})\,dt$$

Proof. Since $(x,t) \mapsto f(x)$, $(x,t) \mapsto t$ are $\mathcal{M}\otimes\mathcal{B}_{\mathbb{R}}$ measurable,

$$\{(x,t) : f(x) > t\} \in \mathcal{M}\otimes\mathcal{B}_{\mathbb{R}}$$

so Tonelli ensure (ii) and (i) follows by monotonicity of $\varnothing$.
Finally,

$$\begin{aligned}
\int_{[0,+\infty)} \mu(\{x : f(x) > t\})\, d\nu(t) &= \int_{[0,+\infty)} \int_X 1_{\{(x,t):f(x)>t\}}(x,t)\, d\mu(x)\, d\nu(t) \\
&= \int_X \int_{[0,+\infty)} 1_{\{(x,t):f(x)>t\}}(x,t)\, d\nu(t)\, d\mu(x) \\
&= \int_X \nu(\{t : f(x) > t\} \cap [0,+\infty))\, d\mu(x) \\
&= \int_X \nu([0, f(x))\, d\mu(x) \\
&= \int_X \varnothing(f(x))\, d\mu(x)
\end{aligned}$$

■