# MATH 209 Lecture Notes

Chenjia Lin

## Introduction

In this course, we study important results from measure theory. We begin by motivating the definition of measures with the Banach-Tarski Paradox, slowly making our way up to the Radon-Nikodym Theorem.

## Contents

# 1 Measures, Cantor sets, and the Banach-Tarski Paradox

In this lecture, we informally introduce the notion of measure and discuss a potential issue that will inform how we establish our framework for measure theory.

Let $X$ be a set and $A$ some subset of $X$. Roughly speaking, the *measure* of $A$, denoted $\mu(A)$, describes the "size" of $A$ relative to $X$. Note that there are many ways to "measure" the size of $A$ and many different sizes of sets. For instance, measures can be either real or vector-valued, positive and/or negative, or even $\infty$ or $-\infty$ (but not both within the same framework in which we define measures). In this course, we will only consider real-valued measures that are sometimes negative and $\infty$ or $-\infty$.

In general, we require a measure $\mu$ to satisfy either of the following properties:

1. *Finitely additive:* If $A$ and $B$ are disjoint subsets of $X$, then $\mu(A \cup B) = \mu(A) + \mu(B)$.
2. *$\sigma$-additive:* If $\{A_n\}_{n=1}^{\infty}$ is a family of pairwise disjoint sets, then $\mu\left(\sum_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n)$.

**Example 1.1.** The following are some examples of measures:

- *Counting measure:* the number of elements in a subset $A \subseteq X$.
- *Discrete measure:* we define $\mu(\{x\})$ for each $x \in X$ and $\mu(A) := \sum_{x \in A} \mu(\{x\})$.
- *Lebesgue measure in* $\mathbb{R}$: given an interval $I \subseteq \mathbb{R}$, we define $\mu(I) := \mathrm{vol}(I)$. If $A = \bigcup_{n=1}^{\infty} I_n$ is a countable union of disjoint intervals, then $\mu(A) = \sum_{i=1}^{\infty} \mu(I_n)$.

In particular, the middle-third *Cantor set* $\mathcal{C}$ is an example of a set in $\mathbb{R}$ with measure zero. Most notably, $\mathcal{C}$ has the following properties:

1. $\mathcal{C}$ is closed and bounded—hence compact—
2. $\mathrm{int}(\mathcal{C}) = \varnothing$,
3. and $\mathcal{C}$ has no *isolated points*—points with some neighborhood $U$ such that $\mathcal{C} \cap U = \{x\}$.

More generally, we define a Cantor set to be one that is homeomorphic to $\mathcal{C}$, such as the set constructed by removing middle-fourths from $[0, 1]$ instead of middle-thirds. In fact, we have the following fact about Cantor sets:

**Theorem 1.2.** A set $C \subseteq \mathbb{R}$ is a Cantor set if and only if

1. $C$ is compact,
2. $\mathrm{int}(C) = \varnothing$,
3. and $C$ has no isolated points.

We can also construct a Cantor set in $\mathbb{R}^2$ by subdividing a square into 16 smaller squares and removing all regions except the squares' corners with each iteration. However, we cannot use Theorem 1.2 to characterize Cantor sets in $\mathbb{R}^2$. For instance, a circle in $\mathbb{R}^2$ (the 1-sphere) satisfies all three properties listed above but is not homeomorphic to the Cantor set constructed from squares.

In (roughly) defining the measure of a subset of $X$, we did not address the question of whether $\mu(A)$ exists for any arbitrary $A \subseteq X$. As we will see in later lectures, we will run into issues if we tried to define a measure for every subset of $X$. To resolve this, we will only define measures for certain subsets of $X$. This collection of subsets is known as a *$\sigma$-algebra* of $X$.

To see how defining measures for every subset can go wrong, consider the following result:

**Theorem 1.3** (Banach-Tarski Paradox). If $A, B \subseteq \mathbb{R}^3$ are bounded sets with nonempty interiors, then we have decompositions $A = A_1 \cup \cdots \cup A_n$ and $B = B_1 \cup \cdots \cup B_n$ such that $A_i$ is a shift and/or rotation of $B_i$ for $1 \leq i \leq n$.

To clarify, a decomposition of $A$ is a collection of pairwise disjoint subsets $A_1, \ldots, A_n$ whose union is equal to $A$. Note that shifts and rotations are specific types of *isometries*, and more generally, we say $A = A_1 \cup \cdots \cup A_n$ *decomposes into* $B = B_1 \cup \cdots \cup B_n$ if there exist isometries $\varphi_1, \ldots, \varphi_n$ such that $B_i = \varphi(A_i)$ for $1 \leq i \leq n$. We write $X \simeq Y$ if $X$ and $Y$ are isometric, and we write $X \cong Y$ if $X$ decomposes into $Y$. We can verify that $\cong$ also defines an equivalence relation.[1]

One popular way of phrasing the Banach-Tarski Paradox is that a sphere can be decomposed into finitely many pieces that can be reassembled into two spheres of the same size as the one we started with using only shifts and rotations. The point of the paradox is that volumes are not invariant even under shifts and rotations, hence we cannot define volumes for arbitrary sets. In other words, volume is not a fool-proof way to assign a measure to any set.

## 2 Proving the Banach-Tarski Paradox

We ended the last class by introducing the Banach-Tarski Paradox and its implications in measure theory.

**Theorem** (Banach-Tarski Paradox). If $A, B \subseteq \mathbb{R}^3$ are bounded sets with nonempty interiors, then we have decompositions $A = A_1 \cup \cdots \cup A_n$ and $B = B_1 \cup \cdots \cup B_n$ such that $A_i$ is a shift and/or rotation of $B_i$ for $1 \leq i \leq n$.

The following corollary follows from the Banach-Tarski Paradox.

**Corollary 2.1.** There does not exist a finitely additive measure defined for all $A \subseteq \mathbb{R}^3$ that is invariant under isometries and for which the ball has positive and finite measure.

In other words, there is no "reasonable" extension of volume to all subsets in $\mathbb{R}^3$. As such, one way to circumvent this issue is by considering a $\sigma$-algebra of $\mathbb{R}^3$. Note that the Banach-Tarski Paradox holds in $\mathbb{R}^n$ for $n \geq 3$, but fails in say $\mathbb{R}^2$. There is indeed a finitely additive measure defined for all subsets of $\mathbb{R}^2$—in fact, it is based on area—such that the ball has positive and finite measure. The difference between $\mathbb{R}^2$ and $\mathbb{R}^3$ lies in the group of isometries that are available in the two spaces, and we will see this manifest in our proof of the paradox.

**Definition 2.2.** A *group* is an ordered pair $(G, \cdot)$ consisting of a set $G$ and a binary operation $\cdot : G \times G \to G$ such that

1. $(a \cdot b) \cdot c = a \cdot (b \cdot c)$ for all $a, b, c \in G$,
2. there exists an *identity element* $1$ such that $a \cdot 1 = 1 \cdot a = a$ for all $a \in G$,
3. and for each $a \in G$, there exists $a^{-1} \in G$ such that $a \cdot a^{-1} = a^{-1} \cdot a = 1$.

We say that a group $(G, \cdot)$ is *free* if no *nontrivial cancellations* exist. A *trivial cancellation* occurs when an element $a$ is adjacent to its inverse $a^{-1}$ in some expression, such as the following:

$$a^{-1} \cdot \underbrace{b \cdot b^{-1}}_{\text{cancel}} \cdot a^{-1} = a^{-1} \cdot 1 \cdot a^{-1} = a^{-1} \cdot a^{-1}.$$

To clarify, if say $a^{-1} \cdot a^{-1}$ from above (or some other product of elements in $G$) were equal to the identity and $a^{-1} \neq 1$, then $a^{-1} \cdot a^{-1}$ would be a nontrivial cancellation, and we conclude that $G$ is not a free group.

[1] The notation here was adopted from Chris Wilson's notes for MATH 209.

**Example 2.3.** If $(G, \cdot)$ is a commutative group, then $G$ is not a free group because an expression like $a^{-1} \cdot b^{-1} \cdot a \cdot b$ is equal to 1 without trivial cancellations.

**Definition 2.4.** Let $x$ and $y$ belong to a group $(G, \cdot)$. The *subgroup generated by $x$ and $y$* is the set of all finite products of $x$, $y$, $x^{-1}$, and $y^{-1}$.

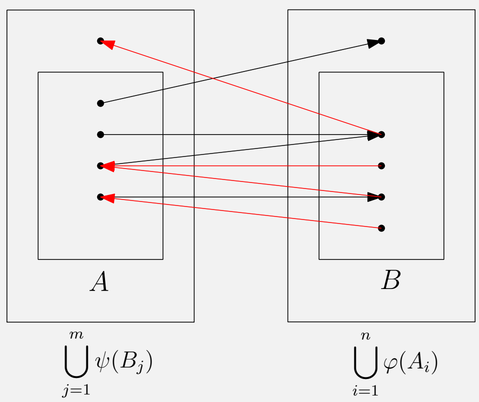
Note that if $(G, \cdot)$ is a noncommutative group, then the subgroup generated by $x, y \in G$ where $x \neq y^{-1}$ is a free group. The distinction between $\mathbb{R}^2$ and $\mathbb{R}^3$ is that rotations in $\mathbb{R}^2$ do not admit free groups and $\mathbb{R}^3$ does—by rotating about different sets of axes at some specific rational angles. We will use this fact for now in our proof of the paradox.

**Lemma 2.5.** Suppose $A = A_1 \sqcup \cdots \sqcup A_n$ and $B = B_1 \sqcup \cdots \sqcup B_m$ and there exist isometries $\varphi_i : A_i \to \varphi_i(A_i)$ and $\psi_j : B_j \to \psi_j(B_j)$ for $1 \leq i \leq n$ and $1 \leq j \leq m$ such that

$$B \subseteq \bigcup_{i=1}^{n} \varphi_i(A_i) \text{ and } A \subseteq \bigcup_{j=1}^{m} \psi_i(B_j).$$

Then $A \cong B$.

*Proof.* We want to construct a new decompositions of $A$ and $B$ so that $A \cong B$. Consider the graph with points in $A$ and $B$ as nodes and edges formed by connecting $x \in A'_i$ with $\varphi_i(x)$ (black edges) and $y \in B'_i$ with $\psi(y)$ (red edges).




The figure above gives a rough sketch of our construction but without the subsets given from each decomposition. Note that the images $\varphi_1(A_1), \ldots, \varphi_n(A_n)$ may intersect, and this is shown wherever two points in $A$ get mapped to the same point. From our original graph, we will perform the following:

1. If multiple edges of the same color converge onto the same point, then remove edges until one edge remains.

2. If $x \in A_i$ and $\varphi_i(x) \notin B$, then remove this corresponding edge.

We end up with a graph with nodes only inside $A$ or $B$. If we consider the paths formed by alternating black and red edges, notice that such paths are either infinite in one direction or cyclic with an even number of edges. For each cycle we remove (without loss of generality) all of the red edges, and for each infinite path we remove every other path (not the first edge of the sequence). We end up with a bijection between $A$ and $B$ that we can use to decompose $A$ and $B$ and therefore show $A \cong B$. $\square$

The following is an brief outline of our proof of the Banach-Tarski Paradox.

1. Show Banach-Tarski for the 2-sphere in $\mathbb{R}^3$.

2. Extend Banach-Tarski for spheres to balls.

3. Prove Banach-Tarski for bounded sets with nonempty interiors by applying the theorem for balls finitely many times and Lemma 2.5.

*Proof of the Banach-Tarski Paradox.* Let $\alpha$ and $\beta$ denote rotations of the sphere in $\mathbb{R}^3$ that generate a free group. Let $G$ be the group generated by $\alpha$ and $\beta$, and consider the partitioning of $G$ into subgroups $G_1$, $G_2$, $G_3$, $G_4$, and $G_5$ defined by

$$\begin{aligned}
&G_1 = \{1\}, \\
&G_2 := \{\text{words starting with } \alpha\}, \\
&G_3 := \{\text{words starting with } \beta\}, \\
&G_4 := \{\text{words starting with } \alpha^{-1}\}, \\
&G_5 := \{\text{words starting with } \beta^{-1}\}.
\end{aligned}$$

Now observe that $\alpha^{-1} \cdot G_2 = G_1 \cup G_2 \cup G_3 \cup G_5$ and $\beta^{-1} \cdot G_3 = G_1 \cup G_2 \cup G_3 \cup G_4$, so we have

$$G = (\alpha^{-1} \cdot G_2) \cup G_4 = (\beta^{-1} \cdot G_3) \cup G_5.$$

As an aside, we begin to see a Banach-Tarski-esque property appear in the partition of $G$ above. We will now consider how the groups above act on points on the sphere. If $S \subseteq \mathbb{R}^3$ denotes the 2-sphere and $x \in S$, then the *orbit of $x$ under $G$* is the image $\mathcal{O} := \{g \cdot x : g \in G\}$. Just as we $G = G_1 \cup \cdots \cup G_5$, we see that $\mathcal{O} = \mathcal{O}_1 \cup \cdots \cup \mathcal{O}_5$, where $\mathcal{O}_i = G_i \cdot x$ for $1 \leq i \leq 5$. And similarly, we have

$$\mathcal{O} = (\alpha^{-1} \cdot \mathcal{O}_2) \cup \mathcal{O}_4 = (\beta^{-1} \cdot \mathcal{O}_3) \cup \mathcal{O}_5.$$

As a word of caution, the fact that $G$ is a free group that can be decomposed into $G_1, \ldots, G_5$ does not guarantee that $\mathcal{O}$ is free—in other words, $\mathcal{O}_1, \ldots, \mathcal{O}_5$ is not necessarily a decomposition of $\mathcal{O}$. For instance, if $x$ also lies on an axis of rotation for some $g \in G$, then $x = g \cdot x$ and we see that $x$ admits a nontrivial expression. As such, $\mathcal{O}_1, \ldots, \mathcal{O}_5$ are not pairwise disjoint. Now for some terminology: we say that an orbit $\mathcal{O}$ is "good" if $g_1 x \neq g_2 x$ for all distinct $g_1, g_2 \in g$ and "bad" if there exist distinct $g_1, g_2 \in G$ such that $g_1 x = g_2 x$. We see that if $x \in S$ such that $\mathcal{O}$ is a good orbit, then $\mathcal{O}$ partitions nicely as described above.

Now let $N := \{x \in S : \mathcal{O}_x \text{ is a bad orbit}\}$, and observe that $N$ is countable; notice that $G$ describes countably many rotations and there are only two fixed points for each rotation. Hence, there are countably many $x \in S$ that are fixed points of some rotation in $G$, which corresponds to countably many points in $N$.

We shift our attention now to $S \setminus N$, which is the set of points on $S$ that produce good orbits. We will see that $S \setminus N$ partitions nicely into five disjoint subsets. Let $\mathcal{A} := \{\mathcal{O}_x : x \in S \setminus N\}$ denote the collection of all orbits $\mathcal{O}_x$ given by $G$ for each $x \in S \setminus N$, and let $A$ be a set that contains exactly one point from each orbit in $\mathcal{A}$. To clarify, $A$ is not equal to $S \setminus N$ because two points on $S \setminus N$ can lie on the same orbit. We define $A_i := G_i \cdot A$ for $1 \leq i \leq 5$, and note that $S \setminus N = A_1 \cup \cdots \cup A_5$. We have $A_1 \cup \cdots \cup A_5 \subseteq S \setminus N$ because the rotation of points on $S \setminus N$ will remain on $S$. To see that $S \setminus N \subseteq A_1 \cup \cdots \cup A_5$, let $x \in S$ and consider $y \in A \cap \mathcal{O}_x$. We know there exists $g \in G$ such that $y = g \cdot x$, hence $x = g^{-1} \cdot y$. Notice then that $g^{-1}$ lies in $G$ is belongs to some $G_i$. And so, we have $S \setminus N \subseteq A_1 \cup \cdots \cup A_5$ and $S \setminus N = A_1 \cup \cdots \cup A_5$. And like our decomposition of $G$, we have

$$S \setminus N = (\alpha^{-1} \cdot A_2) \cup A_4 = (\beta^{-1} \cdot A_3) \cup A_5.$$

Great! We now have $S \setminus N \cong$ (two copies of $S \setminus N$). We want to show $S \cong$ (two copies of $S$), and this will follow if we have $S \cong S \setminus N$.

To show $S \cong S \setminus N$, let us first consider a simpler fact $C \cong C \setminus \{x\}$ involving a circle $C$ and a point $x \in C$. To see that this is true, let $\rho$ be some irrational angle measure and $R := \{x, \rho x, \rho^2 x, \ldots\}$ the set of all rotations of $x$ by $\rho$. Now consider the complement $C \setminus R$ and rotation $\rho R$ of $R$ by $\rho$, and notice that $(C \setminus R) \cup (\rho \cdot R) = C \setminus \{x\}$. Since $(C \setminus R) \cup R = C$, we see that $C \cong C \setminus \{x\}$. To show that $S \cong S \setminus N$, consider a set of circles such that each circle contains exactly one point in $N$ and the planes that the circles lie on are all parallel. Let $M$ denote the set of all such circles, and pick an axis of rotation that is perpendicular to each of these planes. Now let $R := \{M, \rho M, \rho^2 M, \dots\}$ be the set of rotations of $M$ by some irrational angle measure $\rho$, and we have $(S \setminus R) \cup (\rho \cdot R) = S \setminus N$, just like in the case with circles. Therefore, we have $S \setminus N \cong S$.

To extend the paradox from spheres to balls, notice that $B \setminus \{0\}$ is equal to the union of concentric spheres. Applying the paradox to $B \setminus \{0\}$, we have $B \setminus \{0\} \cong$ (two copies of $B \setminus \{0\}$). To show that $B \cong$ (two copies of $B$), we will show $B \cong B \setminus \{0\}$. Let $C$ be a circle in $B$ that passes through the origin and observe that

$$B = C \cup (B \setminus C) \cong (C \setminus \{0\}) \cup (B \setminus C) = B \setminus \{0\}.$$

And so, we have $B \cong$ (two copies of $B$).

We want to finally extend the paradox to bounded sets $A, B \subseteq \mathbb{R}^3$ with nonempty interior. If $A$ and $B$ have nonempty interiors, then they each contain some closed ball. We can take the ball in $A$ and replicate it sufficiently many times so that it can be shifted to cover $B$. We do the same with the ball in $B$ to cover $A$, and by Lemma 2.5 we have $A \cong B$. □

As mentioned earlier in the lecture, the Banach-Tarski Paradox does not hold in $\mathbb{R}^2$ because area is a finitely additive measure defined for all subsets of $\mathbb{R}^2$ that is invariant under isometries (not proven in this course).

# 3 Measure spaces and Borel sets

A takeaway from the Banach-Tarski Paradox and our informal introduction to measures is that not all subsets of a space $X$ are measureable. Our resolution is to specify the exact sets that are measureable using a *$\sigma$-algebra*.

**Definition 3.1.** A *$\sigma$-algebra* on a set $X$ is a collection $\Sigma$ of subsets of $X$ such that

1. $\varnothing, X \in \Sigma$,
2. $\bigcup_{i=1}^{\infty} A_i$ and $\bigcap_{i=1}^{\infty} A_i$ belong to $\Sigma$ for $\{A_i\}_{i=1}^{\infty} \subseteq \Sigma$,
3. and $X \setminus A \in \Sigma$ if $A \in \Sigma$.

**Definition 3.2.** If $\Sigma$ is a $\sigma$-algebra on $X$, then a *measure* $\mu$ on $X$ is a function where

1. $\mu(\varnothing) = 0$
2. and $\mu\Big(\bigcup_{i=1}^{\infty} A_n\Big) = \sum_{i=1}^{\infty} \mu(A_n)$ for pairwise disjoint subsets $A_i \in \Sigma$.

**Remark.** We will assume that measures are $\sigma$-additive unless noted otherwise.

We say $X$ if a *measure space* if there exists a $\sigma$-algebra $\Sigma$ and measure $\mu$ on $X$, and we write $(X, \Sigma, \mu)$ to bundle these features together. In most contexts, a measure space $X$ may also have some additional structure, such as a topology describing which sets are open.

**Definition 3.3.** Let $(X, \Sigma, \mu)$ be a measure space and $\mathcal{U} \subseteq \Sigma$. The *$\sigma$-algebra generated by $\mathcal{U}$* is the smallest $\sigma$-algebra on $X$ containing $\mathcal{U}$.

We know that the $\sigma$-algebra generated by $\mathcal{U}$ exists because it is the intersection of all $\sigma$-algebras containing $\mathcal{U}$. We can easily verify that the intersection of $\sigma$-algebras is indeed a $\sigma$-algebra.

**Definition 3.4.** Let $X$ be a topological space with $\sigma$-algebra $\Sigma$ and measure $\mu$. A *Borel set* is a set in $\Sigma$ that is generated by open sets of $X$.

Note that if $\{U_\lambda\}_{\lambda\in\Lambda} \subseteq \Sigma$ is a collection of open sets, then the $\sigma$-algebra generated by $\{U_\lambda\}_{\lambda\in\Lambda}$ contains the complements $X \setminus U_\lambda$, which are closed subsets of $X$. It's actually pretty difficult to decide whether some arbitrary set in $X$ is Borel because we need to see how its generated.

To get a clearer picture of Borel sets, we will first adopt some notation. We denote an open set in $X$ by $G$ and a closed set by $F$. We say a set is $G_\delta$ if it is a countable intersection of open sets, and $F_\sigma$ if it is a countable union of closed sets.

**Example 3.5.** The set of rational numbers $\mathbb{Q}$ is an $F_\sigma$ set, as $\mathbb{Q}$ can be expressed as the countable union of singletons, which are closed. If $\mathbb{Q} = \bigcup_{n=1}^{\infty}\{q_n\}$, then observe that

$$\mathbb{R} \setminus \mathbb{Q} = \mathbb{R} \setminus \left(\bigcup_{n=1}^{\infty}\{q_n\}\right) = \bigcap_{n=1}^{\infty} \mathbb{R} \setminus \{q_n\},$$

which shows $\mathbb{R}\setminus\mathbb{Q}$ is a $G_\delta$ set.

Note that $\mathbb{Q}$ is not $G_\delta$; suppose instead that $\mathbb{Q} = \bigcap_{n=1}^{\infty} G_n$, where $G_n$ are open sets in $\mathbb{R}$. By assumption, we know that the intersection $\bigcap_{n=1}^{\infty} G_n$ is nonempty. If $G_1$ is open in $X$, then there exists an closed interval $I_1 \subseteq G_1$ so that $x_1 \notin I_1$. Now define closed intervals $I_n \subseteq G_n$ where $I_n \subseteq I_{n-1}$ and $x_n \notin I_n$ for $n \geq 1$. We end up with nested closed intervals $I_n \subseteq I_{n-1} \subseteq \cdots \subseteq I_2 \subseteq I_1$ in $\mathbb{R}$, and we have that $\bigcap_{n=1}^{\infty} I_n$ is nonempty. By construction, we have $\bigcap_{n=1}^{\infty} I_n \subseteq \bigcap_{n=1}^{\infty} G_n = \mathbb{Q}$. But notice that $\bigcap_{n=1}^{\infty} I_n$ also does not contain any rationals, so it must contain an element of $\mathbb{R}\setminus\mathbb{Q}$. However, we now arrive at a contradiction and we conclude that $\mathbb{Q}$ is not $G_\delta$.

**Example 3.6.** Observe that

$$(0,1] = \bigcap_{n=1}^{\infty}\left(0, 1+\frac{1}{n}\right) = \bigcup_{n=1}^{\infty}\left[\frac{1}{n}, 1\right],$$

which implies $(0,1]$ is both a $G_\delta$ and $F_\sigma$ set.

Not all Borel sets are $G_\delta$ or $F_\sigma$—in other words, strictly countable unions and intersections of closed and open sets—but we can add more variety to the types of sets we can consider. We say that a set if $F_{\sigma\delta}$ if it is the countable intersection of $F_\sigma$ sets, so it looks like

$$\bigcap_{i=1}^{\infty} F_{\sigma,i} = \bigcap_{i=1}^{\infty}\bigcup_{j=1}^{\infty} F_{i,j}$$

where $F_{i,j}$ is closed. Similarly, we say that a set is $G_{\delta\sigma}$ if it is the countable union of $G_\delta$ sets, which looks like

$$\bigcup_{i=1}^{\infty} G_{\delta,i} = \bigcup_{i=1}^{\infty}\bigcap_{j=1}^{\infty} G_{i,j}$$

where $G_{i,j}$ is open. Note that $F_{\sigma\sigma}$ and $G_{\delta\delta}$ are actually no different from $F_\sigma$ and $G_\delta$, respectively. We can inductively define sets that are $F_{\sigma\delta\sigma\delta}$ or $F_{\delta\sigma\delta\sigma}$, and we can organize these sets into the following diagram.

$$\begin{array}{cc}
G & F \\
G_\delta & F_\sigma \\
G_{\delta\sigma} & F_{\sigma\delta} \\
G_{\delta\sigma\delta} & F_{\sigma\delta\sigma} \\
\vdots & \vdots
\end{array}$$

We have the following inclusion relations:

$$
\begin{gathered}
G \subseteq G_\delta \subseteq G_{\delta\sigma} \subseteq G_{\delta\sigma\delta} \subseteq \cdots, \\
F \subseteq F_\sigma \subseteq F_{\sigma\delta} \subseteq F_{\sigma\delta\sigma} \subseteq \cdots, \\
F \subseteq G_\delta \subseteq F_{\sigma\delta} \subseteq G_{\delta\sigma\delta} \subseteq \cdots, \\
G \subseteq F_{\mathcal{S}} \subseteq G_{\delta\sigma} \subseteq F_{\sigma\delta\sigma} \subseteq \cdots.
\end{gathered}
$$

We can keep extending the diagram from above, but we mix up those types of sets even further. We say that a set is $\omega$ if it is the countable union or intersection of sets that are $G_{\delta\sigma\delta\cdots}$ and/or $F_{\sigma\delta\sigma\cdots}$. We continue this iterative process to define sets of types $\omega+1, \omega+2, \cdots, 2\omega, 3\omega, \cdots, \omega^2, \omega^3, \cdots$ and types beyond polynomials in $\omega$. The process actually terminates after uncountably many steps, and we end up generating a $\sigma$-algebra.

The takeaway from this series of constructions is that they are contained by the set of all Borel sets on $(X, \Sigma, \mu)$, which shows how crazy Borel sets can be. The notion of Borel sets can also be extended to functions.

**Definition 3.7.** Let $(X, \Sigma, \mu)$ be a measure space. We say a function $f : X \to \mathbb{R}$ is a *Borel function* if the preimage of an open set in $\mathbb{R}$ is Borel.

**Definition 3.8.** Let $(X, \Sigma, \mu)$ be a measure space. We say $f : X \to \mathbb{R}$ is $\Sigma$*-measurable* if $f^{-1}(G) \in \Sigma$ for all open $G \subseteq \mathbb{R}$. More generally, if $(Y, \Sigma', \mu)$ is another measure space, then we say $f : X \to Y$ is $\Sigma\Sigma'$*-measureable* if $f^{-1}(A) \in \Sigma$ for all $A \in \Sigma'$.

If we know that the preimage of every open set is Borel, then the preimage of every closed set is also Borel. If we know that the preimage of countably many sets is Borel, then the preimage of their union—which is the union of the preimages—is also Borel. Instead of looking at all sets in a space $X$, it suffices to consider the generating sets in the $\sigma$-algebra.

**Example 3.9.** The following are examples of Borel functions.

1. Continuous functions; the preimage of an open set is open, hence Borel.

2. Suppose we have a sequence $(f_n)$ of continuous functions $f_n : X \to \mathbb{R}$ that converges pointwise to $f$. We claim that $f$ is Borel. To see this, observe that

$$
\begin{aligned}
f^{-1}(G) &= \{x \in X : f(x) \in G\}, \\
&= \bigcup_{q \in \mathbb{Q}^+} \{x \in X : \text{there exists } N \text{ such that } \operatorname{dist}(f_n(x), \mathbb{R} \setminus G) > q \text{ for all } n \geq N\},
\end{aligned}
$$

where we define

$$\operatorname{dist}(f_n(x), \mathbb{R} \setminus G) = \inf_{y \in R \setminus G} |f_n(x) - y|.$$

Intuitively, the complicated set consists of $x \in X$ where $(f_n(x))$ is eventually at least a distance of $q$ away from the boundary of $G$, and notice that this construction excludes $x \in X$ where $(f_n(x))$ converges to a point on the boundary of $G$. Now let $H_q := \{y \in \mathbb{R} : \operatorname{dist}(y, R \setminus G) > q\}$; notice that $H_q$ is open, and we have

$$f^{-1}(G) = \bigcup_{q \in \mathbb{Q}^+} \bigcup_N \bigcap_{n \geq N} f_n^{-1}(H_q),$$

which shows $f^{-1}(G)$ is a $G_{\delta\sigma}$ set. In this context, we say $f$ is a *Baire-1* function.

3. Following the previous example, we say $f$ is a Baire-2 function if there is a sequence $\{f_1, f_2, \ldots\}$ of Baire-2 functions $f_n$ that converge to $f$. And just as we had sets of type $\omega$ and $\omega^3$, we can also define Baire-$\omega$ functions.

## 4 Baire Category Theorem and Banach-Mazur Games

In this lecture, we introduce the Baire Category Theorem as a corollary to some results about *Banach-Mazur Games*. First, consider the following definitions and examples.

**Definition 4.1.** Let $X$ be a topological space. A subset $A \subset X$ is *nowhere dense* if for all nonempty open sets $G \subseteq X$, there exists a nonempty open set $H \subseteq G$ such that $A \cap H = \varnothing$.

**Example 4.2.** The middle-third Cantor set $\mathcal{C}$ is nowhere dense. Recall that the Cantor set is constructed by removing middle-thirds (open intervals) from segments in each iteration, and notice that if $I$ is an open interval intersecting $\mathcal{C}$, then $I$ will contain one of the deleted middle-thirds.

Roughly speaking, a nowhere dense set is one that has many "holes"—articulated by the existence of an open subset $H \subseteq G$. As we saw above, the Cantor is full of holes.

**Definition 4.3.** A subset $A \subseteq X$ is of *first category* if $A = \bigcup_{n=1}^{\infty} A_n$ where $A_n$ are nowhere dense. We say $A$ is of *second category* if it is not of the first category. We say $A$ is *residual* if $X \setminus A$ is of the first category.

**Definition 4.4.** A set $X$ is a *Baire space* if every nonempty open set is of second category.

A first category set $A$ can be thought of as being very small and its complement $X \setminus A$ very large. A second category set can be neither very small nor very large—somewhere between medium and large.

**Theorem 4.5** (Baire Category Theorem)**.** If $(X, \rho)$ is a complete metric space, then $X$ is a Baire space.

To make sense of the Baire Category Theorem, recall that $\mathbb{Q} \subseteq \mathbb{R}$ is not a $G_\delta$ set. To realize this using the Baire Category Theorem, observe the following:

- $\mathbb{Q}$ is a first category set because it is the countable union of rational singletons, which are nowhere dense.
- If $A$ is a dense $G_\delta$ set, then $A = \bigcap_{n=1}^{\infty} A_n$, where $A_n$ are open and dense. We have $X \setminus A = \bigcup_{n=1}^{\infty} X \setminus A_n$. Notice that $X \setminus A_n$ is nowhere dense, so we see that $X \setminus A$ is of first category. Thus, $A$ is residual.

Now suppose that $\mathbb{Q}$ is a $G_\delta$ set. Since $\mathbb{Q}$ is dense and countable, we see that $\mathbb{Q}$ is both of first category and residual. Note that $\mathbb{R} = \mathbb{Q} \cup (\mathbb{R} \setminus \mathbb{Q})$, so we see that $\mathbb{R}$ is the union of two first category sets. Hence, $\mathbb{R}$ is of first category. But $\mathbb{R}$ is a second category set according to the Baire Category Theorem, so we arrive at a contradiction. In general, we have showed that a countable dense set cannot be $G_\delta$.

We will prove the Baire Category Theorem using a stronger result from analyzing Banach-Mazur Games. The following are the instructions for the game in $\mathbb{R}$:

1. Let $A \subseteq \mathbb{R}$ be some arbitrary subset.
2. Player 1 begins the game by picking some closed interval $I_1 \subseteq \mathbb{R}$, and Player 2 proceeds by constructing an closed interval $I_2 \subseteq I_2$.
3. Players 1 and 2 take turns and we end up with a sequence of nested closed sets $\cdots I_n \subseteq I_{n-1} \subseteq \cdots \subseteq I_1$.
4. By the Nested Interval Theorem in $\mathbb{R}$, $\bigcap_{n=1}^{\infty} I_n$ is nonempty. Player 1 wins if $\bigcap_{n=1}^{\infty} I_n$ intersects $A$ and Player 2 wins otherwise.

We might require the length of $I_n$ to be less than half of that of $I_{n-1}$. In that case, $\bigcap_{n=1}^{\infty} I_n$ is a single point. This additional detail is not important at the moment. As with any game, we would like to know if either player has a winning strategy. To get some intuition, consider this first example:

**Example 4.6.** Suppose $A$ is the middle-third Cantor set. In an attempt to win, Player 1 will construct $I_1$ so that it intersects $A$. But for Player 2 to win, it suffices for him/her to take $I_2$ to be one of the deleted intervals of the Cantor set because $I_3$ can never be constructed in a way that intersects $A$.

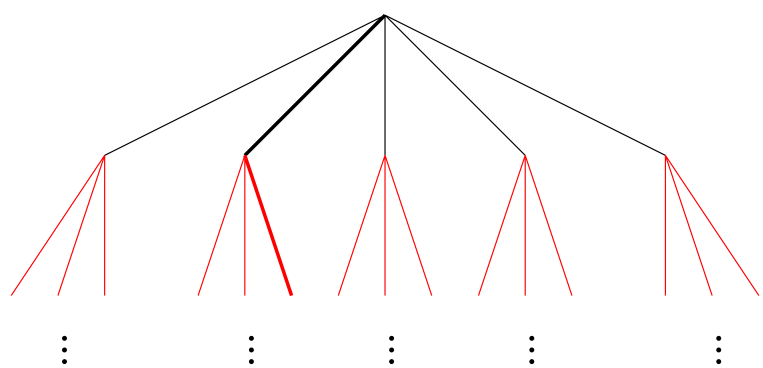
**Example 4.7.** Suppose $A = \mathbb{Q}$. Note that Player 2 will always win because $\mathbb{R} \setminus \mathbb{Q}$ is dense and he/she can construct intervals that eventually converge to an irrational number. Alternatively, we can adopt the same strategy in our proof of $\mathbb{Q}$ is not $G_\delta$ (Example 3.5).

Notice that Player 2's winning strategy in the first example seems to reflect the fact that the Cantor set is nowhere dense, as seen in Example 4.2. From the two examples above, we suspect the following is true:

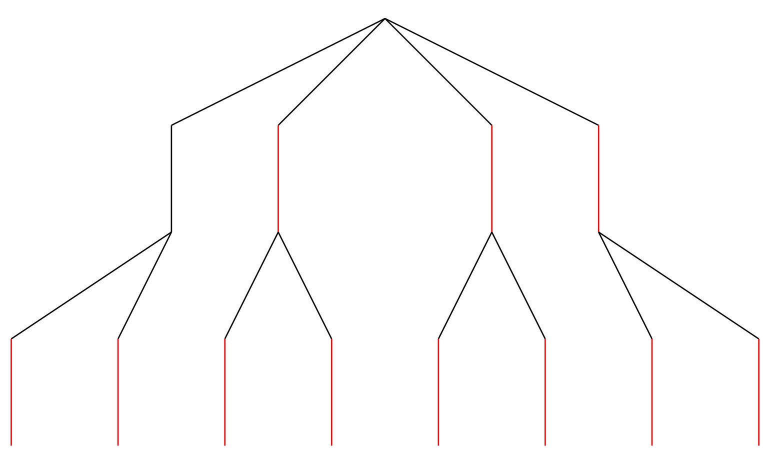
**Proposition 4.8.** Player 2 has a winning strategy if $A$ is a first category set.

*Proof.* Suppose $A$ is a first category set. Then there exist nowhere dense open sets $A_n$ for $n \geq 1$ such that $A = \bigcup_{n=1}^{\infty} A_n$. Because $A_n$ is nowhere dense, Player 2 can choose $I_{2n}$ to be disjoint from $A_n$. We have $\bigcap_{n=1}^{\infty} I_n$ must be disjoint from $\bigcup_{n=1}^{\infty} A_n$, so Player 2 wins. $\square$

In fact, the converse is also true! As with any two-player game, we can construct a tree that tracks all possible moves that can be made by the players.



Each edge represents an interval chosen by a player. In the diagram above, black edges represent Player 1's moves while red edges represent those by Player 2. Each branch (an alternating sequence of black and red edges) represents the moves in one complete game. If a player—say Player 2—has a winning strategy, then there is only one way to respond based on the previous move by the opponent.



**Proposition 4.9.** Player 2 has a winning strategy only if $A$ is a first category set.

*Proof.* For simplicity, we assume $A \subseteq \mathbb{R}$. If Player 2 has a winning strategy, then Player 2 chooses exactly one interval in response to each move by Player 1 (see figure from earlier). Deferring the technical details for now, we will construct a sequence of dense sets $B_1, B_2, \ldots$ such that

$$B_1 \supseteq B_2 \supseteq \cdots$$

and $B_n$ is a countable union of intervals $I_{2n}$ that respond to some choice of

$$I_{2n-1} \subseteq I_{2n-2} \subseteq \cdots \subseteq I_1.$$

We define $A_n := X \setminus B_n$, and we have a sequence $A_1, A_2, \ldots$ of nowhere dense sets with $A_1 \subseteq A_2 \subseteq \cdots$. If $x \in \bigcap_{n=1}^{\infty} B_n$, then there exists an infinite sequence of nested intervals $I_1 \supseteq I_2 \supseteq \cdots$ containing $x$. Since $I_1 \supseteq I_2 \supseteq \cdots$ is constructed with Player 2's winning strategy, we know $\left(\bigcap_{n=1}^{\infty} I_n\right) \cap A = \varnothing$ and $x \notin A$. We have that $x \in \bigcap_{n=1}^{\infty} B_n$ implies $x \notin A$, so $x \in A$ implies $x \in \bigcup_{n=1}^{\infty} A_n$ and $A \subseteq \bigcup_{i=1}^{\infty} A_n$. Note that $\bigcup_{i=1}^{\infty} A_n$ of first category because it is a union of nowhere dense sets, and it follows that $A$ is of first category as well.

Now for the technical details. We will construct $B_1$ inductively by appending intervals to its union and $B_n$ for $n \geq 2$ based on $B_{n-1}$. Let $\mathbb{Q} = \{q_1, q_2, \ldots\}$, and suppose $I$ is a choice of interval by Player 1 (on the first move) that contains $q_1$. We define the interval $J_1$ to be Player 2's response to $I$ using his/her strategy, and we define $J_m$ for $m \geq 2$ in the following manner:

1. If $q_{m+1} \in J_m$, then we will not append a new interval and instead repeat the algorithm for $q_{m+2}$.
2. If $q_{m+1} \notin J_m$, then let $I$ be a possible first move by Player 1 that contains $q_{m+1}$ and is disjoint from $J_1, J_2, \ldots, J_m$. Define $J_{m+1}$ to be Player 2's response to $I$.

We have a sequence $J_1, J_2, \ldots$ of disjoint intervals, and note that $\bigcup_{m=1}^{\infty} J_m$ is dense in $\mathbb{R}$ (by appealing to denseness of $\mathbb{Q}$ in $\mathbb{R}$). We define $B_1 := \bigcup_{m=1}^{\infty} J_m$, and we define $B_n$ for $n \geq 2$ to be the union of all intervals $I_{2n}$ from Player 2 that respond to some $I_{2n-1}$—which is Player 1's response to some $I_{2n-2} \subseteq B_{n-1}$. We have $I_{2n} \subseteq I_{2n-1} \subseteq I_{2n-2} \subseteq \cdots$, and so $B_n \subseteq B_{n-1} \subseteq \cdots \subseteq B_1$. Notice that $B_n$ are all dense and because each of the intervals that comprise $B_n$ are disjoint, we see that each $x \in \bigcap_{n=1}^{\infty} B_n$ will belong to exactly one sequence $I_1 \supseteq I_2 \supseteq \cdots$. $\square$

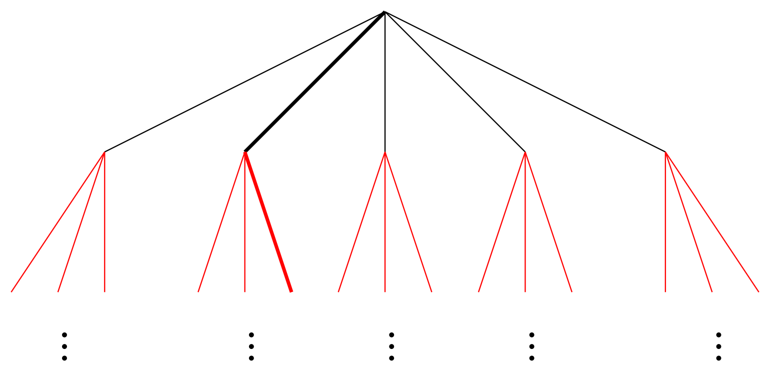
Our proofs of Propositions 4.8 and 4.9 concerned the Banach-Mazur Game on the real line. The propositions will still hold in any complete metric space $(X, \rho)$, whereby instead of intervals, we play the game with closed balls. The upshot of studying the Banach-Mazur Game is a rather quick proof of the Baire Category Theorem.

**Theorem** (Baire Category Theorem)**.** If $(X, \rho)$ is a complete metric space, then $X$ is a Baire space.

*Proof.* Let $G$ be a nonempty open set in $X$. We want to show that $G$ is of second category, which is equivalent to showing that Player 1 has a winning strategy when the Banach-Mazur Game is played with $G$. Player 1's strategy simply consists of choosing the first ball $B_1$ to be inside $G$, as $\bigcap_{n=1}^{\infty} B_n$ is a closed set in $G$. $\square$

# 5 Baire property and Borel sets

Recall that we can track the moves played in a two-player using a tree consisting of nodes and edges. If a player has a winning strategy, then he/she will respond in exactly one way to each of the opponent's moves.



We can treat each branch of the tree as a point in the set of all possible games. We can further define a sense of distance between two points based on tbe number of moves on which the two games agree. In other words, two points are close to each other if their sequence of moves diverge only after a large number of moves. Using this notion of distance, we can define a topology on the set of all games. We won't formulate these ideas rigorously, but they are the necessary underpinnings for the following definition and results.

**Definition 5.1.** A game is *Borel* if the set of all winning branches is a Borel subset of all possible branches.

**Theorem 5.2** (Martin)**.** There exists a winning strategy for either the first or second players in a Borel game.

**Corollary 5.3.** If $B \subseteq \mathbb{R}$ is a Borel set, then there exists an interval $I$ such that either $B$ or $X \setminus B$ is of first category in $I$.

Roughly speaking, Corollary 5.3 tells us that $B$ cannot be "medium-sized" in any interval, where our notion of size here comes from our interpretation of first and second category sets. While the corollary follows from Martin's Theorem, we will offer an proof of the statement without it.

**Definition 5.4.** A set $A$ has the *Baire property* if there is an open set $G$ such that

$$A \triangle G := (A \setminus G) \cup (G \setminus A)$$

is of first category.

**Example 5.5.** The following are sets with the Baire property:

1. First category sets; if $A$ is of first category, then $A \triangle \varnothing = A$ implies that $A$ has the Baire property.
2. Residual sets; observe that $X \triangle A = (X \setminus A) \cup (A \setminus X) = X \setminus A$ is of first category.
3. Open sets; if $A$ is open, then observe that $A \triangle A = \varnothing$ is of first category.
4. Closed sets; if $F$ is closed, then observe that $F \triangle \operatorname{int}(F) = F \setminus \operatorname{int}(F)$ is a nowhere dense set, hence of first category.

**Proposition 5.6.** The collection of all subsets of $X$ with the Baire property forms a $\sigma$-algebra.

*Proof.* Let $\Sigma$ be the collection of all subsets of $X$ with the Baire Property.

1. Note $\varnothing, X \in \Sigma$ because $\varnothing$ and $X$ are open.

2. If $A \in \Sigma$, then there exists an open set $G$ such that $A \triangle G$ is of first category. To show that $X \setminus A \in \Sigma$, observe that

$$(X \setminus A) \triangle \operatorname{int}(X \setminus G) \subseteq \Big((X \setminus A) \triangle (X \setminus G)\Big) \cup \Big((X \setminus G) \triangle \operatorname{int}(X \setminus G)\Big).$$

If $A \triangle G$ is of first category, then $(X \setminus A) \triangle (X \setminus G)$ is also first category. Since $X \setminus G$ is closed, we know $(X \setminus G) \triangle \operatorname{int}(X \setminus G)$ is also of first category. And so, we have that $(X \setminus A) \triangle \operatorname{int}(X \setminus G)$ is of first category and thus $X \setminus A$ has the Baire property.

3. Suppose $\{A_1, A_2, \dots\} \subseteq \Sigma$. Since each $A_n$ has the Baire property, there exists an open set $G_n$ such that $A_n \triangle G_n$ is of first category. Now observe that

$$\left(\bigcup_{n=1}^{\infty} A_n\right) \triangle \left(\bigcup_{n=1}^{\infty} G_n\right) \subseteq \bigcup_{n=1}^{\infty} (A_n \triangle G_n),$$

and because $\bigcup_{n=1}^{\infty}(A_n \triangle G_n)$ is of first category, we see that $(\bigcup_{n=1}^{\infty} A_n) \triangle (\bigcup_{n=1}^{\infty} G_n)$ must be of first category as well and thus $\bigcup_{n=1}^{\infty} A_n \in \Sigma$. $\square$

**Proposition 5.7.** Every Borel set on $X$ has the Baire property.

*Proof.* Proposition 5.6 tells us that the collection of subsets of $X$ with the Baire property forms a $\sigma$-algebra $\Sigma$, and recall from Example 5.5 that open sets have the Baire property. Since Borel sets are generated by the open subsets of $X$, we see that they necessarily belong to $\Sigma$ and therefore have the Baire property. $\square$

**Corollary 5.8.** If $B \subseteq \mathbb{R}$ is a Borel set, then there exists an interval $I$ such that either $B$ or $X \setminus B$ is of first category in $I$.

*Proof.* $\square$

# 6 Applications of Baire Category Theorem and Banach-Mazur Games

The Baire-Category Theorem and Banach-Mazur Games are commonly used to deduce the existence of a particular object satisfying an array of "typical" properties.

Given some space $X$ (such as of functions or points) and a property $P$ (such as continuity or differentiability), let $P(X) \subseteq X$ denote the set of all points in $X$ that satisfy $P$. We say that $P$ is *typical* if $P(X)$ is a residual subset of $X$. If we have a countable collection $\{P_i\}_{i=1}^{\infty}$ of typical properties, then the intersection $\bigcap_{i=1}^{\infty} P_i(X)$ is residual subset of $X$.

To illustrate this idea, consider the following example.

**Example 6.1.** Let $X = C[0,1]$ be the space of continuous functions on $[0,1]$ endowed with the norm

$$\|f - g\| = \max_{x \in [0,1]} |f(x) - g(x)|.$$

We would like to know if there exists a continuous function $f$ that is not monotone on any subinterval of $[0,1]$. Such functions do indeed exist, and we will prove existence using the Baire Category Theorem.

Observe that:

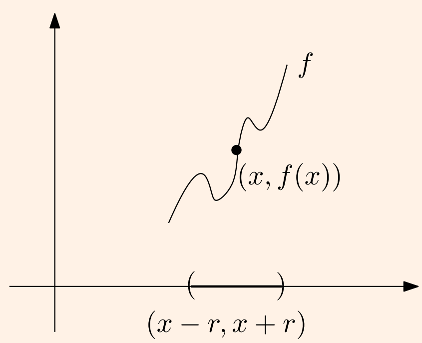
1. For any $I$ interval in $[0,1]$, there exists a function that is not monotone on $I$.
2. The set of functions in $C[0,1]$ that are not monotone on $I$ is dense. In other words, given some function $f \in C[0,1]$ and $\varepsilon > 0$, there exists a function $g \in C[0,1]$ that is not monotone on $I$ and $\|f - g\| < \varepsilon$.
3. If $g$ is not monotone on $I$, then there exists $\varepsilon > 0$ such that $h$ is not monotone on $I$ for $h \in C[0,1]$ such that $\|g - h\| < \varepsilon$.

We want to show the converse of statement (1). Let $p, q \in \mathbb{Q}$ and define $I_{p,q} := [p,q]$. By statement (2), there exists a dense open subset $A_{p,q}$ of $C[0,1]$ where all $f \in A_{p,q}$ are not monotone on $I_{p,q}$. Observe then that the intersection $\bigcap_{p,q\in\mathbb{Q}} A_{p,q}$ is the set of all functions that are not monotone on any rational interval $I_{p,q}$, hence not monotone on any subinterval of $[0,1]$. Since each $A_{p,q}$ is open and dense, the intersection $\bigcap_{p,q\in\mathbb{Q}} A_{p,q}$ is residual (see first proof of $\mathbb{Q}$ is not $G_\delta$). And so, there are residually many functions in $C[0,1]$ that are nowhere monotone.

To emphasize, there is no such thing as a "typical continuous function"—we must specify some property about continuous functions and the set of functions that satisfy the property must be residual. The example above demonstrates an application of the Baire Category Theorem to show existence. We will now consider another example but with Banach-Mazur Games.

**Example 6.2.** Again, consider $X = C[0,1]$. For some $x \in [0,1]$, we say a function $f \in C[0,1]$ is *locally monotone* if there exists $r > 0$ such that either

- $0 < y - x < r$ implies $f(x) < f(y)$ and $0 < x - y < r$ implies $f(x) > f(y)$,
- $0 < y - x < r$ implies $f(x) > f(y)$ and $0 < x - y < r$ implies $f(x) < f(y)$.




Does there exist $f \in C[0,1]$ that is locally monotone nowhere? Although hard to construct, there indeed exists such a continuous function. Unlike the previous example, however, local monotonicity requires us to deal with the uncountably many points on $[0,1]$ rather than monotonicity on subintervals of $[0,1]$. As such, the Baire Category Theorem is not particularly useful here and we would resort to designing a Banach-Mazur Game.

1. Let $A \subseteq C[0,1]$ denote the set of functions that are locally monotone somewhere.
2. Player 1 starts by choosing a ball $B(f_1, r_1)$, and Player 2 responds by choosing a ball $B(f_2, r_2)$ such that $B(f_2, r_2) \subseteq B(f_1, r_1)$.
3. Players 1 and 2 take turns and we end up with a sequence of nested balls

$$\cdots \subseteq B(f_n, r_1) \subseteq B(f_{n-1}, r_{n-1}) \subseteq \cdots \subseteq B(f_1, r_1).$$

   To simplify notation, let $B_n = B(f_n, r_n)$.

4. Since $C[0,1]$ is complete, $\bigcap_{n=1}^{\infty} B_n$ is nonempty. We say Player 1 wins if $\bigcap_{n=1}^{\infty} B_n$ intersects $A$ and Player 2 wins otherwise.

Recall that Player 2 has a winning strategy if and only if $A$ is of first category. If we can create a winning strategy for Player 2, then it follows that $A$ is of first category and its complement—the set of nowhere locally monotone functions in $C[0,1]$—is residual. Our winning strategy for Player 2 proceeds as follows:

1.

**Corollary 6.3.** A nowhere locally monotone function $f \in C[0,1]$ is nowhere differentiable.

*Proof.* If $f \in C[0,1]$ is nowhere locally monotone, then it follows that the derivative $f'$ is neither positive nor negative at any $x \in [0,1]$. We will show that a derivative of 0 is also unattainable on $[0,1]$; □

# 7 Carathéodory Extension Theorem

Suppose we have a set $X$, some collection of subsets $\Sigma_0 \subseteq \mathcal{P}(X)$, and a function $\mu_0 : \Sigma_0 \to [0,\infty]$. We would like to extend $\Sigma_0$ and $\mu_0$ to some $\sigma$-algebra $\Sigma$ and measure $\mu$ on $X$. As long as $\Sigma_0$ and $\mu_0$ satisfy some specific properties, the Carathéodory Extension Theorem can be used to obtain a $\sigma$-algebra and measure.

**Theorem 7.1.** Let $X$ be a set. Suppose $\Sigma_0 \subseteq \mathcal{P}(X)$ is a collection of subsets of $X$ and $\mu_0 : \Sigma_0 \to [0,\infty]$ a function such that

1. if $A, B \in \Sigma_0$, then $A \setminus B$ and $A \cap B$ belong to $\Sigma_0$,
2. $\mu_0$ is finitely additive on $\Sigma_0$,
3. and $\mu_0$ is a *relative outer measure*—in other words,
$$\mu_0\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} \mu(A_n)$$
for all $\{A_n\}_{n=1}^{\infty} \subseteq \Sigma_0$.

Then there exists a $\sigma$-algebra $\Sigma$ and measure $\mu : \Sigma \to [0,\infty]$ such that $\Sigma_0 \subseteq \Sigma$ and $\mu_0 = \mu$ on $\Sigma_0$.

The following is an outline of our proof:

1. Define an *outer measure* $\mu$ on all subsets of $X$. We say $\mu$ is an outer measure if
$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} \mu(A_n)$$
for $\{A_n\}_{n=1}^{\infty} \subseteq \mathcal{P}(X)$.
2. Define a $\sigma$-algebra $\Sigma$ such that the outer measure $\mu$ is a measure on $\Sigma$.
3. Show that $(\Sigma_0, \mu_0)$ do indeed extend to $(\Sigma, \mu)$.

*Proof of Step 1.* For some $A \subseteq X$, we define $\mu : \mathcal{P}(X) \to [0, \infty]$ by

$$\mu(A) := \inf \left\{ \sum_{n=1}^{\infty} \mu_0(A_n) : \{A_n\}_{n=1}^{\infty} \subseteq \Sigma_0 \text{ is a cover of } A \right\}.$$

If $A$ admits no cover $\{A_n\}_{n=1}^{\infty} \subseteq \Sigma_0$, then we define $\mu(A) := \infty$. We will show that $\mu$ is an outer measure—in other words, if $\{A_n\}_{n=1}^{\infty} \subseteq \mathcal{P}(X)$, then

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} \mu(A_n).$$

We will do so by showing that

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} \mu(A_n) + \varepsilon$$

for all $\varepsilon > 0$. By definition of $\mu(A_n)$, there exists a cover $\{A_{n,m}\}_{m=1}^{\infty} \subseteq \Sigma_0$ such that

$$\left(\sum_{m=1}^{\infty} \mu_0(A_{n,m})\right) - \mu(A_n) < \frac{\varepsilon}{2^n}.$$

We have

$$\sum_{n=1}^{\infty} \left(\left(\sum_{m=1}^{\infty} \mu_0(A_{n,m})\right) - \mu(A_n)\right) = \left(\sum_{n=1}^{\infty}\sum_{m=1}^{\infty} \mu_0(A_{n,m})\right) - \sum_{n=1}^{\infty} \mu(A_n) < \varepsilon,$$
$$\left(\sum_{n=1}^{\infty}\sum_{m=1}^{\infty} \mu_0(A_{n,m})\right) < \sum_{n=1}^{\infty} \mu(A_n) + \varepsilon.$$

Notice that

$$\bigcup_{n=1}^{\infty} A_n \subseteq \bigcup_{n=1}^{\infty}\bigcup_{m=1}^{\infty} A_{m,n},$$

and by the definition of $\mu\left(\bigcup_{n=1}^{\infty} A_n\right)$, we have

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \left(\sum_{n=1}^{\infty}\sum_{m=1}^{\infty} \mu_0(A_{n,m})\right) < \sum_{n=1}^{\infty} \mu(A_n) + \varepsilon.$$

And so, this yields

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} \mu(A_n)$$

as desired. □

*Proof of Step 2.* We will now construct a $\sigma$-algebra $\Sigma$ based on our measure $\mu$. We say $A \in \Sigma$ if and only if

$$\mu(S \cap A) + \mu(S \setminus A) = \mu(S)$$

for all $S \subseteq X$. We check that $\Sigma$ is a $\sigma$-algebra with the following steps.

1. First, note that it follows from our definition of $\mu$ that $\mu(\varnothing) = 0$. For some $S \subseteq X$, notice that

$$\mu(S \cap \varnothing) + \mu(S \setminus \varnothing) = \mu(\varnothing) + \mu(S) = 0 + \mu(S) = \mu(S).$$

As such, we have $\varnothing \in \Sigma$. Now observe that

$$\mu(S \cap X) + \mu(S \setminus X) = \mu(S) + \mu(\varnothing) = \mu(S) + 0 = \mu(S),$$

which implies $X \in \Sigma$.

2. Let $A \in \Sigma$. To see that $X \setminus A \in \Sigma$, observe that for $S \subseteq X$, we have

$$\mu(S \cap (X \setminus A)) + \mu(S \setminus (X \setminus A)) = \underbrace{\mu(S \setminus A) + \mu(S \cap A) = \mu(S)}_{\text{since } A\in\Sigma}.$$

3. Let $\{A_n\}_{n=1}^{\infty} \subseteq \Sigma$. To show that $\bigcup_{n=1}^{\infty} A_n \in \Sigma$, we will first show $A_1 \cup A_2 \in \Sigma$ if $A_1 \cap A_2 = \varnothing$. In other words, if $S \subseteq X$, then we want to show

$$\mu\Big(S \cap (A_1 \cup A_2)\Big) + \mu\Big(S \setminus (A_1 \cup A_2)\Big) = \mu(S).$$

Because $A_1, A_2 \in \Sigma$, we have

$$\begin{aligned}\mu\Big(S \cap (A_1 \cup A_2)\Big) &= \mu\Big((S \cap (A_1 \cup A_2)) \cap A_1\Big) + \mu\Big((S \cap (A_1 \cup A_2)) \setminus A_1\Big),\\ &= \mu(S \cap A_1) + \mu(S \cap A_2)\end{aligned}$$

and thus

$$\mu\Big(S \cap (A_1 \cup A_2)\Big) + \mu\Big(S \setminus (A_1 \cup A_2)\Big) = \mu(S \cap A_1) + \mu(S \cap A_2) + \mu\Big(S \setminus (A_1 \cup A_2)\Big)$$

In addition, observe that

$$\begin{aligned}\mu(S) &= \mu(S \cap A_1) + \mu(S \setminus A_2),\\ &= \mu(S \cap A_1) + \mu\Big((S \cap A_2) \cup \big(S \setminus (A_1 \cup A_2)\big)\Big),\\ \mu\Big((S \cap A_2) \cup \big(S \setminus (A_1 \cup A_2)\big)\Big) &= \mu\Big(\Big((S \cap A_2) \cup \big(S \setminus (A_1 \cup A_2)\big)\Big) \cap A_2\Big)\\ &\quad + \mu\Big(\Big((S \cap A_2) \cup \big(S \setminus (A_1 \cup A_2)\big)\Big) \setminus A_2\Big),\\ &= \mu(S \cap A_2) + \mu(S \setminus (A_1 \cup A_2)),\end{aligned}$$

which implies

$$\mu(S) = \mu(S \cap A_1) + \mu(S \cap A_2) + \mu(S \setminus (A_1 \cup A_2)).$$

Hence, we see that

$$\mu\Big(S \cap (A_1 \cup A_2)\Big) + \mu\Big(S \setminus (A_1 \cup A_2)\Big) = \mu(S)$$

and therefore $A_1 \cup A_2 \in \Sigma$ if $A_1 \cap A_2 = \varnothing$. If instead $A \cap B \neq \varnothing$, then the same fact still holds (though we won't prove it here). And so, it follows that the finite union of sets in $\Sigma$ also belongs to $\Sigma$.

To show that $\bigcup_{n=1}^{\infty} A_n \in \Sigma$, we will assume without loss of generality that the sets in $\{A_n\}_{n=1}^{\infty}$ are pairwise disjoint (since we can take $\bigcup_{n=1}^{\infty} A_n$ and divide it into an infinite sequence of disjoint subsets). We want to show

$$\mu\left(S \cap \left(\bigcup_{n=1}^{\infty} A_n \in \Sigma\right)\right) + \mu\left(S \setminus \left(\bigcup_{n=1}^{\infty} A_n \in \Sigma\right)\right) = \mu(S)$$

for $S \subseteq X$. First, observe that because $\mu$ is an outer measure and the $A_n$ are pairwise disjoint, we have

$$\mu\left(S \cap \left(\bigcup_{n=1}^{\infty} A_n\right)\right) = \mu\left(\bigcup_{n=1}^{\infty}(S \cap A_n)\right) \le \sum_{n=1}^{\infty} \mu(S \cap A_n).$$

Notice that

$$\begin{aligned}
\mu\left(\bigcup_{n=1}^{\infty}(S \cap A_n)\right) &= \mu\left(\bigcup_{n=1}^{N}(S \cap A_n)\right) + \mu\left(\bigcup_{n=N+1}^{\infty}(S \cap A_n)\right), \\
&= \sum_{n=1}^{N} \mu(S \cap A_n) + \mu\left(\bigcup_{n=N+1}^{\infty}(S \cap A_n)\right), \\
&\ge \sum_{n=1}^{N} \mu(S \cap A_n).
\end{aligned}$$

Since the inequality above holds for all $N \in \mathbb{N}$, we have

$$\mu\left(\bigcup_{n=1}^{\infty}(S \cap A_n)\right) \ge \sum_{n=1}^{\infty} \mu(S \cap A_n).$$

And so, we have

$$\mu\left(\bigcup_{n=1}^{\infty}(S \cap A_n)\right) = \sum_{n=1}^{\infty} \mu(S \cap A_n).$$

Now observe that

$$\begin{aligned}
\mu(S) &= \mu\left(\left(\bigcup_{n=1}^{\infty}(S \cap A_n)\right) \cup \left(S \setminus \left(\bigcup_{n=1}^{\infty}(S \cap A_n)\right)\right)\right), \\
&= \sum_{n=1}^{N} \mu(S \cap A_n) + \mu\left(\left(\bigcup_{n=N+1}^{\infty}(S \cap A_n)\right) \cup \left(S \setminus \left(\bigcup_{n=1}^{\infty}(S \cap A_n)\right)\right)\right), \\
&\ge \sum_{n=1}^{N} \mu(S \cap A_n) + \mu\left(S \setminus \left(\bigcup_{n=1}^{\infty}(S \cap A_n)\right)\right)
\end{aligned}$$

Since the inequality above holds for all $N \in \mathbb{N}$, we have

$$\mu(S) \ge \sum_{n=1}^{\infty} \mu(S \cap A_n) + \mu\left(S \setminus \left(\bigcup_{n=1}^{\infty}(S \cap A_n)\right)\right).$$

Because $\mu$ is an outer measure, the reverse inequality holds as well. We conclude that

$$\mu(S) = \sum_{n=1}^{\infty} \mu(S \cap A_n) + \mu\left(S \setminus \left(\bigcup_{n=1}^{\infty}(S \cap A_n)\right)\right)$$

and therefore

$$\mu\left(S \cap \left(\bigcup_{n=1}^{\infty} A_n \in \Sigma\right)\right) + \mu\left(S \setminus \left(\bigcup_{n=1}^{\infty} A_n \in \Sigma\right)\right) = \mu(S)$$

And so, we have shown that $\Sigma$ defines a $\sigma$-algebra. $\square$

*Proof of Step 3.* (to be completed) □

# 8 Lebesgue measure and measureability

Let $\Sigma_0 \subseteq \mathcal{P}(\mathbb{R}^n)$ be the collection of all closed rectangles in $\mathbb{R}^n$ and their finite unions/intersections, and let $\mu_0 : \Sigma_0 \to [0, \infty]$ be the function that measures the volume of sets in $\Sigma_0$. By the Carathéodory Extension Theorem, we obtain a $\sigma$-algebra $\Sigma \subseteq \mathcal{P}(\mathbb{R}^n)$ and measure $\mu : \Sigma \to [0, 1]$ such that $\Sigma_0 \subseteq \Sigma$ and $\mu = \mu_0$ on $\Sigma_0$. We refer to $\mu$ as the *Lebesgue measure on* $\mathbb{R}^n$. As a reminder, we have

$$\mu(A) := \inf \left\{ \sum_{n=1}^{\infty} \mu_0(A_n) : \{A_n\}_{n=1}^{\infty} \subseteq \Sigma_0 \text{ is a cover of } A \right\}$$

for $A \in \Sigma$.

**Example 8.1.** Let $(\mathbb{R}^n, \Sigma, \mu)$ be the Lebesgue measure space on $\mathbb{R}^n$ (via the Carathéodory Extension Theorem). The following sets are Lebesgue measureable:

1. Null (measure zero) sets; we say $A \subseteq \mathbb{R}^n$ is a *null set* if for all $\varepsilon > 0$, there exists a countable cover $\{A_n\}_{n=1}^{\infty}$ (by balls, rectangles, open sets, etc.) such that

$$\sum_{n=1}^{\infty} \mathrm{vol}(A_n) < \varepsilon.$$

   In other words, $\mu(A) = 0$. We see that $A \in \Sigma$ because for $S \subseteq \mathbb{R}^n$, we have

$$\mu(S \cap A) + \mu(S \setminus A) \leq \mu(A) + \mu(S) = 0 + \mu(S)$$

   and $\mu(S \cap A) + \mu(S \setminus A) \geq \mu(S)$ because $\mu$ is an outer measure. And so, we have

$$\mu(S \cap A) + \mu(S \setminus A) = \mu(S).$$

2. Borel sets; recall that Borel sets are generated by open sets and open sets belong to $\Sigma$.

**Remark.** Note that if $\mu$ is a measure obtained from the Carathéodory Extension Theorem, it is true more generally that null and Borel sets are measureble (assuming a topology on our measure space).

The following is an alternate characterization of Lebesgue measureable sets.

**Theorem 8.2.** A subset $A \subseteq \mathbb{R}^n$ is Lebesgue measureable if and only if there exists a Borel set $B \subseteq \mathbb{R}^n$ such that $A \triangle B$ is of measure zero.

**Theorem 8.3.**

# 9 Convergence theorems of sequences and integrals

**Definition 9.1.** Suppose $(X, \Sigma, \mu)$ is a measure space and $P$ is some property about points in $X$. We say that $P$ is true *almost everywhere* if

$$\mu\big(\{x \in X : x \text{ does not satisfy } P\}\big) = 0.$$

**Theorem 9.2** (Egoroff's Theorem).

**Theorem 9.3** (Bounded Convergence Theorem).

**Theorem 9.4** (Fatou's Lemma).

**Theorem 9.5** (Monotone Convergence Theorem).

**Lemma 9.6.**

**Theorem 9.7** (Dominated Convergence).

## 10 Product measures via Carathéodory Extension Theorem

Suppose we have measure spaces $(X_1, \Sigma_1, \mu_1)$ and $(X_2, \Sigma_2, \mu_2)$. We would like to define a measure $\mu_1 \times \mu_2$ on $X_1 \times X_2$, and one natural suggestion would be

$$\mu_1 \times \mu_2(A \times B) := \mu_1(A) \cdot \mu_2(B). \tag{1}$$

However, this does not define a measure; if $X_1 = X_2 = \mathbb{R}$ and $\mu_1$ and $\mu_2$ are the Lebesgue measure, then take $A \subseteq \mathbb{R}$ to be a non-measureable set and $B := \{0\}$. Note that $A \times \{0\} \subseteq \mathbb{R} \times \{0\}$, and because $\mathbb{R} \times \{0\} \subseteq \mathbb{R}^2$ has measure zero, it follows that $A \times \{0\}$ has measure zero as well. In other words, we *expect* $\mu_1 \times \mu_2(A \times \{0\}) = 0$. However, the quantity $\mu_1(A) \cdot \mu_2(\{0\})$ is not defined because $A$ is not measureable in $\mathbb{R}$. Hence, we see that (1) does not define a measure on $X_1 \times X_2$.

Fortunately, we can create a measure on $X_1 \times X_2$ by a carefully applying the Carathéodory Extension Theorem. Suppose we have finitely many disjoint "rectangles" $A_1 \times B_1, \cdots, A_n \times B_n$. Define $E := \bigcup_{i=1}^n A_i \times B_i$ and a primitive measure $\mu_0$ on $X_1 \times X_2$ by

$$\mu_0(E) := \sum_{i=1}^n \mu_1(A_i) \cdot \mu_2(B_i).$$

Step 1 of the Carathéodory Extension Theorem defines the outer measure $\mu_1 \times \mu_2$ by

$$\mu_1 \times \mu_2(F) := \inf \left\{ \sum_{i=1}^\infty \mu_0(A_i \times B_i) : A_i \in \Sigma_1, B_i \in \Sigma_2, \text{ and } F \subseteq \bigcup_{i=1}^\infty A_i \times B_i \right\}$$

for $F \subseteq X_1 \times X_2$. Step 2 defines a $\sigma$-algebra $\Sigma$ by $E \in \Sigma$ if and only if

$$\mu_1 \times \mu_2(S) = \mu_1 \times \mu_2(S \setminus E) + \mu_1 \times \mu_2(S \cap E)$$

for all $S \subseteq X_1 \times X_2$, and Step 3 tells us $\Sigma_1 \times \Sigma_2 \subseteq \Sigma$ and $\mu_1 \times \mu_2 = \mu_0$ on $\Sigma_1 \times \Sigma_2$. Altogether, we have that the outer measure $\mu_1 \times \mu_2$ is a measure on $\Sigma$.

# 11 Fubini's Theorem

**Theorem 11.1.** Let $(X_1, \Sigma_1, \mu_1)$ and $(X_2, \Sigma_2, \mu_2)$ be measure spaces, and let $\mu_1 \times \mu_2$ denote the product measure (via Carathéodory Extension Theorem). If $f \in L^1(X \times Y)$, then

1. the function $f^y : X_1 \to \mathbb{R}$ defined by $f^y(x) = f(x, y)$ belongs to $L^1(X_1)$ for almost all $y \in X_2$,
2. the integral function $\displaystyle\int_{X_1} f^y(x)\, \mathrm{d}\mu_1$ belongs to $L^1(X_2)$, and

$$\int_{X_2} \left( \int_{X_1} f(x, y)\, \mathrm{d}\mu_1 \right) \mathrm{d}\mu_2 = \int_{X_1} \left( \int_{X_2} f(x, y)\, \mathrm{d}\mu_2 \right) \mathrm{d}\mu_1 = \int_{X_1 \times X_2} f(x, y)\, \mathrm{d}(\mu_1 \times \mu_2).$$

# 12 Construction of a non-Lebesgue-measureable set in $\mathbb{R}$

Suppose $\mathbb{R}$ is endowed with the Lebesgue measure and Borel $\sigma$-algebra. We will construct a set $E$ in $\mathbb{R}$ that is not Lebesgue-measureable.

To begin, consider $\mathbb{R}$ as a vector space over $\mathbb{Q}$, and let $\{x_\lambda\}_{\lambda \in \Lambda}$ be the basis of $\mathbb{R}$. We have that for each $x \in \mathbb{R}$, there exist a finite subset $\{x_{\lambda_1}, \dots, x_{\lambda_n}\} \subseteq \{x_\lambda\}_{\lambda \in \Lambda}$ and coefficients $q_i \in \mathbb{Q}$ for $1 \le i \le n$ such that

$$x = \sum_{i=1}^{n} q_i x_{\lambda_i}.$$

Let $x_\alpha \in \{x_\lambda\}_{\lambda \in \Lambda}$ be some basis vector and define $E \subseteq \mathbb{R}$ to be the set

$$E := \left\{ x \in \mathbb{R} : x = \sum_{i=1}^{n} q_i x_{\lambda_i} \text{ where } x_{\lambda_i} \neq x_\alpha \text{ for } 1 \le i \le n \right\}.$$

In other words, $E = \operatorname{span}\big(\{x_\lambda\}_{\lambda \in \Lambda} \setminus \{x_\alpha\}\big)$. Since $\mathbb{Q}$ is countable, let $\mathbb{Q} = \{q_1, q_2, \dots\}$. We define

$$E_i := E + q_i \cdot x_\alpha,$$

and observe that $E_i \cap E_j = \varnothing$ for $i \neq j$ and $\bigcup_{i=1}^{\infty} E_i = \mathbb{R}$. Note that $\mu(E_i) = \mu(E_j)$ for all $i, j \in \mathbb{N}$, and we have

$$\mu(\mathbb{R}) = \mu\left( \bigcup_{i=1}^{\infty} E_i \right) = \sum_{i=1}^{\infty} \mu(E_i).$$

Because $\mu(\mathbb{R}) = \infty$, we must have $\mu(E_i) > 0$.

Now suppose we restrict our $E_i$'s to $[0, 1]$; we have

$$[0,1] \cap \left( \bigcup_{i=1}^{\infty} E_i \right) = \bigcup_{i=1}^{\infty} [0,1] \cap E_i = [0,1].$$

As before, we have $\mu\big([0,1] \cap E_i\big) = \mu\big([0,1] \cap E_j\big)$ for all $i, j \in \mathbb{N}$ and

$$\sum_{i=1}^{\infty} \mu\big([0,1] \cap E_i\big) = \mu\big([0,1]\big) = 1.$$

We see that $\mu\big([0,1] \cap E_i\big) \neq 0$, but if $\mu\big([0,1] \cap E_i\big) > 0$, then our infinite sum above does not converge. And so, we cannot assign a Lebesgue measure to $[0,1] \cap E_i$.

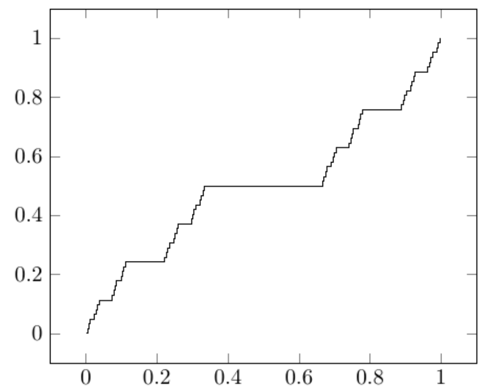
## 13 Measureable functions and Lebesgue measure zero

Recall that a function $f : X \to \mathbb{R}$ on a measure space $(X, \Sigma, \mu)$ is *measureable* if the pre-image $f^{-1}(B)$ of a Borel set $B$ in $\mathbb{R}$ is measureable in $X$. Is it true that if $f$ is measureable, then the image $f(A)$ of a measureable set $A \subseteq X$ is measurable in $\mathbb{R}$? Unfortunately, this is not true.

First, we will define the familiar Cantor function (Devil's Staircase) using ternary expansions of numbers in $\mathbb{R}$, namely if $x \in \mathbb{R}$, then $x = \sum_{n=1}^{\infty} a_n/3^n$ where $a_n \in \{0, 1, 2\}$. Let $\mathcal{C}$ denote the middle-third Cantor set, and notice that $x =\in C$ if and only if $a_n \in \{0, 2\}$ for all $n \in \mathbb{N}$. We define the Cantor function $f : [0, 1] \to \mathbb{R}$ by

$$f(x) = f\left(\sum_{n=1}^{\infty} \frac{a_n}{3^n}\right) := \sum_{n=1}^{\infty} \frac{a_n}{2^{n+1}}.$$

The following is the graph of $f$ (stolen online).




We won't prove this fact here, but know that the Cantor function is continuous and surjective. In other words, $f(\mathcal{C}) = [0, 1]$. Recall (from somewhere) that $\mu(\mathcal{C}) = 0$ and $\mu([0, 1]) = 1$ by definition of the Lebesgue measure in $\mathbb{R}$. One observation here is that although $f$ is continuous, the measure of $\mathcal{C}$ is not invariant under $f$. Now suppose $E \subseteq [0, 1]$ is some non-measureable subset; notice that $f^{-1}(E) \cap \mathcal{C} \subseteq \mathcal{C}$ has measure zero, yet $f\big(f^{-1}(E) \cap \mathcal{C}\big) = E$, which is non-measurable.

## 14 $L^p$ space and Hölder and Minkowski Inequalities

**Definition 14.1.** Let $(X, \Sigma, \mu)$ be a measure space where $\mu$ is $\sigma$-finite. For $1 \leq p < \infty$, we define $L^p$ to be the collection of functions $f : X \to \mathbb{R}$ such that

$$\int_X |f|^p \, \mathrm{d}\mu < \infty.$$

We define $L^\infty$ to be the collection of functions where there exists finite $M > 0$ such that $|f(x)| \leq M$ almost everywhere. We define the $p$-norm $(1 \leq p < \infty)$ on $X$ by

$$\|f\|_p := \left(\int_X |f|^p \, \mathrm{d}\mu\right)^{1/p}$$

and the $\infty$-norm by

$$\|f\|_\infty = \inf\{M \in \mathbb{R} : |f(x)| \leq M \text{ almost everywhere.}\}$$

**Proposition 14.2.** Suppose $(X, \Sigma, \mu)$ is a measure space where $\mu$ is $\sigma$-finite. If $1 \leq p < q \leq \infty$, then

$$L^P(X, \mu) \supseteq L^q(X, \mu)$$

*Proof.* Let $f \in L^q(X, \mu)$ and define $A := \{x \in X : |f(x)| \geq 1\}$. If $1 \leq p < q \leq \infty$, then

$$\begin{aligned}\int_X |f|^p \, \mathrm{d}\mu &= \int_A |f|^p \, \mathrm{d}\mu + \int_{X \setminus A} |f|^p \, \mathrm{d}\mu, \\ &\leq \int_A |f|^q \, \mathrm{d}\mu + \int_{X \setminus A} |f|^p \, \mathrm{d}\mu.\end{aligned}$$

Because $\mu$ is $\sigma$-finite, we have that $\displaystyle\int_{X \setminus A} |f|^p \, \mathrm{d}\mu < \infty$. As a result, we see that

$$\int_X |f|^p \, \mathrm{d}\mu < \infty,$$

which implies $f \in L^p(X, \mu)$. □

**Theorem 14.3** (Young's Inequality)**.** If $a, b > 0$ and $p, q \in (1, \infty)$ such that $\dfrac{1}{p} + \dfrac{1}{q} = 1$, then

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

*Proof.* First, recall that a function $\varphi : \mathbb{R} \to \mathbb{R}$ is *convex* if $\varphi''(x) \geq 0$ for all $x \in \mathbb{R}$, from which it follows that

$$\varphi(\lambda x + (1 - \lambda)y) \leq \lambda\varphi(x) + (1 - \lambda)\varphi(y)$$

for $\lambda \in \mathbb{R}$ and $x, y \in \mathbb{R}$. In particular, $\exp(x)$ is a convex function.

For $a, b > 0$, observe that $ab = \exp(\log(a) + \log(b))$, which implies

$$ab = \exp\left(\frac{\log(a^p)}{p} + \frac{\log(b^q)}{q}\right).$$

By convexity of $\exp(x)$, we have

$$\begin{aligned}\exp\left(\frac{\log(a^p)}{p} + \frac{\log(b^q)}{q}\right) &\leq \frac{1}{p}\exp(\log(a^p)) + \left(1 - \frac{1}{p}\right)\exp(\log(b^q)), \\ &= \frac{1}{p}\exp(\log(a^p)) + \frac{1}{q}\exp(\log(b^q)), \\ &= \frac{a^p}{p} + \frac{b^q}{q}.\end{aligned}$$

□

**Theorem 14.4** (Hölder's Inequality)**.** If $f \in L^p(X, \mu)$ and $g \in L^q(X, \mu)$ where $\dfrac{1}{p} + \dfrac{1}{q} = 1$, then

$$\|fg\|_1 \leq \|f\|_p \cdot \|g\|_q.$$

*Proof.* First, note that for $a, b \geq 0$ and $0 \leq \lambda \leq 1$, we have

$$a^\lambda b^{1-\lambda} \leq \lambda a + (1-\lambda) b$$

by concavity of $\log(x)$. Let $a = \dfrac{|f(x)|^p}{\|f\|_p^p}$, $b = \dfrac{|g(x)|^q}{\|g\|_q^q}$, and $\lambda = 1/p$. Notice that $1 - \lambda = 1/q$. By the inequality above, we have

$$\frac{|f(x)|}{\|f\|_p} \cdot \frac{|g(x)|}{\|g\|_q} \leq \frac{1}{p} \cdot \frac{|f(x)|^p}{\|f\|_p^p} + \frac{1}{q} \cdot \frac{|g(x)|^q}{\|g\|_q^q},$$

which implies

$$\begin{aligned}
\int_X \frac{|f|}{\|f\|_p} \cdot \frac{|g|}{\|g\|_q} \, \mathrm{d}\mu &\leq \int_X \frac{1}{p} \cdot \frac{|f|^p}{\|f\|_p^p} \, \mathrm{d}\mu + \int_X \frac{1}{q} \cdot \frac{|g|^q}{\|g\|_q^q} \, \mathrm{d}\mu, \\
\frac{1}{\|f\|_p \|g\|_q} \int_X \|fg\| \, \mathrm{d}\mu &\leq \frac{1}{p} \cdot \frac{1}{\|f\|_p^p} \int_X |f|^p \, \mathrm{d}\mu + \frac{1}{q} \cdot \frac{1}{\|g\|_q^q} \int_X |g|^q \, \mathrm{d}\mu, \\
\frac{1}{\|f\|_p \|g\|_q} \cdot \|fg\|_1 &\leq \frac{1}{p} \cdot 1 + \frac{1}{q} \cdot 1.
\end{aligned}$$

Since $\dfrac{1}{p} + \dfrac{1}{q} = 1$, we have

$$\|fg\|_1 \leq \|f\|_p \|g\|_q.$$

□

**Theorem 14.5** (Minkowski's Inequality)**.** If $f, g \in L^p(X, \mu)$ and $1 \leq p \leq \infty$, then

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p.$$

*Proof.* If $f, g \in L^p(X, \mu)$, then by the convexity of $|a|^p$, we have

$$\begin{aligned}
\left| \frac{1}{2} f(x) + \left(1 - \frac{1}{2}\right) g(x) \right|^p &\leq \frac{1}{2} |f(x)|^p + \left(1 - \frac{1}{2}\right) |g(x)|^p, \\
\left| \frac{f(x) + g(x)}{2} \right|^p &\leq \frac{|f(x)|^p + |g(x)|^p}{2}, \\
\frac{1}{2} \|f + g\|_p^p &\leq \frac{1}{2} \left( \|f\|_p^p + \|g\|_p^p \right),
\end{aligned}$$

which implies $f + g \in L^p(X, \mu)$. Now observe that

$$\begin{aligned}
|f(x) + g(x)|^p &\leq |f(x) + g(x)|^{p-1} \cdot (|f(x)| + |g(x)|), \\
\int_X |f + g|^p \, \mathrm{d}\mu &\leq \int_X |f + g|^{p-1} \cdot |f| \, \mathrm{d}\mu + \int_X |f + g|^{p-1} \cdot |g| \, \mathrm{d}\mu
\end{aligned}$$

For $p > 1$, let $q \in \mathbb{R}$ such that $\dfrac{1}{p} + \dfrac{1}{q} = 1$. Then we have $p = (p-1)q$. Notice that if $h \in L^q(X, \mu)$, we have

$$\int_X |h|^p \, \mathrm{d}\mu = \int_X \left( |h|^{p-1} \right)^q \, \mathrm{d}\mu,$$

so $|h|^{p-1} \in L^q(X, \mu)$ and $\|h\|_p^{p/q} = \|h^{p-1}\|_q$. Since $f + g \in L^p(X, \mu)$, applying Hölder's Inequality to the inequality above gives us

$$
\begin{aligned}
\int_X |f+g|^p \, \mathrm{d}\mu &\leq \left\| |f+g|^{p-1} \right\|_q \cdot \|f\|_p + \left\| |f+g|^{p-1} \right\|_q \cdot \|g\|_p, \\
\|f+g\|_p^p &\leq \left\| |f+g|^{p-1} \right\|_q \left( \|f\|_p + \|g\|_p \right), \\
&= \|f+g\|_p^{p/q} \left( \|f\|_p + \|g\|_p \right).
\end{aligned}
$$

It follows then that

$$
\frac{\|f+g\|_p}{\|f+g\|^{p/q}} = \|f+g\|_p \leq \|f\|_p + \|g\|_p. \qquad \square
$$

## 15 Dense subsets of $L^p$

**Theorem 15.1** (Absolute Continuity of Lebesgue Integral)**.** If $f \in L^1(X, \mu)$, then for all $\varepsilon > 0$ there exists $\delta > 0$ such that

$$
\int_A |f| \, \mathrm{d}\mu < \varepsilon
$$

for all subsets $A \subseteq X$ where $\mu(A) < \delta$.

*Proof.* For $M \in \mathbb{N}$, define

$$
\{f \leq M\} := \{x \in X : f(x) \leq M\}
$$

and let $f_M(x) := f(x) \cdot \chi_{\{f \leq M\}}(x)$. Notice that $\lim_{M \to \infty} f_M(x) = f(x)$ and $|f_M(x)| \leq |f(x)|$ for all $x \in X$, so the Dominated Convergence Theorem tells us

$$
\lim_{M \to \infty} \int_X |f_M - f| \, \mathrm{d}\mu = 0.
$$

Then for $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that

$$
\int_X |f_M - f| \, \mathrm{d}\mu < \frac{\varepsilon}{2}
$$

for all $M \geq N$. Observe that

$$
\begin{aligned}
\int_X |f_M - f| \, \mathrm{d}\mu &= \int_{\{f > M\}} (|f| - |f_M|) \, \mathrm{d}\mu + \int_{\{f \leq M\}} (|f| - |f_M|) \, \mathrm{d}\mu, \\
&= \int_{\{f > M\}} |f| \, \mathrm{d}\mu + 0,
\end{aligned}
$$

which shows that

$$
\int_{\{f > M\}} |f| \, \mathrm{d}\mu < \frac{\varepsilon}{2}.
$$

Now let $\delta := \dfrac{\varepsilon}{2M}$ where $M \geq N$ and take $A \subseteq X$ such that $\mu(A) < \delta$. We have

$$\begin{aligned}
\int_A |f| \, \mathrm{d}\mu &= \int_{A \cap \{f > M\}} |f| \, \mathrm{d}\mu + \int_{A \cap \{f \leq M\}} |f| \, \mathrm{d}\mu, \\
&\leq \int_{\{f > M\}} |f| \, \mathrm{d}\mu + \int_{A \cap \{f \leq M\}} M \, \mathrm{d}\mu, \\
&< \frac{\varepsilon}{2} + M \cdot \mu(A), \\
&= \frac{\varepsilon}{2} + \frac{\varepsilon}{2}, \\
&= \varepsilon.
\end{aligned}$$

□

**Lemma 15.2.** Suppose $(X, \|\cdot\|)$ is a normed linear space. Then $X$ is complete if and only if every absolutely convergent sum converges—in other words, if $(x_n) \subseteq X$ is a sequence where $\sum_{n=1}^{\infty} \|x_n\| < \infty$ then $\sum_{n=1}^{\infty} x_n$ converges.

*Proof.* First, suppose that $X$ is complete. Then let $(x_n) \subseteq X$ be a sequence such that $\sum_{n=1}^{\infty} \|x_n\|$. converges. To show that $\sum_{n=1}^{\infty} x_n$ converges, it suffices to show that its sequence of partial sums is Cauchy. For $k, \ell \in \mathbb{N}$ where $k < \ell$, observe that

$$\left\| \sum_{n=1}^{\ell} x_n - \sum_{n=1}^{k} x_n \right\| = \left\| \sum_{n=k+1}^{\ell} x_n \right\| \leq \sum_{n=k+1}^{\ell} \|x_n\|.$$

Because $\sum_{n=1}^{\infty} \|x_n\|$ converges, for $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that

$$\left| \sum_{n=1}^{\ell} \|x_n\| - \sum_{n=1}^{k} \|x_n\| \right| = \sum_{n=k+1}^{\ell} \|x_n\| < \varepsilon$$

for $N \leq k < \ell$. It follows then that $\left\| \sum_{n=1}^{\ell} x_n - \sum_{n=1}^{k} x_n \right\| < \varepsilon$, which shows that $\sum_{n=1}^{\infty} x_n$ converges.

To prove the converse, let $(x_n)$ be a Cauchy sequence in $X$. Then for all $k \in \mathbb{N}$, there exists $N_k \in \mathbb{N}$ such that $\|x_n - x_m\| < \frac{1}{2^k}$ for all $n, m \geq N_k$. Define $x_{N_0} := 0$ and consider the sequence

$$\{x_{N_1}, x_{N_2} - x_{N_1}, x_{N_3} - x_{N_2}, \dots\},$$

where without loss of generality, we assume $N_{k-1} \leq N_k$. Observe that $\|x_{N_k} - x_{N_{k-1}}\| < \frac{1}{2^{k-1}}$, so we have

$$\begin{aligned}
\sum_{k=1}^{\ell} \|x_{N_k} - x_{N_{k-1}}\| &< \|x_{N_1}\| + \frac{1}{2} + \frac{1}{2^2} + \cdots + \frac{1}{2^{\ell-1}}, \\
\lim_{\ell \to \infty} \sum_{k=0}^{\ell} \|x_{N_k} - x_{N_{k-1}}\| &< \lim_{\ell \to \infty} \left( \|x_{N_1}\| + \frac{1}{2} + \frac{1}{2^2} + \cdots + \frac{1}{2^{\ell-1}} \right), \\
&= \|x_{N_1}\| + 1.
\end{aligned}$$

Since $\sum_{k=1}^{\infty} \|x_{N_k} - x_{N_{k-1}}\|$ converges, we see that $\sum_{k=1}^{\infty} (x_{N_k} - x_{N_{k-1}})$ converges to some $x \in X$. We will show that $(x_n)$ converges to $x$; if $\sum_{k=1}^{\infty} (x_{N_k} - x_{N_{k-1}}) = x$, then

$$\lim_{\ell \to \infty} \left\| \sum_{k=1}^{\ell} (x_{N_k} - x_{N_{k-1}}) - x \right\| = \lim_{\ell \to \infty} \|x_{N_\ell} - x\| = 0.$$

As such, for $\varepsilon > 0$ there exists $L \in \mathbb{N}$ such that $\|x_{N_\ell} - x\| < \varepsilon/2$ for $\ell \geq L$. Now consider $N_k \geq L$ such that

$$|x_n - x_{N_k}| < \frac{1}{2^k} < \frac{\varepsilon}{2}$$

for all $n \geq N_k$. And so, if $n \geq N_k \geq L$, we have

$$\begin{aligned}\|x - x_n\| &\leq \|x_n - x_{N_k}\| + \|x_{N_k} - x\|,\\ &\leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2},\\ &< \varepsilon,\end{aligned}$$

which shows that $(x_n)$ converges to $x$. Therefore, we conclude that $X$ is complete. □

**Theorem 15.3.** If $(X, \Sigma, \mu)$ is a measure space and $1 \leq p < \infty$, then $L^p(X, \mu)$ is complete.

*Proof.* To show that $L^p(X, \mu)$ is complete, it suffices to show that if $(f_n) \subseteq L^p(X, \mu)$ is a sequence where $\sum_{n=1}^{\infty} \|f_n\|_p$ converges, then $\sum_{n=1}^{\infty} f_n$ converges in $L^p(X, \mu)$ (Lemma 15.2).

First, define

$$g_k(x) := \sum_{n=1}^{k} |f_n(x)|.$$

By Minkowski's Inequality, we have

$$\|g_k\|_p = \left\| \sum_{n=1}^{k} |f_n| \right\|_p \leq \sum_{n=1}^{k} \|f_n\|_p$$

and thus

$$\|g_k\|_p \leq \sum_{n=1}^{\infty} \|f_n\|_p.$$

Notice that $(g_k(x))$ is a monotone increasing sequence bounded above by $\sum_{n=1}^{\infty} |f_n(x)|$ at each $x \in X$, so there exists some function $g$ such that $(g_k(x))$ converges pointwise to $g(x)$. Note that $g < \infty$ almost everywhere. We have

$$|g_k(x)| \leq \sum_{n=1}^{\infty} |f_n(x)|,$$

and because $\sum_{n=1}^{\infty} |f_n(x)|$. □

# 16 Covering Theorems of Vitali and Besicovitch

**Definition 16.1.** If $A \subseteq \mathbb{R}^n$ and $\mathcal{B}$ is a collection of balls covering $A$, then we say that $\mathcal{B}$ is a *Vitali cover* if for every $x \in A$ and $\varepsilon > 0$, there exists a ball $B \in \mathcal{B}$ containing $x$ such that $\operatorname{diam}(B) \leq \varepsilon$..

**Theorem 16.2** (Vitali Covering Theorem in $\mathbb{R}^n$). If $\mathcal{B}$ is a Vitali cover of $A \subseteq \mathbb{R}^n$, then there exists a subcollection $\{B_k\}_{k=1}^\infty \subseteq \mathcal{B}$ that covers $A$ almost everywhere.

*Proof.* Without loss of generality, suppose that $A$ is bounded. Otherwise, take $\{R_k\}_{k=1}^\infty$ to the countable collection of lattice squares that cover $\mathbb{R}^n$ and consider the intersection

$$A \cap \left( \bigcup_{k=1}^\infty R_k \right) = \bigcup_{k=1}^\infty (A \cap R_k).$$

Notice that $A \cap R_k$ is bounded, and we have expressed $A$ as a countable union of bounded sets.

Assuming $A$ is bounded, let $\mathcal{B} := \{B_\lambda\}_{\lambda \in \Lambda}$. For simplicity, we will also assume that each $B \in \mathcal{B}$ intersects $A$ and has radius less than $n$.We will construct a finite subcollection through the following steps:

1. Take $B_1 \in \mathcal{B}$ such that
$$0.9 \cdot \sup \underbrace{\{\operatorname{rad}(B) : B \in \mathcal{B}\}}_{\text{denote by } S_0} \leq \operatorname{rad}(B_1).$$
2. After constructing $B_1, B_2, \ldots, B_{k-1}$, we take $B_k \in \mathcal{B}$ such that
$$0.9 \cdot \sup \underbrace{\{\operatorname{rad}(B) : B \in \mathcal{B} \text{ and } B_n \text{ is pairwise disjoint from } B_1, \ldots, B_{k-1}\}}_{\text{denote by } S_{k-1}} \leq \operatorname{rad}(B_k)$$

Having obtained $\{B_k\}_{k=1}^\infty$, observe that if $x \in A \setminus \bigcup_{k=1}^\infty$ and $B \in \mathcal{B}$ covers $x$, then $B \cap B_k \neq \varnothing$ for some $k \in \mathbb{N}$. Suppose instead $B \cap B_n = \varnothing$ for all $n \in \mathbb{N}$. Notice that $(\operatorname{rad}(B_k)) \to 0$, so there exists a minimal $K \in \mathbb{N}$ such that $\operatorname{rad}(B_k) \leq \operatorname{rad}(B)$ for all $k \geq K$. We have

$$\operatorname{rad}(B_K) \leq \operatorname{rad}(B) \leq \operatorname{rad}(B_{K-1}).$$

By construction of $\{B_k\}_{k=1}^\infty$, we have

$$0.9 \cdot \sup(S_{K-1}) \leq \operatorname{rad}(B_K).$$

Notice that $\operatorname{rad}(B) \leq \sup(S_{K-1})$ since

$\square$