GRADUATE REAL ANALYSIS A
LECTURE NOTES

BY
OMER TAMUZ
WITH
LUCAS ABOUNADER AND SANTIAGO ADAMS

California Institute of Technology

2023
Acknowledgments

We are indebted to Kimberly Golubeva for her work on these lecture notes.
## Contents

1 Measures
   1.1 Measuring and why it is hard .................................................. 5
   1.2 Algebras, σ-algebras and measures ........................................... 6
   1.3 Outer measures ................................................................. 9
   1.4 10/6: Carathéodory and extending measures (by Lucas Abounader and Santiago Adams) 11
   1.5 10/9: Borel measures on \( \mathbb{R} \) ......................................... 14

2 Integration .............................................................................. 15
   2.1 10/11: Measurable functions .................................................. 15
   2.2 10/13: Integration of nonnegative functions ............................. 15
   2.3 10/16: Dominated convergence (by Joey Litvin and Bharathan Sundar) 15
   2.4 10/18: Modes of convergence ................................................... 15
   2.5 10/20: Measures on \( \mathbb{R}^n \) (by Jonah Yoshida and Wei Hou) ........ 15

3 Differentiation ...................................................................... 16
   3.1 10/23: Signed measures ......................................................... 16
   3.2 10/25: Lebesgue-Radon-Nikodym (by Mohammed Said Alhalimi & Eric Paul) 16
   3.3 10/27: Regular signed measures and differentiation ................. 16
   3.4 10/30: Total variation ............................................................ 16
   3.5 11/1: Bounded variation (by Tal Hershko and Elizabeth Xiao) .... 16

4 Topological vector spaces ......................................................... 17
   4.1 11/3: Normed vector spaces .................................................. 17
   4.2 11/6: Linear functionals ......................................................... 17
   4.3 11/8: Baire Category Theorem (by Holly Krynicki and Lara San Martin Suarez) 17
   4.4 11/10: Seminorms (by Edward Hou and Shiyang Shen) ............ 17
   4.5 11/13: Hilbert spaces I .......................................................... 17
   4.6 11/15: Hilbert spaces II (by Samuel Goodman and Brian Yang) ... 17

5 \( L^p \) spaces ......................................................................... 18
   5.1 11/17: Hölder’s & Minkowski’s inequalities .............................. 18
   5.2 11/20: Distribution functions .................................................. 18
   5.3 11/22: Interpolation (by Eric Ma) ............................................ 18

6 Radon Measures ................................................................... 19
   6.1 11/27: Riesz Representation Theorem .................................. 19
   6.2 11/29: Regularity and approximation ..................................... 19
   6.3 12/1: Dual of \( C_0 \) (by Gu ZhaoXing and Zhaojun Chen) ......... 19

7 Exercises ............................................................................ 20
Disclaimer

This a not a textbook. These are lecture notes.
1 Measures

1.1 Measuring and why it is hard

What is the area of the open unit disk \( D = \{(x, y) : x^2 + y^2 < 1\} \)? This is an old question, with a (maybe) simple answer. An interesting question that came much later is: What do we formally mean by “area”? For the unit disk there are a few different formal definitions, that all give the same answer. One of them is the following: Start by saying that a rectangle \([a, b] \times [c, d] \subset \mathbb{R}^2\) has area \((b - a)(d - c)\). Say that the disjoint union of a finite number of rectangles has area which is equal to the sum of the areas of the rectangles (here we already need to check that this is consistent, and it is). Finally, let the area of the disk equal the supremum of all areas of finite disjoint unions of rectangles that are contained in it.

While this gives a good answer for disks and other nice (i.e., open) shapes like disks. It has some nice properties:

1. The area of the unit square is one.
2. The area of the unit disk is \(\pi\).
3. The area of a subset \(A\) is equal to the area of its translate \(A + (x, y)\).
4. The area of a countable disjoint union \(A_1 \cup A_2 \cup \cdots\) of open sets is equal to the sum of the areas of \(A_1, A_2, \ldots\).

But this notion fails it fails for sets that are not as nice as \(D\). For example, what is the area of the unit disk minus its rational points: \(D \setminus \mathbb{Q}^2\)? According to this definition the answer is zero. The same holds for the complement of this set inside \(D\), and so we do not have additivity for all sets. The same issue appears more generally in \(\mathbb{R}^n\).

Ideally, what we want is a map \(\mu : 2^{\mathbb{R}^n} \to [0, \infty]\) with the following properties:

1. \(\mu([0,1]) = 1\).
2. If \(A_1, A_2, \ldots\) are pairwise disjoint subsets of \(\mathbb{R}^2\) then
   \[\mu(\bigcup_n A_n) = \sum_n \mu(A_n).\]
3. \(\mu(A + x) = \mu(A)\) for any \(x \in \mathbb{R}^n\) and subset \(A\) of \(\mathbb{R}^n\).

Unfortunately, such a map does not exist. To see this, consider already \(n = 1\). For simplicity, we will in fact consider \(S^1 = \{e^{2\pi i x} : x \in \mathbb{R}\} \subset \mathbb{C}\), and show that there does not exist a measure \(\mu : 2^{S^1} \to [0, 1]\) such that

1. \(\mu(S^1) = 1\).
2. If \(A_1, A_2, \ldots\) are pairwise disjoint subsets of \(S^1\) then
   \[\mu(\bigcup_n A_n) = \sum_n \mu(A_n).\]
Table 1: The construction of \((A_n)_n\). Each column corresponds to some \(A_n\), and each row is an equivalence class. It is easy to see in this table that \(A_n \cap A_m\) is empty when \(n \neq m\).

3. \(\mu(A \cdot z) = \mu(A)\) for any \(z \in S^1\) and subset \(A\) of \(S^1\).

For \(w \in [0,1)\) define the map \(T_w : S^1 \to S^1\) by \(T_w(z) = z \cdot e^{2\pi i w}\). This is simply a rotation by \(2\pi w\) radians. The third property is equivalent to \(\mu(T_w(A)) = \mu(A)\) for all \(A \subset S^1\).

Define an equivalence relation on \(S^1\) by \(e^{2\pi i x} \sim e^{2\pi i y}\) if \(x - y \in \mathbb{Q}\). Equivalently, \(z \sim z'\) if there is \(w \in \mathbb{Q} \cap [0,1)\) such that \(T_w(z) = z'\). Denote by \([x] = \{x \cdot e^{2\pi i q} : q \in \mathbb{Q} \cap [0,1)\}\) the equivalence class of \(x\). Choose a representative for each equivalence class, and let \(V\) be the set of representatives.

Let \(W = \{w_1, w_2, \ldots\} = \mathbb{Q} \cap [0,1)\) be an enumeration of \(\mathbb{Q} \cap [0,1)\), and let \(A_n = T_{w_n}(V)\). We claim that \(A_n \cap A_m = \emptyset\) if \(n \neq m\). To see this, suppose that \(T_{w_n}(z) \in A_n\) and \(T_{w_m}(z') \in A_m\) for \(z, z' \in A\). If \([z] = [z']\) then \(z = z'\), since both are representatives, and hence \(T_{w_n}(z) = T_{w_m}(z')\). If \([z] \neq [z']\) then \([T_{w_n}(z)] = [z] \neq [z'] = [T_{w_m}(z')]\), and so \(T_{w_m}(z) \neq T_{w_m}(z')\) (see Table 1). It follows that

\[
\mu(\cup_n A_n) = \sum_n \mu(A_n).
\]

Note the union \(\cup_n A_n\) is equal to \(S^1\), since \(T_{w_1}(z) + T_{w_2}(z), \ldots\) \([z]\), and so the left-hand side is equal to 1. But \(\mu(A_n) = \mu(A)\) by translation invariance, and so the right-hand side cannot equal 1.

Note that if we relax the countble additivty requirement and only require finite additivity, then such measures do exist. For \(n = 2\) we can even have a measure that is also invariant to rotations. However, for \(n \geq 3\) it is impossible, even with finite additivity, to be invariant to both translations and rotations. This is a consequence of the Banach-Tarski paradox.

### 1.2 Algebras, \(\sigma\)-algebras and measures

In the previous lecture we proved that we could not measure the length / area / volume of general sets in \(\mathbb{R}^n\). Our solution will be to restrict ourselves to measures that are only defined on a subset \(\Sigma \subset 2^{\mathbb{R}^n}\) of all sets.

For simplicity, consider the case of \(\mathbb{R}\). Which set \(\Sigma\) should we choose? We want to make it as big as possible, so that we have many sets we can measure, but we cannot make it too big. We want to include open (say) intervals in \(\Sigma\), since we know their measure. We want our measure to be additive, and so we want to include countable unions of open intervals. And likewise we want to include their complements. More generally, once we have included a set
in $\Sigma$, we want to also include its complement, and we want to include all possible countable unions of all sets we have already included.

If we are more generally trying to measure a set $\Omega$ (rather than $\mathbb{R}^n$ specifically) then we want $\Sigma$ to have the following properties:

1. $\Omega \in \Sigma$.
2. If $A \in \Sigma$ then $\Omega \setminus A \in \Sigma$.
3. If $A_1, A_2, \ldots \in \Sigma$ then $\bigcup A_n \in \Sigma$.

A collection of sets $\Sigma$ that satisfies these properties is called a $\sigma$-algebra. We say that $\Sigma$ is an algebra if it closed to finite unions, rather than countable unions. The pair $(\Omega, \Sigma)$ is called a measurable space.

Some (trivial) examples of $\sigma$-algebras:

1. $\Sigma = \{\emptyset, \Omega\}$. This is the trivial $\sigma$-algebra. It is contained in every $\sigma$-algebra on $\Omega$.
2. $\Sigma = 2^\Omega$. This $\sigma$-algebra contains every $\sigma$-algebra on $\Omega$.
3. $\Sigma$ is finite if and only if there is a partition $\Omega = A_1 \cup A_2 \cup \ldots \cup A_n$ such that $\Sigma$ consists of all unions of partitions elements.
4. If $\Omega$ is countable then every $\sigma$-algebra is of the form above, but where the partition is not necessarily finite.
5. For $\Omega = \Omega_1 \times \Omega_2$, $\Sigma_1 = \{A_1 \times \Omega_2 : A_1 \subseteq \Omega_1\}$.

For a non-trivial example (which is very important in probability), let $\Omega = \mathbb{R}^N$, and let the tail $\sigma$-algebra be given by

$$T = \{A \subseteq \Omega : \forall \omega \in A \forall n \in \mathbb{N} \forall x \in \mathbb{R} \ (\omega_1, \ldots, \omega_{n-1}, x, \omega_{n+1}, \ldots) \in A\}.$$

An example of $A \in T$ is

$$A = \{\omega : \lim_n \omega_n = 17\}.$$ 

Returning to our case of $\Omega = \mathbb{R}$, we want to include in $\Sigma$ the open intervals. So, for $\Sigma$ to be a $\sigma$-algebra, we have to include in it many more sets, because once we have added the open intervals we also have countable unions of open intervals, then we have the complements of these, then we have countable unions of these, etc.

To construct our $\sigma$-algebra, we make the following observation:

**Lemma 1.1.** Let $(\Sigma_i)_{i \in I}$ be a collection of $\sigma$-algebras of a set $\Omega$. Then $\bigcap_i \Sigma_i$ is a $\sigma$-algebra.
It follows that for any collection $\Theta \subset 2^\Omega$ there is a minimal $\sigma$-algebra that contains $\Theta$: the intersection of all $\sigma$-algebras that contain $\Theta$. Note that this set is non-empty because it includes $2^\Omega$. We denote it by $\mathcal{M}(\Theta)$.

When $\Omega$ is a topological space, the minimal $\sigma$-algebra that includes all the open sets is called the Borel $\sigma$-algebra of $\Omega$, and is denoted by $\mathcal{B}_\Omega$. This is the $\sigma$-algebra for which we will construct a measure on $\mathbb{R}$. For the case of $\mathbb{R}$, it is useful to note that $\mathcal{B}_\mathbb{R}$ is also equal to $\mathcal{M}(\Theta)$, where $\Theta$ is (for example) the collection of the half-open intervals: $\Theta = \{(a, b) : a < b \in \mathbb{R}\}$.

Given a $\sigma$-algebra $\Sigma$ over a set $\Omega$, a measure is a map $\mu : \Sigma \to [0, \infty]$ such that

1. $\mu(\emptyset) = 0$.

2. If $A_1, A_2, \ldots \in \Sigma$ are pairwise disjoint, then

$$\mu(\bigcup_n A_n) = \sum_n \mu(A_n).$$

We say that $\mu$ is finite if $\mu(A) < \infty$ for all $A \in \Sigma$. We say that it is $\sigma$-finite if there exist $A_1, A_2, \ldots \in \Sigma$ such that $\bigcup_n A_n = \Omega$ and $\mu(A_n) < \infty$ for all $n$. In this course we will focus on finite and $\sigma$-finite measures. Maps that satisfy only finite additivity are called finitely additive measures.

Our goal in the next two lectures will be to prove the following theorem:

**Theorem 1.2.** There exists a (unique) measure $\mu : \mathcal{B}_\mathbb{R} \to [0, \infty]$ such that $\mu((a, b)) = b - a$ for all $b \geq a$.

Note that this theorem implies that this measure is translation invariant, i.e., that $\mu(A) = \mu(A + x)$ for all $A \in \mathcal{B}$ and $x \in \mathbb{R}$, since for a fixed $x$, the map $\mu_x(A) = \mu(A + x)$ is also a measure such that $\mu_x([a, b]) = b - a$.

The following are basic and useful observation about measures:

**Claim 1.3.** Let $\mu : \Sigma \to [0, \infty]$ be a measure. Then

1. If $A \subseteq B \subseteq \Sigma$ then $\mu(A) \leq \mu(B)$.

2. If $A_1, A_2, \ldots \in \Sigma$ then $\mu(\bigcup_n A_n) \leq \sum_n \mu(A_n)$.

3. If $A_1 \subseteq A_2 \subseteq \cdots \in \Sigma$ then $\mu(\bigcup_n A_n) = \lim_n \mu(A_n)$.

4. If $A_1 \supseteq A_2 \supseteq \cdots \in \Sigma$ then $\mu(\bigcap_n A_n) = \lim_n \mu(A_n)$, provided $\mu(A_n) < \infty$ for some $n$.

The proof of the first claim follows from the fact that $B = A \cup B \setminus A$ and (finite) additivity. To see the third claim, define $A_0 = \emptyset$ and $B_n = A_n \setminus A_{n-1}$. Then $B_1, B_2, \ldots$ are pairwise disjoint and $\bigcup_n B_n = \bigcup_n A_n$. We can now apply additivity to conclude that $\mu(\bigcup_n B_n) = \sum_n \mu(B_n)$. Since $\sum_n \mu(B_n) = \lim_n \sum_{m=1}^n \mu(B_m) = \lim_n \mu(A_n)$ (by additivity) we are done.
1.3 Outer measures

Let $\Theta$ be the set of open rectangles in the plane:

$$\Theta = \{(a,b) \times (c,d) \subseteq \mathbb{R}^2\}.$$  

We denote the area of a rectangle $R = (a,b) \times (c,d)$ by $\rho(R) = (b-a) \cdot (d-c)$. Given any subset $A \subseteq \mathbb{R}^2$, a natural upper bound to its area is what is known as its outer measure:

$$\mu^*(A) = \inf \left\{ \sum_n \rho(R_n) : R_1, R_2, \ldots \in \Theta, A \subseteq \bigcup_n R_n \right\}.$$  \hspace{1cm} (1.1)

This defines a map $\mu^*: 2^{\mathbb{R}^2} \to [0,\infty]$. One can verify that—unlike the map proposed in Lecture 1.1—this one gives the desired answer for $D \setminus \mathcal{Q}^2$, as well to its complement in $D, D \cap \mathcal{Q}^2$. We will prove the former later in this lecture. To see the latter, enumerate $A = D \cap \mathcal{Q}^2 = \{(q_1,p_1),(q_2,p_2),\ldots\}$, fix any $Q \ni \epsilon > 0$ and let $R_n = [q_n - \epsilon 2^{-n/2}, q_n + \epsilon 2^{-n/2}] \times [p_n - \epsilon 2^{-n/2}, p_n + \epsilon 2^{-n/2}]$, so that $\rho(R_n) = 4\epsilon 2^{-n}$. Then $A \subseteq \bigcup_n R_n$, and $\sum_n \rho(R_n) = 4\epsilon$.

More generally, let $\Omega$ be a nonempty set, and let $\Theta$ be a collection of subsets of $\Omega$ that includes $\emptyset$ and $\Omega$. Let $\rho: \Theta \to [0,\infty]$ be such that $\rho(\emptyset) = 0$, and define $\mu^*: 2^\Omega \to [0,\infty]$ as in (1.1). Then

**Claim 1.4.**

1. $\mu^*(\emptyset) = 0$.

2. For all $A \subseteq B \subseteq \Omega$ it holds that $\mu^*(A) \leq \mu^*(B)$.

3. For all $A_1, A_2, \ldots \subseteq \Omega$, $\mu^*(\bigcup_n A_n) \leq \sum_n \mu^*(A_n)$.

**Proof.** (1) and (2) are left to the reader. For (3), denote $A = \bigcup_n A_n$ and fix $\epsilon > 0$. Then for each $A_n$ there is a sequence $E_1^n, E_2^n, E_3^n, \ldots \in \Theta$ such that $A \subseteq \bigcup_k E_k^n$ and $\sum_k \rho(E_k^n) \leq \mu^*(A_n) + \epsilon 2^{-n}$. Then $A \subseteq \bigcup_n \bigcup_k E_k^n$ and $\sum_{n,k} \rho(E_k^n) \leq \epsilon$. \hfill \qedsymbol

Any map $\mu^*: 2^\Omega$ that satisfies the properties of this claim is called an outer measure. Returning to our example of $A = D \setminus \mathcal{Q}^2$, we note that $A \subseteq D$, and so $\mu^*(A) \leq \mu^*(D)$. On the other hand, $D = (D \setminus \mathcal{Q}^2) \cup (D \cap \mathcal{Q}^2)$, and so

$$\mu^*(D) \leq \mu^*(D \setminus \mathcal{Q}^2) + \mu^*(D \cap \mathcal{Q}^2) = \mu^*(D \setminus \mathcal{Q}^2).$$

It will be useful to define our outer measures $\mu^*$ using a $\rho$ that has more structure. In the case of $\Omega = \mathbb{R}$, we can indeed extend $\rho$ beyond intervals to the algebra generated by the intervals. In particular, let $\Theta_0$ be any interval of the form $(-\infty, b)$. Let $\mathcal{A}$ be the algebra of sets generated by $\Theta_0$. This includes the complements of the sets in $\mathcal{A}$, i.e., intervals of the form $(a, \infty)$, intersections of these sets, i.e., intervals of the form $(a, b]$, and finite unions of such intervals.

Let $\rho: \mathcal{A} \to [0,\infty]$ assign to each $A \in \mathcal{A}$ its length in the obvious way. Then

**Claim 1.5.**

1. $\rho(\emptyset) = 0$. 

2. If $R_1, R_2, \ldots \in \mathcal{A}$ are pairwise disjoint and $R = \cup_n R_n$ is in $\mathcal{A}$ then $\rho(R) = \sum_n \rho(R_n)$.

The proof of (1) is immediate. The proof of (2) takes some work, but we will skip it.

A map $\rho$ from an algebra to $[0, \infty]$ that satisfies these properties is called a premeasure. Consider now the outer measure $\mu^*$ defined by a premeasure $\rho$.

**Claim 1.6.** For every $R \in \mathcal{A}$, $\mu^*(R) = \rho(R)$.

**Proof.** Clearly $\mu^*(R) \leq \rho(R)$, since $R \subseteq \cup_n R_n$ for $R = R_1 = R_2 = \cdots$. For the other direction, choose $R_1, R_2, \ldots \in \mathcal{A}$ such that $R \subseteq \cup_n R_n$ and $\sum_n \rho(R_n) \leq \mu^*(R) + \varepsilon$. Define $S_n = R \cap R_n$. Then $R = \cup_n S_n$. Let $T_n = S_n \setminus \cup_{k=1}^{n-1} S_n$. Then again $R = \cup_n T_n$, and now $T_1, T_2, \ldots$ are pairwise disjoint. Since they are also in $\mathcal{A}$, we have that

$$\rho(R) = \sum_n \rho(T_n) \leq \sum_n \rho(R_n) \leq \mu^*(R) + \varepsilon,$$

and so $\rho(R) \leq \mu^*(R)$.

The next claim is an even more important property of $\mu^*$.

**Claim 1.7.** For every $R \in \mathcal{A}$ it holds for every $A \subseteq \Omega$ that

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

**Proof.** Since $A \cap R$ and $A \cap R^c$ are disjoint, it follows by subadditivity that

$$\mu^*(A) \leq \mu^*(A \cap R) + \mu^*(A \cap R^c).$$

For the other direction, find $R_1, R_2, \ldots \in \mathcal{A}$ such that $A \subseteq \cup_n R_n$ and $\sum_n \rho(R_n) \leq \mu^*(A) + \varepsilon$. Let $S_n = R_n \cap R$ and $T_n = R_n \cap R^c$. Then

$$\sum_n \rho(R_n) = \sum_n \rho(S_n) + \sum_n \rho(T_n),$$

by the countable additivity of $\rho$. Note that for $S_n \subseteq \cup_n S_n$ and $A \cap R^c \subseteq \cup_n T_s$. Hence

$$\mu^*(A) + \varepsilon \geq \sum_n \rho(S_n) + \sum_n \rho(T_n) \geq \mu^*(A \cap R) + \mu^*(A \cap R^c)$$

This claim can be interpreted as follows: A set $R \in \mathcal{A}$ has the nice property that if we define the two maps $\mu^*_1(A) = \mu^*(A \cap R)$ and $\mu^*_2(A) = \mu^*(A \cap R^c)$ then $\mu^* = \mu^*_1 + \mu^*_2$, as we expect from measures.

More generally, given any outer measure $\mu^*$ defined on $\Omega$ using some $\Theta$ and $\rho$, we will say that $S \subseteq \Omega$ is $\mu^*$-measurable if

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

for all $A \subseteq \Omega$. Note that for $A$ that contains $S$, this translates to

$$\mu^*(S) = \mu^*(A) - \mu^*(A \cap S^c)$$

which is close to what we tried to do in the first lecture.
1.4 10/6: Carathéodory and extending measures (by Lucas Abounader and Santiago Adams)

In this lecture we will state (and prove!) Carathéodory’s Extension Theorem, which allows us to take an arbitrary outer measure $\mu^*: 2^{\Omega} \to [0, \infty]$ and construct a measure $\mu = \mu^*|:\Sigma \subset 2^{\Omega} \to [0, \infty]$. In particular, we will find a good way to exclude the pathological sets (that confounded us in Lecture 1.1) in specifying the $\sigma$-algebra $\Sigma$, by looking only at the “nice” $\mu^*$-measurable sets. Since premeasures induce outer measures (by 1.1), we will ultimately be able to construct measures from arbitrary premeasures as well.

For those that have previously seen a construction of the Lebesgue measure, recall that one starts with the more primitive concept of Lebesgue outer measure (and, preceding this, the specification of a desired premeasure on elementary sets); this procedure is now presented in a slightly more abstract setting. The punchline here (in a way, justifying abstraction) is some guarantee of uniqueness, given the premeasure. Thus, when we apply the results proven in this lecture to construct, e.g., Lebesgue-Stieltjes, Hausdorff, or product measures, we can pat ourselves on the back, satisfied with the knowledge that the fruits of our labor are the unique objects we’re looking for.

- Recap of outer measure ($\mu^*: 2^{\Omega} \to [0, \infty]$ satisfying $\mu^*(\emptyset) = 0$, monotonicity for $A \subseteq B$, and countable subadditivity).

- Recap of $\mu^*$-measurable ($\forall A, \mu^*(A) = \mu^*(A \cap R) + \mu^*(A \cap R^c) = \mu^*(A \cap R) + \mu^*(A \setminus R)$), highlighting that we want the set $R$ to be able to split any set $A$ in a way that recovers $\mu^*(A)$.

**Theorem 1.8** (Carathéodory). If $\mu^*$ is an outer measure on $\Omega$, then the collection $\Sigma$ of $\mu^*$-measurable subsets of $\Omega$ is a $\sigma$-algebra, and the restriction of $\mu^*$ to $\Sigma$ is a measure.

**Proof.** We start with the claim that $\Sigma$ is more weakly an algebra:

- (nonempty). $\emptyset \in \Sigma$. In fact, for any null set $N$ ($\mu^*(N) = 0$), we have $N \in \Sigma$:

  \[
  \mu^*(A) \leq \mu^*(A \cap N) + \mu^*(A \setminus N) = \mu^*(A \setminus N) \leq \mu^*(A),
  \]

  applying monotonicity ($A \cap N \subseteq N \implies \mu^*(A \cap N) \leq 0$) and subadditivity in succession.

- (complements). Note the definition of $\Sigma$ is symmetric with respect to $R$ and $R^c$.

- (finite union). [Not going to go through this part in entirety; a bit tedious]. Suppose $A, B \in \Sigma$. For any $E \in 2^X$, we want to show

  \[
  \mu^*(E) = \mu^*(E \cap (A \cup B)) + \mu^*(E \setminus (A \cup B)). \tag{1.2}
  \]

  It helps to classify points in $E$ based upon their inclusions in $A$ and $B$. That is, with the partitions

  \[
  E_{00} := E \setminus (A \cup B); \quad E_{10} := (E \setminus B) \cap A; \quad E_{01} := (E \setminus A) \cap B; \quad E_{11} := E \cap A \cap B,
  \]

  11
we can convert 1.2 into the equivalent
\[ \mu^*(E_{00} \cup E_{10} \cup E_{01} \cup E_{11}) = \mu^*(E_{10} \cup E_{01} \cup E_{11}) + \mu^*(E_{00}). \] (1.3)

Using \( A \) to split the sets \( E \) and \( E \setminus E_{00} \) (applying \( \mu^* \)-measurability):
\[ \mu^*(E_{00} \cup E_{10} \cup E_{01} \cup E_{11}) = \mu^*(E_{10} \cup E_{11}) + \mu^*(E_{00} \cup E_{01}) \] (1.4)
and
\[ \mu^*(E_{10} \cup E_{01} \cup E_{11}) = \mu^*(E_{10} \cup E_{11}) + \mu^*(E_{01}), \] (1.5)
while using \( B \) to split \( E \setminus A \) we obtain
\[ \mu^*(E_{00} \cup E_{01}) = \mu^*(E_{01}) + \mu^*(E_{00}). \] (1.6)

Collecting 1.4, 1.5 and 1.6 we obtain 1.3, as desired.

Now we extend finite unions to countable ones:

• (\( \sigma \)-algebra). It suffices to consider only countable disjoint unions. (Why? Given \((A_j)_{j \in \mathbb{N}}\), replace with \((B_j)_{j \in \mathbb{N}}\) defined by \(B_n = A_n \setminus \bigcup_{j=1}^{n-1} A_j\).) So we want
\[ \mu^*(A) = \mu^*(A \cap \bigcup_{j=1}^{\infty} E_j) + \mu^*(A \setminus \bigcup_{j=1}^{\infty} E_j) \]
for \((E_j)_{j=1}^{\infty} \subset \Sigma\) a collection of pairwise disjoint sets. \( \leq \) is given by subadditivity, so we only need to prove \( \geq \). Given \( N \), define \( B_N = \bigcup_{j=1}^{N} E_j \), and note that \( B_N \) is \( \mu^* \)-measurable.

For the difference,
\[ B_N \subseteq \bigcup_{j=1}^{N} E_j \implies \mu^*(A \setminus \bigcup_{j=1}^{N} E_j) \leq \mu^*(A \setminus B_N). \]

For the intersection, we want
\[ \mu^*(A \cap \bigcup_{j=1}^{\infty} E_j) \geq \lim_{N \to \infty} \mu^*(A \cap B_N). \]

Since the \( B_N \) are \( \mu^* \)-measurable, for any \( N \)
\[ \mu^*(A \cap B_{N+1}) = \mu^*(A \cap B_N) + \mu^*(A \cap E_{N+1} \setminus B_N), \]
leading to the expressions
\[ \mu^*(A \cap B_N) = \sum_{n=0}^{N-1} \mu^*(A \cap E_{n+1} \setminus B_n) \overset{\text{lim}}{\longrightarrow} \lim_{N \to \infty} \mu^*(A \cap B_N) = \sum_{n=0}^{\infty} \mu^*(A \cap E_{n+1} \setminus B_n). \]

Finally we note that \( \bigcup_{n=0}^{\infty} (A \cap E_{n+1} \setminus B_n) \) is just \( A \cap \bigcup_{n=1}^{\infty} E_n \), so by countable subadditivity
\[ \mu^*(A \cap \bigcup_{n=1}^{\infty} E_n) \leq \sum_{n=0}^{\infty} \mu^*(A \cap E_{n+1} \setminus B_n), \]
as desired.
The second part of the theorem states that $\mu = \mu^*|_{\Sigma}$ is in fact a measure. The only thing $\mu^*$ is missing is (σ-)additivity.

- (additivity). We need to show
  \[ \mu^*(\bigcup_{j=1}^{\infty} E_j) \geq \sum_{j=1}^{\infty} \mu^*(E_j) \]
  for $E_1, E_2, \cdots \in \Sigma$ disjoint. This turns out to be equivalent to finite additivity, as demonstrated:
  \[(\forall N): \sum_{j=1}^{N} \mu^*(E_j) \leq \mu^*(\bigcup_{j=1}^{N} E_j) \leq \mu^*(\bigcup_{j=1}^{\infty} E_j) = \sum_{j=1}^{\infty} \mu^*(E_j) \leq \mu^*(\bigcup_{j=1}^{\infty} E_j),\]
  and the finite case follows directly from the key $\mu^*$-measurability property of $\Sigma$: For $E, F \in \Sigma$ disjoint,
  \[ \mu^*(E \cup F) = \mu^*(E) + \mu^*(F). \]
  \[\square\]

- Recap of premeasure; $\mu_0 : \mathcal{A} \to [0, \infty]$ on an algebra $\mathcal{A}$ with $\mu_0(\emptyset) = 0$, and $\mu_0(\bigcup_{j=1}^{\infty} R_j) = \sum_{j=1}^{\infty} \mu_0(R_j)$ whenever $R_1, R_2, \cdots \in \mathcal{A}$ disjoint.

- Recall that
  \[ \mu^*_\rho(A) = \inf \left\{ \sum_{j=1}^{\infty} \rho(R_j) : R_j \in \Theta, A \subseteq \bigcup_{j=1}^{\infty} R_j \right\} \]
  is an outer measure for any function $\rho : \Theta \to [0, \infty]$ with $\emptyset, X \in \Theta$ and $\rho(\emptyset) = 0$, and thus for any premeasure $\mu_0$.

**Theorem 1.9.** Let $\mathcal{A}$ be an algebra of subsets of $\Omega$, let $\mu_0 : \mathcal{A} \to [0, \infty]$ be a premeasure, and let $\mathcal{B} = \mathcal{M}(\mathcal{A})$ be the σ-algebra generated by $\mathcal{A}$. Then $\mu = \mu^*|_{\mathcal{B}}$ is a measure on $\mathcal{B}$ that extends $\mu_0$, and this measure is the unique extension if $\mu_0$ is σ-finite. [If time] Even if $\mu_0$ is not σ-finite, given another extension $\nu$ of $\mu_0$, then we have $\nu(E) \leq \mu(E)$ for all $E$ (so $\mu$ is maximal) and $\nu(E) = \mu(E)$ whenever $\mu(E) < \infty$.

**Proof.**

- Equation 1.1 and Carathéodory’s Theorem (Theorem 1.8) present the construction of $\mu : \Sigma \to [0, \infty]$ on a (potentially larger!) σ-algebra $\Sigma \supseteq \mathcal{A}$ of $\mu^*$-measurable sets.

- Suppose that $E \in \mathcal{B}$, and take any cover $\bigcup_{j=1}^{\infty} A_j \supseteq E$ where $A_j \in \mathcal{A}$. Then if $\nu$ is another extension, $\nu(E) \leq \sum_{j=1}^{\infty} \nu(A_j) = \sum_{j=1}^{\infty} \mu_0(A_j)$, so $\nu(E) \leq \mu(E)$. Let $A = \bigcup_{j=1}^{\infty} A_j$; then
  \[ \nu(A) = \lim_{n \to \infty} \nu(\bigcup_{j=1}^{n} A_j) = \lim_{n \to \infty} \mu(\bigcup_{j=1}^{n} A_j) = \mu(A). \]
If $\mu(E) < \infty$ we can pick $A_j$ such that $\mu(A) < \mu(E) + \varepsilon$, so $\mu(A \setminus E) < \varepsilon$ and

$$\mu(E) \leq \mu(A) = \nu(A) = \nu(E) + \nu(A \setminus E) \leq \nu(E) + \mu(A \setminus E) \leq \nu(E) + \varepsilon.$$ 

$\varepsilon$ is arbitrary, hence $\nu(E) = \mu(E)$. This proves the last claim.

- Finally, suppose $\mu_0$ is $\sigma$-finite, and take a countable partition $\bigcup_{j=1}^{\infty} A_j = \Omega$ with $\mu_0(A_j) < \infty$ (we can use the familiar trick and assume the $A_j$ are disjoint). Then for any $E$:

$$\mu(E) = \sum_{j=1}^{\infty} \mu(E \cap A_j) = \sum_{j=1}^{\infty} \nu(E \cap A_j) = \nu(E).$$

As a Corollary, we can prove Theorem 1.2. Take the premeasure $\rho$ on the algebra $\mathcal{A}$ generated by $\{(a, b] : a, b \in \mathbb{R}\}$, use Theorem 1.9 to obtain a measure $\mu$ on $\Sigma = \mathcal{M}(\mathcal{A})$, and note that $\mathcal{M}(\mathcal{A}) = \mathcal{B}_\mathbb{R}$.

1.5 10/9: Borel measures on $\mathbb{R}$
2 Integration

2.1 10/11: Measurable functions

2.2 10/13: Integration of nonnegative functions

2.3 10/16: Dominated convergence (by Joey Litvin and Bharathan Sundar)

2.4 10/18: Modes of convergence

2.5 10/20: Measures on $\mathbb{R}^n$ (by Jonah Yoshida and Wei Hou)
3 Differentiation

3.1 10/23: Signed measures

3.2 10/25: Lebesgue-Radon-Nikodym (by MohammedSaid Alhalimi & Eric Paul)

3.3 10/27: Regular signed measures and differentiation

3.4 10/30: Total variation

3.5 11/1: Bounded variation (by Tal Hershko and Elizabeth Xiao)
4 Topological vector spaces

4.1 11/3: Normed vector spaces
4.2 11/6: Linear functionals
4.3 11/8: Baire Category Theorem (by Holly Krynicki and Lara San Martin Suarez)
4.4 11/10: Seminorms (by Edward Hou and Shiyang Shen)
4.5 11/13: Hilbert spaces I
4.6 11/15: Hilbert spaces II (by Samuel Goodman and Brian Yang)
5 \( L^p \) spaces

5.1 11/17: Hölder’s & Minkowski’s inequalities

5.2 11/20: Distribution functions

5.3 11/22: Interpolation (by Eric Ma)
6 Radon Measures

6.1 11/27: Riesz Representation Theorem

6.2 11/29: Regularity and approximation

6.3 12/1: Dual of $C_0$ (by Gu ZhaoXing and Zhaojun Chen)
7 Exercises

1. **Independent fair coin tosses.** Measures are important in probability, where they model the chance of uncertain outcomes. Probability measures are measures that assign unit measure to the entire space.

For example, let $\Omega = \{-1, +1\}^N$ for some $N \in \mathbb{N}$. We think of elements of $\Omega$ as functions $\omega: \{1, \ldots, N\} \rightarrow \{-1, +1\}$. The i.i.d. (independent and identically distributed) Bernoulli measure $\mu: 2^\Omega \rightarrow [0, 1]$ is given by $\mu(A) = 2^{-N}|A|$. To see why this captures the idea of independent tosses, define for $i \leq N$ the map $T_i: \Omega \rightarrow \Omega$ by

$$[T_i(\omega)](j) = \begin{cases} \omega(j) & \text{if } j \neq i \\ -\omega(j) & \text{if } j = i. \end{cases}$$

The measure $\mu$ is $T_i$-invariant: $\mu(T_i(A)) = \mu(A)$ for all $A \subseteq \Omega$, i.e., the probability of an outcome does not change if we consider the outcome in which the $i$th coin has the opposite sign.

Let $\mathcal{X} = \{-1, +1\}^\mathbb{N}$. We think of elements of $\mathcal{X}$ as functions $\omega: \mathbb{N} \rightarrow \{-1, +1\}$. We endow $\mathcal{X}$ with the topology of pointwise convergence: $\lim_n \omega_n = \omega$ if for all $i \in \mathbb{N}$ it holds that $\lim_n \omega_n(i) = \omega(i)$ (note that this implies that the sequence $\omega_n(i)$ eventually stabilizes on $\omega(i)$). With this topology, $\mathcal{X}$ is also known as the Cantor set.

(a) Prove that this topology is also the topology generated by the metric $D: \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}_{\geq 0}$ given by

$$D(\omega, \theta) = \inf\{2^{-N} : (\omega(1), \ldots, \omega(N)) = (\theta(1), \ldots, \theta(N))\}.$$

(b) Prove that $U \subset \mathcal{X}$ is clopen (both closed and open) if and only if there is some $N \in \mathbb{N}$ and $A \subset \{-1, +1\}^N$ such that

$$U = \{\omega \in \mathcal{X} : (\omega(1), \ldots, \omega(N)) \in A\}.$$

(c) Prove that $U \subset \mathcal{X}$ is open if and only if it is a countable union of clopen sets.

(d) We would like to extend our i.i.d. measure from the finite setting to the infinite setting. I.e., we would like to find a probability measure $\mu: 2^\mathcal{X} \rightarrow [0, 1]$ such that $\mu(T_i(A)) = \mu(A)$ for all $A \subseteq \mathcal{X}$ and $i \in \mathbb{N}$. Prove that this is impossible.

(e) Let $\mathcal{A}$ be the collection of clopen sets of $\mathcal{X}$. Prove that it is an algebra.

(f) Define $\rho: \mathcal{A} \rightarrow [0, 1]$ as follows. Given $A \subset \{-1, +1\}^\mathbb{N}$ and

$$U = \{\omega \in \mathcal{X} : (\omega(1), \ldots, \omega(N)) \in A\},$$

let $\rho(U) = 2^{-N}|A|$. Show that $\rho$ is a premeasure.

(g) Use Carathéodory’s Theorem to show that there is a probability measure on the Borel $\sigma$-algebra of $\mathcal{X}$ that is invariant to $\{T_i\}_{i \in \mathbb{N}}$. 

20
2. **Countable additivity from dominance.** Let \((\Omega, \Sigma)\) be a measurable space, and let \(\mu : \Sigma \to [0, \infty)\) be a measure. We say that \(\mu\) dominates a map \(\nu : \Sigma \to [0, \infty)\) if \(\mu(A) \geq \nu(A)\) for all \(A \in \Sigma\).

Show that if \(\mu\) is a finite measure (i.e., \(\mu(\Omega) < \infty\)), \(\nu\) is a finitely additive measure \(\nu\), and \(\mu\) dominates \(\nu\), then \(\nu\) is in fact a measure.

3. **Criticality of \(\sigma\)-finite hypothesis to Theorem 1.9.** Let \(\mathcal{A}\) be the collection of all subsets in \(\mathbb{R}\) that can be expressed as finite unions of half-open intervals \([a, b)\). Let \(\mu_0 : \mathcal{A} \to [0, \infty] \) be the function such that \(\mu_0(E) = \infty\) if \(E \neq \emptyset\) and \(\mu_0(\emptyset) = 0\).

(a) Show that \(\mu_0\) is a premeasure.

(b) Show that \(\mathcal{M}(\mathcal{A})\) is the Borel \(\sigma\)-algebra \(\mathcal{B}_\mathbb{R}\).

(c) Show that the extension \(\mu : \mathcal{B}_\mathbb{R} \to [0, \infty]\) of \(\mu_0\) defined by 1.1 and Theorem 1.8 assigns an infinite measure to any non-empty Borel set.

(d) Show that the counting measure \(\mu_c(A) = |A|\) is another extension of \(\mu_0\) on \(\mathcal{B}_\mathbb{R}\).

4. **Completing the measure extension.** Let \(\mu_0 : \mathcal{A} \to [0, \infty]\) be a premeasure which is \(\sigma\)-finite and let \(\mu : \Sigma \to [0, \infty]\) be its (unique!) extension on the \(\sigma\)-algebra \(\Sigma\) of \(\mu^*\)-measurable sets (recall that \(\Sigma\) may be even larger than \(\mathcal{M}(\mathcal{A})\), the \(\sigma\)-algebra generated by \(\mathcal{A}\)).

(a) Show that if \(E \in \Sigma\), then there is an \(F \in \mathcal{M}(\mathcal{A})\) containing \(E\) such that \(\mu(F \setminus E) = 0\) (thus \(F\) consists of the union of \(E\) and a null set). Furthermore, show that \(F\) can be chosen to be a countable intersection \(F = \bigcap_{j=1}^{\infty} F_j\) of sets \(F_j\), each of which is a countable union \(F_j = \bigcup_{k=1}^{\infty} F_{j,k}\) of sets \(F_{j,k} \in \mathcal{A}\) (i.e., \(F\) is an element of \(\mathcal{A}_{\sigma\delta}\)).

(b) If \(E \in \Sigma\) has finite measure, and \(\varepsilon > 0\), show that there exists \(F \in \mathcal{A}\) such that \(\mu(E \Delta F) \leq \varepsilon\), where

\[ E \Delta F := (E \cup F) \setminus (E \cap F) = (E \setminus F) \cup (F \setminus E) \]

is the symmetric difference of \(E\) and \(F\).

(c) Conversely, if \(E\) is a set such that for every \(\varepsilon > 0\) there exists \(F \in \mathcal{A}\) such that \(\mu^*(E \Delta F) \leq \varepsilon\), show that \(E \in \Sigma\).