X

Measure Theory

Part X: Contents

35 Measure spaces

Here is an outline of where we are going next. Our *goal* over the next few chapters is to develop the machinery to state (and in some cases prove) the law of large numbers and the central limit theorem. For these purposes, the scant amount of work we did in Calculus 101 is going to be awfully insufficient: integration over $\mathbb{R}$ (or even $\mathbb{R}^n$) is just not going to cut it.

This chapter will develop the theory of "measure spaces", which you can think of as "spaces equipped with a notion of size". We will then be able to integrate over these with the so-called Lebesgue integral (which in some senses is almost strictly better than the Riemann one).

§35.1 Letter connotations

There are a lot of "types" of objects moving forward, so here are the letter connotations we'll use throughout the next several chapters. This makes it easier to tell what the "type" of each object is just by which letter is used.

- Measure spaces denoted by Ω, their elements denoted by ω.
- Algebras and σ-algebras denoted by script $\mathscr{A}$, $\mathscr{B}$, ... Sets in them denoted by early capital Roman A, B, C, D, E, ...
- Measures (i.e. functions assigning sets to reals) denoted usually by μ or ρ.
- Random variables (functions sending worlds to reals) denoted usually by late capital Roman X, Y, Z, ...
- Functions from $\mathbb{R} \to \mathbb{R}$ by Roman letters like f and g for pdf's and F and G for cdf's.
- Real numbers denoted by lower Roman letters like x, y, z.

§35.2 Motivating measure spaces via random variables

To motivate *why* we want to construct measure spaces, I want to talk about a (real) **random variable**, which you might think of as

- the result of a coin flip,
- the high temperature in Boston on Saturday,
- the possibility of rain on your 18.725 date next weekend.

Why does this need a long theory to develop well? For a simple coin flip one intuitively just thinks "50% heads, 50% tails" and is done with it. The situation is a little trickier with temperature since it is continuous rather than discrete, but if all you care about is that one temperature, calculus seems like it might be enough to deal with this.

But it gets more slippery once the variables start to "talk to" each other: the high temperature tells you a little bit about whether it will rain, because e.g. if the temperature is very high it's quite likely to be sunny. Suddenly we find ourselves wishing we could talk about conditional probability, but this is a whole can of worms — the relations between these sorts of things can get very complicated very quickly.

The big idea to getting a formalism for this is that:

Our measure spaces Ω will be thought of as a space of entire worlds, with each $\omega \in \Omega$ representing a world. Random variables are functions from worlds to $\mathbb{R}$.

This way, the space of "worlds" takes care of all the messy interdependence.

Then, we can assign "measures" to sets of worlds: for example, to be a fair coin means that if you are only interested in that one coin flip, the "fraction" of worlds in which that coin showed heads should be $\frac{1}{2}$. This is in some ways backwards from what you were told in high-school: officially, we start with the space of worlds, rather than starting with the probabilities.

It will soon be clear that there is no way we can assign a well-defined measure to every single one of the 2^Ω subsets. Fortunately, in practice, we won't need to, and the notion of a σ-algebra will capture the idea of "enough measur-*able* sets for us to get by".

Remark 35.2.1 (Random seeds) — Another analogy if you do some programming: each $\omega \in \Omega$ is a *random seed*, and everything is determined from there.

§35.3 Motivating measure spaces geometrically

So, we have a set Ω of possible points (which in the context of the previous discussion can be thought of as the set of worlds), and we want to assign a *measure* (think volume) to subsets of points in Ω. We will now describe some of the obstacles that we will face, in order to motivate *how* measure spaces are defined (as the previous section only motivated *why* we want such things).

If you try to do this naïvely, you basically immediately run into set-theoretic issues. A good example to think about why this might happen is if $\Omega = \mathbb{R}^2$ with the measure corresponding to area. You can define the area of a triangle as in high school, and you can then try and define the area of a circle, maybe by approximating it with polygons. But what area would you assign to the subset $\mathbb{Q}^2$, for example? (It turns out "zero" is actually a working answer.) Or, a unit disk is composed of infinitely many points; each of the points better have measure zero, but why does their union have measure π then? Blah blah blah.

We'll say more about this later, but you might have already heard of the **Banach-Tarski paradox** which essentially shows there is no good way that you can assign a measure to every single subset of $\mathbb{R}^3$ and still satisfy basic sanity checks. There are just too many possible subsets of Euclidean space.

However, the good news is that most of these sets are not ones that we will ever care about, and it's enough to define measures for certain "sufficiently nice sets". The adjective we will use is *measurable*, and it will turn out that this will be way, way more than good enough for any practical purposes.

We will generally use A, B, ... for measurable sets and denote the entire family of measurable sets by curly $\mathscr{A}$.

§35.4 σ-algebras and measurable spaces

Here's the machine code.

Definition 35.4.1. A **measurable space** consists of a space Ω of points, and a **σ-algebra** $\mathscr{A}$ of subsets of Ω (the "measurable sets" of Ω). The set $\mathscr{A}$ is required to satisfy the following axioms:

- $\mathscr{A}$ contains $\varnothing$ and Ω.
- $\mathscr{A}$ should be closed under complements and *countable* unions/intersections. (Hint on nomenclature: σ usually indicates some sort of "countably finite" condition.)

(Complaint: this terminology is phonetically confusing, because it can be confused with "measure space" later. The way to think about is that "measur*able* spaces have a σ-algebra, so we *could* try to put a measure on it, but we *haven't*, yet.")

Though this definition is how we actually think about it in a few select cases, for the most part, and we will usually instantiate $\mathscr{A}$ in practice in a different way:

Definition 35.4.2. Let Ω be a set, and consider some family of subsets $\mathscr{F}$ of Ω. Then the **σ-algebra generated by $\mathscr{F}$** is the smallest σ-algebra $\mathscr{A}$ which contains $\mathscr{F}$.

As is commonplace in math, when we see "generated", this means we sort of let the definition "take care of itself". So, if $\Omega = \mathbb{R}$, maybe I want $\mathscr{A}$ to contain all open sets. Well, then the definition means it should contain all complements too, so it contains all the closed sets. Then it has to contain all the half-open intervals too, and then.... Rather than try to reason out what exactly the final shape $\mathscr{A}$ looks like (which basically turns out to be impossible), we just give up and say "$\mathscr{A}$ is all the sets you can get if you start with the open sets and apply repeatedly union/complement operations". Or even more bluntly: "start with open sets, shake vigorously".[1]

I've gone on too long with no examples.

Example 35.4.3 (Examples of measurable spaces)

The first two examples actually say what $\mathscr{A}$ is; the third example (most important) will use generation.

(a) If Ω is any set, then the power set $\mathscr{A} = 2^\Omega$ is obviously a σ-algebra. This will be used if Ω is countably finite, but it won't be very helpful if Ω is huge.

(b) If Ω is an uncountable set, then we can declare $\mathscr{A}$ to be all subsets of Ω which are either countable, or which have countable complement. (You should check this satisfies the definitions.) This is a very "coarse" algebra.

(c) If Ω is a topological space, the **Borel σ-algebra** is defined as the σ-algebra generated by all the open sets of Ω. We denote it by $\mathscr{B}(\Omega)$, and call the space a **Borel space**. As warned earlier, it is basically impossible to describe what it looks like, and instead you should think of it as saying "we can measure the open sets".

[1] As will be mentioned later in Section 36.4, an explicit construction using transfinite induction is possible. That construction is also useful for, for example, proving $|\mathscr{B}(\mathbb{R})| = |\mathbb{R}|$.

Question 35.4.4. Show that the closed sets are in $\mathscr{B}(\Omega)$ for any topological space Ω. Show that $[0, 1)$ is also in $\mathscr{B}(\mathbb{R})$.

§35.5 Measure spaces

Definition 35.5.1. Measurable spaces $(\Omega, \mathscr{A})$ are then equipped with a function $\mu \colon \mathscr{A} \to [0, +\infty]$ called the **measure**, which is required to satisfy the following axioms:

- $\mu(\varnothing) = 0$
- **Countable additivity**: If A_1, A_2, ... are disjoint sets in $\mathscr{A}$, then

$$\mu\left(\bigsqcup_n A_n\right) = \sum_n \mu(A_n).$$

The triple $(\Omega, \mathscr{A}, \mu)$ is called a **measure space**. It's called a **probability space** if $\mu(\Omega) = 1$.

Exercise 35.5.2 (Weaker equivalent definitions)**.** I chose to give axioms for $\mathscr{A}$ and μ that capture how people think of them in practice, which means there is some redundancy: for example, being closed under complements and unions is enough to get intersections, by de Morgan's law. Here are more minimal definitions, which are useful if you are trying to prove something satisfies them to reduce the amount of work you have to do:

(a) The axioms on $\mathscr{A}$ can be weakened to (i) $\varnothing \in \mathscr{A}$ and (ii) $\mathscr{A}$ is closed under complements and countable unions.

(b) The axioms on μ can be weakened to (i) $\mu(\varnothing) = 0$, (ii) $\mu(A \sqcup B) = \mu(A) + \mu(B)$, and (iii) for $A_1 \supseteq A_2 \supseteq \cdots$, we have $\mu\left(\bigcap_n A_n\right) = \lim_n \mu(A_n)$.

Remark 35.5.3 — Here are some immediate remarks on these definitions.

- If $A \subseteq B$ are measurable, then $\mu(A) \le \mu(B)$ since $\mu(B) = \mu(A) + \mu(B - A)$.
- In particular, in a probability space all measures are in $[0, 1]$. On the other hand, for general measure spaces we'll allow $+\infty$ as a possible measure (hence the choice of $[0, +\infty]$ as codomain for μ).
- We want to allow at least countable unions / additivity because with finite unions it's too hard to make progress: it's too hard to estimate the area of a circle without being able to talk about limits of countably infinitely many triangles.

We *don't* want to allow uncountable unions and additivity, because uncountable sums basically never work out. In particular, there is a nice elementary exercise as follows:

Exercise 35.5.4 (Tricky)**.** Let S be an uncountable set of positive real numbers. Show that some finite subset $T \subseteq S$ has sum greater than 10^{2019}. Colloquially, "uncountably many positive reals cannot have finite sum".

So countable sums are as far as we'll let the infinite sums go. This is the reason why we considered σ-algebras in the first place.

Example 35.5.5 (Measures)

We now discuss measures on each of the spaces in our previous examples.

(a) If $\mathscr{A} = 2^\Omega$ (or for that matter any $\mathscr{A}$) we may declare $\mu(A) = |A|$ for each $A \in \mathscr{A}$ (even if $|A| = \infty$). This is called the **counting measure**, simply counting the number of elements.

This is useful if Ω is countably infinite, and optimal if Ω is finite (and nonempty). In the latter case, we will often normalize by $\mu(A) = \frac{|A|}{|\Omega|}$ so that Ω becomes a probability space.

(b) Suppose Ω was uncountable and we took $\mathscr{A}$ to be the countable sets and their complements. Then

$$\mu(A) = \begin{cases} 0 & A \text{ is countable} \\ 1 & \Omega \setminus A \text{ is countable} \end{cases}$$

is a measure. (Check this.)

(c) Elephant in the room: defining a measure on $\mathscr{B}(\Omega)$ is hard even for $\Omega = \mathbb{R}$, and is done in the next chapter. So you will have to hold your breath. Right now, all you know is that by declaring my *intent* to define a measure $\mathscr{B}(\Omega)$, I am hoping that at least every open set will have a volume.

§35.6 A hint of Banach-Tarski

I will now try to convince you that $\mathscr{B}(\Omega)$ is a necessary concession, and for general topological spaces like $\Omega = \mathbb{R}^n$, there is no hope of assigning a measure to 2^Ω. (In the literature, this example is called a Vitali set.)

Example 35.6.1 (A geometric example why $\mathscr{A} = 2^\Omega$ is unsuitable)

Let Ω denote the unit circle in $\mathbb{R}^2$ and $\mathscr{A} = 2^\Omega$. We will show that any measure μ on Ω with $\mu(\Omega) = 1$ will have undesirable properties.

Let $\sim$ denote an equivalence relation on Ω defined as follows: two points are equivalent if they differ by a rotation around the origin by a rational multiple of π. We may pick a representative from each equivalence class, letting X denote the set of representatives. Then

$$\Omega = \bigsqcup_{\substack{q \in \mathbb{Q} \\ 0 \le q < 2}} (X \text{ rotated by } q\pi \text{ radians}).$$

Since we've only rotated X, each of the rotations should have the same measure m. But $\mu(\Omega) = 1$, and there is no value we can assign that measure: if $m = 0$ we get $\mu(\Omega) = 0$ and $m > 0$ we get $\mu(\Omega) = \infty$.

Remark 35.6.2 (Choice) — Experts may recognize that picking a representative (i.e. creating set X) technically requires the Axiom of Choice. That is why, when people talk about Banach-Tarski issues, the Axiom of Choice almost always gets honorable mention as well.

Stay tuned to actually see a construction for $\mathscr{B}(\mathbb{R}^n)$ in the next chapter.

§35.7 Measurable functions

Prototypical example for this section: For $S \subseteq \Omega$, $\mathbf{1}_S \colon \Omega \to \mathbb{R}$ is a measurable function if and only if S is a measurable set.

In the past, when we had topological spaces, we considered continuous functions. The analog here is:

Definition 35.7.1. Let $(X, \mathscr{A})$ and $(Y, \mathscr{B})$ be measurable spaces (or measure spaces). A function $f \colon X \to Y$ is **measurable** if for any open set $S \subseteq Y$ (i.e. $S \in \mathscr{B}$) we have $f^{\text{pre}}(S)$ is measurable (i.e. $f^{\text{pre}}(S) \in \mathscr{A}$).

Apart from the obvious symmetry with the definition of continuous function, as we will see in Section 37.2, this definition is such that for a nonnegative function $f \colon \Omega \to \mathbb{R}_{\ge 0}$, $\int_\Omega f \, d\mu$ exists if and only if f is measurable.

Remark 35.7.2 — By symmetry, you might have guessed that a function $f : X \to Y$ is measurable if for any measurable $S \subseteq Y$, we have $f^{\text{pre}}(S) \subseteq X$ is measurable.

Nevertheless, this definition doesn't work the way we expect — even continuous function can fail this definition.

Example 35.7.3 (Continuous function with non-measurable preimage of measurable set)

Let $f \colon [0,1] \to [0,1]$ be the Devil's Staircase (or Cantor function). This function is continuous, and has the property that, let $C \subseteq [0,1]$ be the Cantor set, then $|C| = 0$, yet $f^{\text{img}}(C) = [0,1]$ with measure 1.

Let $g \colon [0,1] \to [0,2]$ be defined by $g(x) = f(x) + x$. Then,

- For each open interval (a,b) that is removed from the Cantor set C, then $|g^{\text{img}}((a,b))| = |(a,b)|$.
- $g^{\text{img}}(C)$ has measure 1.

Note that g is bijective, let $h \colon [0,2] \to [0,1]$, $h = g^{-1}$. Then h is continuous, however:

- $h^{\text{pre}}(C) = g^{\text{img}}(C)$ has measure 1, so it has some non-measurable subset,
- C has measure 0, so every subset of C is (Lebesgue) measurable,
- thus, $h^{\text{pre}}(D)$ is non-measurable for some measurable subset $D \subseteq C$.

In practice, most functions you encounter will be continuous anyways, and in that case we are fine.

Proposition 35.7.4 (Continuous implies Borel measurable)

Suppose X and Y are topological spaces and we pick the Borel measures on both. A function $f \colon X \to Y$ which is continuous as a map of topological spaces is also measurable.

Proof. Follows from the fact that pre-images of open sets are open, thus Borel measurable. □

§35.8 On the word "almost"

In later chapters we will begin seeing the phrase "almost everywhere" and "almost surely" start to come up, and it seems prudent to take the time to talk about it now.

Definition 35.8.1. We say that property P occurs **almost everywhere** or **almost surely** if the set

$$\{\omega \in \Omega \mid P \text{ does not hold for } \omega\}$$

has measure zero.

For example, if we say "$f = g$ almost everywhere" for some functions f and g defined on a measure space Ω, then we mean that $f(\omega) = g(\omega)$ for all $\omega \in \Omega$ other than a measure-zero set.

There, that's the definition. The main thing to now update your instincts on is that

In measure theory, we basically only care about things up to almost-everywhere.

Here are some examples:

- If $f = g$ almost everywhere, then measure theory will basically not tell these functions apart. For example, $\int_\Omega f \; d\omega = \int_\Omega g \; d\omega$ will hold for two functions agreeing almost everywhere.
- As another example, if we prove "there exists a unique function f such that so-and-so", the uniqueness is usually going to be up to measure-zero sets.

You can think of this sort of like group isomorphism, where two groups are considered "basically the same" when they are isomorphic, except this one might take a little while to get used to.[2]

§35.9 A few harder problems to think about

Problem 35A†. Let $(\Omega, \mathscr{A}, \mu)$ be a probability space. Show that the intersection of countably many sets of measure 1 also has measure 1.

Problem 35B (On countable σ-algebras)**.** Let $\mathscr{A}$ be a σ-algebra on a set Ω. Suppose that $\mathscr{A}$ has countable cardinality. Prove that $|\mathscr{A}|$ is finite and equals a power of 2.

[2]In fact, some people will even define functions on measure spaces as *equivalence classes* of maps, modded out by agreement outside a measure zero set.

36 Constructing the Borel and Lebesgue measure

It's very difficult to define in one breath a measure on the Borel space $\mathscr{B}(\mathbb{R}^n)$. It is easier if we define a weaker notion first. There are two such weaker notions that we will define:

- A **pre-measure**: satisfies the axioms of a measure, but defined on *fewer* sets than a measure: they'll be defined on an "algebra" rather than the full-fledged "σ-algebra".
- An **outer measure**: defined on 2^Ω but satisfies weaker axioms.

It will turn out that pre-measures yield outer measures, and outer measures yield measures.

§36.1 Pre-measures

Prototypical example for this section: Let $\Omega = \mathbb{R}^2$. Then we take $\mathscr{A}_0$ generated by rectangles, with μ_0 the usual area.

The way to define a pre-measure is to weaken the σ-algebra to an algebra.

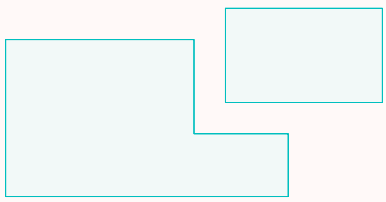
Definition 36.1.1. Let Ω be a set. We define notions of an **algebra**, which is the same as σ-algebra except with "countable" replaced by finite everywhere.

That is: an algebra $\mathscr{A}_0$ on Ω is a nonempty subset of 2^Ω, which is closed under complement and *finite* union. The smallest algebra containing a subset $\mathscr{F} \subseteq 2^\Omega$ is the **algebra generated by $\mathscr{F}$**.

In practice, we will basically always use generation for algebras.

Example 36.1.2

When $\Omega = \mathbb{R}^n$, we can let $\mathcal{L}_0$ be the algebra generated by $[a_1, b_1] \times \cdots \times [a_n, b_n]$. A typical element might look like:



Unsurprisingly, since we have *finitely* many rectangles and their complements involved, in this case we actually *can* unambiguously assign an area, and will do so soon.

Definition 36.1.3. A **pre-measure** μ_0 on a algebra $\mathscr{A}_0$ is a function $\mu_0 \colon \mathscr{A}_0 \to [0, +\infty]$ which satisfies the axioms

- $\mu_0(\varnothing) = 0$, and

- **Countable additivity**: if $A_1, A_2, \dots$ are disjoint sets in $\mathscr{A}_0$ and *moreover* the disjoint union $\bigsqcup A_i$ is contained in $\mathscr{A}_0$ (not guaranteed by algebra axioms!), then

$$\mu_0\left(\bigsqcup_n A_n\right) = \sum_n \mu_0(A_n).$$

Example 36.1.4 (The pre-measure on $\mathbb{R}^n$)

Let $\Omega = \mathbb{R}^2$. Then, let $\mathcal{L}_0$ be the algebra generated by rectangles $[a_1, a_2] \times [b_1, b_2]$. We then let

$$\mu_0\left([a_1, a_2] \times [b_1, b_2]\right) = (a_2 - a_1)(b_2 - b_1)$$

the area of the rectangle. As elements of $\mathcal{L}_0$ are simply *finite* unions of rectangles and their complements (picture drawn earlier), it's not difficult to extend this to a pre-measure λ_0 which behaves as you expect — although we won't do this.

Since we are sweeping something under the rug that turns out to be conceptually important, I'll go ahead and blue-box it.

Proposition 36.1.5 (Geometry sanity check that we won't prove)

For $\Omega = \mathbb{R}^n$ and $\mathcal{L}_0$ the algebra generated by rectangular prisms, one can define a pre-measure λ_0 on $\mathcal{L}_0$.

From this point forwards, we will basically do almost no geometry[1] whatsoever in defining the measure $\mathscr{B}(\mathbb{R}^n)$, and only use set theory to extend our measure. So, Proposition 36.1.5 is the only sentry which checks to make sure that our "initial definition" is sane.

To put the point another way, suppose an **insane scientist**[2] tried to define a notion of area in which every rectangle had area 1. Intuitively, this shouldn't be possible: every rectangle can be dissected into two halves and we ought to have $1 + 1 \neq 1$. However, the only thing that would stop them is that they couldn't extend their pre-measure on the algebra $\mathcal{L}_0$. If they somehow got past that barrier and got a pre-measure, nothing in the rest of the section would prevent them from getting an entire *bona fide* measure with this property. Thus, in our construction of the Lebesgue measure, most of the geometric work is captured in the (omitted) proof of Proposition 36.1.5.

§36.2 Outer measures

Prototypical example for this section: Keep taking $\Omega = \mathbb{R}^2$; see the picture to follow.

The other way to weaken a measure is to relax the countable additivity, and this yields the following:

Definition 36.2.1. An **outer measure** μ^* on a set Ω is a function $\mu^* \colon 2^\Omega \to [0, +\infty]$ satisfying the following axioms:

- $\mu^*(\varnothing) = 0$;

[1] White lie. Technically, we will use one more fact: that open sets of $\mathbb{R}^n$ can be covered by countably infinitely many rectangles, as in Exercise 36.5.1. This step doesn't involve any area assignments, though.

[2] Because "mad scientists" are overrated.

- if $E \subseteq F$ and $E, F \in 2^\Omega$ then $\mu^*(E) \le \mu^*(F)$;
- for any subsets E_1, E_2, ... of Ω we have

$$\mu^*\left(\bigcup_n E_n\right) \le \sum_n \mu^*(E_n).$$

(I don't really like the word "outer measure", since I think it is a bit of a misnomer: I would rather call it "fake measure", since it's not a measure either.)

The reason for the name "outer measure" is that you almost always obtain outer measures by approximating them from "outside" sets. Officially, the result is often stated as follows (as Problem 36A†).

> For a set Ω, let $\mathcal{E}$ be *any* subset of 2^Ω and let $\rho \colon \mathcal{E} \to [0, +\infty]$ be *any* function. Then
> $$\mu^*(E) = \inf\left\{\sum_{n=1}^{\infty} \rho(E_n) \mid E_n \in \mathcal{E},\ E \subseteq \bigcup_{n=1}^{\infty} E_n\right\}$$
> is an outer measure.

However, I think the above theorem is basically always wrong to use in practice, because it is *way too general*. As I warned with the insane scientist, we really do want some sort of sanity conditions on ρ: otherwise, if we apply the above result as stated, there is no guarantee that μ^* will be compatible with ρ in any way.

So, I think it is really better to apply the theorem to pre-measures μ_0 for which one *does* have some sort of guarantee that the resulting μ^* is compatible with μ_0. In practice, this is always how we will want to construct our outer measures.

Theorem 36.2.2 (Constructing outer measures from pre-measures)
Let μ_0 be a pre-measure on an algebra $\mathscr{A}_0$ on a set Ω.

(a) The map $\mu^* \colon 2^\Omega \to [0, +\infty]$ defined by

$$\mu^*(E) = \inf\left\{\sum_{n=1}^{\infty} \mu_0(A_n) \mid A_n \in \mathscr{A}_0,\ E \subseteq \bigcup_{n=1}^{\infty} A_n\right\}$$

is an outer measure.

(b) Moreover, this measure agrees with μ_0 on sets in $\mathscr{A}_0$.

Intuitively, what is going on is that $\mu^*(A)$ is the infimum of coverings of A by countable unions of elements in $\mathscr{A}_0$. Part (b) is the first half of the compatibility condition I promised; the other half appears later as Proposition 36.3.2.

Proof of Theorem 36.2.2. As alluded to already, part (a) is a special case of Problem 36A† (and proving it in this generality is actually easier, because you won't be distracted by unnecessary properties).

We now check (b), that $\mu^*(A) = \mu_0(A)$ for $A \in \mathscr{A}_0$. One bound is quick:

> **Question 36.2.3.** Show that $\mu^*(A) \le \mu_0(A)$.

For the reverse, suppose that $A \subseteq \bigcup_n A_n$. Then, define the sets

$$\begin{aligned} B_1 &= A \cap A_1 \\ B_2 &= (A \cap A_2) \setminus B_1 \\ B_3 &= (A \cap A_3) \setminus (B_1 \cup B_2) \\ &\vdots \end{aligned}$$

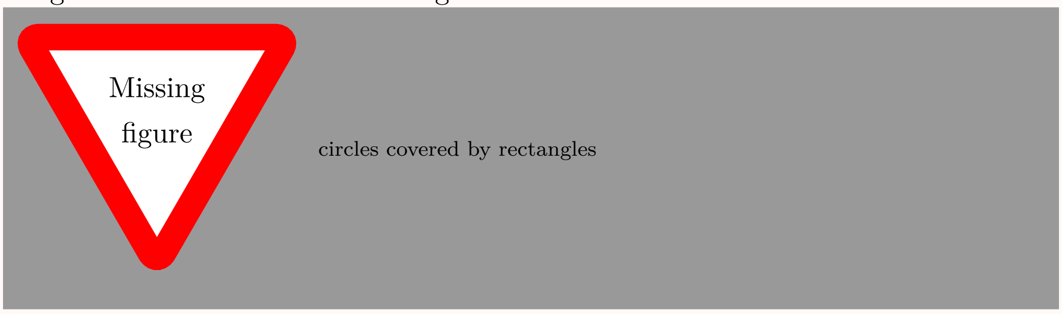
and so on. Then the B_n are disjoint elements of $\mathscr{A}_0$ with $B_n \subset A_n$, and we have rigged the definition so that $\bigsqcup_n B_n = A$. Thus by definition of pre-measure,

$$\mu_0(A) = \sum_n \mu_0(B_n) \le \sum_n \mu_0(A_n)$$

as desired. $\square$

Example 36.2.4

Let $\Omega = \mathbb{R}^2$ and λ_0 the pre-measure from before. Then $\lambda^*(A)$ is, intuitively, the infimum of coverings of the set A by rectangles. Here is a picture you might use to imagine the situation with A being the unit disk.



§36.3 Carathéodory extension for outer measures

We will now take any outer measure and turn it into a proper measure. To do this, we first need to specify the σ-algebra on which we will define the measure.

Definition 36.3.1. Let μ^* be an outer measure. We say a set A is **Carathéodory measurable with respect to μ^***, or just **μ^*-measurable**, if the following condition holds: for any set $E \in 2^\Omega$,

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

This definition is hard to motivate, but turns out to be the right one. One way to motivate is this: it turns out that in $\mathbb{R}^n$, it will be equivalent to a reasonable geometric condition (which I will state in Proposition 36.4.3), but since that geometric definition requires information about $\mathbb{R}^n$ itself, this is the "right" generalization for general measure spaces.

Since our goal was to extend our $\mathscr{A}_0$, we had better make sure this definition lets us measure the initial sets that we started with!

Proposition 36.3.2 (Carathéodory measurability is compatible with the initial $\mathscr{A}_0$)

Suppose μ^* was obtained from a pre-measure μ_0 on an algebra $\mathscr{A}_0$, as in Theorem 36.2.2. Then every set in $\mathscr{A}_0$ is μ^*-measurable.

This is the second half of the compatibility condition that we get if we make sure our initial μ_0 at least satisfies the pre-measure axioms. (The first half was (b) of Theorem 36.2.2.)

Proof. Let $A \in \mathscr{A}_0$ and $E \in 2^\Omega$; we wish to prove $\mu^*(E) = \mu^*(E \cap A) + \mu^*(E \setminus A)$. The definition of outer measure already requires $\mu^*(E) \le \mu^*(E \cap A) + \mu^*(E \setminus A)$ and so it's enough to prove the reverse inequality.

By definition of infimum, for any $\varepsilon > 0$, there is a covering $E \subset \bigcup_n A_n$ with $\mu^*(E)+\varepsilon \ge \sum_n \mu_0(A_n)$. But

$$\sum_n \mu_0(A_n) = \sum_n (\mu_0(A_n \cap A) + \mu_0(A_n \setminus A)) \ge \mu^*(E \cap A) + \mu^*(E \setminus A)$$

with the first equality being the definition of pre-measure on $\mathscr{A}_0$, the second just being by definition of μ^* (since $A_n \cap A$ certainly covers $E \cap A$, for example). Thus $\mu^*(E) + \varepsilon \ge \mu^*(E \cap A) + \mu^*(E \setminus A)$. Since the inequality holds for any $\varepsilon > 0$, we're done. □

To add extra icing onto the cake, here is one more niceness condition which our constructed measure will happen to satisfy.

Definition 36.3.3. A **null set** of a measure space $(\Omega, \mathscr{A}, \mu)$ is a set $A \in \mathscr{A}$ with $\mu(A) = 0$. A measure space $(\Omega, \mathscr{A}, \mu)$ is **complete** if whenever A is a null set, then all subsets of A are in $\mathscr{A}$ as well (and hence null sets).

This is a nice property to have, for obvious reasons. Visually, if I have a bunch of dust which I *already* assigned weight zero, and I blow away some of the dust, then the remainder should still have an assigned weight — zero. The extension theorem will give us σ-algebras with this property.

Theorem 36.3.4 (Carathéodory extension theorem for outer measures)

If μ^* is an outer measure, and $\mathscr{A}^{\text{cm}}$ is the set of μ^*-measurable sets with respect to μ^*, then $\mathscr{A}^{\text{cm}}$ is a σ-algebra on Ω, and the restriction μ^{cm} of μ^* to $\mathscr{A}^{\text{cm}}$ gives a *complete* measure space.

(Phonetic remark: you can think of the superscript $^{\text{cm}}$ as standing for either "Carathéodory measurable" or "complete". Both are helpful for remembering what this represents. This notation is not standard but the pun was too good to resist.)

Thus, if we compose Theorem 36.2.2 with Theorem 36.3.4, we find that every pre-measure μ_0 on an algebra $\mathscr{A}_0$ naturally gives a σ-algebra $\mathscr{A}^{\text{cm}}$ with a complete measure μ^{cm}, and our two compatibility results (namely (b) of Theorem 36.2.2, together with Proposition 36.3.2) means that $\mathscr{A}^{\text{cm}} \supset \mathscr{A}_0$ and μ^{cm} agrees with μ.

Here is a table showing the process, where going down each row of the table corresponds to restriction process.

		Construct order	Notes
2^Ω	μ^*	Step 2	μ^* is outer measure obtained from μ_0
$\mathscr{A}^{\text{cm}}$	μ^{cm}	Step 3	$\mathscr{A}^{\text{cm}}$ defined as μ^*-measurable sets, $(\mathscr{A}^{\text{cm}}, \mu^{\text{cm}})$ is complete.
$\mathscr{A}_0$	μ_0	Step 1	μ_0 is a pre-measure

§36.4 Defining the Lebesgue measure

This lets us finally define the Lebesgue measure on $\mathbb{R}^n$. We wrap everything together at once now.

Definition 36.4.1. We create a measure on $\mathbb{R}^n$ by the following procedure.

- Start with the algebra $\mathcal{L}_0$ generated by rectangular prisms, and define a *pre-measure* λ_0 on this $\mathcal{L}_0$ (this was glossed over in the example).
- By Theorem 36.2.2, this gives the **Lebesgue outer measure** λ^* on $2^{\mathbb{R}^n}$, which is compatible on all the rectangular prisms.
- By Carathéodory (Theorem 36.3.4), this restricts to a complete measure λ on the σ-algebra $\mathcal{L}(\mathbb{R}^n)$ of λ^*-measurable sets (which as promised contains all rectangular prisms).[3]

The resulting complete measure, denoted λ, is called the **Lebesgue measure**.

The algebra $\mathcal{L}(\mathbb{R}^n)$ we obtained will be called the **Lebesgue σ-algebra**; sets in it are said to be **Lebesgue measurable**.

Here is the same table from before, with the values filled in for the special case $\Omega = \mathbb{R}^n$, which gives us the Lebesgue algebra.

		Construct order	Notes
$2^{\mathbb{R}^n}$	λ^*	Step 2	λ^* is Lebesgue outer measure
$\mathcal{L}(\mathbb{R}^n)$	λ	Step 3	Lebesgue σ-algebra (complete)
$\mathcal{L}_0$	λ_0	Step 1	Define pre-measure on rectangles

Of course, now that we've gotten all the way here, if we actually want to *compute* any measures, we can mostly gleefully forget about how we actually constructed the measure and just use the properties. The hard part was to showing that there *is* a way to assign measures consistently; actually figuring out what that measure's value is *given that it exists* is often much easier. Here is an example.

Example 36.4.2 (The Cantor set has measure zero)

The standard **middle-thirds Cantor set** is the subset [0, 1] obtained as follows: we first delete the open interval $(1/3, 2/3)$. This leaves two intervals $[0, 1/3]$ and $[2/3, 1]$ from which we delete the middle thirds again from both, i.e. deleting $(1/9, 2/9)$ and $(7/9, 8/9)$. We repeat this procedure indefinitely and let C denote the result. An illustration is shown below.

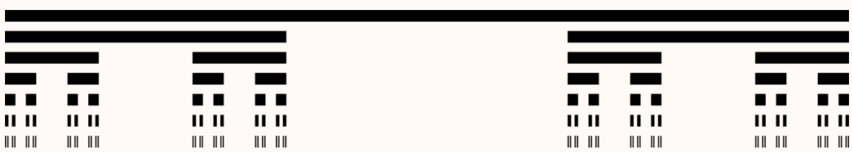


Image from [1207]

It is a classic fact that C is uncountable (it consists of ternary expansions omitting the digit 1). But it is measurable (it is an intersection of closed sets!) and we contend it has measure zero. Indeed, at the nth step, the result has measure $(2/3)^n$ leftover.

[3] If I wanted to be consistent with the previous theorems, I might prefer to write $\mathcal{L}^{\text{cm}}$ and λ^{cm} for emphasis. It seems no one does this, though, so I won't.

So $\mu(C) \le (2/3)^n$ for every n, forcing $\mu(C) = 0$.

This is fantastic, but there is one elephant in the room: how are the Lebesgue σ-algebra and the Borel σ-algebra related? To answer this question briefly, I will state two results (but another answer is given in the next section). The first is a geometric interpretation of the strange Carathéodory measurable hypothesis.

Proposition 36.4.3 (A geometric interpretation of Lebesgue measurability)

A set $A \subseteq \mathbb{R}^n$ is Lebesgue measurable if and only if for every $\varepsilon > 0$, there is an open set $U \supset A$ such that

$$\lambda^*(U \setminus A) < \varepsilon$$

where λ^* is the Lebesgue outer measure.

I want to say that this was Lebesgue's original formulation of "measurable", but I'm not sure about that. In any case, we won't need to use this, but it's good to see that our definition of Lebesgue measurable has a down-to-earth geometric interpretation.

Question 36.4.4. Deduce that every open set is Lebesgue measurable. Conclude that the Lebesgue σ-algebra contains the Borel σ-algebra. (A different proof is given later on.)

However, the containment is proper: there are more Lebesgue measurable sets than Borel ones. Indeed, it can actually be proven using transfinite induction (though we won't) that $|\mathscr{B}(\mathbb{R})| = |\mathbb{R}|$.[4] Using this, one obtains:

Exercise 36.4.5. Show the Borel σ-algebra is not complete. (Hint: consider the Cantor set. You won't be able to write down an example of a non-measurable set, but you can use cardinality arguments.) Thus the Lebesgue σ-algebra strictly contains the Borel one.

Nonetheless, there is a great way to describe the Lebesgue σ-algebra, using the idea of completeness.

Definition 36.4.6. Let $(\Omega, \mathscr{A}, \mu)$ be a measure space. The **completion** $(\Omega, \overline{\mathscr{A}}, \overline{\mu})$ is defined as follows: we let

$$\overline{\mathscr{A}} = \{A \cup N \mid A \in \mathscr{A}, N \text{ subset of null set}\}.$$

and $\overline{\mu}(A \cup N) = \mu(A)$. One can check this is well-defined, and in fact $\overline{\mu}$ is the unique extension of μ from $\mathscr{A}$ to $\overline{\mathscr{A}}$.

This looks more complicated than it is. Intuitively, all we are doing is "completing" the measure by telling $\overline{\mu}$ to regard any subset of a null set as having measure zero, too.

Then, the saving grace:

Theorem 36.4.7 (Lebesgue is completion of Borel)

For $\mathbb{R}^n$, the Lebesgue measure is the completion of the Borel measure.

Proof. This actually follows from results in the next section, namely Exercise 36.5.1 and part (c) of Carathéodory for pre-measures (Theorem 36.5.5). □

[4]See `https://math.stackexchange.com/a/70891` for a sketch.

§36.5 A fourth row: Carathéodory for pre-measures

Prototypical example for this section: The fourth row for the Lebesgue measure is $\mathscr{B}(\mathbb{R}^n)$.

In many cases, $\mathscr{A}^{\text{cm}}$ is actually bigger than our original goal, and instead we only need to extend μ_0 on $\mathscr{A}_0$ to μ on $\mathscr{A}$, where $\mathscr{A}$ is the σ-algebra generated by $\mathscr{A}_0$. Indeed, our original goal was to get $\mathscr{B}(\mathbb{R}^n)$, and in fact:

Exercise 36.5.1. Show that $\mathscr{B}(\mathbb{R}^n)$ is the σ-algebra generated by the $\mathcal{L}_0$ we defined earlier.

Fortunately, this restriction is trivial to do.

Question 36.5.2. Show that $\mathscr{A}^{\text{cm}} \supset \mathscr{A}$, so we can just restrict μ^{cm} to $\mathscr{A}$.

We will in a moment add this as the fourth row in our table.

However, if this is the end goal, than a somewhat different Carathéodory theorem can be stated because often one more niceness condition holds:

Definition 36.5.3. A pre-measure or measure μ on Ω is **σ-finite** if Ω can be written as a countable union $\Omega = \bigcup_n A_n$ with $\mu(A_n) < \infty$ for each n.

Question 36.5.4. Show that the pre-measure λ_0 we had, as well as the Borel measure $\mathscr{B}(\mathbb{R}^n)$, are both σ-finite.

Actually, for us, σ-finite is basically always going to be true, so you can more or less just take it for granted.

Theorem 36.5.5 (Carathéodory extension theorem for pre-measures)
Let μ_0 be a pre-measure on an algebra $\mathscr{A}_0$ of Ω, and let $\mathscr{A}$ denote the σ-algebra generated by $\mathscr{A}_0$. Let $\mathscr{A}^{\text{cm}}$, μ^{cm} be as in Theorem 36.3.4. Then:

(a) The restriction of μ^{cm} to $\mathscr{A}$ gives a measure μ extending μ_0.

(b) If μ_0 was σ-finite, then μ is the unique extension of μ_0 to $\mathscr{A}$.

(c) If μ_0 was σ-finite, then μ^{cm} is the completion of μ, hence the unique extension of μ_0 to $\mathscr{A}^{\text{cm}}$.

Here is the updated table, with comments if μ_0 was indeed σ-finite.

		Construct order	Notes
2^Ω	μ^*	Step 2	μ^* is outer measure obtained from μ_0
$\mathscr{A}^{\text{cm}}$	μ^{cm}	Step 3	$(\mathscr{A}^{\text{cm}}, \mu^{\text{cm}})$ is completion $(\mathscr{A}, \mu)$, $\mathscr{A}^{\text{cm}}$ defined as μ^*-measurable sets
$\mathscr{A}$	μ	Step 4	$\mathscr{A}$ defined as σ-alg. generated by $\mathscr{A}_0$
$\mathscr{A}_0$	μ_0	Step 1	μ_0 is a pre-measure

And here is the table for $\Omega = \mathbb{R}^n$, with Borel and Lebesgue in it.

		Construct order	Notes
$2^{\mathbb{R}^n}$	λ^*	Step 2	λ^* is Lebesgue outer measure
$\mathcal{L}(\mathbb{R}^n)$	λ	Step 3	Lebesgue σ-algebra, completion of Borel one
$\mathscr{B}(\mathbb{R}^n)$	μ	Step 4	Borel σ-algebra, generated by $\mathcal{L}_0$
$\mathcal{L}_0$	λ_0	Step 1	Define pre-measure on rectangles

Going down one row of the table corresponds to restriction, while each of $\mu_0 \to \mu \to \mu^{\text{cm}}$ is a unique extension when μ_0 is σ-finite.

Proof of Theorem 36.5.5. For (a): this is just Theorem 36.2.2 and Theorem 36.3.4 put together, combined with the observation that $\mathscr{A}^* \supset \mathscr{A}_0$ and hence $\mathscr{A}^* \supset \mathscr{A}$. Parts (b) and (c) are more technical, and omitted. $\square$

§36.6 From now on, we assume the Borel measure

explain why

§36.7 A few harder problems to think about

Problem 36A† (Constructing outer measures from arbitrary ρ)**.** For a set Ω, let $\mathcal{E}$ be *any* subset of 2^Ω and let $\rho\colon \mathcal{E} \to [0, +\infty]$ be *any* function. Prove that

$$\mu^*(E) = \inf\left\{ \sum_{n=1}^{\infty} \rho(E_n) \mid E_n \in \mathcal{E},\ E \subseteq \bigcup_{n=1}^{\infty} E_n \right\}$$

is an outer measure.

Problem 36B (The insane scientist)**.** Let $\Omega = \mathbb{R}^2$, and let $\mathcal{E}$ be the set of (non-degenerate) rectangles. Let $\rho(E) = 1$ for every rectangle $E \in \mathcal{E}$. Ignoring my advice, the insane scientist uses ρ to construct an outer measure μ^*, as in Problem 36A†.

(a) Find $\mu^*(S)$ for each subset S of $\mathbb{R}^2$.

(b) Which sets are μ^*-measurable?

You should find that no rectangle is μ^*-measurable, unsurprisingly foiling the scientist.

Problem 36C. A function $f\colon \mathbb{R} \to \mathbb{R}$ is continuous. Must f be measurable with respect to the Lebesgue measure on $\mathbb{R}$?

37 Lebesgue integration

On any measure space $(\Omega, \mathscr{A}, \mu)$ we can then, for a function $f\colon \Omega \to [0, \infty]$ define an integral

$$\int_\Omega f \, d\mu.$$

This integral may be $+\infty$ (even if f is finite). As the details of the construction won't matter for us later on, we will state the relevant definitions, skip all the proofs, and also state all the properties that we actually care about. Consequently, this chapter will be quite short.

§37.1 The definition

The construction is done in four steps.

Definition 37.1.1. If A is a measurable set of Ω, then the **indicator function** $\mathbf{1}_A\colon \Omega \to \mathbb{R}$ is defined by

$$\mathbf{1}_A(\omega) = \begin{cases} 1 & \omega \in A \\ 0 & \omega \notin A. \end{cases}$$

Step 1 (Indicator functions) — For an indicator function, we require

$$\int_\Omega \mathbf{1}_A \, d\mu := \mu(A)$$

(which may be infinite).

We extend this linearly now for nonnegative functions which are sums of indicators: these functions are called **simple functions**.

Step 2 (Simple functions) — Let $A_1, \dots, A_n$ be a finite collection of measurable sets. Let $c_1, \dots, c_n$ be either nonnegative real numbers or $+\infty$. Then we define

$$\int_\Omega \left(\sum_{i=1}^n c_i \mathbf{1}_{A_i} \right) d\mu := \sum_{i=1}^n c_i \mu(A_i).$$

If $c_i = \infty$ and $\mu(A_i) = 0$, we treat $c_i \mu(A_i) = 0$.

One can check the resulting sum does not depend on the representation of the simple function as $\sum c_i \mathbf{1}_{A_i}$. In particular, it is compatible with the previous step.

Conveniently, this is already enough to define the integral for $f\colon \Omega \to [0, +\infty]$. Note that $[0, +\infty]$ can be thought of as a topological space where we add new open sets $(a, +\infty]$ for each real number a to our usual basis of open intervals. Thus we can equip it with the Borel sigma-algebra.[1]

[1] We *could* also try to define a measure on it, but we will not: it is a good enough for us that it is a measurable space.

Step 3 (Nonnegative functions) — For each measurable function $f\colon \Omega \to [0, +\infty]$, let

$$\int_\Omega f \; d\mu := \sup_{0 \le s \le f} \left(\int_\Omega s \; d\mu \right)$$

where the supremum is taken over all *simple* s such that $0 \le s \le f$. As before, this integral may be $+\infty$.

That is,

We define the integral $\int_\Omega f \; d\mu$ by approximating it from below with simple functions, for which we know how to integrate.

One can check this is compatible with the previous definitions. At this point, we introduce an important term.

Definition 37.1.2. A measurable (nonnegative) function $f\colon \Omega \to [0, +\infty]$ is **absolutely integrable** or just **integrable** if $\int_\Omega f \; d\mu < \infty$.

Warning: I find "integrable" to be *really* confusing terminology. Indeed, *every* measurable function from Ω to $[0, +\infty]$ can be assigned a Lebesgue integral, it's just that this integral may be $+\infty$. So the definition is far more stringent than the name suggests. Even constant functions can fail to be integrable:

Example 37.1.3 (We really should call it "finitely integrable")

The constant function 1 is *not* integrable on $\mathbb{R}$, since $\int_\mathbb{R} 1 \; d\mu = \mu(\mathbb{R}) = +\infty$.

For this reason, I will usually prefer the term "absolutely integrable". (If it were up to me, I would call it "finitely integrable", and usually do so privately.)

Remark 37.1.4 (Why don't we approximate the integral from above?) — For bounded functions on measure spaces with $|\Omega| < \infty$, we can equivalently define

$$\int_\Omega f \; d\mu := \inf_{0 \le f \le s} \left(\int_\Omega s \; d\mu \right)$$

where the infimum is taken over all simple s such that $f \le s$. However, if the functions are unbounded or $|\Omega| = \infty$, the situation is not that simple:

- The function $f(x) = x^{-2}$ defined over $\Omega = (1, \infty)$ is absolutely integrable, yet for all simple s such that $f \le s$ we have $\int_\Omega s \; d\mu = \infty$.
- The function $f(x) = x^{-0.5}$ defined over $\Omega = (0, 1)$ is absolutely integrable, yet there's no simple s such that $f \le s$ and s is finite almost everywhere.

Finally, this lets us integrate general functions.

Definition 37.1.5. In general, a measurable function $f\colon \Omega \to [-\infty, \infty]$ is **absolutely integrable** or just **integrable** if $|f|$ is.

Since we'll be using the first word, this is easy to remember: "absolutely integrable" requires taking absolute values.

Step 4 (Absolutely integrable functions) — If $f \colon \Omega \to [-\infty, \infty]$ is absolutely integrable, then we define

$$f^+(x) = \max\{f(x), 0\}$$
$$f^-(x) = \min\{f(x), 0\}$$

and set

$$\int_\Omega f \, d\mu = \int_\Omega |f^+| \, d\mu - \int_\Omega |f^-| \, d\mu$$

which in particular is finite.

That said, calling it "finitely integrable" here would also make it as easy to remember:

Exercise 37.1.6. Show that $\int_\Omega |f| d\mu < \infty$ if and only if $\int_\Omega |f^+| d\mu < \infty$ and $\int_\Omega |f^-| d\mu < \infty$.

You may already start to see that we really like nonnegative functions: with the theory of measures, it is possible to integrate them, and it's even okay to throw in $+\infty$'s everywhere. But once we start dealing with functions that can be either positive or negative, we have to start adding finiteness restrictions — actually essentially what we're doing is splitting the function into its positive and negative part, requiring both are finite, and then integrating.

To finish this section, we state for completeness some results that you probably could have guessed were true. Fix $\Omega = (\Omega, \mathscr{A}, \mu)$, and let f and g be measurable real-valued functions such that $f(x) = g(x)$ almost everywhere.

- (Almost-everywhere preservation) The function f is absolutely integrable if and only if g is, and if so, their Lebesgue integrals match.
- (Additivity) If f and g are absolutely integrable then

  $$\int_\Omega f + g \, d\mu = \int_\Omega f \, d\mu + \int_\Omega g \, d\mu.$$

  The "absolutely integrable" hypothesis can be dropped if f and g are nonnegative.
- (Scaling) If f is absolutely integrable and $c \in \mathbb{R}$ then cf is absolutely integrable and

  $$\int_\Omega cf \, d\mu = c \int_\Omega f \, d\mu.$$

  The "absolutely integrable" hypothesis can be dropped if f is nonnegative and $c > 0$.
- (Monotoncity) If f and g are absolutely integrable and $f \le g$, then

  $$\int_\Omega f \, d\mu \le \int_\Omega g \, d\mu.$$

  The "absolutely integrable" hypothesis can be dropped if f and g are nonnegative.

There are more famous results like monotone/dominated convergence that are also true, but we won't state them here as we won't really have a use for them in the context of probability. (They appear later on in a bonus chapter.)

§37.2 An equivalent definition

The Lebesgue integral can also be defined as follows — which should be more intuitive on the various choices of the definitions we made in the steps.

In this definition,

The integral $\int_\Omega f \, d\mu$ is just the volume of the region under the graph of f.

Let us define it:

Step 1 (The region under the graph) — For a nonnegative function $f \colon \Omega \to \mathbb{R}$, define the **region under the function** f, $R(f)$, to be $\{(x,y) \in \Omega \times \mathbb{R}, 0 \le y \le f(x)\}$.

Remark 37.2.1 — It should be clear why we only define this for nonnegative function initially — for general function f, the only way we could sensibly define the region would be something like the following:

$$R^+(f) = \{(x,y) \in \Omega \times \mathbb{R}, f(x) \ge 0, 0 \le y \le f(x)\},$$
$$R^-(f) = \{(x,y) \in \Omega \times \mathbb{R}, f(x) \le 0, 0 \ge y \ge f(x)\}.$$

Nevertheless, notice that $R^+(f)$ is simply the region under the function $f^+(x) = \max\{f(x), 0\}$, and $R^-(f)$ has the same measure as the region under the function $f^-(x) = \min\{f(x), 0\}$, so defining $\int_\Omega f \, d\mu$ for nonnegative functions first would actually simplify the definition.

Step 2 (Making $\Omega \times \mathbb{R}$ into a measure space) — We define a pre-measure on $\Omega \times \mathbb{R}$ the obvious way: if $X \subseteq \Omega$ and $Y \subseteq \mathbb{R}$ are measurable subsets respectively, then assign $|X \times Y| = |X| \times |Y|$.

The pre-measure can be extended to a measure, as we have done in the previous chapter.

Step 3 (Nonnegative functions) — For each function $f \colon \Omega \to [0, +\infty]$, let

$$\int_\Omega f \, d\mu := |R(f)|.$$

The integral is well-defined whenever $R(f)$ is measurable.

As promised in Section 35.7, the definition of measurable function satisfies:

A nonnegative function f is measurable if and only if we can "measure" the region below the graph of f.

The last step is exactly the same as in the previous section.

Step 4 (Absolutely integrable functions) — If $f \colon \Omega \to [-\infty, \infty]$ is absolutely integrable, then we define

$$\int_\Omega f \, d\mu = \int_\Omega |f^+| \, d\mu - \int_\Omega |f^-| \, d\mu.$$

§37.3 Relation to Riemann integrals (or: actually computing Lebesgue integrals)

For closed intervals, this actually just works out of the box.

Theorem 37.3.1 (Lebesgue integral generalizes Riemann integral)
Let $f\colon [a,b] \to \mathbb{R}$ be a Riemann integrable function (where $[a,b]$ is equipped with the Borel measure). Then f is also Lebesgue integrable and the integrals agree:

$$\int_a^b f(x) \, dx = \int_{[a,b]} f \, d\mu.$$

Note that a Riemann integrable function *must be bounded*, which means if you try to construct a function $f\colon [0,1] \to \mathbb{R}$ in the same vein as Problem 37B† by

$$f(x) = \begin{cases} \frac{\sin(1/x)}{x} & x > 0 \\ 0 & x = 0 \end{cases}$$

the function f will in fact *not* be Riemann integrable! Although of course, the improper Riemann integral $\lim_{\varepsilon \to 0^+} \int_\varepsilon^1 f(x) \, dx$ exists.

Thus in practice, we do all theory with Lebesgue integrals (they're nicer), but when we actually need to compute $\int_{[1,4]} x^2 \, d\mu$ we just revert back to our usual antics with the Fundamental Theorem of Calculus.

Example 37.3.2 (Integrating x^2 over $[1,4]$)
Reprising our old example:

$$\int_{[1,4]} x^2 \, d\mu = \int_1^4 x^2 \, dx = \frac{1}{3} \cdot 4^3 - \frac{1}{3} \cdot 1^3 = 21.$$

This even works for *improper* integrals, if the functions are nonnegative. The statement is a bit cumbersome to write down, but here it is.

Theorem 37.3.3 (Improper integrals are nice Lebesgue ones)
Let $f \ge 0$ be a *nonnegative* continuous function defined on $(a,b) \subseteq \mathbb{R}$, possibly allowing $a = -\infty$ or $b = \infty$. Then

$$\int_{(a,b)} f \, d\mu = \lim_{\substack{a' \to a^+ \\ b' \to b^-}} \int_{a'}^{b'} f(x) \, dx$$

where we allow both sides to be $+\infty$ if f is not absolutely integrable.

The right-hand side makes sense since $[a', b'] \subsetneq (a, b)$ is a compact interval on which f is continuous. This means that improper Riemann integrals of nonnegative functions can just be regarded as Lebesgue ones over the corresponding open intervals.

It's probably better to just look at an example though.

Example 37.3.4 (Integrating $1/\sqrt{x}$ on $(0, 1)$)

For example, you might be familiar with improper integrals like

$$\int_0^1 \frac{1}{\sqrt{x}} \, dx := \lim_{\varepsilon \to 0^+} \int_\varepsilon^1 \frac{1}{\sqrt{x}} \, dx = \lim_{\varepsilon \to 0^+} \left(2\sqrt{1} - 2\sqrt{\varepsilon}\right) = 2.$$

(Note this appeared before as Problem 30C*.) In the Riemann integration situation, we needed the limit as $\varepsilon \to 0^+$ since otherwise $\frac{1}{\sqrt{x}}$ is not defined as a function $[0, 1] \to \mathbb{R}$. However, it is a *measurable nonnegative* function $(0, 1) \to [0, +\infty]$, and hence

$$\int_{(0,1)} \frac{1}{\sqrt{x}} \, d\mu = 2.$$

If f is not nonnegative, then all bets are off. Indeed Problem 37B† is the famous counterexample.

§37.4 A few harder problems to think about

Problem 37A* (The indicator of the rationals)**.** Take the indicator function $\mathbf{1}_\mathbb{Q} \colon \mathbb{R} \to \{0, 1\} \subseteq \mathbb{R}$ for the rational numbers.

(a) Prove that $\mathbf{1}_\mathbb{Q}$ is not Riemann integrable.

(b) Show that $\int_\mathbb{R} \mathbf{1}_\mathbb{Q}$ exists and determine its value — the one you expect!

Problem 37B† (An improper Riemann integral with sign changes)**.** Define $f \colon (1, \infty) \to \mathbb{R}$ by $f(x) = \frac{\sin(x)}{x}$. Show that f is not absolutely integrable, but that the improper Riemann integral

$$\int_1^\infty f(x) \, dx := \lim_{b \to \infty} \int_1^b f(x) \, dx$$

nonetheless exists.

38 Swapping order with Lebesgue integrals

§38.1 Motivating limit interchange

Prototypical example for this section: $\mathbf{1}_{\mathbb{Q}}$ *is good!*

One of the issues with the Riemann integral is that it behaves badly with respect to convergence of functions, and the Lebesgue integral deals with this. This is therefore often given as a poster child on why the Lebesgue integral has better behaviors than the Riemann one.

We technically have already seen this: consider the indicator function $\mathbf{1}_{\mathbb{Q}}$, which is not Riemann integrable by Problem 37A*. But we can readily compute its Lebesgue integral over $[0,1]$, as

$$\int_{[0,1]} \mathbf{1}_{\mathbb{Q}} \, d\mu = \mu\left([0,1] \cap \mathbb{Q}\right) = 0$$

since it is countable.

This *could* be thought of as a failure of existence for the Riemann integral.

Example 38.1.1 ($\mathbf{1}_{\mathbb{Q}}$ is a limit of finitely supported functions)

We can define the sequence of functions g_1, g_2, ... by

$$g_n(x) = \begin{cases} 1 & (n!)x \text{ is an integer} \\ 0 & \text{else.} \end{cases}$$

Then each g_n is piecewise continuous and hence Riemann integrable on $[0,1]$ (with integral zero), but $\lim_{n\to\infty} g_n = \mathbf{1}_{\mathbb{Q}}$ is not.

The limit here is defined in the following sense:

Definition 38.1.2. Let f and $f_1, f_2, \dots : \Omega \to \mathbb{R}$ be a sequence of functions. Suppose that for each $\omega \in \Omega$, the sequence

$$f_1(\omega),\ f_2(\omega),\ f_3(\omega),\ \dots$$

converges to $f(\omega)$. Then we say f_1, f_2, ... **converges pointwise** to the limit f, written $\lim_{n\to\infty} f_n = f$.

We can define $\liminf_{n\to\infty} f_n$ and $\limsup_{n\to\infty} f_n$ similarly.

By "the Lebesgue integral has better behavior", we means the following:

Proposition 38.1.3

If $f_1, f_2, \dots : \Omega \to \mathbb{R}$ are measurable functions, then $\liminf_{n\to\infty} f_n$ and $\limsup_{n\to\infty} f_n$ are measurable.

When f_n are all nonnegative, this means $\int_\Omega \liminf_{n\to\infty} f_n d\mu$ and $\int_\Omega \limsup_{n\to\infty} f_n d\mu$ exists. (If they can be negative, the behavior is not that nice. Problem 37B† gives an example.)

Unfortunately, even if the integral exists, we can't always exchange pointwise limit with Lebesgue integral.

Why would we want to? For instance, if we face this problem:

> Compute $\lim_{k\to\infty} \int_1^\infty \frac{1}{k} e^{-x^2}\, dx$.

While the integral $\int e^{-x^2} dx$ is not computable by elementary means, we would like to say the limit is simply 0 (why wouldn't it be?)

Unfortunately, pointwise convergence is actually a fairly weak notion of convergence.

Example 38.1.4

In all of these examples, we cannot interchange the limit and the integral without changing the result.

- The sequence $f_k(x) = \frac{\sin(x)}{x} \cdot \mathbf{1}_{(1,k)}$ converges pointwise to $f(x) = \frac{\sin(x)}{x} \cdot \mathbf{1}_{(1,\infty)}$ as $k \to \infty$, and the limit $\lim_{k\to\infty} \int f_k(x)\, dx$ exists, but f is not integrable.
- Similarly, $f_k(x) = \frac{\sin(1/x)}{x} \cdot \mathbf{1}_{(1/k,\infty)}$ converges pointwise to $f(x) = \frac{\sin(1/x)}{x} \cdot \mathbf{1}_{(0,\infty)}$ as $k \to \infty$, the limit $\lim_{k\to\infty} \int f_k(x)\, dx$ exists and is finite, but f is not integrable.
- The sequence $f_k(x) = \frac{\mathbf{1}_{(0,k)}}{k}$ converges pointwise to $f(x) = 0$ as $k \to \infty$, for every k then $\int f_k(x)\, dx = 1$, but $\int f(x)\, dx = 0$.

  Note that, in this case, the convergence is actually uniform!
- We don't even need k in the denominator — the sequence $f_k(x) = \mathbf{1}_{(0,k)}$ also converges pointwise to $f(x) = 0$, but this time, for every k then $\int f_k(x)\, dx = \infty$!
- The sequence $f_k(x) = k \cdot \mathbf{1}_{(0,1/k)}$ converges pointwise to $f(x) = 0$ as $k \to \infty$. But similar to above, $\int f_k(x)\, dx = 1$ for every k, but $\int f(x)\, dx = 0$.

The last example is similar in behavior to an example known as the Witch's hat.[a]

[a] https://www.geogebra.org/m/dv7ctmed has an animation.

As such, the convergence theorems stated below is an attempt to classify all the possible anomalies, and to show that in "usual" cases, interchanging limit and integral just works.

As mentioned earlier, we choose to use the Lebesgue integral instead of the Riemann integral, because in such cases, the Lebesgue integral will usually just exist.

§38.2 Overview

The three big-name results for exchanging pointwise limits with Lebesgue integrals is:

- Fatou's lemma: the most general statement possible, for any nonnegative measurable functions.
- Monotone convergence: "increasing limits" just work.
- Dominated convergence (actually Fatou-Lebesgue): limits that are not too big (bounded by some absolutely integrable function) just work.

§38.3 Fatou's lemma

In all the above examples, we see that:

- The failure of the interchange of limit and integral is caused by the functions in the sequence have too much room to "wiggle around", and
- as such, the integrals $\int f_k(x)dx$ are all greater than the integral of the limit $\int f(x)dx$.

Of course, by negating all the functions $f_k(x)$ we can get $\lim_{k\to\infty} \int f_k(x)dx < \int f(x)dx$. But, as it turns out, for nonnegative functions, *this sort of behavior is the only behavior possible*. In other words,

For nonnegative functions, if limit of integral is not equal to integral of limit, the former one is always larger.

Lemma 38.3.1 (Fatou's lemma, preliminary version)
Let $f_1, f_2, \dots : \Omega \to [0, +\infty]$ be a sequence of *nonnegative* measurable functions, converging pointwise to f. Then f is nonnegative, measurable, and

$$\int_\Omega f \; d\mu \le \lim_{n\to\infty} \left(\int_\Omega f_n \; d\mu \right).$$

Here we allow either side to be $+\infty$.

As it turns out, this lemma can significantly be generalized as follows. If you compare the two statements, you can see the two lim operators are changed to lim inf — when the sequence actually converges, lim inf and lim equals.

Lemma 38.3.2 (Fatou's lemma)
Let $f_1, f_2, \dots : \Omega \to [0, +\infty]$ be a sequence of *nonnegative* measurable functions. Then $\liminf_{n\to\infty} f_n \colon \Omega \to [0, +\infty]$ is measurable and

$$\int_\Omega \left(\liminf_{n\to\infty} f_n \right) \; d\mu \le \liminf_{n\to\infty} \left(\int_\Omega f_n \; d\mu \right).$$

Here we allow either side to be $+\infty$.

Notice that there are *no extra hypothesis* on f_n other than nonnegative: which makes this quite surprisingly versatile if you ever are trying to prove some general result.

§38.4 Everything else

The big surprise is how quickly all the "big-name" theorem follows from Fatou's lemma. Here is the so-called "monotone convergence theorem".

Corollary 38.4.1 (Monotone convergence theorem)

Let f and $f_1, f_2, \dots : \Omega \to [0, +\infty]$ be a sequence of *nonnegative* measurable functions such that $\lim_n f_n = f$ and $f_n(\omega) \le f(\omega)$ for each n. Then f is measurable and

$$\lim_{n\to\infty} \left(\int_\Omega f_n \, d\mu \right) = \int_\Omega f \, d\mu.$$

Here we allow either side to be $+\infty$.

Proof. We have

$$\begin{aligned} \int_\Omega f \, d\mu &= \int_\Omega \left(\liminf_{n\to\infty} f_n \right) \, d\mu \\ &\le \liminf_{n\to\infty} \int_\Omega f_n \, d\mu \\ &\le \limsup_{n\to\infty} \int_\Omega f_n \, d\mu \\ &\le \int_\Omega f \, d\mu \end{aligned}$$

where the first $\le$ is by Fatou lemma, and the second by the fact that $\int_\Omega f_n \le \int_\Omega f$ for every n. This implies all the inequalities are equalities and we are done. □

You can see how short the proof is — proving $\limsup_{n\to\infty} \int_\Omega f_n \, d\mu \le \int_\Omega f \, d\mu$ is the easy half, and the difficult half is automatically taken care of by Fatou's lemma.

Remark 38.4.2 (The monotone convergence theorem does not require monotonicity!) — In the literature it is much more common to see the hypothesis $f_1(\omega) \le f_2(\omega) \le \dots \le f(\omega)$ rather than just $f_n(\omega) \le f(\omega)$ for all n, which is where the theorem gets its name. However as we have shown this hypothesis is superfluous! This is pointed out in https://mathoverflow.net/a/296540/70654, as a response to a question entitled "Do you know of any very important theorems that remain unknown?".

Example 38.4.3 (Monotone convergence gives $\mathbf{1}_\mathbb{Q}$)

This already implies Example 38.1.1. Letting g_n be the indicator function for $\frac{1}{n!}\mathbb{Z}$ as described in that example, we have $g_n \le \mathbf{1}_\mathbb{Q}$ and $\lim_{n\to\infty} g_n(x) = \mathbf{1}_\mathbb{Q}(x)$, for each individual x. So since $\int_{[0,1]} g_n \, d\mu = 0$ for each n, this gives $\int_{[0,1]} \mathbf{1}_\mathbb{Q} = 0$ as we already knew.

The most famous result, though is the following.

Corollary 38.4.4 (Fatou–Lebesgue theorem)

Let f and $f_1, f_2, \dots : \Omega \to \mathbb{R}$ be a sequence of measurable functions. Assume that $g \colon \Omega \to \mathbb{R}$ is an *absolutely integrable* function for which $|f_n(\omega)| \le |g(\omega)|$ for all $\omega \in \Omega$. Then the inequality

$$\begin{aligned} \int_\Omega \left(\liminf_{n\to\infty} f_n \right) \, d\mu &\le \liminf_{n\to\infty} \left(\int_\Omega f_n \, d\mu \right) \\ &\le \limsup_{n\to\infty} \left(\int_\Omega f_n \, d\mu \right) \le \int_\Omega \left(\limsup_{n\to\infty} f_n \right) \, d\mu. \end{aligned}$$

Proof. There are three inequalities:

- The first inequality follows by Fatou on $g + f_n$ which is nonnegative.
- The second inequality is just $\liminf \le \limsup$. (This makes the theorem statement easy to remember!)
- The third inequality follows by Fatou on $g - f_n$ which is nonnegative. □

Exercise 38.4.5. Where is the fact that g is absolutely integrable used in this proof?

Corollary 38.4.6 (Dominated convergence theorem)

Let $f_1, f_2, \dots : \Omega \to \mathbb{R}$ be a sequence of measurable functions such that $f = \lim_{n\to\infty} f_n$ exists. Assume that $g \colon \Omega \to \mathbb{R}$ is an *absolutely integrable* function for which $|f_n(\omega)| \le |g(\omega)|$ for all $\omega \in \Omega$. Then

$$\int_\Omega f \, d\mu = \lim_{n\to\infty} \left(\int_\Omega f_n \, d\mu \right).$$

In other words,

If there's only finite "space" for the functions f_k to "wiggle around", then no anomaly can happen.

Proof. If $f(\omega) = \lim_{n\to\infty} f_n(\omega)$, then $f(\omega) = \liminf_{n\to\infty} f_n(\omega) = \limsup_{n\to\infty} f_n(\omega)$. So all the inequalities in the Fatou-Lebesgue theorem become equalities, since the leftmost and rightmost sides are equal. □

Note this gives yet another way to verify Example 38.1.1. In general, the dominated convergence theorem is a favorite cliché for undergraduate exams, because it is easy to create questions for it. Here is one example showing how they all look.

Example 38.4.7 (The usual Lebesgue dominated convergence examples)

Suppose one wishes to compute

$$\lim_{n\to\infty} \left(\int_{(0,1)} \frac{n \sin(n^{-1}x)}{\sqrt{x}} \right) dx$$

then one starts by observing that the inner term is bounded by the absolutely integrable function $x^{-1/2}$. Therefore it equals

$$\begin{aligned} \int_{(0,1)} \lim_{n\to\infty} \left(\frac{n \sin(n^{-1}x)}{\sqrt{x}} \right) dx &= \int_{(0,1)} \frac{x}{\sqrt{x}} \, dx \\ &= \int_{(0,1)} \sqrt{x} \, dx = \frac{2}{3}. \end{aligned}$$

We can also say something else about the behavior of the anomalies — that is, when $|\Omega| < \infty$, the anomaly only happens in a set of *small measure*.

Theorem 38.4.8 (Egorov's theorem)

Let $f_1, f_2, \ldots : \Omega \to \mathbb{R}$ be a sequence of measurable functions, on a measure space Ω with $|\Omega| < \infty$, such that $f = \lim_{n\to\infty} f_n$ exists and is finite almost everywhere. Then, for any $\varepsilon > 0$, we can find a subset $U \subseteq \Omega$, such that the remainder has small measure:

$$|\Omega \setminus U| < \varepsilon,$$

and the convergence is uniform on U: the sequence

$$f_1|_U, f_2|_U, \ldots$$

converges to f_U uniformly.

This is because of the following theorem.

Theorem 38.4.9 (Uniform convergence theorem)

Let $f_1, f_2, \ldots : \Omega \to \mathbb{R}$ be a sequence of integrable functions, on a measure space Ω with $|\Omega| < \infty$, such that $\lim_{n\to\infty} f_n = f$, and the convergence is uniform. Then f is integrable and,

$$\lim_{n\to\infty} \left(\int_\Omega f_n \, d\mu \right) = \int_\Omega f \, d\mu.$$

In other words,

The fact that $\int f \, d\mu \neq \lim \int f_k \, d\mu$ is only caused by $\int_{\Omega\setminus U} f \, d\mu \neq \lim \int_{\Omega\setminus U} f \, d\mu$.

Example 38.4.10 (Removing a set of small measure will allow interchanging the integral and the limit)

We take a few examples from Example 38.1.4, and see what happens if we remove a set of small measure here.

- Consider the sequence $f_k(x) = k \cdot \mathbf{1}_{(0,1/k)}$. If, for any $\varepsilon > 0$, we delete a segment $(0, \varepsilon)$ from the domain of f_k, then we will have that f_k converges uniformly to f as $k \to \infty$, and that $\lim_{k\to\infty} \int f_k(x) \, dx = \int f(x) \, dx = 0$.
- Similarly, the sequence $f_k(x) = \frac{\sin(1/x)}{x} \cdot \mathbf{1}_{(1/k,1)}$ converges pointwise to $f(x) = \frac{\sin(1/x)}{x} \cdot \mathbf{1}_{(0,1)}$, and if we delete a segment $(0, \varepsilon)$, then everything checks out.

Remark 38.4.11 — Just because we only need to delete a set of small measure, doesn't mean the set is concentrated in a small interval. The reader is invited to construct a sequence $f_k \colon [0,1] \to \mathbb{R}^+$ that converges pointwise to f, but in order to make the convergence uniform, a dense subset of $[0,1]$ need to be removed. (Hint: take any discontinuous everywhere nonnegative function f, and set $f_k = \min(k, f)$.)

§38.5 Fubini and Tonelli

TO BE WRITTEN

§38.6 A few harder problems to think about

problems

39 Bonus: A hint of Pontryagin duality

In this short chapter we will give statements about how to generalize our Fourier analysis (a bonus chapter Chapter 14) to a much wider class of groups G.

§39.1 LCA groups

Prototypical example for this section: $\mathbb{T}$, $\mathbb{R}$.

Earlier we played with $\mathbb{R}$, which is nice because in addition to being a topological space, it is also an abelian group under addition. These sorts of objects which are both groups and spaces have a name.

Definition 39.1.1. A group G is a **topological group** is a Hausdorff[1] topological space equipped also with a group operation $(G, \cdot)$, such that both maps

$$\begin{aligned} G \times G \to G \quad &\text{by} \quad (x, y) \mapsto xy \\ G \to G \quad &\text{by} \quad x \mapsto x^{-1} \end{aligned}$$

are continuous.

For our Fourier analysis, we need some additional conditions.

Definition 39.1.2. A **locally compact abelian (LCA) group** G is one for which the group operation is abelian, and moreover the topology is *locally compact*: for every point p of G, there exists a compact subset K of G such that $K \ni p$, and K contains some open neighborhood of p.

Our previous examples all fall into this category:

Example 39.1.3 (Examples of locally compact abelian groups)

- Any finite group Z with the discrete topology is LCA.
- The circle group $\mathbb{T}$ is LCA and also in fact compact.
- The real numbers $\mathbb{R}$ are an example of an LCA group which is *not* compact.

These conditions turn out to be enough for us to define a measure on the space G. The relevant theorem, which we will just quote:

[1] Some authors omit the Hausdorff condition.

Theorem 39.1.4 (Haar measure)
Let G be a locally compact abelian group. We regard it as a measurable space using its Borel σ-algebra $\mathscr{B}(G)$. There exists a measure $\mu\colon \mathscr{B}(G) \to [0,\infty]$, called the **Haar measure**, satisfying the following properties:

- $\mu(gS) = \mu(S)$ for every $g \in G$ and measurable S. That means that μ is "translation-invariant" under translation by G.
- $\mu(K)$ is finite for any compact set K.
- if S is measurable, then $\mu(S) = \inf\{\mu(U) \mid U \supseteq S \text{ open}\}$.
- if U is open, then $\mu(U) = \sup\{\mu(S) \mid S \supseteq U \text{ measurable}\}$.

Moreover, it is unique up to scaling by a positive constant.

Remark 39.1.5 — Note that if G is compact, then $\mu(G)$ is finite (and positive). For this reason the Haar measure on a LCA group G is usually normalized so $\mu(G) = 1$.

For this chapter, we will only use the first two properties at all, and the other two are just mentioned for completeness. Note that this actually generalizes the chapter where we constructed a measure on $\mathscr{B}(\mathbb{R}^n)$, since $\mathbb{R}^n$ is an LCA group!

So, in short: if we have an LCA group, we have a measure μ on it.

§39.2 The Pontryagin dual

Now the key definition is:

Definition 39.2.1. Let G be an LCA group. Then its **Pontryagin dual** is the abelian group

$$\widehat{G} := \{\text{continuous group homomorphisms } \xi\colon G \to \mathbb{T}\}.$$

The maps ξ are called **characters**. It can be itself made into an LCA group.[2]

Example 39.2.2 (Examples of Pontryagin duals)

- $\widehat{\mathbb{Z}} \cong \mathbb{T}$, since group homomorphisms $\mathbb{Z} \to \mathbb{T}$ are determined by the image of 1.
- $\widehat{\mathbb{T}} \cong \mathbb{Z}$. The characters are given by $\theta \mapsto n\theta$ for $n \in \mathbb{Z}$.
- $\widehat{\mathbb{R}} \cong \mathbb{R}$. This is because a nonzero continuous homomorphism $\mathbb{R} \to S^1$ is determined by the fiber above $1 \in S^1$. (Algebraic topologists might see covering projections here.)
- $\widehat{\mathbb{Z}/n\mathbb{Z}} \cong \mathbb{Z}/n\mathbb{Z}$, characters ξ being determined by the image $\xi(1) \in \mathbb{T}$.
- $\widehat{G \times H} \cong \widehat{G} \times \widehat{H}$.

[2]If you must know the topology, it is the **compact-open topology**: for any compact set $K \subseteq G$ and open set $U \subseteq \mathbb{T}$, we declare the set of all ξ with $\xi^{\text{img}}(K) \subseteq U$ to be open, and then take the smallest topology containing all such sets. We won't use this at all.

Exercise 39.2.3 ($\widehat{Z} \cong Z$, for those who read Section 18.1)**.** If Z is a finite abelian group, show that $\widehat{Z} \cong Z$, using the results of the previous example. You may now recognize that the bilinear form $\cdot \colon Z \times Z \to \mathbb{T}$ is exactly a choice of isomorphism $Z \to \widehat{Z}$. It is not "canonical".

True to its name as the dual, and in analogy with $(V^\vee)^\vee \cong V$ for vector spaces V, we have:

Theorem 39.2.4 (Pontryagin duality theorem)
For any LCA group G, there is an isomorphism

$$G \cong \widehat{\widehat{G}} \qquad \text{by} \qquad x \mapsto (\xi \mapsto \xi(x)) .$$

The compact case is especially nice.

Proposition 39.2.5 (G compact $\iff$ $\widehat{G}$ discrete)
Let G be an LCA group. Then G is compact if and only if $\widehat{G}$ is discrete.

Proof. Problem 39B. $\square$

§39.3 The orthonormal basis in the compact case

Let G be a compact LCA group, and work with its Haar measure. We may now let $L^2(G)$ be the space of square-integrable functions to $\mathbb{C}$, i.e.

$$L^2(G) = \left\{ f \colon G \to \mathbb{C} \quad \text{such that} \quad \int_G |f|^2 < \infty \right\}.$$

Thus we can equip it with the inner form

$$\langle f, g \rangle = \int_G f \cdot \overline{g}.$$

In that case, we get all the results we wanted before:

Theorem 39.3.1 (Characters of $\widehat{G}$ form an orthonormal basis)
Assume G is LCA and compact (so $\widehat{G}$ is discrete). Then the characters

$$(e_\xi)_{\xi \in \widehat{G}} \qquad \text{by} \qquad e_\xi(x) = e(\xi(x)) = \exp(2\pi i \xi(x))$$

form an orthonormal basis of $L^2(G)$. Thus for each $f \in L^2(G)$ we have

$$f = \sum_{\xi \in \widehat{G}} \widehat{f}(\xi) e_\xi$$

where

$$\widehat{f}(\xi) = \langle f, e_\xi \rangle = \int_G f(x) \exp(-2\pi i \xi(x)) \, d\mu.$$

The sum $\sum_{\xi \in \widehat{G}}$ makes sense since $\widehat{G}$ is discrete. In particular,

- Letting $G = Z$ for a finite group G gives "Fourier transform on finite groups".
- The special case $G = \mathbb{Z}/n\mathbb{Z}$ has its own Wikipedia page: the "discrete-time Fourier transform".
- Letting $G = \mathbb{T}$ gives the "Fourier series" earlier.

§39.4 The Fourier transform of the non-compact case

If G is LCA but not compact, then Theorem 39.3.1 becomes false. On the other hand, it's still possible to define $\widehat{G}$. We can then try to write the Fourier coefficients anyways: let

$$\widehat{f}(\xi) = \int_G f \cdot \overline{e_\xi} \, d\mu$$

for $\xi \in \widehat{G}$ and $f \colon G \to \mathbb{C}$. The results are less fun in this case, but we still have, for example:

Theorem 39.4.1 (Fourier inversion formula in the non-compact case)

Let μ be a Haar measure on G. Then there exists a unique Haar measure ν on $\widehat{G}$ (called the **dual measure**) such that: whenever $f \in L^1(G)$ and $\widehat{f} \in L^1(\widehat{G})$, we have

$$f(x) = \int_{\widehat{G}} \widehat{f}(\xi)\xi(x) \, d\nu$$

for almost all $x \in G$ (with respect to μ). If f is continuous, this holds for all x.

So while we don't have the niceness of a full inner product from before, we can still in some situations at least write f as integral in sort of the same way as before.

In particular, they have special names for a few special G:

- If $G = \mathbb{R}$, then $\widehat{G} = \mathbb{R}$, yielding the "(continuous) Fourier transform".
- If $G = \mathbb{Z}$, then $\widehat{G} = \mathbb{T}$, yielding the "discrete time Fourier transform".

§39.5 Summary

We summarize our various flavors of Fourier analysis from the previous sections in the following table. In the first part G is compact, in the second half G is not.

Name	Domain G	Dual $\widehat{G}$	Characters
Binary Fourier analysis	$\{\pm 1\}^n$	$S \subseteq \{1, \dots, n\}$	$\prod_{s \in S} x_s$
Fourier transform on finite groups	Z	$\xi \in \widehat{Z} \cong Z$	$e(i\xi \cdot x)$
Discrete Fourier transform	$\mathbb{Z}/n\mathbb{Z}$	$\xi \in \mathbb{Z}/n\mathbb{Z}$	$e(\xi x/n)$
Fourier series	$\mathbb{T} \cong [-\pi, \pi]$	$n \in \mathbb{Z}$	$\exp(inx)$
Continuous Fourier transform	$\mathbb{R}$	$\xi \in \mathbb{R}$	$e(\xi x)$
Discrete time Fourier transform	$\mathbb{Z}$	$\xi \in \mathbb{T} \cong [-\pi, \pi]$	$\exp(i\xi n)$

You might notice that the **various names are awful**. This is part of the reason I got confused as a high school student: every type of Fourier series above has its own Wikipedia article. If it were up to me, we would just use the term "G-Fourier transform", and that would make everyone's lives a lot easier.

§39.6 A few harder problems to think about

Problem 39A. If G is compact, so $\widehat{G}$ is discrete, describe the dual measure ν.

Problem 39B. Show that an LCA group G is compact if and only if $\widehat{G}$ is discrete. (You will need the compact-open topology for this.)