# 2 The Lebesgue Measure on $\mathbb{R}^n$

## 2.1 Defining the Lebesgue Measure

### 2.1.0 The empty set

Define $\lambda(\varnothing) = 0$.

### 2.1.1 Special rectangles

We can take a closed interval $[a, b] \subset \mathbb{R}$, and then we can take a rectangle or box as

$$I = [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_n, b_n] \subset \mathbb{R}^n = \{x \in \mathbb{R}^n : a_i \leq x_i \leq b_i\}.$$

Then we define

$$\lambda(I) = (b_1 - a_1) \ldots (b_n - a_n) = \prod_{i=1}^{n} (b_i - a_i).$$

**Exercise 2.1.** *Let $I \subset \mathbb{R}^n$ be a special rectangle. Prove that the following conditions are equivalent:*

1. $\lambda(I) = 0$
2. $I^\circ = \varnothing$
3. *$I$ is contained in an affine subspace of $\mathbb{R}^n$ having dimension smaller than $n$. (An affine subspace is a set $\{x_0 + x : x \in E\}$ where $E$ is a subspace and $x_0$ is a fixed point.)*

We will call these "rectangles" even if they are one-dimensional, or one-hundred-dimensional. This is mostly because the pictures we're going to draw are all two-dimensional, and that's mostly because those are easy to draw.

### 2.1.2 Special Polygons

A *special polygon* is a finite union of special rectangles, each of which has nonzero measure. All of the sides or edges must be perpendicular to a coordinate axis.

We can define the measure of a special polygon straightforwardly. If $I_1, \ldots, I_N$ are special rectangles with disjoint interiors, and $P = \bigcup_{k=1}^{N} I_k$ is a special polygon, then $\lambda(P) = \sum_{k=1}^{N} \lambda(I_k)$.

This is really the only reasonable definition: if we chop our special polygon into pieces, we want the measure of the pieces to add up to the measure of the polygon.

There are two compatibility conditions we should need to check, but they're tedious and boring and straightforward to check so we'll just state them here.

**Fact 2.2.**

- *Every special polygon can be expressed as the union of finitely many special rectangles with disjoint interiors.*
- *If $P$ is a special polygon, and $P = \bigcup_{k=1}^{n} I_k = \bigcup_{\ell=1}^{m} J_\ell$ are two different ways of writing $P$ as a union of special rectangles with disjoint interiors, then $\sum_{k=1}^{n} \lambda(I_k) = \sum_{\ell=1}^{m} J_\ell$.*

**Proposition 2.3.**

- *If $P_1 \subseteq P_2$ then $\lambda(P_1) \leq \lambda(P_2)$.*
- *If $P_1$ and $P_2$ have disjoint interiors, then $\lambda(P_1 \cup P_2) = \lambda(P_1) + \lambda(P_2)$.*

### 2.1.3 Open Sets

Here we want to define the measure of any open set. We will follow this by defining the measure of compact sets, and then extend to arbitrary sets by squeezing them between open and compact sets.

**Definition 2.4.** If $\varnothing \neq G \subseteq \mathbb{R}^n$ is an open set, we define

$$\lambda(G) = \sup \{\lambda(P) : P \subseteq G, P \text{ is a special polygon}\}.$$

We know the set we're taking the supremum over is non-empty, since $G$ has some interior and thus contains some rectangle. If the set of polygon measures is bounded then $\lambda(G)$ is a real number; if the set is unbounded, then we write $\lambda(G) = \infty$.

**Proposition 2.5.** *If $G$ is open and $P$ is a special polygon with $P \subset G$, then there is another special polygon $P'$ with $P \subset P' \subset G$ and $\lambda(P) < \lambda(P')$.*

*Proof.* Since $P$ is closed and $G$ is open, we have $G \cap P^C$ open. Let $x \in G \cap P^C$; then there is an $r$ such that $x \in B_r(x) \subset G \cap P^C$, and we can let $I$ be a closed special rectangle contained in $B_r(x)$. Then set $P' = P \cup I$; $P'$ is a special rectangle, and clearly $P \subset P' \subset G$. □

**Exercise 2.6.** *If $G$ is a bounded open set, prove that $\lambda(G) < \infty$.*

**Proposition 2.7.** *Let $G \subseteq \mathbb{R}^n$ be an open set. Then*

1. $0 \leq \lambda(G) \leq \infty$
2. $\lambda(G) = 0$ *if and only if* $G = \varnothing$.

3. $\lambda(\mathbb{R}^n) = \infty$.

4. *If* $G_1 \subset G_2$ *are open sets, then* $\lambda(G_1) \leq \lambda(G_2)$.

5. *If* $G_k$ *is open for* $k \in \mathbb{N}$*, then*
$$\lambda\left(\bigcup_{k=1}^{\infty} G_k\right) \leq \sum_{k=1}^{\infty} \lambda(G_k).$$

6. *If* $G_k$ *are* disjoint *open sets, then*
$$\lambda\left(\bigsqcup_{k=1}^{\infty} G_k\right) = \sum_{k=1}^{\infty} \lambda(G_k).$$

7. *If* $P$ *is a special polygon, then* $\lambda(P) = \lambda(P^\circ)$.

*Proof.* 1. By definition.

2. If $G \neq \varnothing$ then $G$ contains some nontrivial special polygon $P$. Then $\lambda(G) \geq \lambda(P) > 0$.

3. Exercise.

4. This is basically a property of suprema. If $G_1 \subset G_2$, then any special polygon contained in $G_1$ is also contained in $G_2$. Thus
$$\{P \subset G_1\} \subset \{P \subset G_2\}$$
$$\{\lambda(P) : P \subset G_1\} \subset \{\lambda(P) : P \subset G_2\}$$
$$\sup\{\lambda(P) : P \subset G_1\} \subset \sup\{\lambda(P) : P \subset G_2\}$$

since any upper bound for the larger set is also an upper bound for the smaller set.

5. This one is trickier than it looks. First note that the union is in fact an open set.

   Let $P$ be a special polygon such that $P \subset \bigcup_{k=1}^{\infty} G_k$. Since $P$ is compact, we know there is a Lebesgue number $\varepsilon > 0$ such that, for every $x \in P$, there is a $k$ with $B_\varepsilon(x) \subset G_k$. (See lemma 1.27).

   We know $P$ is a union of non-overlapping rectangles; we can always further subdivide those rectangles, and thus we can assume that $P = \bigcup_{j=1}^{n} I_j$ with each $I_j$ a special rectangle of diameter less than $2\varepsilon$. If we let $x_j$ be the center of the rectangle $I_j$, then we have $I_j \subset B_\varepsilon(x_j) \subset G_k$ for some $k$.

Now we can divide our rectangles up according to their open sets. For each $k$, define $P_k$ to be the union of all $I_j$ such that $I_j \subset G_k$ and $I_J \not\subset G_i$ for $i < k$. (This second condition is just to make sure we don't double-count any rectangle). But every rectangle is contained in at least one of these open sets, so $P = \bigcup_{k=1}^{\infty} P_k$.

Most of these $P_k$ are empty, since there are only finitely many special rectangles running around. But each non-empty $P_k$ is a special polygon, with $P_k \subset G_k$. And we know the $P_k$ have disjoint interiors. Thus we know that

$$\lambda(P) = \sum_{k=1}^{N} \lambda(P_k) \leq \sum_{k=1}^{N} \lambda(G_k) \leq \sum_{k=1}^{\infty} \lambda(G_k).$$

But we've shown that for any special polygon $P \subset \bigcup G_k$, we have $\lambda(P) \leq \sum \lambda(G_k)$. Thus we have

$$\lambda\left(\bigcup_{k=1}^{\infty} G_k\right) = \sup\left\{\lambda(P) : P \subset \bigcup G_k\right\} \leq \sum_{k=1}^{\infty} \lambda(G_k).$$

6. We already know that $\lambda \bigsqcup G_k \leq \sum \lambda(G_k)$ by property 5. So we just need to show the reverse, that $\sum \lambda(G_k) \leq \lambda \bigsqcup G_k$.

   For each $k$, let $P_k$ be a special polygon with $P_k \subset G_k$. Then the $P_k$ are disjoint, and for any $n \in \mathbb{N}$ we have

$$\sum_{k=1}^{n} \lambda(P_k) = \lambda\left(\bigcup_{k=1}^{n} P_k\right) \leq \lambda\left(\bigsqcup_{k=1}^{\infty} G_k\right)$$

   Since this is true for any special polygons $P_k \subset G_k$, we know that the union is an upper bound for any polygons; thus it's an upper bound for the supremum, and we get

$$\sum_{k=1}^{n} \lambda(G_k) \leq \lambda\left(\bigsqcup_{k=1}^{\infty} G_k\right).$$

   Then this statement is true for any finite sum on the left, so it must still be true in the limit; so we have

$$\sum_{k=1}^{\infty} \lambda(G_k) \leq \lambda\left(\bigsqcup_{k=1}^{\infty} G_k\right).$$

   And this is what we needed to show.

7. It's easy to see that $\lambda(P^\circ) \leq \lambda(P)$ (though not completely trivial). If $Q$ is a special polygon with $Q \subset P^\circ$, then $Q \subset P$ and thus $\lambda(Q) \leq \lambda(P)$. This is true for any $Q$, and thus we have
$$\lambda(P^\circ) = \sup_{Q \subset P^\circ} \lambda(Q) \leq \lambda(P).$$
Now we need to prove the other direction. We'll start by proving it for special rectangles. If $I$ is a special rectangle, then for any $\varepsilon > 0$ we can find a rectangle $I' \subset I^\circ$ such that $\lambda(I') > \lambda(I) - \varepsilon$ (by simply shrinking each dimension by $\sqrt[n]{\varepsilon}/2$ or something like that).

   This tells us that $\lambda(I^\circ) > \lambda(I) - \varepsilon$. But this is true for any $\varepsilon > 0$, so we have $\lambda(I^\circ) \geq \lambda(I)$.

   Now if $P$ is a special polygon written as a union of non-overlapping special rectangles $I_k$, then $\bigcup_{k=1}^n I_k^\circ$ is a disjoint union contained in $P^\circ$. Thus
$$\lambda(P) = \sum_{k=1}^n \lambda(I_k) \leq \sum_{k=1}^n \lambda(I_k^\circ) \leq \lambda(P)^\circ.$$
□

**Exercise 2.8.** *Prove that every nonempty open subset of* $\mathbb{R}$ *can be written as a countable disjoint union of open intervals* $G = \bigcup k(a_k, b_k)$*, and this expression is unique.*

*Then conclude that* $\lambda(G) = \sum_k (b_k - a_k)$*.*

*Remark* 2.9. In $\mathbb{R}$ we can use this as our construction, but it doesn't really generalize to $\mathbb{R}^n$ easily. You *can* make that work, but it's even more painful.

### 2.1.4 Compact Sets

If $K \subset \mathbb{R}^n$ is compact, then define
$$\lambda(K) = \inf\{\lambda(G) : K \subset G, G \text{ open}\}.$$

There's something we immediately have to check: if $K$ is a compact special polygon, does this new definition match the old one?

This is a little hard to talk about, so we'll introduce some very temporary notation. We'll use $\alpha$ for the definition of measure we gave in 2.1.2 that applies specifically to special polygons. And we'll use $\beta$ for the definition that applies to any compact set. We'll prove they're both the same, and then we can go back to calling both of them $\lambda$ instead.

It's any to see that $\alpha(P) \leq \beta(P)$ for any special polygon $P$. Whenever $P \subset G$, then $\alpha(P) \leq \lambda(G)$. Therefore, $\alpha(P) \leq \inf\{\lambda(G)\} = \beta(P)$.

Conversely, we want to show that $\beta(P) \leq \alpha(P)$. Suppose $P = \bigcup_{k=1}^n I_k$ is a union of non-overlapping rectangles. For any $\varepsilon > 0$ we can pick special rectangles $I_k' \subset I_k^\circ$ such that $\lambda(I_k') < \lambda(I_k) + \varepsilon/n$.

Then if we set $G = \bigcup_{k=1}^n I_k'^\circ$ we have $P \subset G$, and thus

$$\begin{aligned} \beta(P) \leq \lambda(G) &\leq \sum_{k=1}^n \lambda(I_k'^\circ) \\ &< \sum_{k=1}^n \lambda(I_k) + \varepsilon = \alpha(P) + \varepsilon. \end{aligned}$$

Since this is true for any $\varepsilon > 0$, we have $\beta(P) \leq \alpha(P)$.

We want to prove several properties of the measure of these compact sets. But mostly we can leverage the results we already proved about open sets.

**Proposition 2.10.** *1.* $0 \leq \lambda(K) < \infty$

*2. If* $K_1 \subset K_2$ *then* $\lambda(K_1) \leq \lambda(K_2)$.

*Proof.* Exercise ☐

*3.* $\lambda(K_1 \cup K_2) \leq \lambda(K_1) + \lambda(K_2)$.

*Proof.* If $K_1 \subset G_1$ and $K_2 \subset G_2$ then $K_1 \cup K_2 \subset G_1 \cup G_2$, and thus

$$\lambda(K_1 \cup K_2) \leq \lambda(G_1 \cup G_2) \leq \lambda(G_1) + \lambda(G_2).$$

Thus

$$\lambda(K_1 \cup K_2) \leq \inf \lambda(G_1) + \lambda(G_2) = \lambda(K_1) + \lambda(K_2).$$

☐

*4. If* $K_1$ *and* $K_2$ *are disjoint, then* $\lambda(K_1 \cup K_2) = \lambda(K_1) + \lambda(K_2)$.

*Proof.* Wince $K_1$ and $K_2$ are compact, there is a $\varepsilon > 0$ such that for every $x \in K_1, y \in K_2$, then $d(x, y) \geq \varepsilon$. (This is the Lebesgue number again, with the open sets being $K_1^C$ and $K_2^C$.) Then if we let $G$ be an open set containing $K_1 \cup K_2$, we can write

$$\begin{aligned} G_1 &= G \cap \bigcup_{x \in K_1} B_{\varepsilon/2}(x) \\ G_2 &= G \cap \bigcup_{x \in K_2} B_{\varepsilon/2}(x). \end{aligned}$$

Then we have $K_i \subset G_i$ and $G_1 \cap G_2 = \varnothing$. So we have

$$\lambda(K_1) + \lambda(K_1) \leq \lambda(G_1) + \lambda(G_2) = \lambda(G_1 \cup G_2) \leq \lambda(G).$$

Since this holds for any $G \supset K_1 \cup K_2$, we have $\lambda(K_1) + \lambda(K_2) \leq \lambda(K_1 \cup K_2)$. Since the opposte inequality follows from part (3), that proves equality.

□

*Remark* 2.11. We didn't try to prove any results about infinite unions of compact sets. Why not?

Here we should mention one very important example: the Cantor set. (It is sometimes known as the ternary Cantor set to distinguish from some generalizations.)

**Definition 2.12.** We first define a family of open intervals contained in $[0, 1]$. We define $G_1 = (\frac{1}{3}, \frac{2}{3})$; then $[0, 1] \setminus G_1$ is two closed intervals of length one third. We remove the middle third of each of these: we define $G_2 = (\frac{1}{9}, \frac{2}{9}) \cup (\frac{7}{9}, \frac{8}{9})$. Then $[0, 1] \setminus (G_1 \cup G_2) =$ is four closed intervals of length $1/9$. We can iterate this construction to get an infinite sequence of disjoint open sets $G_1, G_2, \dots$.

We define the (ternary) *Cantor set* to be the set

$$C = [0, 1] \setminus \bigcup_{k=1}^{\infty} G_k.$$

**Fact 2.13.** *The Cantor set $C$ is uncountable.*

**Exercise 2.14.** *Prove that $C$ is compact. Then prove that $\lambda(C) = 0$.*

### 2.1.5 Inner and Outer Measure

We would like to extend our definition of measure to cover any set. We don't quite have the ability to do that yet, but we can define two quanties that do apply to any set.

**Definition 2.15.** Let $A \subseteq \mathbb{R}^n$. Then we define

- The *outer measure* of $A$

$$\lambda^*(A) = \inf\{\lambda(G) : A \subset G \text{ open}\}$$

- The *inner measure* of $A$

$$\lambda_*(A) = \sup\{\lambda(K) : A \supset K \text{ compact}\}.$$

Notice that these are basically concepts we've seen before; outer measure is how we defined the measure of a compact set, and inner measure is basically how we defined the measure of an open set. So this entire set of definitions has a sort of push-pull quality.

**Proposition 2.16.** *1.* $\lambda_*(A) \leq \lambda^*(A)$.

*Proof.* If $K \subset A \subset G$, then $K \subset G$, and thus $\lambda(K) \leq \lambda(G)$. □

*2. If* $A \subseteq B$ *then* $\lambda^*(A) \leq \lambda^*(B)$ *and* $\lambda_*(A) \leq \lambda_*(B)$.

*3.* $\lambda^* \left(\bigcup_{k=1}^{\infty} A_k\right) \leq \sum_{k=1}^{\infty} \lambda^*(A_k)$.

*Proof.* We basically want to cover each $A_k$ with an open set. If $\varepsilon > 0$, then for each $k$ we can find a $G_k \supseteq A_k$ such that $\lambda(G_k) < \lambda^*(A_k) + \varepsilon 2^{-k}$. Then we have

$$\begin{aligned}\lambda^* \left(\bigcup_{k=1}^{\infty} A_k\right) &\leq \lambda \left(\bigcup_{k=1}^{\infty} G_k\right) \leq \sum_{k=1}^{\infty} \lambda(G_k) \\ &< \sum_{k=1}^{\infty}(\lambda^*(A_k) + \varepsilon 2^{-k}) = \sum_{k=1}^{\infty} \lambda^*(A_k) + \varepsilon.\end{aligned}$$

Since this holds for any $\varepsilon > 0$, we're done.

□

*4. If the* $A_k$ *are disjoint, then*

$$\lambda_* \left(\bigcup_{k=1}^{\infty} A_k\right) \geq \sum_{k=1}^{\infty} \lambda_*(A_k).$$

*Proof.* Exercise. □

*5. If* $A$ *is open or compact, then* $\lambda^*(A) = \lambda_*(A) = \lambda(A)$.

*Proof.* If $A$ is open, then clearly $\lambda^*(A) = \lambda(A)$. If $P$ is any special polygon with $P \subset A$, then $P$ is compact, so $\lambda(P) \leq \lambda_*(A)$; and thus $\lambda(A) \leq \lambda_*(A)$.

But then $\lambda(A) \leq \lambda_*(A) \leq \lambda^*(A) = \lambda(A)$, so all the numbers are equal.

Now suppose $A$ is compact. Then $\lambda_*(A) = \lambda(A)$ clearly, and $\lambda(A) = \lambda^*(A)$ because that's the definition of $\lambda(A)$. □

### 2.1.6 Sets with Finite Outer Measure

Recall we want to assign a measure to every possible set. In the last subsubsection 2.1.5 we defined two "measure-like" numbers that apply to any set. But which one should we use?

It turns out that very strange things can happen in general, which we will see later. But all of those strangenesses are avoided if our two measures are in fact the same.

**Definition 2.17.** Let $A \subseteq \mathbb{R}^n$ be a set with finite outer measure. We say that $A$ is *measurable* and belongs to $\mathcal{L}_0$ if $\lambda^*(A) = \lambda_*(A)$, and in that case we define the *measure* of $A$ to be $\lambda(A) = \lambda^*(A) = \lambda_*(A)$.

**Proposition 2.18.** *The family $\mathcal{L}_0$ contains all open sets with finite measure and all compact sets. Our new definition of measure belongs to every previous definition of measure we've given.*

**Lemma 2.19.** *If $A, B \in \mathcal{L}_0$ are disjoint, then $A \cup B \in \mathcal{L}_0$ and $\lambda(A \cup B) = \lambda(A) + \lambda(B)$.*

*Proof.*

$$\begin{aligned}\lambda^*(A \cup B) &\leq \lambda^*(A) + \lambda^*(B) = \lambda(A) + \lambda(B)\\ &= \lambda_*(A) + \lambda_*(B) \leq \lambda_*(A \cup B)\\ &\leq \lambda^*(A \cup B).\end{aligned}$$

□

We want to be able to tell whether a set is measurable in an easy-to-compute way. The main tool for this is the following theorem on approximation, which says that if we can approximate our set with open sets and compact sets that are "close together" then our set is in $\mathcal{L}_0$.

**Theorem 2.20** (Approximation of Measure)**.** *Let $A \subseteq \mathbb{R}^n$ such that $\lambda^*(A) < \infty$. Then $A \in \mathcal{L}_0$ if and only if: for every $\varepsilon > 0$ there is a compact set $K$ and an open $G$ such that $K \subseteq A \subseteq G$ and $\lambda(G \setminus K) < \varepsilon$.*

*Proof.* If $A \in \mathrm{L}_0$, then that means that the inner measure and outer measure of $A$ are the same. But we can always approximate the outer measure well with an open set, and the inner measure with a compact set. So for any $\varepsilon > 0$ we can find $G \subseteq A, K \subseteq A$ such that

$$\begin{aligned}\lambda(G) &< \lambda^*(A) + \varepsilon/2 = \lambda(A) + \varepsilon/2\\ \lambda(K) &> \lambda^*(A) - \varepsilon/2 = \lambda(A) - \varepsilon/2.\end{aligned}$$

Since $K$ and $G \setminus K$ are disjoint, we have $\lambda(G) = \lambda(K) + \lambda(G \setminus K)$. Rearranging this gives

$$\begin{aligned}\lambda(G \setminus K) &= \lambda(G) - \lambda(K) \\ &< \lambda(A) + \varepsilon/2 - \lambda(A) + \varepsilon/2 = \varepsilon.\end{aligned}$$

Conversely, suppose that for any $\varepsilon > 0$ there exist $K \subseteq A \subseteq G$ with $\lambda(G \setminus K) < \varepsilon$. Fix an epsilon, and then choose such sets $G$ and $K$. We have that

$$\begin{aligned}\lambda^*(A) \leq \lambda(G) &= \lambda(K) + \lambda(G \setminus K) \\ &< \lambda(K) + \varepsilon \leq \lambda_*(A) + \varepsilon.\end{aligned}$$

Since this holds for any $\varepsilon > 0$, we conclude that $\lambda^*(A) \leq \lambda_*(A) \leq \lambda^*(A)$. Thus the outer and inner measures are equal, and $A \in \mathcal{L}_0$ by definition.

□

We want to figure out how we can combine $\mathcal{L}_0$ sets to get other $\mathcal{L}_0$ sets. We start by looking at our binary operations, and then we'll figure out how to work with countably many sets at once.

**Proposition 2.21.** *If $A, B \in \mathcal{L}_0$ then $A \cup B, A \cap B, A \setminus B \in \mathcal{L}_0$ as well.*

*Proof.* We'll start with the set difference, using the theorem on approximation.

Fix $\varepsilon > 0$, and then we can write $K_1 \subseteq A \subseteq G_1, K_2 \subseteq B \subseteq G_2$ with $\lambda(G_i \setminus K_i) < \varepsilon/2$. Then set $K = K_1 \setminus G_2$ and $G = G_1 \setminus K_2$.

$G$ is clearly open, and $K$ is closed and thus compact. Further, we have $K \subseteq A \setminus B \subseteq G$, and $G \setminus K \subseteq (G_1 \setminus K_1) \cup (G_2 \setminus K_2)$. Thus $\lambda(G \setminus K) < \varepsilon$, and thus $A \setminus B \in \mathcal{L}_0$.

Given this fact, we can prove the other two claims with minimal work. $A \cap B = A \setminus (A \setminus B)$ is a difference of differences of $\mathcal{L}_0$ sets, and thus is in $\mathcal{L}_0$. And $A \cup B = (A \setminus B) \cup B$ is a disjoint union of $\mathcal{L}_0$ sets, and thus is $\mathcal{L}_0$ by lemma 2.19. □

**Theorem 2.22** (Countable additivity)**.** *Let $A_k \in \mathcal{L}_0$, and set $A = \bigcup_{k=1}^\infty A_k$. Assume $\lambda^*(A) < \infty$. Then $A \in \mathcal{L}_0$, and*

$$\lambda(A) \leq \sum_{k=1}^\infty \lambda(A_k).$$

*Further, if the $A_k$ are disjoint, then*

$$\lambda(A) = \sum_{k=1}^\infty \lambda(A_k).$$

*Proof.* If the $A_k$ are disjoint, this is easy. We know that

$$\begin{aligned}\lambda^*(A) &\leq \sum_{k=1}^{\infty} \lambda^*(A_k) \\ &= \sum_{k=1}^{\infty} \lambda_*(A_k) \leq \lambda_*(A) \leq \lambda^*(A).\end{aligned}$$

If the $A_k$ are not disjoint, we can't do anything this simple. The first inequality holds, but we don't actually have an inequality on the inner measure. But if we can reduce this to a question about a disjoint union, then we can use the previous result and things become much simpler.

Define a new family of sets as follows. We take $B_1 = A_1$, and then for each $k > 1$ we define

$$B_k = A_k \setminus \left( \bigcup_{i=1}^{k-1} A_k \right).$$

Then each $B_k \in \mathcal{L}_0$, and clearly the $B_k$ are disjoint. Each $B_k$ is a subset of the corresponding $A_k$, and $\bigcup_{k=1}^{\infty} B_k = \bigcup_{k=1}^{\infty} A_k = A$. Then we can use our result on disjoint unions to see that $A \in \mathcal{L}_0$, and

$$\lambda(A) = \sum_{k=1}^{\infty} \lambda(B_k) \leq \sum_{k=1}^{\infty} \lambda(A_k).$$

□

### 2.1.7 Measurable Sets

**Definition 2.23.** Let $A \subset \mathbb{R}^n$. We say that $A$ is *(Lebesgue) measurable* if, for every $M \in \mathcal{L}_0$, then $A \cap M \in \mathcal{L}_0$. If $A$ is measurable, we define the *(Lebesgue) measure* of $A$ to be

$$\lambda(A) = \sup\{\lambda(A \cap M) : M \in \mathcal{L}_0\}.$$

We denote the set of all measurable subsets of $\mathbb{R}^n$ with the symbol $\mathcal{L}$.

We now have another (final!) definition of measure; so we need to make sure it's the same as our previous definitions.

**Proposition 2.24.** *Let $A \subseteq \mathbb{R}^n$ with $\lambda^*(A) < \infty$. Then $A \in \mathcal{L}_0$ if and only if $A \in \mathcal{L}$. And if $A \in \mathcal{L}$, then our two definitions of measure coincide.*

*Proof.* If $A \in \mathcal{L}_0$, then $A \cap M \in \mathcal{L}_0$ for any $M \in \mathcal{L}_0$, and thus $A \in \mathcal{L}$.

Conversely, suppose $A \in \mathcal{L}$. We know that $B_k(0) \in \mathcal{L}_0$ since it's open, so we know that $A_k = A \cap B_k(0) \in \mathcal{L}_0$. But $A = \bigcup_{k=1}^{\infty} A_k$, and theorem 2.22 tells us that $A \in \mathcal{L}_0$.

Now we need to prove that the measure formulas coincide; for the rest of this proof we'll use $\lambda'$ for our final definition of measure given in Definition 2.23.

Suppose $A \in \mathcal{L}_0 \subset \mathcal{L}$. Then since for any $M \in \mathcal{L}_0$, we know that $A \cap M \subseteq A$, and so $\lambda(A \cap M) \leq \lambda(A)$, and thus $\lambda'(A) \leq \lambda(A)$. But conversely, $A \in \mathcal{L}_0$, so we must have $\lambda(A) \leq \lambda(A \cap A) \leq \lambda'(A)$. Thus $\lambda' = \lambda$. □

## 2.2 Basic Properties of the Lebesgue Measure

Now that we have finally given a complete definition of Lebesgue measure, we want to collect all the properties that apply to it. Many of these are properties we've seen already at various earlier stages of the construction, but we need to see they still hold at this completed stage. Some other properties are basically new.

**Proposition 2.25.** *1. $A \in \mathcal{L}$ if and only if $A^C \in \mathcal{L}$.*

*2. If $A_k \in \mathcal{L}$, then $\bigcup_{k=1}^\infty A_k \in \mathcal{L}$ and $\bigcap_{k=1}^\infty A_k \in \mathcal{L}$ .*

*3. If $A, B \in \mathcal{L}$ then $A \setminus B \in \mathcal{L}$.*

*Proof.* 1. For any $M \in \mathcal{L}_0$, we know that $A^c \cap M = M \setminus A = M \setminus (A \cap M)$. This is a difference of $\mathcal{L}_0$ sets, and thus is in $\mathcal{L}_0$. Therefore $A^C \in \mathcal{L}$.

2. If $A_k \in \mathcal{L}$ and $A = \bigcup_{k=1}^\infty A_k$, then for any $M$ we have that $A \cup M = \bigcup_{k=1}^\infty A_k \cap M$. Since $\lambda^*(A \cap M) \leq \lambda(M)$ is finite, theorem 2.22 tells us that $A \cap M \in \mathcal{L}_0$. Thus $A \in \mathcal{L}$.

   The result on intersections follows from De Morgan's Laws: we know that
   $$\bigcap_{k=1}^\infty A_k = \left( \bigcup_{k=1}^\infty A_k^C \right)^C .$$
   Since complements and countable unions preserve measurability, this is a measurable set.

3. $A \setminus B = A \cap B^C$ is measurable.

□

**Proposition 2.26.** *[Countable Additivity] If $A_k$ are measurable, then*
$$\lambda\left( \bigcup_{k=1}^\infty A_k \right) \leq \sum_{k=1}^\infty \lambda(A_k).$$
*If the union is disjoint, then*
$$\lambda\left( \bigcup_{k=1}^\infty A_k \right) = \sum_{k=1}^\infty \lambda(A_k).$$

*Proof.* Let $A = \bigcup_{k=1}^{\infty} A_k$. Then by theorem 2.22 we know that

$$\lambda(A \cup M) \leq \sum_{k=1}^{\infty} \lambda(A_k \cap M) \leq \sum_{k=1}^{\infty} \lambda(A_k).$$

Thus $\sum_{k=1}^{\infty} \lambda(A_k)$ is an upper bound for $\lambda(A \cap M)$, and so we have $\lambda(A) \leq \sum_{k=1}^{\infty} \lambda(A_k)$.

Now suppose the sets are disjoint; we just need to prove the opposite inequality. For any $n \in \mathbb{N}$ we can choose sets $M_1, \dots, M_n \in \mathcal{L}_0$, and define $M = \bigcup_{k=1}^{n} M_k$. Then

$$\begin{aligned}\lambda(A) \geq \lambda(A \cap M) &= \sum_{k=1}^{\infty} \lambda(A_k \cap M) \\ &\geq \sum_{k=1}^{n} \lambda(A_k \cap M) \geq \sum_{k=1}^{n} \lambda(A_k \cap M_k).\end{aligned}$$

Since $\lambda(A_k \cap M_k) \leq \lambda(A_k)$, we conclude that $\lambda(A) \geq \sum_{k=1}^{n} \lambda(A_k)$. Since this is true for any $n \in \mathbb{N}$, we must have $\lambda(A) \geq \sum_{k=1}^{\infty} \lambda(A_k)$, as desired.

□

**Proposition 2.27.** *Suppose $A_1, A_2, \dots$ are measurable sets. Then:*

1. *If $A_1 \subseteq A_2 \subseteq \dots$, then*
$$\lambda\left(\bigcup_{k=1}^{\infty} A_k\right) = \lim_{k\to\infty} \lambda(A_k).$$
2. *If $A_1 \supseteq A_2 \supseteq \dots$, and further if $\lambda(A_1) < \infty$, then*
$$\lambda\left(\bigcap_{k=1}^{\infty} A_k\right) = \lim_{k\to\infty} \lambda(A_k).$$

*Proof.* 1. We can write $\bigcup A_k$ as a disjoint union; in this case this is very easy, since we have
$$\bigcup_{k=1}^{\infty} A_k = A_1 \cup \bigcup_{k=2}^{\infty} (A_k \setminus A_{k-1}).$$

Then countable additivity implies that

$$\begin{aligned}\lambda\left(\bigcup_{k=1}^{\infty} A_k\right) &= \lambda(A_1) + \sum_{k=2}^{\infty} \lambda(A_k \setminus A_{k-1}) \\ &= \lim_{n\to\infty} \lambda(A_1) + \sum_{k=2}^{n} \lambda(A_k \setminus A_{k-1}) \\ &= \lim_{n\to\infty} \lambda\left(A_1 \cup \bigcup_{k=2}^{n} (A_k \setminus A_{k-1})\right) = \lim_{n\to\infty} \lambda(A_n).\end{aligned}$$

2. Exercise.

□

**Proposition 2.28.** *1. All open sets and all closed sets are measurable.*

*2. If $\lambda^*(A) = 0$, then $A$ is measurable and $\lambda(A) = 0$.*

*Proof.* 1. If $G$ is open, then we can write $G = \bigcup_{k=1}^{\infty}(G \cap B_k(0))$ as a countable union of bounded open sets. Each bounded open set has finite outer measure and thus is measurable; and we know a countable union of measurable sets is measurable. Thus $G$ is measurable.

If $F$ is closed, then $F^C$ is open and thus measurable. So $F$ is measurable.

2. We know that $0 \leq \lambda_*(A) = 0 \leq \lambda^*(A) = 0$. Thus $A \in \mathcal{L}_0$ and so $A$ is measurable, and $\lambda(A) = 0$.

□

**Proposition 2.29** (Approximation)**.** *Let $A \subseteq \mathbb{R}^n$. Then $A$ is measurable if and only if: for every $\varepsilon > 0$ there exist $F \subseteq A \subseteq G$ such that $\lambda(G \setminus F) < \varepsilon$.*

*Proof.* First, suppose $A$ has the approximation property as described. We're going to approximate $A$ with a clearly measurable set and then show the remainder is so small that it must also be measurable.

For any $k \in \mathbb{N}$ we can find $F_k \subseteq A \subseteq G_k$, with $\lambda(G_k \setminus F_k) < \frac{1}{k}$. Let $B = \bigcup_{k=1}^{\infty} F_k$. Then $B$ is a countable union of measurable sets and thus measurable, and $B \subseteq A$.

Further, we know that $A \setminus B \subseteq G_k \setminus B \subseteq G_k \setminus F_k$, and thus $\lambda^*(A \setminus B) \leq \lambda(G_k \setminus F_k) < \frac{1}{k}$. Since this holds for each $k$, we see that $\lambda^*(A \setminus B) = 0$, and thus $A \setminus B$ is measurable. We conclude that $A = B \cup (A \setminus B)$ is a union of measurable sets and thus measurable.

Conversely, suppose $A$ is a measurable subset of $\mathbb{R}^n$. If we take any *finite measure* subset, we know we can approximate it; so we'll build a sequence of these approximations that approximate all of $A$.

For each $k$, define $E_k = B_k(0) \setminus B_{k-1}(0)$, which you can visualize like a washer or annulus centered at zero. Since each $E_k$ is bounded, we know that $A \cap E_K \in \mathcal{L}_0$, and thus we can find a compact set $K_k$ and an open set $G_k$ such that $K_k \subseteq A \cap E_k \subseteq G_k$ and $\lambda(G_k \setminus K_k) < \varepsilon 2^{-k}$.

We define $F = \bigcup_{k=1}^{\infty} K_k$ and $G = \bigcup_{k=1}^{\infty} G_k$. It's clear that $G_k$ is open. It's less trivial to see that $F$ is closed, but we can check that it contains all of its limit points; if $x \in \overline{F}$ then $x$ must be a limit point of some finite union $\bigcup_{k=1}^{n} K_k$, and this is a finite union and thus closed, so $x \in \bigcup_{k=1}^{n} K_k \subseteq F$.

So we have $F \subseteq A \subseteq G$ are closed and open respectively. And we can see that

$$G \setminus F = \bigcup_{k=1}^{\infty} (G_k \setminus F) \subseteq \bigcup_{k=1}^{\infty} (G_k \setminus K_k)$$

and so

$$\begin{aligned} \lambda(G \setminus F) &\leq \sum_{k=1}^{\infty} \lambda(G_k \setminus K_k) \\ &< \sum_{k=1}^{\infty} \varepsilon 2^{-k} = \varepsilon. \end{aligned}$$

□

**Proposition 2.30.** *1. If $A$ is measurable, then $\lambda_*(A) = \lambda^*(A) = \lambda(A)$.*

*2. If $A \subseteq B$ and $B$ is measurable, then $\lambda^*(A) + \lambda_*(B \setminus A) = \lambda(B)$.*

*Proof.* 1. If $\lambda^*(A) < \infty$, then this follows from section 2.1.6. So suppose $A$ is measurable, and $\lambda^*(A) = \infty$.

If $\lambda(A) < \infty$, then we could find $F \subseteq A \subseteq G$ with $\lambda(G \setminus F) < 1$, and then we'd have that

$$\lambda(G) = \lambda(G \setminus A) + \lambda(A) \leq \lambda(G \setminus F) + \lambda(A) < 1 + \lambda(A) < \infty$$

which is a contradiction.

Now we just need to show that $\lambda_*(A) = \infty$. We know that $\lambda(A \cap B_k(0)) < \infty$, and for any $k$ we have

$$\lambda(A \cap B_k(0)) = \lambda_*(A \cap B_k(0)) \leq \lambda_*(A).$$

But we know that $\lim_{k\to\infty} \lambda(A \cap B_k(0)) = \lambda(A) = \infty$ since this is a union of an ascending chain. Thus we also must have that $\lambda_*(A) = \infty$.

2. For any open $G \supseteq A$, we know that

$$\begin{aligned} \lambda(G) + \lambda_*(B \setminus A) \geq \lambda(B \cap G) + \lambda_*(B \setminus A) &\geq \lambda(B \cap G) + \lambda_*(B \setminus G) \\ &= \lambda(B \cap G) + \lambda(B \setminus G) = \lambda(B). \end{aligned}$$

This holds for any $G$, so we have $\lambda(B) \leq \lambda^*(A) + \lambda_*(B \setminus A)$.

Conversely, for any compact $K \subseteq B \setminus A$, we can do basically the same thing:

$$\begin{aligned} \lambda^*(A) + \lambda(K) &\leq \lambda^*(B \setminus K) + \lambda(K) \\ &= \lambda(B \setminus K) + \lambda(K) = \lambda(B). \end{aligned}$$

Thus $\lambda^*(A) + \lambda_*(B \setminus A) \leq \lambda(B)$.

$\square$

**Proposition 2.31** (Carathéodory)**.** *A set $A$ is measureable if and only if for every $E \subseteq \mathbb{R}^n$, we have that*

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

*Remark* 2.32. This proposition provides another way to construct measure; we could have used the outer measure only and avoided inner measure. But this presentation would have been somewhat less concrete, and made some other steps kind of tricky.

*Proof.* Notice first that this equation is partly cheating. For *any* set $A$, measurable or not, we know that

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

by the countable subadditivity of outer measure as proven in 2.16. So in either direction we're really just looking at the opposite inequality.

Suppose $A$ is measurable. If $E \subset G$ open, then

$$\lambda(G) = \lambda(G \cap A) + \lambda(G \cap A^c) \geq \lambda^*(E \cap A) + \lambda^*(E \cap A^c).$$

Thus

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

by definition of outer measure.

Conversely, suppose that $\lambda^*(E) = \lambda^*(E \cap A) + \lambda^*(E \cap A^c)$ for any $E$. Then in particular, for any *finitely measurable* $M \in \mathcal{L}_0$ we have

$$\lambda(M) = \lambda^*(M) = \lambda^*(M \cap A) + \lambda^*(M \cap A^c).$$

But we also know that

$$\lambda(M) = \lambda_*(M \cap A) + \lambda^*(M \cap A^c)$$

from proposition 2.30, since we can take $M \cap A^c = M \setminus (M \cap A)$.

But subtracting these equations gives that $0 = \lambda^*(M \cap A) - \lambda_*(M \cap A)$, and thus $\lambda^*(M \cap A) = \lambda_*(M \cap A)$; and this is precisely what it means to say that $M \cap A \in \mathcal{L}_0$. since this holds for any measurable $M$, then $A \in \mathcal{L}$ by definition. $\square$

## 2.3 Abstract Measure Spaces

At this point I want to take a moment and discuss which of the properties of the Lebesgue measure generalize, and are necessary for it to be "a measure".

We first want to talk about the properties that measurable sets have to have.

**Definition 2.33.** Let $X$ be any set. We define an *algebra* of subsets of $X$ to be a subset $\mathcal{M} \subseteq 2^X$ of the power set of $X$ that satisfies the following properties:

- $\varnothing \in \mathcal{M}$
- If $A, B \in \mathcal{M}$ then $A \cup B \in \mathcal{M}$.
- If $A \in \mathcal{M}$ then $A^C = X \setminus A \in \mathcal{M}$.

It's easy to see that an algebra of sets must be closed under any finite unions, and also under finite intersections and under set difference.

All these statements are true of the Lebesgue measurable sets. But the measurable sets have one extra property:

**Definition 2.34.** Let $\mathcal{M} \subseteq 2^X$ be an algebra. Then it is a $\sigma$-*algebra* if it is also closed under countable unions (and thus intersections): if $A_1, A_2, \ldots, \in \mathcal{M}$ then $\bigcup_{k=1}^{\infty} A_k \in \mathcal{M}$.

**Example 2.35.**

- The power set $2^X$ is a $\sigma$-algebra.
- $\{\varnothing, X\}$ is a $\sigma$-algebra. In fact, this is a sub-$\sigma$-algebra of any $\sigma$-algebra.
- The measurable sets $\mathcal{L} \subset 2^{\mathbb{R}^n}$ are a $\sigma$-algebra.
- Let $X$ be any set, ane let $\mathcal{M}_0$ be the set of all sets $A$ such that either $A$ is finite or $A^C$ is finite. Then $\mathcal{M}_0$ is an algebra but not a $\sigma$-algebra.
- Let $X$ be any set, ane let $\mathcal{M}_1$ be the set of all sets $A$ such that either $A$ is countable or $A^C$ is countable. Then $\mathcal{M}_0$ is an a $\sigma$-algebra.
- Any finite algebra is a $\sigma$-algebra for basically dumb reasons.

From this we want to find a way to *build* $\sigma$-algebras. There is one lemma which will be very useful for this:

**Exercise 2.36.** *Let $X$ be a set, and $\mathcal{M}_i \subset 2^X$ be a $\sigma$-algebra for each $i$ in some index set $I$. Prove that $\bigcap_{i \in I} \mathcal{M}_i$ is a $\sigma$-algebra.*

Notice this is a little weird. We're not intersecting subsets of $X$ to get a new subset of $X$; we're intersecting collections of subsets of $X$ to get a new collection of subsets of $X$.

Now suppose $\mathcal{N} \subset 2^X$ is any collection of subsets—not necessarily an algebra. We can consider the family of $\sigma$-algebras that contain $\mathcal{N}$. Clearly there are some such $\sigma$-algebras, since $2^X$ is itself a $\sigma$-algebra. If we take the intersection of all these $\sigma$-algebras, we will get a new $\sigma$-algebra:

$$\mathcal{M} = \bigcap_{\mathcal{P} \supseteq \mathcal{N}} \mathcal{P}.$$

Then $\mathcal{M}$ will contain $\mathcal{N}$, and it is contained in any $\sigma$-algebra that contains $\mathcal{N}$, so it is the smallest $\sigma$-algebra containing $\mathcal{N}$. We say that $\mathcal{M}$ is the $\sigma$-algebra *generated* by $\mathcal{N}$.

An important note is that this is, and essentially must be, non-constructive. There are sets in $\mathcal{M}$ that we can't build by a countable chain of unions or intersections of elements of $\mathcal{N}$. In fact, a set in $\mathcal{M}$ can be a countable union of countable intersections of countable unions of countable intersections of ...

If we want to construct the $\sigma$-algebra $\mathcal{M}$ explicitly, we need to do some sort of transfinite induction, which is cumbersome and we just don't want to do it. But it's clear (non-constructively) that $\mathcal{M}$ must exist, and we're satisfied with that.

So far this tells us that we *can* generate $\sigma$-algebras, but doens't tell us what we want to do with them, or which $\sigma$-algebras we want. But if we want to build a measure, we definitely want to be able to measure all the "reasonable" sets.

**Definition 2.37.** The class of *Borel sets* in $\mathbb{R}^n$, denoted $\mathcal{B}$, is the $\sigma$-algebra generated by the collection of open sets. Clearly $\mathcal{B} \subseteq \mathcal{L}$. We sometimes write $\mathcal{B}_n$ when we need to specify the dimension.

**Exercise 2.38.** *Prove that the class of Borel sets is* also *the $\sigma$-algebra generated by the collection of special rectangles.*

*Thus $\mathcal{B}$ is the smallest $\sigma$-algebra that contains all the sets we obviously want to be able to measure.*

The Borel sets in $\mathbb{R}^n$ are not actually all the Lebesgue measurable sets. But they are close.

**Definition 2.39.** If $A \subseteq \mathbb{R}^n$ is measurable with $\lambda(A) = 0$, then $A$ is a *null* set. $A$ is null if and only if $\lambda^*(A) = 0$.

**Theorem 2.40.** *Suppose $A \subseteq \mathbb{R}^n$ is measurable. Then we can write $A = E \cup N$ such that $E$ and $N$ are disjoint, $E$ is a Borel set, and $N$ is a null set.*

*Proof.* For every $k \in \mathbb{N}$ there is a closed set $F_k \subseteq A$ such that $\lambda(A \setminus F_k) < \frac{1}{k}$, by our approximation property 2.29. Set $E = \bigcup_{k=1}^{\infty} F_k$. Then $E$ is not necessarily closed, but it is certainly Borel since it's a countable union of closed, Borel sets.

Further, $E \subseteq A$. Then for any $k$, we have

$$\lambda(A \setminus E) \leq \lambda(A \setminus F_k) < \frac{1}{k}.$$

Thus $\lambda(A \setminus E) = 0$ and thus $A \setminus E$ is null. □

In fact, we proved something much stronger than the theorem statement. The set $E$ is not only Borel, it is specifically a countable union of closed sets; we call such sets $F_\sigma$ sets. Dually, a countable intersection of open sets is called a $G_\delta$ set.

**Exercise 2.41.** *Prove that if $N \subseteq \mathbb{R}^n$ is null, then there is a Borel null set $N'$ such that $N \subseteq N'$. In particular, prove that $N'$ can be chosen to be a $G_\delta$ set.*

**Theorem 2.42.** *Let $E \subseteq \mathbb{R}^n$ be Borel, and let $f : E \to \mathbb{R}^m$ be continuous. If $A$ is Borel in $\mathbb{R}^m$, then $f^{-1}(A)$ is Borel in $\mathbb{R}^n$.*

*Proof.* This proof has to be a little weird again, because we have to use the universal property of Borel sets; we can't actually study the structure of $f^{-1}(A)$ and see that it's Borel—first because we don't know what it "should" look like, and second because we don't know what $A$ looks like.

So we'll define a class of subsets: let

$$\mathcal{M} = \{A : A \subset \mathbb{R}^m, \quad f^{-1}(A) \in \mathcal{B}_n\}.$$

If we can show that $\mathcal{B}_m \subseteq \mathcal{M}$ then we have proven what we want to prove. But $\mathcal{B}_m$ is the smallest $\sigma$-algebra containing all the open sets in $\mathbb{R}^m$; so we want to prove that $\mathcal{M}$ is *a* $\sigma$-algebra containing all the open sets in $\mathbb{R}^m$.

First we claim that $\mathcal{M}$ is a $\sigma$-algebra. We have to check the three axioms:

1. $f^{-1}(\varnothing) = \varnothing \in \mathcal{B}_n$, so $\varnothing \in \mathcal{M}$.

2. Suppose $A_k \in \mathcal{M}$ for all $k \in \mathbb{N}$. Then for each $k$ we know that $f^{-1}(A_j) \in \mathcal{B}_n$. Thus

$$f^{-1}\left(\bigcup_{k=1}^{\infty} A_k\right) = \bigcup_{k=1}^{\infty} f^{-1}(A_k) \in \mathcal{B}_n$$

since $\mathcal{B}_n$ is a $\sigma$-algebra and this is a countable union of $\mathcal{B}_n$ sets. Thus $\bigcup_{k=1}^{\infty} f^{-1}(A_k) \in \mathcal{M}$.

3. Suppose $A \in \mathcal{M}$. Then $f^{-1}(A) \in \mathcal{B}_n$, and thus
$$f^{-1}(A^C) = \{x \in E : f(x) \notin A\} = E \setminus \{x \in E : f(x) \in A\} = E \setminus f^{-1}(A)$$
is a difference of $\mathcal{B}_n$ sets, and thus is in $\mathcal{B}_n$. So $A^C \in \mathcal{M}$.

Thus $\mathcal{M}$ satisfies the three axioms of a $\sigma$-algebra: it contains the null set, and is closed under countable unions and under complements. So $\mathcal{M}$ is a $\sigma$-algebra.

So now we just need to show that $\mathcal{M}$ contains all the open sets. So let $G \subseteq \mathbb{R}^m$ be an open set. Then we can write $f^{-1}(G) = E \cap H$ where $H \subseteq \mathbb{R}^n$ is open. Thus $H \in \mathcal{B}_n$, and we know $E \in \mathcal{B}_n$, so $f^{-1}(G) = H \cap E \in \mathcal{B}_n$. So $G \in \mathcal{M}$.

□

*Remark* 2.43. We know that $\mathcal{M}$ contains all the Borel sets; but it might contain far, far more—and whether it does depends on the specific function. In the extreme case where $f$ is constant, then $\mathcal{M}$ is the largest possible $\sigma$-algebra, containing every possible subset of $\mathbb{R}^m$.

**Corollary 2.44.** *Let $E \subseteq \mathbb{R}^n, F \subseteq \mathbb{R}^m$ be Borel, and let $f : E \to F$ be a homeomorphism. Then $f$ gives a bijection between Borel sets in $E$ and in $F$. That is, If $B \subseteq E$, then $B \in \mathcal{B}_n$ if and only if $f(B) \in \mathcal{B}_m$.*

*Proof.* This follows because $f$ and $f^{-1}$ are both continuous. The previous theorem shows that if $f(B)$ is Borel, then so is $B$; considering the function $f^{-1}$ instead shows that if $B$ is Borel, then so is $(f^{-1})^{-1}(B) = f(B)$. □

We still haven't defined an actual measure, though. Clearly we want to use $\sigma$-algebras to define the class of measurable sets; but what does an actaul measure look like?

**Definition 2.45.** A *measure space* consists of three objects:

- A nonempty set $X$
- A $\sigma$-algebra $\mathcal{M} \subseteq 2^X$
- A function $\mu : \mathcal{M} \to [0, \infty]$ such that $\mu(\varnothing) = 0$, and if $A_1, A_2, \dots$ are disjoint then
$$\mu\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} \mu(A_k).$$

We say the function $\mu$ is a measure.

**Exercise 2.46.** *Prove the following facts about abstract measures:*

1. *If $A, B \in \mathcal{M}$ and $A \subseteq B$, then $\mu(A) \leq \mu(B)$.*
2. *If $A_1, A_2, \dots \in \mathcal{M}$, then*
$$\mu\left(\bigcup_{k=1}^{\infty} A_k\right) \leq \sum_{k=1}^{\infty} \mu(A_k).$$
3. *If $A_1 \subseteq A_2 \subseteq \dots$ are in $\mathcal{M}$ then*
$$\mu\left(\bigcup_{k=1}^{\infty} A_k\right) = \lim_{k\to\infty} \mu(A_k).$$

**Example 2.47.**
- We can take $X = \mathbb{R}^n$, $\mathcal{M} = \mathcal{L}$, and $\mu = \lambda$. This is the Lebesgue measure.
- We can take $X = \mathbb{R}^n$, $\mathcal{M} = \mathcal{B}$ the set of Borel sets, and $\mu = \lambda$. This is the same measure, but allows fewer sets to be measurable. In particular, many sets which are null under the Lebesgue measure are unmeasurable here.
- Take $X$ to be any set, $\mathcal{M} = 2^X$, and $\mu(A) = \infty$ if $A \neq \varnothing$.
- The counting measure: Take $X$ to be a non-empty set, $\mathcal{M} = 2^X$, and
$$\mu(A) = \begin{cases} \#A & A \text{ finite} \\ \infty & A \text{ infinite} \end{cases}$$
- The Dirac measure: let $X$ be any non-empty set and $\mathcal{M} = 2^X$. Fix some $x_0 \in X$ and define $\mu(A) = \chi_A(x_0)$. We usually call this measure the *Dirac measure* and write it $\delta_{x_0}$. It is also sometimes called the Diract delta function, despite not being a function on $X$.

Most of what we'll prove about the Lebesgue measure is actually true in any abstract measure space; in particular, our definition of integral will work for any measure.

**Definition 2.48.** Let $X, \mathcal{M}, \mu$ be a measure space. We can define a new measure space called teh *completion* of $(X, \mathcal{M}, \mu)$. We define a $\sigma$-algebra $\overline{\mathcal{M}}$ by the property that $A \in \overline{\mathcal{M}}$ if and only if there are $B, C \in \mathcal{M}$ with $B \subseteq A \subseteq C$ and $\mu(C \setminus B) = 0$. Clearly $\mathcal{M} \subseteq \overline{\mathcal{M}}$.

Then in this situation we have $\mu(C) = \mu(B)$, so define $\overline{\mu}(A) = \mu(B)$. It's not too hard to show that $\overline{\mathcal{M}}$ is a $\sigma$-algebra and $\overline{\mu}$ is a measure.

**Exercise 2.49.** *Prove that, if $E \subseteq A \in \overline{\mathcal{M}}$ and $\overline{\mu}(A) = 0$, then $E \in \overline{\mathcal{M}}$ and $\overline{\mu}(E) = 0$.*

**Definition 2.50.** We say a measure space $X, \mathcal{M}, \mu$ is *complete* if whenever $E \subseteq A \in \mathcal{M}$ and $\mu(A) = 0$, the $E \in \mathcal{M}$.

We observe that the Lebesgue measure is complete; it is in fact the completion of a measure defined on the Borel sets.

Given a measure space, we can find sub-measure-spaces. Suppose $(X, \mathcal{M}, \mu)$ is a measure space, and $B \in \mathcal{M}$ is a measureable subset of $X$. Then we can define a new measure space $(B, \mathcal{M}_B, \mu_B)$ by taking $\mathcal{M}_B = \{A \cap B : A \in \mathcal{M}\}$, and defining $\mu_B(A) = \mu(A)$.

This just means that $A$ is measurable in $B$ if it's the intersection of a measurable set with $B$, and the measure is inherited from the larger space.

**Example 2.51.** We know that $[0, 1] \in \mathcal{L}$, so we can define a measure space $([0, 1], \mathcal{L}_{[0,1]}, \lambda_{[0,1]})$, where measurable sets are the intersections of Lebesgue measurable sets with the closed interval. This space has total measure one, and does *exactly what you think it should do.*

**Example 2.52.** If our measure space is $\mathcal{R}^2$, then $\mathbb{R}$ is a Lebesgue-measurable subspace of $\mathcal{R}^2$, so we can look at the measure on $\mathbb{R}$ induced by the measure on $\mathcal{R}^2$. But this isn't really a useful measure! In this case, the induced $\sigma$-algebra is exactly the collection of usual Lebesgue measurable subsets of $\mathbb{R}$. But the measure of any set will be 0.

Finally, we are prepared to make some notes on the topic of probability.

**Definition 2.53.** A *probability space* is a measure space $(\Omega, \mathcal{F}, P)$ (where $\Omega$ is a set, $\mathcal{F}$ is a $\sigma$-algebra of subsets of $\Omega$, and $P$ is a measure) such that $P(\Omega) = 1$.

We say that the elements of $\mathcal{F}$, which are subsets of $\Omega$, are *events*, and the *probability* of an event $A \in \mathcal{F}$ is $P(A)$.

**Example 2.54.** The space $[0, 1]$ with the (induced) Lebesgue measure is a probability space. In fact, $[0, 1] \times [0, 1] \times \cdots \times [0, 1]$ is a probability space.

**Example 2.55.** The space $[0, 2]$ with the regular Lebesgue measure is a measure space but *not* a probability space, since $\lambda([0, 2]) = 2 \neq 1$. But if we define $\mu(A) = \frac{1}{2}\lambda(A)$, then $\mu$ is a measure and so $([0, 2], \mathcal{L}_{[0,2]}, \mu)$ is a probability space.

**Definition 2.56.** Suppose $(\Omega, \mathcal{F}, P)$ is a probability space, and $B \in \mathcal{F}$ with $P(B) > 0$. We define the *conditional probability of $A$ given $B$* by

$$P(A|B) = \frac{P(A \cup B)}{P(B)}.$$

**Exercise 2.57.** *Prove that $P(A|B)$ is a probability measure on $B$.*