Measure Theory

Jacob Denson

September 11, 2017

Table Of Contents

In the analysis of the motion of a particle under random, brownian motion, each possible motion of the particle can be represented as a point ω in some space Ω. During the probability that the motion behaves in certain ways, like whether it is positive, can be reduced to measuring the size of an arbitrary subset of Ω. These ideas are also covered by measure theory.

Chapter 1

The Lebesgue Measure

Area is one of the most primitive measurements in geometry. Every elementary school student knows that the area of a circle of radius r is πr^2, and that the area of a rectangle is equal to the product of the lengths of its two distinct side lengths. But given a general shape in the plane, it suddenly becomes very difficult to determine a shape's area. In the work of the ancient Greeks, especially Archimedes, we find two methods of finding more complex areas:

- If a shape can be cut up into finitely many components, and then rearranged into the form of a different shape by a series of rigid motions, then the new shape has the same area as the old.
- We can obtain upper bounds on the area of the shape by enveloping the shape by another shape with an already known area, and lower bounds by finding shapes enclosed by the shape.

But with the advent of Cartesian coordinates and the subsequent modern introduction of set theory, shapes are identified with subsets of $\mathbf{R}^2$, and it is no longer clear how to define the area of a general subset of $\mathbf{R}^2$. More generally, it isn't clear how to define the lengths of subsets of the real line $\mathbf{R}$, or volumes of general subsets of $\mathbf{R}^3$. These are all *measures* of size in their relative dimensions, and so we call the general study of these obects *measure theory*. In line with this unification, we will let $\mu(S)$ denote the length, area, and volume of a given shape S.

You might argue that these paradoxes are irrelevant in terms of any subset that occurs in modern mathematics, but this is not so. For instance, one often wants to measure the length of continuous curves in the plane, which can be considered another measure on space. In 1890, Peano constructed a continuous curve which covers the entire plane. The length of this curve is certainly suspect, because it is constructed as the limit of piecewise differentiable curves whose lengths tend to infinity. Modern developments even brought the ancient methods of area and volume into question. In 1924, Stefan Banach and Alfred Tarski showed that one can decompose a sphere into a finite number of components, which by a number of rigid motions can be rearranged into two copies of the original sphere! We seem to have produced two equal things from one thing – a feat not far from biblical miracle (though we have to use oranges, not loaves and fishes). In duplicating the sphere, Banach and Tarski showed that even the old methods of area geometry do not stand up to the techniques of modern mathematics, and a reanalysis of the entire field was necessary.

One key idea of measure theory is that the old methods of geometry continue to work, provided that we only concentrate on certain 'nice' subsets of space which obey our intuitions about size. Stefan Banach and Alfred Tarski engineered one of the first partitions of space which do not obey geometric intuition – we call these **unmeasurable sets**, because trying to measure their size causes problems. The main theorems of measure theory only work when we work with **measurable sets**. Indeed, we can reinterpret the methods of ancient geometry into two principles of measure theory:

- If a *measurable* set S has a decomposition into disjoint, *measurable* subsets $S_1,\dots,S_n$, then $\mu(S) = \sum \mu(S_i)$, and if S is mapped into a shape T by a rigid motion, then $\mu(S) = \mu(T)$.
- A set S is measurable, then there is a sequence of shapes $S_1^*, S_2^*, \dots$ of measurable sets, each containing S, which each can be decomposed into a collection of intervals/squares/boxes with $\mu(S_i^*) \to \mu(S)$.

We will begin by extending the notion of sets to fairly general subsets of $\mathbf{R}^n$. We do this not only because this is classically how measure theory was introduced, but also because it brings to light the many intricate parts of the theory which we consider when we build measures on more general 'measure spaces'. We recall that we shall use μ uniformly for the length of a subset of $\mathbf{R}$, the area of a subset of $\mathbf{R}^2$, the volume of a subset of $\mathbf{R}^3$, and higher dimensional variants.

1.1 Measuring Elementary Sets

Let's begin by using basic ideas of Euclidean geometry to find a basic class of sets which we can measure the size of without introducing paradoxes. The length of an **interval** with start point a and end point b, either closed, open, or half open, is $b - a$. We know from elementary geometry that the area of a rectangle is the product of the length of the intervals that define it, and we can generalize this to defining the measure of a general box, formed from the product of intervals. That is, the measure of $I_1 \times \dots \times I_n$, where I_i is an interval starting at a_i and ending at b_i is

$$(b_1 - a_1) \dots (b_n - a_n) = \mu(I_1) \dots \mu(I_n)$$

In general, the easiest sets to measure the area of are those covered by boxes, and we will show this leads to a system of area with a well defined theory.

Lemma 1.1. *If we decompose a box R into the disjoint union of finitely many boxes R_i, then the measure of the box is the sum of the measures of the boxes in the decomposition, in the sense that $\mu(R) = \sum \mu(R_i)$*

Proof. We proceed by a grid decomposition. Suppose first that the rectangular decomposition forms a grid, in the sense that we can index the decomposition as $R_{i_1 \dots i_n}$, where $R_{i_1 \dots i_n} = I^1_{i_1} \times \dots \times I^n_{i_n}$, and the endpoint of $I^k_{i_n}$ is the startpoint of $I^k_{i_n+1}$. Then

$$\begin{aligned} \sum_{i_1,\dots,i_n} \mu(R_{i_1 \dots i_n}) &= \sum_{i_1,\dots,i_n} \mu(I^1_{i_1}) \dots \mu(I^n_{i_n}) \\ &= \prod_{k=1}^{n} \sum_j \mu(I^k_j) \end{aligned}$$

and the theorem is implied in this case by showing that $\sum_j \mu(I^k_j) = \mu(I_k)$, where $R = I_1 \times \dots \times I_n$. But this follows because the sum $\sum_j \mu(I^k_j)$ is a telescoping sum, with the highest indexed interval's endpoint equal to the endpoint of I_k, and the lower indices startpoint equal to the startpoint of

I_k. In general, it suffices to break a general decomposition into a further decomposition forming a grid, in such a way that the sum of the boxes in the first decomposition is equal to the sum of the boxes in the second. This is proven by forming the grid in each dimension, applying another telescoping sum type argument along each dimension. $\square$

A similar grid decomposition like argument proves the following.

Lemma 1.2. *If a family of boxes $R_1, \dots, R_n$ covers R, then $\mu(R) \leqslant \sum \mu(R_i)$.*

Lemma 1.3. *If $R_1, \dots, R_n$ and $S_1, \dots, S_m$ are two disjoint families of boxess with $\bigcup R_i = \bigcup S_i$, then $\sum \mu(R_i) = \sum \mu(S_i)$.*

Alternatively, these theorems can be shown using a discretization argument. We begin by showing that

$$\mu(I) = \lim_{N \to \infty} \frac{|\mathbf{Z}/N \cap I|}{N}$$

From this, it is easy to argue that for any box R,

$$\mu(R) = \lim_{N \to \infty} \frac{|\mathbf{Z}^n/N \cap R|}{N^n}$$

But now if we write $R = \cup R_i$ as the union of disjoint intervals, then

$$\mu(R) = \lim_{N \to \infty} \frac{|\mathbf{Z}^n/N \cap R|}{N^n} = \sum_i \lim_{N \to \infty} \frac{|\mathbf{Z}^n/N \cap R_i|}{N^n} = \sum_i \mu(R_i)$$

and this proves the theorem. One might be tempted to define the measure of an arbitrary subset of $\mathbf{R}^n$ by the formula

$$\mu(E) = \lim_{N \to \infty} \frac{|\mathbf{Z}^n/N \cap E|}{N^n}$$

however, this definition runs into problems. One can find sets where this limit doesn't exist, and even if the limit does exist, we might not even have translation invariance. For instance, with respect to this function $\mathbf{Q} \cap [0,1]$ has length 1, but $\mathbf{Q} + \sqrt{2}$ has length 0. The definition is valid for all *Jordan measurable sets*. A more suitable way to obtain a continuous measure from some kind of discrete measure is by the theory of Monte Carlo integration, which we won't cover here.

The above lemmas guarantee that if $E \subset \mathbf{R}^n$ is the disjoint union of boxes $R_1, \dots, R_n$, then the definitions

$$\mu(E) = \sum \mu(R_i)$$

is well defined. We call a set like E an **elementary set**. We shall find that the *algebraic structure* of the family of sets a measure is defined over is interesting. One often has to consider the areas of the set formed from the union of two sets, or the intersection, and it is useful to know that we can measure the size of a set if it is the union of measurable sets, or the intersection of measurable sets. The next lemma is very useful in that regard

Lemma 1.4. *If R is the finite union of boxes, then R is the finite union of disjoint boxes.*

Proof. One can only prove this by a grid decomposition argument. □

Since the union of two sets which are finite unions of boxes is also a finite union of boxes, we conclude that the union of two elementary sets is also elementary.

Lemma 1.5. *The intersection of two box is a box.*

Proof. Note first that if I is an interval with start point a and endpoint b, and J is an interval with start point c and endpoint d, then $I \cap J$ is either empty, or an interval with startpoint $\max(a, c)$, and endpoint $\min(b, d)$. But then give $R = I_1 \times \dots \times I_n$ and $S = J_1 \times \dots \times J_n$, then

$$R \cap S = (I_1 \cap J_1) \times \dots \times (I_n \cap J_n)$$

which is a box. □

If $E = \bigcup R_i$ and $F = \bigcup S_i$ are elementary sets, then $E \cap F = \bigcup (R_i \cap S_j)$ is also an elementary set. Unfortunately, this family is not closed under the complement operation, because we do not allow unbounded intervals, but it is 'almost' closed under the complement operation.

Lemma 1.6. *If R and S are boxes, then $R - S = R \cap S^c$ is a finite union of rectangles.*

Corollary 1.7. *If E and F are elementary sets, then $E - F$ is an elementary set.*

Proof.
$$(R_1 \cup \cdots \cup R_n) \cap (S_1 \cup \cdots \cup S_m)^c = \cup(R_i \cap S_j)$$
□

A family of sets is an **algebra** if it is closed under the union and complement operation. It is easy to see that an algebra is closed under intersections because $E \cap F = (E^c \cup F^c)^c$, as well as the set subtraction operation. What we have just proven is that the set of elementary subsets contained in $[0,1]^n$ is an algebra. The fact that elementary sets are not an algebra is one of the reasons we have to enlarge the family of sets we measure the size of.

Let's look at some elementary properties of the behaviour of μ on elementary sets:

- It is also easy to see that if $E_1, \dots, E_n$ are disjoint elementary sets, then $\mu(E_1 \cup \cdots \cup E_n) = \sum \mu(E_i)$, so μ is *finitely additive* on elementary subsets.

- For any elementary set E, $\mu(E+x) = \mu(E)$, so μ is *translation invariant*.

- It follows that if E and F are elementary sets, with $E \subset F$, then $\mu(E) \leqslant \mu(F)$, because

  $$\mu(F) = \mu((F-E) \cup E) = \mu(F-E) + \mu(E) \geqslant \mu(E)$$

  thus μ is a *monotone* function on sets.

These properties uniquely define the function μ we constructed up to a scalar factor. Since all the properties are intuitive to us, this tells us we're going in the right direction!

1.2 Jordan Measurable Sets

In the last section, we constructed a consistant measure μ on the family of elementary sets. However, this family is certainly limited. We cannot even use this quantity to measure the area of a circle, or the volume of a sphere. However, we have really only applied the first ancient technique of measuring area, forming disjoint unions of simple shapes. We haven't used the method of approximating shapes by simple sets from above and below. If E is an arbitrary set, and there is a constant $C \geqslant 0$ such that for all ε, there are $E_1 \subset E$ and $E \subset E_2$ with

$$C - \varepsilon \leqslant \mu(E_1) \leqslant \mu(E_2) \leqslant C + \varepsilon$$

then it would be reasonable to define the area of E to be C. We call such a set **Jordan measurable**. More specifically, for any subset of $\mathbf{R}^n$, we define

$$\mu_*(E) = \sup\{\mu(E_1) : E_1 \text{ elementary}, E_1 \subset E\}$$

$$\mu^*(E) = \inf\{\mu(E_1) : E_1 \text{ elementary}, E_2 \supset E\}$$

the **inner** and **outer** measures of the set E. We say a *bounded* set E is Jordan measurable if $\mu_*(E) = \mu^*(E)$.

Theorem 1.8. *A set is Jordan measurable if and only if for every $\varepsilon > 0$, there is an elementary set A with $\mu^*(A \triangle E) \leqslant \varepsilon$.*

Proof. Consider a set E satisfying the second condition. Then there is an elementary set F with $A \triangle E \subset F$ and $\mu(F) \leqslant 2\varepsilon$, and so

$$\mu^*(E) \leqslant \mu(A \cup F) \leqslant \mu(A) + \mu(F) \leqslant \mu(A) + 2\varepsilon$$

since $\mu_*(A \triangle E) \leqslant \mu^*(A \triangle E) \leqslant \varepsilon$, we can find $F \subset A \triangle E$ with $\mu(F) \leqslant \varepsilon$, and

$$\mu_*(E) \geqslant \mu_*(A - F) \geqslant \mu_*(A) - \mu_*(F) \geqslant \mu_*(A) - \varepsilon$$

Then we let $\varepsilon \to 0$. Conversely, if E is Jordan measurable, there are elementary sets F_1, F_2 with $F_1 \subset E \subset F_2$ and $\mu(F_2) - \mu(F_1) = \mu(F_2 - F_1) < \varepsilon$. Then

$$\mu^*(F_2 \triangle E) = \mu^*(F_2 - E) \leqslant \mu^*(F_2 - F_1) \leqslant \varepsilon$$

This shows the condition holds, and also that we can choose the set A to be a superset of E in the proof above. □

Jordan measurable sets satisfy the same algebraic operations are the family of elementary sets.

- If $\mu^*(A - E), \mu^*(B - F) < \varepsilon$, then

  $$\mu^*((A \cap B) - (E \cap F)) \leqslant \mu^*(A - E) + \mu^*(B - F) \leqslant 2\varepsilon$$

  hence $E \cap F$ is measurable if E and F are measurable.

- If $\mu^*(A - E), \mu^*(B - F) < \varepsilon$, then since

$$(A \triangle B) \triangle (E \triangle F) = (A \triangle E) \triangle (B \triangle F) \subset (A \triangle E) \cup (B \triangle F)$$

  Hence $\mu((A \triangle B) \triangle (E \triangle F)) \leqslant \varepsilon$, and the symmetric difference of two sets is measurable.

- Since $E - F = (E \triangle F) \cap E$, the set theoretic minus of two Jordan measurable sets is Jordan measurable.

- To prove $E \cup F$ is Jordan measurable, we can assume E and F are disjoint, because $(E \cup F) = (E \cap F) \cup (E - F) \cup (F - E)$. Note that in this case $\mu_*(E) + \mu_*(F) \leqslant \mu_*(E + F)$, because an interior estimate of E and an interior estimate of F combine as disjoint sets to give an interior estimate of $\mu_*(E + F)$. Now since E and F are measurable, for any $\varepsilon > 0$, we can find elementary sets $E \subset E^*$ and $F \subset F^*$ such that $\mu(E^*) \leqslant \mu_*(E) + \varepsilon$ and $\mu(F^*) \leqslant \mu_*(F) + \varepsilon$. But then

$$\mu(E^* \cup F^*) \leqslant \mu(E^*) + \mu(F^*) \leqslant \mu_*(E) + \mu_*(F) + 2\varepsilon \leqslant \mu_*(E \cup F) + 2\varepsilon$$

  This shows $E \cup F$ is Jordan measurable.

hence, restricted to a particular bounded Jordan measurable set of space, the class of Jordan measurable sets is an algebra. Since $\mu^*(E \cup F) \leqslant \mu^*(E) + \mu^*(F)$ and $\mu_*(E \cup F) \geqslant \mu_*(E) + \mu_*(F)$ holds for all disjoint sets E and F, we conclude that Jordan measurable sets have finite additivity. This implies monotonicity. The translation invariance of the measure follows from the translation invariance of the measure on elementary sets.

Example. *Let R be a closed box in $\mathbf{R}^n$, and $f : R \to \mathbf{R}$ a continuous function. Then the graph $\Gamma(f)$ of f in $\mathbf{R}^{n+1}$ is Jordan measurable, and has Jordan measure zero. This follows because, since R is compact, f is uniformly continuous, and therefore for any ε there are finitely many disjoint boxes $S_1, \dots, S_m$ covering R such that if two points x and y lie in the same box, $|f(x) - f(y)| < \varepsilon$. But this implies that if we fix a point x_i in S_1, then the sets $S_i \times [x_i - \varepsilon, x_i + \varepsilon]$ cover the graph, and so*

$$\mu^*(\Gamma(f)) \leqslant \sum \mu(S_i \times [x_i - \varepsilon, x_i + \varepsilon]) \leqslant 2\varepsilon \sum \mu(S_i) \leqslant 2\varepsilon\mu(R)$$

let $\varepsilon \to 0$ to obtain that $\Gamma(f)$ has upper measure zero, and thus lower measure zero. The set $X = \{(x, t) : 0 \leqslant f(x) \leqslant t\}$ is also Jordan measurable. Given the *same S_i, the sets $S_i \times [0, x_i - \varepsilon)$ are contained in X, and $S_i \times [0, x_i + \varepsilon]$ contain X, and the difference between these two sets is exactly the sets we used to show the measure of the boundary is zero, hence X is Jordan measurable because it can be approximated from above and below.*

Example. *A triangle is not an elementary set, but it is a Jordan measurable set. First, consider a right triangle with sides parallel to the x and y axis. Then, by a translation, we can write one point as $(x, 0)$ and another as $(0, y)$. For each N, consider the disjoint sequence of rectangles*

$$[0, x/N) \times \{0\} \cup [x/N, 2x/N) \times [0, y/N] \cup \cdots \cup [\frac{N-1}{N}x, x] \times [0, \frac{N-1}{N}y]$$

These rectangles are contained in the triangle, and so

$$\mu_*(T) \geqslant \sum_{i=0}^{N-1} \frac{x}{N}\frac{iy}{N} = \frac{xy}{N^2}\frac{(N-1)N}{2} = \frac{N-1}{N}\frac{xy}{2}$$

Letting $N \to \infty$, we find $\mu_*(T) \geqslant xy/2$. On the other hand, consider the disjoint sequence of rectangles

$$[0, x/N) \times [0, y/N] \cup [x/N, 2x/N) \times [0, 2y/N] \cup \cdots \cup [\frac{N-1}{N}x, x] \times [0, y]$$

which contains the triangle, so

$$\mu^*(T) \leqslant \sum_{i=1}^{N} \frac{x}{N}\frac{iy}{N} = \frac{N+1}{N}\frac{xy}{2}$$

Letting $N \to \infty$, we conclude $\mu^*(T) \geqslant xy/2$. Equating estimates, we determine the triangle is Jordan measurable with area $xy/2$. If only one of the sides is horizontal, we may split the triangle into two right triangles with the other side perpendicular to the y axis, so this shape is measurable. If one coordinate is $(x, 0)$, and the other

DO MORE EXERCISES FROM TAO'S BOOK

The Jordan measure is intrinsically connected with the Riemann integral. Given an interval,

TALK ABOUT RIEMANN INTEGRAL, SHOW SET IS JORDAN MEASURABLE IFF ITS BOUNDARY HAS MEASURE ZERO?

1.3 Lebesgue Measure

If we are able to stick with Jordan measurable sets, we should, because it is here that integration theory works in the best way. However, not all sets are Jordan measurable, and often when studying fractal sets one runs into unmeasurable sets. What's more, even if a set $E_1, E_2, \dots$ is Jordan measurable, their union $\bigcup E_i$ and their intersection $\bigcap E_i$ need not be measurable, even if these sets are bounded. In terms of Riemann integrability, this causes problems with understanding the pointwise limit of functions: A sequence of uniformly bounded Riemann integrable functions $f_n : [0,1] \to \mathbf{R}$ which converges pointwise to a bounded function $f : [0,1] \to \mathbf{R}$ need not be Riemann integrable. If we replace pointwise convergence with uniform convergence, then f will be Riemann integrable, but this relates to the fact that uniform convergence allows one to conver f with finitely many rectangles (ELABORATE HERE).

To obtain a family of sets with a well defined measure theory which satisfies countable additivity, we must tinker with how we defined Jordan measure. Recall that to obtain Jordan measure, we took outer and inner estimates of arbitrary sets by covers of *finitely many* rectangles. We obtain Lebesgue integrability if we replace the finiteness with infinitely many rectangles. We consider the values

$$\mu^*(E) = \inf\{\sum\}$$

Definition. If A is a set of real numbers, then it's Lebesgue measure is

$$m(A) = \inf \left\{ \sum_{k=1}^\infty m(I_k) : \bigcup_{k=1}^\infty I_k \supset A \right\}$$

The end goal of this passage is to find out what it takes to prove that if $\{A_i\}$ is a disjoint collection of sets, then $m(\bigcup A_i) = \sum m(A_i)$. This will give us intuition in the abstract case. In this case, one side of the equality is fairly easy to show.

Theorem 1.9. *If $\{A_i\}$ is a countable collection of sets, then $m(\bigcup A_i) \leqslant \sum m(A_i)$.*

Proof. If any A_i has infinite length, then the theorem is trivial. Thus assume all A_i have finite measure. Fix some $\varepsilon > 0$. For each A_i, pick a countable set $\mathcal{I}_i$ of open intervals such that $\sum_{I \in \mathcal{I}_i} I \leqslant m(A_i) + \varepsilon/2^k$. Then $\bigcup \mathcal{I}_i$ is a countable collection of open intervals covering $\bigcup A_i$, and so

$$m(\bigcup A_i) \leqslant \sum_{i=1}^{\infty} \sum_{I \in \mathcal{I}} m(I) \leqslant \sum_{i=1}^{\infty} [m(A_i) + \varepsilon/2^k] = \sum_{i=1}^{\infty} m(A_i) + \varepsilon$$

The proof is completed since ε was arbitrary. □

Lemma 1.10. *If $A \subset B$, $m(A) \leqslant m(B)$.*

Proof. Any cover of B is a cover of A. □

Let us check the m is well defined, when passing from lengths of intervals to approximations of arbitrary sets.

Lemma 1.11. *For any interval $I = (a, b)$, $m(I) = b - a$*

Proof. First, we will verify that $m([a, b]) = b - a$. Let $\mathcal{I}$ be a collection of open intervals such that $\bigcup \mathcal{I} \supset [a, b]$. Without loss of generality, we may choose a finite subcover, since $[a, b]$ is compact. Using this finiteness, construct a sequence $(a_1, b_1), \ldots, (a_n, b_n)$ from $\mathcal{I}$ such that $b_i \geqslant a_{i+1}$ for each i, $a_1 \leqslant a$, and $b_n \geqslant b$. Then

$$\begin{aligned}
\sum_{I \in \mathcal{I}} m(I) \geqslant \sum_{i=1}^{n} m((a_i, b_i)) &= \sum_{i=1}^{n} b_i - a_i \\
&\geqslant (b_n - a_n) + \sum_{i=1}^{n-1} (a_{i+1} - a_i) \\
&= (b_n - a_n) + (a_n - a_1) = b_n - a_1 \geqslant b - a
\end{aligned}$$

Thus $m([a, b]) \geqslant b - a$. Now, fix $\varepsilon > 0$. Choose the cover

$$\mathcal{I} = \{(a - \varepsilon, a + \varepsilon), (a, b), (b - \varepsilon, b + \varepsilon)\}$$

Now $\bigcup \mathcal{I} = (a - \varepsilon, b + \varepsilon) \supset [a, b]$, so

$$m([a, b]) \leqslant m((a, b)) + m((a - \varepsilon, a + \varepsilon)) + m((b - \varepsilon, b + \varepsilon)) = b - a + 4\varepsilon$$

Since ε was arbitrary, $m([a, b]) \leqslant b - a$, and so $m([a, b]) = b - a$.

Surely, $m((a,b)) \leqslant m([a,b]) = b-a$. But also, by Lemma (1.1),

$$m([a,b]) \leqslant m((a,b)) + m(\{a\}) + m(\{b\}) = m((a,b))$$

since the length of a single point is zero. $\square$

Now we want to know that measuring the union is the same as measuring the component parts, as our intuition would tell us. However, Banach and Tarski have warned us that this won't be true of all sets. One side of the inequality can be shown for all sets, but we must specialize to obtain equality – defining exactly what it means for a set to be measurable, as we were discussing above.

Definition. A set A is **measurable** (in the manner of Lebesgue), if for any other set B, $m(B) = m(A \cap B) + m(A^c \cap B)$.

It is simple to verify that $\mathbf{R}$ is a measurable set, and if A is measurable, then so is A^c. More complicated is the fact that open intervals are measurable.

Lemma 1.12. *For any real number a, (a,∞) is measurable.*

Proof. Let A be an arbitrary set. Let $\mathcal{I}$ be a countable collection of intervals such that $\sum_{I\in\mathcal{I}} m(I) \leqslant m(A)+\varepsilon$. Then, for each I, either $I\cap(a,\infty)$ is empty or an interval, as is $I\cap(-\infty,a]$, and the measure of I is equal to the measure of the sum. Thus

$$m(A\cap(a,\infty))+m(A\cap(-\infty,a]) \leqslant \sum m(I_k\cap(a,\infty))+\sum m(I_k\cap(-\infty,a]) = \sum m(I_k) \leqslant m(A)+\varepsilon$$

So $m(A\cap(a,\infty)) + m(A\cap(-\infty,a]) \leqslant m(A)$, and we have already proved the inequality the other way. $\square$

Lemma 1.13. *If A and B are measurable, then so is $A\cup B$.*

Proof. Let S be an arbitrary subset of the reals. Then

$$\begin{aligned}
m(S) &= m(S\cap A) + m(S\cap A^c)\\
&= m(S\cap A) + m(S\cap A^c\cap B) + m(S\cap A^c\cap B^c)\\
&= m(S\cap A) + m(S\cap B\cap A^c) + m(S\cap[A\cup B]^c)\\
&= m([S\cap A]\cup[S\cap B\cap A^c]) + m(S\cap[A\cup B]^c)\\
&= m(S\cap[A\cup B]) + m(S\cap(A\cup B)^c)
\end{aligned}$$

One may get from the second last equation to the third last equation by applying the measurability of A to the first measured set. □

Corollary 1.14. *If A and B are measurable, then $A \cap B$ and $A - B$ are measurable.*

Proof. $A \cap B = (A^c \cup B^c)^c$, and $A - B = A \cap B^c$. □

Corollary 1.15. *All open intervals are measurable.*

Corollary 1.16. *If A is any set, and $E_1, \dots, E_n$ is a finite collection of disjoint measurable sets, then*

$$m(A \cap (\bigcup_{k=1}^{n} E_k)) = \sum_{k=1}^{n} m(A \cap E_k)$$

What we have shown here is that the set of measurable sets is a Boolean algebra. We can go one further.

Lemma 1.17. *If $E_1, E_2, \dots$ is a countable collection of disjoint measurable sets, then $E = E_1 \cup E_2 \cup \dots$ is measurable.*

Proof. Define $F_n = \bigcup_{k=1}^{n} E_n$. Then F_n is measurable, and $F_n^c \supset E^c$. Hence

$$\begin{aligned} m(A) &= m(A \cap F_n) + m(A \cap F_n^c) \geqslant m(A \cap F_n) + m(A \cap E^c) \\ &= \sum_{k=1}^{n} m(A \cap E_k) + m(A \cap E^c) \end{aligned}$$

Since n was arbitrary,

$$m(A) \geqslant \sum_{k=1}^{\infty} m(A \cap E_k) + m(A \cap E^c) \geqslant m(A \cap E) + m(A \cap E^c)$$

But $m(A) \leqslant m(A \cap E) + m(A \cap E^k)$, so E is measurable. □

Corollary 1.18. *The countable union of measurable sets are measurable.*

Proof. Simply modify measurable sets by elementary set operations so they are disjoint, and then take their union. □

From this theorem, we can determine that every open set in **R** is open, as every open set is the countable union of open intervals.

We can know show what we originally set out to solve.

Theorem 1.19. *If $\{A_i\}$ is a countable collection of pairwise disjoint measurable sets, then*

$$m(\bigcup A_i) = \sum m(A_i)$$

Proof. The calculations in the above theorem show that, letting $A = \bigcup A_i$,

$$m(A) \geqslant \sum_{k=1}^{\infty} m(A \cap A_i) + m(A \cap A^c) = \sum_{k=1}^{\infty} m(A_i)$$

but this shows equality, since the other direction of inequality always holds. □

The function m, restricted to measurable sets, will now be known as the Lebesgue measure on **R**. It is the first in a line of a general class of functions known as measures, defined on subsets of a space and measuring these subset's size. The notions of measurable set will be abstracted to the properties proved above. That is, we can measure the union, intersection, and complement of all measurable sets. Most theorems in measure theory are actually be proved in general just as easily as on the Lebesgue measure we have just described.

1.4 Appendix: Banach Tarski

Let us consider the sphere. A nice property of this object is that it is invariant under any rotation - that is, if you take a point, and rotate it around the origin, you will never end up at a point off the unit sphere. Mathematically, we say that the orthogonal group $O(3)$ acts on the sphere S^1.

The core technique of this proof can be executed in a simpler form on free groups. Consider the free group $F_{\{a,b\}}$ on two characters. Let $S(a)$ be the set of all sequences whose simplest form begins with a, and define $S(b)$, $S(a^{-1})$, and $S(b^{-1})$ similarily. We have the following equalities:

$$F_{\{a,b\}} = S(a) \cup S(b) \cup S(a^{-1}) \cup S(b^{-1})$$

$$F_{\{a,b\}} = S(a) \cup aS(a^{-1})$$

$$F_{\{a,b\}} = S(b) \cup bS(b^{-1})$$

Thus we have partitions $F_{\{a,b\}}$ into four sets. By 'rotating' two of these partitions, we obtain two copies of the group.

The trick to the Banach-Tarski paradox on the sphere is to find subsets of the orthogonal group that behave like $F_{\{a,b\}}$. We will say a subset X of euclidean space can be **paradoxically decomposed**, if it can be expressed as the disjoint union of subsets, $X_1 \cup \cdots \cup X_n$, and, under group actions $g_1, \ldots, g_n \in O(3)$, we may express $X = g_1 X_1 \cup \cdots \cup g_i X_i$, and $X = g_{i+1} X_{i+1} \cup \cdots \cup g_n X_n$, for some i.

Lemma 1.20. *There is a subgroup of $O(3)$ isomorphic to $F_{\{a,b\}}$.*

Proof. Map a to a rotation horizontally by $\sqrt{2}\pi$ radians, and map b to a rotation vertically by $\sqrt{2}\pi$ radians. This induces a homomorphism from $F_{\{a,b\}}$ to $O(3)$. We claim this homomorphism is injective. □

Theorem 1.21 (Banach Tarski)**.** *The sphere may be paradoxically decomposed.*

Chapter 2

Abstract Measures

Let us now begin to describe measures in their abstract generality.

Definition. Let X be an arbitrary set. A σ-algebra on X is a family of subsets of X, called measurable sets, such that

(1) X is measurable.

(2) If A is measurable, A^c is measurable.

(3) If $\mathcal{C}$ is a countable family of measurable sets, then $\bigcup \mathcal{C}$ is measurable.

What we have shown is that there is a set function defined on a σ algebra of $\mathbf{R}$ containing all borel sets, and agreeing with the notion of length on all open intervals.

Chapter 3

Extending measures

To construct the Lebesgue measure, we began with an intuitive notion of area over the set of intervals on the real line, and then extended the notion of length to a bigger class of sets, the measurable sets. This chapter attempts to purify this strategy to work on arbitrary measure spaces.

Definition. Let Ω be a set. $\mathcal{A} \subset \mathcal{P}(X)$ is a **semi-algebra** over X if

1. $\varnothing, \Omega \in \mathcal{A}$.
2. If $A, B \in \mathcal{A}$, then $A \cap B \in \mathcal{A}$.
3. If $A, B \in \mathcal{A}$, then there exists a finite, disjoint collection $\mathcal{C} \subset \mathcal{A}$ such that $B - A = \bigcup \mathcal{C}$.

Definition. Fix some semi-algebra $\mathcal{A}$. A **pre-measure** is a function $\mu : \mathcal{A} \to \mathbf{R}^+$, such that

1. If $\mathcal{B} \subset \mathcal{A}$ is a finite, disjoint collection of sets, then
$$\mu\left(\bigcup \mathcal{B}\right) \geqslant \sum_{B \in \mathcal{B}} \mu(B)$$

2. If $\mathcal{B} \subset \mathcal{A}$, is at most countable, and if $A \in \mathcal{A}$ satisfies $A \subset \bigcup \mathcal{B}$, then
$$\mu(A) \leqslant \sum_{B \in \mathcal{B}} \mu(B)$$

The definition of a pre-measure can be phrased in different ways. The proofs of the following are just an exercise in notation.

Lemma 3.1. *The following are alternate ways to specifc a pre-measure. Fix some semi-algebra $\mathcal{A}$, and set function $\mu : \mathcal{A} \to \mathbf{R}^+$:*

1. *If μ is monotone and countably subadditive, where defined, then μ is a pre-measure.*

2. *If μ is countable additive, where defined, then μ is a pre-measure.*

It turns out that all pre-measures are essentially measures, thanks to a theorem by Caratheodory.

Theorem 3.2 (Caratheodory's Extension Theorem)**.** *Every pre-measure can be extended to a measure.*

Proof. Let $\mu : \mathcal{A} \to \mathbf{R}^+$ be a pre-measure, where $\mathcal{A}$ is a semi-algebra on a set Ω. We shall define a measure $\mu' : \mathcal{A}' \to \mathbf{R}^+$ such that $\mu'|_{\mathcal{A}} = \mu$. Let $\mu^* : \mathcal{P}(\Omega) \to \mathbf{R}^+$ be defined by
$$\mu^*(X) = \sup\left\{ \sum_{B \in \mathcal{B}} \mu(B) : \mathcal{B} \subset \mathcal{A} \text{ is at most countable, } A \subset \bigcup \mathcal{B} \right\}$$
We will begin by studying some properties of μ^*.

1. $\mu^*|_{\mathcal{A}} = \mu$.

   Proof. *Consider any $A \in \mathcal{A}$. If $A \subset \bigcup \mathcal{B}$, then property (2) of being a pre-measure shows that $\mu(A) \leqslant \sum_{B\in\mathcal{B}} \mu(B)$. Thus $\mu^*(A) \geqslant \mu(A)$. But by setting $\mathcal{B} = \{A\}$, we see that $\mu^*(A) \leqslant \mu(A)$.*

2. μ^* is monotone.

   Proof. *Let $A \subset B$ be a pair of arbitrary sets. Then if $\mathcal{B}$ is such that $\bigcup \mathcal{B} \supset B$, then $\bigcup \mathcal{B} \supset A$. To obtain $\mu^*(B)$, we are thus taking the supremum of a smaller set than when we obtain $\mu^*(A)$. Thus $\mu^*(A) \leqslant \mu^*(B)$.*

3. μ^* is countably subadditive.

   Proof. *Consider a countable subset $\mathcal{B} \subset \mathcal{P}(\Omega)$, fix some ordering $\{B_i\}$ on $\mathcal{B}$, and consider any $\varepsilon > 0$. For each $B_n \in \mathcal{B}$, we can find a countable set $\mathcal{C}(B) \subset \mathcal{A}$ such $B \subset \bigcup \mathcal{C}(B)$, and $\sum_{C\in\mathcal{C}(B)} \mu(C) \leqslant \mu^*(B) + \varepsilon/2^n$. But then $\bigcup \mathcal{B} \subset \bigcup\{C \in \mathcal{C}(B) : B \in \mathcal{B}\}$, which is a countable subcollection of $\mathcal{A}$, so*

   $$\mu^*(\bigcup \mathcal{B}) \leqslant \sum_{B\in\mathcal{B}} \sum_{C\in\mathcal{C}(B)} \mu(C) \leqslant \sum_{n=1}^{\infty} \mu^*(B_n) + \frac{\varepsilon}{2^n} \leqslant \sum_{B\in\mathcal{B}} \mu^*(B) + \varepsilon$$

   Letting $\varepsilon \to 0$, we obtain the inequality.

Now let $\mathcal{A}' = \{A \in \mathcal{P}(\Omega) : \forall B \subset \Omega, \mu^*(B) = \mu^*(A \cap B) + \mu^*(A^c \cap B)\}$, and let $\mu' = \mu^*|_{\mathcal{A}'}$. We shall verify that this is the measure we are needing to construct. First, notice that is is obvious that $\mathcal{A}'$ is closed under complement.

4. If $A, B \in \mathcal{A}$, then $A \cap B$, $A \cup B \in \mathcal{A}'$.

   Proof. *For any $E \subset \Omega$,*

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

   By (2), we obtain equality, so $A \cap B \in \mathcal{A}'$.

5. μ' satisfies the countable additivity property of a measure.

   Proof. *Let $\mathcal{B} \subset \mathcal{A}'$ be an at most countable collection of disjoint sets. If $\mathcal{B} = \{A, B\}$, then*

   $$\mu'(A \cup B) = \mu'(A \cap (A \cup B)) + \mu'(A^c \cap (A \cup B)) = \mu^*(A) + \mu^*(B)$$

   By induction, the lemma holds if $\mathcal{B}$ is finite. Fix an enumeration $\{B_i\}$ to $\mathcal{B}$. Then we have that, for any $n \in \mathcal{B}$,

   $$\sum_{k=1}^{n} \mu'(B_k) = \mu\left(\bigcup_{i=1}^{n} B_i\right)$$

   Letting $n \to \infty$, we see that $\sum_{k=1}^{\infty} \mu'(B_k) \leqslant \mu(\bigcup \mathcal{B})$. But then by part (3) of this proof, we obtain equality.

Because of (5), $\mathcal{A}'$ is a boolean algebra.

6. $\mathcal{A}'$ is closed under countable union [This shows $\mathcal{A}'$ is a σ-algebra, and completes the proof that μ' is a measure].

   Proof. *Let $\mathcal{B} \subset \mathcal{A}'$ be a countable collection. Without loss of generality, we may assume this collection is disjoint, by modifying $\mathcal{B}$ using the operations of intersection and complement. Let $C \subset \Omega$ be arbitrary. Fix some enumeration $\{B_i\}$ of $\mathcal{B}$, and let $\mathbf{B}_m = \bigcup_{i=1}^{n} B_i$. Then*

   $$\begin{aligned}
   \mu^*(C) &= \mu^*(C \cap \mathbf{B}_m) + \mu^*(C \cap \mathbf{B}_m^c) \\
   &= \sum_{k=1}^{m} \mu^*(C \cap B_k) + \mu^*\left(C \cap \bigcup_{k=1}^{m} B^k\right) \\
   &\geqslant \sum_{k=1}^{\infty} \mu^*(C \cap B_k) + \mu^*\left(C \cap \left(\bigcup \mathcal{B}\right)^c\right) \\
   &\geqslant \mu^*\left(C \cap \bigcup \mathcal{B}\right) + \mu^*\left(C \cap \left(\bigcup \mathcal{B}\right)^c\right)
   \end{aligned}$$

   By (3), $\bigcup \mathcal{B} \in \mathcal{A}'$.

7. $\mathcal{A} \subset \mathcal{A}'$.

Proof. *Let $A \in \mathcal{A}$, and let $D \subset \Omega$. Then, since $\mathcal{A}$ is a semi-algebra, we can write $A^c = \bigcup \mathcal{B}$, where $\mathcal{B} \subset \mathcal{A}$ is finite. Fix $\varepsilon > 0$. We can find $\mathcal{C} \subset \mathcal{A}$ with $D \subset \bigcup \mathcal{C}$, and $\sum_{C \in \mathcal{C}} \mu(C) \leqslant \mu(D) + \varepsilon$. Then*

$$\begin{aligned}
&\mu^*(A \cap D) + \mu^*(A^c \cap D) \\
&\leqslant \mu^*\left(A \cap \left(\bigcup \mathcal{C}\right)\right) + \mu^*\left(\left(\bigcup \mathcal{B}\right) \cap \left(\bigcup \mathcal{C}\right)\right) && \textit{(monotonicity)} \\
&= \mu\left(\bigcup_{C \in \mathcal{C}} A \cap C\right) + \mu\left(\bigcup_{B \in \mathcal{B}, C \in \mathcal{C}} B \cap C\right) \\
&\leqslant \sum_{B \in \mathcal{B}, C \in \mathcal{C}} \mu(A \cap C) + \mu(B \cap C) && \textit{(Property 2)} \\
&\leqslant \sum_{B \in \mathcal{B}} \mu(C) \leqslant \mu(A) + \varepsilon && \textit{(Property 2 of a pre-measure)}
\end{aligned}$$

Letting $\varepsilon \to 0$, we obtain an inequality, which together with (2) gives us an equality.

This concludes the proof. □

Corollary 3.3. *If μ is σ-finite, then the extension is unique on $\mathcal{A}'$.*

Proof. Let $\nu : \mathcal{A}' \to \mathbf{R}$ be another extension of μ, other than μ' constructed above. Let $A \in \mathcal{A}'$. Fix $\varepsilon > 0$. For any countable covering $\mathcal{B} \subset \mathcal{A}$ of A,

$$\nu(A) \leqslant \sum_{B \in \mathcal{B}} \nu(B) = \sum_{B \in \mathcal{B}} \mu(B)$$

Thus $\nu(A) \leqslant \mu^*(A)$. By symmetry, $\nu(A^c) \leqslant \mu^*(A^c)$. But then $\nu(A) + \nu(A^c) = \nu(\Omega) = \mu^*(\Omega)$, and assuming $\nu(\Omega) = \mu^*(\Omega) < \infty$,

$$\nu(A) = \nu(\Omega) - \nu(A^c) \geqslant \mu^*(\Omega) - \mu^*(A^c) = \mu^*(A)$$

Now assume Ω is σ-finite, partitioned into a countable disjoint collection $\mathcal{B} \subset \mathcal{A}$ with $\mu(B) < \infty$ for each $B \in \mathcal{B}$. Fix some enumeration $\{B_i\}$ of $\mathcal{B}$. Then, for any $A \in \mathcal{A}'$,

$$\mu^*(A) = \sum_{k=1}^{\infty} \mu^*(A \cap B_i) = \sum_{k=1}^{\infty} \nu(A \cap B_i) = \nu(A)$$

And so the two measures are equal. □

Example. *Consider the set of intervals of the form* (a, b) *in* $\mathbf{R}$*, where* $a, b \in \mathbf{R}^{\infty}$*. This is quite easily verified to be a semi-algebra. Define*

$$\mu([a,b]) = \mu((a,b]) = \cdots = \mu((a,b)) = b - a$$

μ *is countably additive, as we have shown in chapter 1, so it extends to a measure which contains all borel subsets of* $\mathbf{R}$*. This is just the Lebesgue measure of* $\mathbf{R}$*.*

Chapter 4

Interpolation

Interpolation is a core part of the 'hard' style of analysis, crunchy quantitative estimates that give strict bounds on quantitities. One basic tool here is interpolation, which 'in essense' enables one to take two results A_0 and A_1, and via combining them together obtain a family of results' between' the two results, of the form A_t for $0 \leqslant t \leqslant 1$.

The most basic example occurs in the theory of real numbers. Suppose $0 \leqslant A_0 \leqslant B_0$ and $0 \leqslant A_1 \leqslant B_1$. If we define $A_t = A_0^{1-t} A_1^t$, and $B_t = B_0^{1-t} B_1^t$, then it is trivial to verify that $A_t \leqslant B_t$, for the power function $x \mapsto x^a$ is order preserving for $a > 0$. For $t = 1/2$, we obtain the geometric mean inequality

$$\sqrt{A_0 A_1} \leqslant \sqrt{B_0 B_1}$$

Another way to see that the bound is trivial is to notice that the points $\log A_t$ are just the straight line from $\log A_0$ to $\log A_1$, and the result is established geometrically once we notice that log is order preserving.

4.1 Interpolation of Scalars

Let's consider some examples. If $A_0 = AX^{1/p}$, and $A_1 = AX^{1/q}$, then $A_t = AX^{(1-t)/p+t/q}$. These deductions are trivial, but we can still learn about the general inequality from them. For instance, if $A_0 = A_1$, then we find a lower bound $A \leqslant B_t$ over all $0 \leqslant t \leqslant 1$, and this bound can be refined to

$$A \leqslant B_t \min(B_0/B_1, B_1/B_0)^{\varepsilon}$$

for $\varepsilon < \min(t, 1-t)$. This tells us that the bound $A \leqslant B_t$ can only be sharp when B_0 and B_1 are roughly equal to one another. If $B_0 = 2^n B_1$, then we can improve the standard bound by a factor of $2^{-|n|\varepsilon}$. Since $2^{-|n|\varepsilon}$ is absolutely summable, it is a good heuristic to imply that the needed interpolation bound is negligable for $|n| \gg 0$.

The inequality $A_t \leqslant B_t$ can be easily generalized to the case where the A_t are defined in such a way that $t \mapsto \log A_t$ is a convex function (we say that the A_t are *log convex*), and $t \mapsto \log B_t$ is concave (though we normally always assume the $B_t are constant$). Thus one can interpolate upper bounds for log convex functions to obtain upper bounds across the domain. However, we cannot use interpolation to lower bound log-concave functions, nor can we extrapolate bounds from interior points (bounding A_0 and $A_{1/2}$ give us no info about A_1), but upper bounding A_0 and lower bounding A_t do gives us a lower bound on A_1. This is just the contraposition of the interpolation inequality.

Application of interpolation relies on the existence of a large class of log convex functions. If f and g are log convex, then fg is log convex because the sum of two convex functions is convex. Similarily,

4.2 Interpolation of functions

If we consider a step function $f = A\chi_Y$, then $\|f\|_p = A\mu(Y)^{1/p}$. Notice this is a log convex function of p, because for any $C > 0$, $\log(C)/p$