Measure Theory

Muchang Bahng

Fall 2025

Contents

In math, we are first taught to solve simple equations like $x^2 - 2x + 4 = 0$ for a certain *number* x, but in real world applications, we must now solve for some *function* f satisfying an equation

$$\mathcal{L}(f) = 0 \tag{1}$$

where $\mathcal{L}$ is some operator on functions. This is usually difficult, and many times a solution does not exist. However, we can find approximate solutions, say

$$\begin{aligned} \mathcal{L}(f) &= 1/2 \\ \mathcal{L}(f) &= 1/4 \\ \mathcal{L}(f) &= 1/8 \\ \ldots &= \ldots \end{aligned}$$

and approximate the solution as

$$f = \lim_{n\to\infty} f_n \tag{2}$$

Given that this limit exists, we can usually define f pointwise using a point-wise limit

$$f(x) = \lim_{n\to\infty} f_n(x) \text{ for all } x \tag{3}$$

but the function in total is very ugly and not Riemann integrable. The classic non-Riemann integrable function is the

$$f(x) = \chi_{\mathbb{R}\setminus\mathbb{Q}}(x) := \begin{cases} 1 & \text{if } x \in \mathbb{R} \setminus \mathbb{Q} \\ 0 & \text{if} x \in \mathbb{Q} \end{cases} \tag{4}$$

Since $\mathbb{Q}$ is countable, we can enumerate $\mathbb{Q} = \{q_n\}_{n=1}^\infty$ and define the sequence of functions

$$f_n = 1 - \chi_{\{q_j\}_{j=1}^n}(x) \tag{5}$$

that start off with the constant function 1 and then "removes" points in $\mathbb{Q}$, setting their image to 0. It is clear that since we are removing points, every function in the sequence has an integral (from 0 to 1) of 1, and therefore the integral of f should also be 1.

$$\int_0^1 f_n\,dx = 1 \implies \int_0^1 f\,dx = \int_0^1 \lim_{n\to\infty} f_n\,dx = \lim_{n\to\infty}\int_0^1 f_n\,dx \tag{6}$$

What is crucial for mathematicians to work with is the capability to take the limit from inside the integral to outside the integral. The problem is that f is not a Riemann integral function.

Definition 0.1 (Riemann Integrable Function)

Given a function $f : [0,1] \longrightarrow \mathbb{R}$, let us consider some partition of $[0,1]$ into intervals $P = \{I_0, I_1, \ldots, I_N\}$, then, for each $I \in P$, we can take the supremum $M_I = \sup_{x\in I} f(x)$ and infimum $m_I = \inf_{x\in I} f(x)$ and bound f by the upper and lower Riemann sums.

$$\sum_{I\in P} m_I |I| \leq \int_0^1 f\,dx \leq \sum_{I\in P} M_I |I| \tag{7}$$

where $|I|$ is the length of interval I. If we take *all* possible partitions, the bound should still hold.

$$m = \sup_P \Big\{ \sum_{I\in P} m_I |I| \Big\} \leq \int_0^1 f\,dx \leq \inf_P \Big\{ \sum_{I\in P} M_I |I| \Big\} = M \tag{8}$$

and if the lower bound is equal to the upper bound $m = M$, then the integral is this number and f is considered Riemann integrable.

Now since $\mathbb{Q}$ is dense in $\mathbb{R}$, for every interval I in every partition P will have $m_I = 0$ and $M_I = 1$ for the Riemann function, meaning that the lower bound will always be 0 and the upper bound will always be 1. So, $\int_0^1 \chi_{\mathbb{R}\setminus\mathbb{Q}}(x)$ can take on any value in $[0, 1]$, which isn't helpful. The fact that we can't integrate this really simple function is a problem. For nice functions, we can partition it so that the base of each Riemann rectangle is a nice interval, while the base of the Riemann function is an "interval with holes." The problem really boils down to measuring what the "length" of this set is. So the problem with the Riemann integral isn't the integral itself, but the fact that we can't give a meaningful size to the set $\mathbb{R} \setminus \mathbb{Q}$. Now mathematicians in the 19th century thought that as long as we solve this problem, we should be good to go, but Banach and Tarski proved that there exists sets that cannot be measured with their famous paradox, which says that you can take any set P, disassemble it into a finite set of pieces, and rearrange it (under isometry and translations) so that it has a different size than the original P. So, we have to exclude some sets that are not measurable. The collection of sets that we *can* measure is called the σ-algebra.

1 Jordan Measure

We would like to generalize the concepts of size, which are specified as length/area/volume depending on the dimension of the space we live in. The most intuitive notion of size are those of segments, rectangles, and boxes, and these are the simplest forms of sets that we will work with.

Such that for "simple" sets A where we know what the area is, the outer measure of A should coincide with the area of A. Let's first start by defining what a "simple" set is.

Definition 1.1 (Box)

An **box** $E \subset \mathbb{R}^n$ is defined recursively as follows.
1. An **interval** $I \subset \mathbb{R}$ is one of the sets $(a, b), [a, b), (a, b], [a, b]$ for $a, b \in \mathbb{R}$.
2. For $n > 1$, an box $E \subset \mathbb{R}^n$ is $E = I_1 \times \ldots \times I_n$ for intervals $I_1, \ldots, I_k$.

Definition 1.2 (Size of a Box)

The **size** of a box $E = I_1 \times \ldots \times I_n \subset \mathbb{R}^n$ is defined as follows.
1. The **length** of an interval I is $\ell(I) \coloneqq b - a$.
2. The **size** of E is $|E| \coloneqq \prod_{i=1}^n (b_i - a_i)$.

Now we can combine these to get an elementary set.

Definition 1.3 (Elementary Set)

An **elementary set** is a set $E \subset \mathbb{R}^n$ that is a finite union of boxes.

We would like to have some nice properties of these elementary sets.

Lemma 1.1 (Boolean Closure of Elementary Sets)

Given two elementary sets $E, F \subset \mathbb{R}^n$,
1. $E \cup F$ is elementary.
2. $E \cap F$ is elementary.
3. $E \setminus F$ is elementary.

Proof.

Lemma 1.2 ()

Let $E \subset \mathbb{R}^n$ be an elementary set. Then, E can be expressed as a finite union of *disjoint* boxes.

Proof.

Definition 1.4 (Elementary Measure)

The **elementary measure** of an elementary set $E \subset \mathbb{R}^n$ is defined as the sum of the sizes of each box in a partition:

$$m(E) := \sum_{i=1}^{k} |B_k| \tag{9}$$

We claim that this sum is invariant depending on the partition, and hence, well defined.

Proof.

This elementary measure clearly extends the notion of size, since

$$m(B) = s(B) \tag{10}$$

whenever B is elementary. Furthermore, we can deduce finite additivity and nonnegativity. These are really trivial but we state them as theorems to establish a pattern.

Lemma 1.3 (Fundamental Properties of Elementary Measure)

The elementary measure satisfies the following.

1. *Nonnegativity.* For any elementary set E, $m(E) \geq 0$.
2. *Finite Additivity* Given $E_1, \ldots, E_n$ are disjoint elementary sets,
$$m(E_1 \cup \ldots \cup E_k) = m(E_1) + \ldots + m(E_k) \tag{11}$$
3. *Monotonicity.* Given elementary sets $E \subset F$, we have
$$m(E) \leq m(F) \tag{12}$$
4. *Finite Subadditivity.* Let $E_1, \ldots, E_n$ be any elementary sets (not necessarily disjoint). Then
$$m(E_1 \cup \ldots \cup E_n) \leq m(E_1) + \ldots + m(E_n) \tag{13}$$
5. *Translation Invariance.* For $x \in \mathbb{R}^n$ and elementary set $E \subset \mathbb{R}^n$,
$$m(E) = m(x + E) \tag{14}$$

It turns out that these properties uniquely determine an elementary measure.

Theorem 1.1 (Uniqueness of Elementary Measure)

1.1 Jordan Measure

Now, we define the outer and inner measure, which are defined for *all* subsets of $\mathbb{R}^n$.

Definition 1.5 (Jordan Outer, Inner Measure)

Let $E \subset \mathbb{R}^n$.

1. The **Jordan inner measure** is defined
$$m_*(E) := \sup_{A \subset E, A \text{ elementary}} m(A) \tag{15}$$
2. The **Jordan outer measure** is defined
$$m_*(E) := \inf_{B \supset E, B \text{ elementary}} m(B) \tag{16}$$

Note that if E is unbounded, then there exists no elementary set that is a superset of E, and so the infimum of such a set is $+\infty$ conventionally.

This is where our first big leap in construction comes in. Before, we have defined elementary boxes, which are pretty much guaranteed to have a well-defined elementary measure. Here, we *begin* with a function on the power set of $\mathbb{R}^n$, and then we will filter the power set to those subsets that behave nicely.

Definition 1.6 (Jordan Measurable Set, Jordan Measure)

Let $E \subset \mathbb{R}^n$ be bounded.[a] If $m_*(E) = m^*(E)$, then E is said to be **Jordan-measurable**, and we define

$$m(E) \coloneqq m_*(E) = m^*(E) \tag{17}$$

as the **Jordan measure** of E.

[a] Note that by convention, we don't consider unbounded sets to be Jordan measurable.

Note first of all that Jordan measure is a generalization of elementary measure, since if E is elementary, then we can set $A = E = B$ to achieve these bounds. Furthermore, by monotonicity, we can never get past them, and will always have

$$m(A) \leq m(E) \leq m(B) \tag{18}$$

where m is the elementary measure. So, we can overload the notation and just write m to denote elementary and Jordan measure. Second, note that the Jordan measure shares the same properties.

Lemma 1.4 (Boolean Closure of Jordan Measurable Sets)

Given two Jordan-measurable sets $E, F \subset \mathbb{R}^n$,
1. $E \cup F$ is elementary.
2. $E \cap F$ is elementary.
3. $E \setminus F$ is elementary.

Proof.

The properties of the Jordan measure parallel those of elementary measure.

Theorem 1.2 (Fundamental Properties of Jordan Measure)

The elementary measure satisfies the following.
1. *Nonnegativity.* For any elementary set E, $m(E) \geq 0$.
2. *Finite Additivity* Given $E_1, \ldots, E_n$ are disjoint elementary sets,
$$m(E_1 \cup \ldots \cup E_k) = m(E_1) + \ldots + m(E_k) \tag{19}$$
3. *Monotonicity.* Given elementary sets $E \subset F$, we have
$$m(E) \leq m(F) \tag{20}$$
4. *Finite Subadditivity.* Let $E_1, \ldots, E_n$ be any elementary sets (not necessarily disjoint). Then
$$m(E_1 \cup \ldots \cup E_n) \leq m(E_1) + \ldots + m(E_n) \tag{21}$$
5. *Translation Invariance.* For $x \in \mathbb{R}^n$ and elementary set $E \subset \mathbb{R}^n$,
$$m(E) = m(x + E) \tag{22}$$

Proof.

Jordan measurable sets are sets that are "almost" elementary, but a few sets already come to mind that are not Jordan measurable.

Example 1.1 (Rationals in Unit Interval)

$\mathbb{Q} \cap [0, 1]$ is not Jordan measurable.

It may be hard to tell directly whether something is Jordan measurable. This is where the "Cauchy criterion" of Jordan measurable sets comes in.

Theorem 1.3 (Equivalent Notions)

E is Jordan measurable iff $\forall \epsilon > 0$, $\exists$ elementary sets $A \subset E \subset B$ s.t. $m(B \setminus A) < \epsilon$.

Proof.

Note how the previous lemma is very similar to this theorem on Riemann integrability.

Example 1.2 (Regions Under Graphs are Jordan Measurable)

Example 1.3 (Triangle is Jordan Measurable)

Example 1.4 (Compact Convex Polytopes are Jordan Measurable)

Example 1.5 (Open and Closed Balls in Euclidean Space are Jordan Measurable)

Example 1.6 (Subsets of Jordan Null Sets have 0 Jordan Measure)

Theorem 1.4 (Uniqueness of Jordan Measure)

Theorem 1.5 (Topological Approximations of Jordan Measurable Sets)

Let $E \subset \mathbb{R}^n$ be a bounded set. Then,

1. E and its closure $\overline{E}$ have the same Jordan outer measure.
2. E and its interior E° have the same Jordan outer measure.
3. E is Jordan measurable iff the topological boundary ∂E has Jordan outer measure 0.

Proof.

Example 1.7 (Bullet Riddled Square)

Show that both sets have a Jordan inner measure 0 and Jordan outer measure 1.

1. $[0,1]^2 \setminus \mathbb{Q}^2$.
2. $[0,1]^2 \cap \mathbb{Q}^2$.

Finally, a little teaser theorem.

Theorem 1.6 (Caratheodory Property)

Let $E \subset \mathbb{R}^n$ be a bounded set, and $F \subset \mathbb{R}^n$ be an elementary set. Show that

$$m^*(E) = m^*(E \cap F) + m^*(E \setminus F) \tag{23}$$

where m^* is the Jordan outer measure.

Proof.

1.2 Riemann Integration

Now, we connect the Riemann integral to the Jordan measure.

Theorem 1.7 ()

If $E \subset [a,b]$ is Jordan measurable, then the indicator function $\mathbb{1}_E$ is Riemann integrable, and

$$\int_a^b \mathbb{1}_E \, dx = m(E) \tag{24}$$

2 Lebesgue Measure

We have seen that there are some common sets that are not Jordan measurable, but a bigger problem is that countable unions and intersections aren't.

Example 2.1 (Countable Union/Intersection of Jordan Measurable Sets are Not J.M.)

We show a few counterexamples.

1. *Countable Union.* Take $\mathbb{N} = \{n\}_{n=1}^\infty$. Each point n has Jordan measure 0, but their union is unbounded so it isn't Jordan measurable.
2. *Bounded Countable Union.* Maybe we can fix this by considering bounded unions. But consider $E = \mathbb{Q} \cap [0,1]$. By density of rationals,
$$m_*(E) = 0 \neq 1 = m^*(E) \tag{25}$$
3. *Intersection.* Consider the Cantor set C, which is bounded, but again
$$m_*(C) = 0 \neq 1 = m^*(C) \tag{26}$$

In some applications, we can just ignore some of these sets as "pathological." But for Riemann integrability, we are integrating over Jordan measurable sets. I turns out that it is really important that we can *at least* guarantee that countable unions and intersections of Jordan measurable sets are measurable. If we don't, then not even uniform convergence—an extremely strong property—can preserve continuity, differentiability, and integrability of a sequence of functions.

Example 2.2 (Disasters of Uniform Limit of Integrable Functions is Not Integrable)

Tao Exercise 1.2.2.

This motivates the definition of a σ-algebra.

Definition 2.1 (σ-Algebra)

A $\boldsymbol{\sigma}$**-algebra** on a set X is a collection of subsets of X satisfying the following:

1. *Contains Empty and Full Set.* $\emptyset, X \in \mathcal{A}$.
2. *Closed Under Countable Unions.* $A_n \in \mathcal{A}$ for all $n \in \mathbb{N}$ implies $\bigcup_n A_n \in \mathcal{A}$.
3. *Closed Under Complements.* $A \in \mathcal{A} \implies A^c \in \mathcal{A}$.
4. *Closed Under Countable Intersections.*[a] $A_n \in \mathcal{A}$ for all $n \in \mathbb{N}$ implies $\bigcap_n A_n \in \mathcal{A}$.

[a] This is actually a consequence of the previous properties. We can utilize the fact that $\cap_{k=1}^\infty A_k = X \setminus \cup_{k=1}^\infty A_k^c$

2.1 Lebesgue Outer Measure

Therefore, we would like the collection of our measurable sets to be a σ-algebra. To do this, we tinker around with the definition of the Jordan measure. Note that by definition, the Jordan outer measure can be equivalently written as

$$\begin{aligned} m^*(S) &:= \inf\{m(E) \mid S \subset E, E \text{ elementary}\} && (27) \\ &= \inf\left\{ \sum_{i=1}^k |B_i| \;\middle|\; S \subset \bigcup_{i=1}^k B_i, B_i \text{ boxes}, k \in \mathbb{N} \right\} && (28) \end{aligned}$$

Note that the *finite* number of boxes k are allowed to vary over all naturals. To define the Lebesgue measure, we change the finite to countable.

Definition 2.2 (Lebesgue Outer Measure)

Given any set $E \subset \mathbb{R}^n$, the **Lebesgue outer measure** is a map

$$m^* : 2^{\mathbb{R}^n} \to [0, +\infty], \qquad m^*(E) = \inf \left\{ \sum_{k=1}^\infty |B_k| \;\middle|\; E \subset \bigcup_{k=1}^\infty B_k \right\} \tag{29}$$

It's a hard definition, but a natural one, since we're taking all these boxes and trying to make them as snug as possible to define the outer measure of an arbitrary set. We first first check that this is indeed a generalization of Jordan measure, in the sense that if E is Jordan measurable, then its Lebesgue outer measure is the same as its Jordan measure.

Theorem 2.1 (Lebesgue Outer Measure Coincides with Interval Length)

m^* satisfies the property that for any interval $I \subset \mathbb{R}$, $m^*(I) = |I|$.

Proof. Let's first consider the case when I is closed. Let $I = [a, b]$. Then, we know from definition that

$$m^*(I) \coloneqq \inf \left\{ \sum_{k=1}^\infty |I_k| \;\middle|\; I \subset \bigcup_{k=1}^\infty I_k \right\} \tag{30}$$

where $I_k = [a_n, b_n]$. We wish to show that the above quantity equals $b - a$.

1. $m^*(I) \leq b - a$. This is pretty easy since we can just set the cover to consist of the single interval I, and since $m^*(I)$ must be the infimum of it, then we must have $m^*(I) \leq b - a$. A technicality is that we must strictly use countable covers, but in this case, we can just fix $\epsilon > 0$ and see

$$I_1 = [a, b], \qquad I_k = [b - \frac{\epsilon}{2^k}, b + \frac{\epsilon}{2^k}] \tag{31}$$

In this case the sums of lengths of all $I_2, \ldots$ is $\epsilon/2 < \epsilon$, and so

$$I_k \leq b - a + \epsilon \qquad \forall \epsilon > 0 \tag{32}$$

2. Proving $m^*(I) \geq b - a$ is harder. In here, we use the "ϵ of room" trick. Take any $\epsilon > 0$. Then there exists a cover $\{I_k\}_k$ s.t.

$$m^*(I) = \epsilon \geq \sum_{k=1}^\infty |I_k| = \sum_{k=1}^\infty b_k - a_k \tag{33}$$

All we wish to show that the RHS $\geq b - a$, but we can't really manipulate the infinite sum. This is where we use the fact that $[a, b]$ is compact[a], and so we can take a finite subcover $\{I_{k_j}\}_{j=1}^n$. Therefore,

$$m^*(I) + \epsilon \geq \sum_{j=1}^n b_{k_j} - a_{k_j} \tag{34}$$

Now we can rearrange this: set the a_{k_j}'s to be increasing, and for simplicity let's reindex them to a_j, b_j. Then, it must be the case that $a_1 < a$.

(a) Consider (a_1, b_1). If $b_1 > b$, we are done.

(b) Otherwise, $b_1 \in (a_2, b_2)$. If $b_2 > b$, then

$$b_2 - a_2 + b_1 - a_1 \geq b_2 - a_1 > b - a \tag{35}$$

(c) If not, then we keep going until we get to (a_n, b_n). If $b_n > b$, then

$$b_n - a_n + b_{n-1} - a_{n-1} + \ldots + b_1 - a_1 \geq b_n - a_1 > b - a \tag{36}$$

[a] since it's closed and bounded

The proof may be a bit unfamiliar since we have used two tricks.

1. *Geometric Sequence of ϵ Trick.* To account for countable collections, we set ϵ to be decreasing geometrically so that the series converges to ϵ.
$$\epsilon = \sum_{n=1}^{\infty} \frac{\epsilon}{2^n} \tag{37}$$
2. *ϵ of Room Trick.* This trick is used often when you need an opposite inequality that isn't as obvious, and it can only be used for inequalities involving a supremum or infimum.
$$\inf\{S\} \geq x \tag{38}$$
By using the definition of sup/inf as the *least* upper/lower bound, we can scooch over by an ϵ to find an element in $s \in S$ that does satisfy the inequality
$$\inf\{S\} + \epsilon > s \geq x \tag{39}$$
and since ϵ was arbitrary, we are done.

It is clear that the Lebesgue outer measure is always less than the Jordan outer measure.
$$m^*(E) \leq m^{(J),*}(E) \tag{40}$$
When are these different?

Example 2.3 (Lebesgue Outer Measure Much Smaller than Jordan Outer Measure)

Consider $\mathbb{Q}$.
1. The Jordan outer measure is $+\infty$ since it is unbounded.
2. However, any countable set of $\mathbb{R}$ has Lebesgue outer measure 0. Just enumerate $E = \{x_1, \ldots\}$. Then, we set $I_k = \left(x_k - \frac{\epsilon}{2^k}, x_k + \frac{\epsilon}{2^k}\right)$. Then,
$$\sum_{k=1}^{\infty} \ell(I_k) = \epsilon \tag{41}$$

What about the inner measure? It turns out that we don't get much if we replace the finite to countable in the Jordan inner measure.

Lemma 2.1 (Axiomatic Properties of Lebesgue Outer Measure)

The Lebesgue outer measure satisfies the following.
1. *Null Empty Set.* $m^*(\emptyset) = 0$.
2. *Monotonicity.* Given sets $E \subset F \subset \mathbb{R}^n$, we have
$$m^*(E) \leq m^*(F) \tag{42}$$
3. *Countable Subadditivity.* For any countable collection of subsets $\{E_k\}$ of $\mathbb{R}^d$,
$$m^*\left(\bigcup_n E_n\right) \leq \sum_n m^*(E_n) \tag{43}$$

Proof. The first two properties are trivial. For the third, let's start by writing out the definition for the outer measure for each E_n.

$$m^*(E_n) := \inf \left\{ \sum_{k=1}^\infty |B_k^{(n)}| \;\middle|\; E_n \subset \bigcup_{k=1}^\infty B_k^{(n)}, B_k^{(n)} \text{ boxes} \right\} \tag{44}$$

Somehow, we want to sum these values over all n and prove that this is greater than the measure of the union. The first realization should be that for a fixed cover $\{B_k^{(n)}\}_k$ of E_n, the collection

$$\{B_k^{(n)}\}_{n,k} \text{ covers } \bigcup_n E_n \tag{45}$$

This gives us a clue as to comparing the collection of covers of each E_n, with the cover of $\cup E_n$. There is no straightforward way to do this,[a] so we want to try and *fix* these collections. We can do this with the ϵ of room trick. For each E_n, we can find a cover $\{B_k^{(n)}\}_k$ s.t.

$$m^*(E_n) + \frac{\epsilon}{2^n} \geq \sum_{k=1}^\infty |B_k^{(n)}| \tag{46}$$

Then, we can take the infinite sum.

$$\sum_{n=1}^\infty \left(m^*(E_n) + \frac{\epsilon}{2^n} \right) \geq \sum_{n=1}^\infty \sum_{k=1}^\infty |B_k^{(n)}| \geq m^* \left(\bigcup_n E_n \right) \tag{47}$$

where the final inequality follows from the fact that $\{B_k^{(n)}\}_{n,k}$ is a cover of $\cup_n E_n$, and so it must be greater than the infimum of all possible covers. All monotonic series converge in $[0, +\infty]$, so we can "take out" the ϵ term to get

$$\epsilon + \sum_{n=1}^\infty m^*(E_n) \geq m^* \left(\bigcup_n E_n \right) \tag{48}$$

which holds for all $\epsilon > 0$, and so we are done.

[a] On one side, we have the sum of a countable number of infimums of some sets, and on the other hand, we have the infimum of the unions of all of these sets.

Theorem 2.2 (Translation Invariance)

m^* is translation invariant. That is, for any $E \subset \mathbb{R}^n$,

$$m^*(E) = m^*(E + a), \qquad E + a := \{x + a \in \mathbb{R}^n \mid a \in E\} \tag{49}$$

Proof. This is straightforward and requires no tricks. Note that

$$\{B_k\} \text{ is a countable cover of } E \iff \{B_k + a\} \text{ is a countable cover of } E + a \tag{50}$$

It is also clear that

$$|B| = |B + a| \tag{51}$$

for any box $B \subset \mathbb{R}^n$, so the sets of sizes that we are taking the infimum over is exactly the same between the two.

The final property claims that we can always drop an outer-measure 0 set and it won't affect the outer measure of the original set. Therefore, when talking about measurability of intervals, we don't have to worry about endpoints, or even whether it is missing a countable number of elements from it!

Lemma 2.2 (Sets of Measure 0 have no Effect)

Suppose $m^*(E) = 0$ and A is any set. Then, $m^*(A \cup E) = m^*(A)$.

Proof. We have

$$m^*(A \cup E) = \underbrace{m^*\big((A \cup E) \cap E\big)}_{=0} + m^* \underbrace{\big((A \cup E) \cap E^c\big)}_{\subset A} \leq m^*(A) \leq m^*(A) \tag{52}$$

But $A \cup E \supset A$, so $m^*(A \cup E) = m^*(A)$.

2.2 Measurable Sets

The next step is to take the outer measure and define *Lebesgue measurable* sets. The problem is that—unlike the Jordan measure—we don't have the inner measure to work with. This turns out to be not much of a problem, since through *Littlewood's first principle*[1], we can define measurability as being well-approximated by an open set.

Definition 2.3 (Measurable Set)

A set $E \subset \mathbb{R}^d$ is **Lebesgue measurable** if it satisfies one of the equivalent properties.

1. *Carathéodory's criterion.*[a] For every set $A \subset \mathbb{R}^n$,

$$m^*(A) = m^*(A \cap E) + m^*(A \cap E^c) \tag{53}$$

   Note that due to countable subadditivity, we are guaranteed to have $\leq$. Therefore, it suffices to prove only for $\geq$. The sets with which this inequality is strict is not measurable, and the measurable sets specifically satisfy equality for countable sets.
2. *Outer Approximately Open.* $\forall \epsilon > 0$, $\exists$ open set $O \supset E$ s.t. $m(O \setminus E) \leq \epsilon$.
3. *Inner Approximately Closed.* $\forall \epsilon > 0$, $\exists$ closed set $F \subset E$ s.t. $m^*(E \setminus F) < \epsilon$.
4. *Outer Exactly G_δ.* $\exists$ a G_δ set G s.t. $E \subset G$ and $m^*(G \setminus E) = 0$.
5. *Inner Exactly F_σ.* $\exists$ a F_σ set F s.t. $F \subset E$ and $m^*(E \setminus F) = 0$.

[a] Colloquially, no matter how nasty a subset A is, E should be nice enough that we can cut E into two pieces C and D.

Proof. Listed.

1. (2) $\implies$ (1). Then for every $k \in \mathbb{N}$, we can find $O_k \supset E$ s.t. $m^*(O_k \setminus E) \leq 1/k$. Define the G_δ set $G = \cap_{k=1}^\infty O_k$. Then, $(G \setminus E) \subset (O_k \setminus E)$ for all $k \implies m^*(G \setminus E) \leq 1/k$ for all k. Therefore $m^*(G \setminus E) = 0$, and $E = G \setminus (G \setminus E)$ is measurable.
2. (1) $\implies$ (2). Assume $m^*(E) < +\infty$. Find a cover $\{I_k\}_{k=1}^\infty$ s.t. $\sum_{k=1}^\infty \ell(I_k) \leq m^*(E) + \epsilon$. Call $O = \cup_k I_k$. Since E is measurable, $m^*(O \setminus E) = m^*(O) - m^*(E) \leq \sum_{k=1}^\infty \ell(I_k) - m^*(E) \leq \epsilon$
3. (1) $\iff$ (3). Straightforward with argument above.
4. (1) $\iff$ (4). Generally, we use the fact that E measurable iff E^c measurable. Find $O \supset E^c$ open, with $m^*(O \setminus E^c) \leq \epsilon$. Then $F = O^c$ is closed, $F \subset E$, and $m^*(E \setminus F) \leq \epsilon$.
5. (1) $\iff$ (5). Same argument as (1) $\iff$ (4).

Depending on the context, it is helpful to use one definition over another when proving measurability. Just remember that the Carathéodory definition is the most general, since it doesn't even assume a topology on a space, and that is the definition that we will use by default. So what kind of sets are measurable?

[1] One of the major themes in measure theory, where we say that measurable sets are well-approximated by open and closed sets.

Example 2.4 (Outer Measure 0 Sets are Lebesgue Measurable)

For any outer measure m^* on X, $E \subset X$ with $m^*(E) = 0$ implies that E is m^*-measurable. Take any A. Then $(A \cap E) \subset E$ and $(A \cap E^c) \subset A$. So by monotonicity,

$$m^*(A \cap E) + m^*(A \cap E^c) \leq m^*(E) + m^*(A) = m^*(A) \tag{54}$$

and this by definition means that E is measurable.

We could continue with more examples, but our main priority is to show that this family of Lebesgue measurable sets is indeed a σ-algebra, and it covers much more than Jordan measurable sets. The path to prove that countable unions are measurable is a long one, and we'll a lot of lemmas.

Lemma 2.3 (Complements of Measurable Sets are Measurable)

If E is measurable, then so is E^c.

Proof. The definition is symmetric in both E and E^c.

Lemma 2.4 (Excision Property)

If $E \subset \mathbb{R}^n$ is measurable with $m^*(E) < +\infty$, and $E \subset F$ for arbitrary set F, then

$$m^*(F \setminus E) = m^*(F) - m^*(E) \tag{55}$$

Proof. By measurability of E, we can see

$$\begin{aligned} m^*(F) &= m^*(F \cap E) + m^*(F \cap E^c) && (56)\\ &= m^*(E) + m^*(F \setminus E) && (57) \end{aligned}$$

This excision property combined with the fact that outer measure 0 sets are always measurable gives us the property of *completeness*.[2] That is, given measurable sets $A \subset B \subset C$ with $m^*(A) = m^*(C)$, B must be measurable. This basically says that if you a set that is squeezed in between two measurable sets of equal measure, then the middle set will also be measurable.

Lemma 2.5 (Finite Unions are Measurable)

A finite union of measurable sets is measurable.

Proof. This proof is basically applying set theory laws, and there's not much more to that. It suffices to prove for E_1, E_2, and the rest follows by induction. Fix any A. Then

$$\begin{aligned} m^*(A) &= m^*(A \cap E_1) + m^*(A \cap E_1^c) && (58)\\ &= m^*(A \cap E_1) + m^*\big(A \cap E_1^c \cap E_2\big) + m^*\big((A \cap E_1^c) \cap E_2^c\big) && (59) \end{aligned}$$

Now we apply the identity $(A \cap E_1^c) \cap E_2^c = A \cap (E_1 \cup E_2)^c$, so the third term can be changed

$$= m^*(A \cap E_1) + m^*\big((A \cap E_1^c) \cap E_2\big) + m^*\big(A \cap (E_1 \cup E_2)^c\big) \tag{60}$$

Then we apply the identity $(A \cap E_1) \cup (A \cap E_1^c \cap E_2) = A \cap (E_1 \cup E_2)$, so we can apply finite subadditivity

[2] Nothing to do with completeness in the sense of real numbers or metric spaces.

on the first two terms to get

$$\geq m^*(A \cap (E_1 \cup E_2)) + m^*\big(A \cap (E_1 \cup E_2)^c\big) \tag{61}$$

which proves that $E_1 \cup E_2$ is measurable.

So we have proved that the set of all measurable sets is closed under finite unions. By definition it works for finite intersections. This makes it into an *algebra*, but we want to upgrade this to a σ-algebra by proving closure under *countable* unions. We first prove a lemma.

Lemma 2.6 (Finite Additivity of Outer Measure on Disjoint Measurable Sets)

Suppose A is any set, $\{E_k\}_{k=1}^n$ disjoint and measurable. Then,

$$m^*\bigg(A \cap \Big(\bigcup_{k=1}^n E_k\Big)\bigg) = \sum_{k=1}^n m^*(A \cap E_k) \tag{62}$$

In particular,

$$m^*\bigg(\bigcup_{k=1}^n E_k\bigg) = \sum_{k=1}^n m^*(E_k) \tag{63}$$

Proof. It should be clear that we prove by induction, and intuitively, this disjointness should be essential for canceling out some measure terms. $n = 1$ is trivial. This takes a bit of fiddling around with where we should start, but if we just look at the LHS, we can try and use Carathéodory, by setting the arbitrary set to be $B = A \cap (\cup_k E_k)$ and writing $m^*(B) = m^*(B \cap E_n) + m^*(B \cap E_n^c)$.

$$\begin{aligned}
m^*\bigg(A \cap \Big(\bigcup_{k=1}^n E_k\Big)\bigg) &= m^*\Bigg(\bigg(A \cap \Big(\bigcup_{k=1}^n E_k\Big)\bigg) \cap E_n\Bigg) + m^*\Bigg(\bigg(A \cap \Big(\bigcup_{k=1}^n E_k\Big)\bigg) \cap E_n^c\Bigg) && (64)\\
&= m^*(A \cap E_n) + m^*\bigg(A \cap \Big(\bigcup_{k=1}^{n-1} E_k\Big)\bigg) && (65)\\
&= \sum_{k=1}^n m^*(A \cap E_k) && (66)
\end{aligned}$$

But note that by disjointness, we have

$$\bigg(A \cap \Big(\bigcup_{k=1}^n E_k\Big)\bigg) \cap E_n = A \cap E_n, \qquad \bigg(A \cap \Big(\bigcup_{k=1}^n E_k\Big)\bigg) \cap E_n^c = A \cap \Big(\bigcup_{k=1}^{n-1} E_k\Big) \tag{67}$$

by the induction hypothesis.

Here is a wrong proof that does an incorrect form of induction. I first assumed that we can just work with a family of two sets E_1, E_2, so I started deriving something like this:

$$\begin{aligned}
m^*(A) &= m^*(A \cap (E_1 \cup E_2)) + m^*(A \cap (E_1 \cup E_2)^c) && (68)\\
&= m^*((A \cap E_1) \cup (A \cap E_2) + \underbrace{m^*((A \setminus E_1) \cap (A \setminus E_2))}_{\geq 0} && (69)\\
&\geq m^*((A \cap E_1) \cup (A \cap E_2)) && (70)
\end{aligned}$$

Note that the E_k's being disjoint means that they are *pairwise* disjoint, and so it is *not* sufficient to prove for only E_1, E_2. So don't do this.

Theorem 2.3 (Countable Unions are Measurable)

Suppose $E_1, E_2, \ldots$ are a countable collection of measurable sets. Then, $E = \cup_{j=1}^\infty E_j$ is measurable.

Proof. They key is to look at disjoint sets. WLOG, one can assume E_j are disjoint. Indeed, we can define new sets[a]

$$E_n' \coloneqq E_n \setminus \bigg(\bigcup_{j=1}^{n-1} E_j \bigg) \tag{71}$$

that are measurable, with $\cup E_n' = \cup E_n$. Now, fix any set A. Define sets $F_n = \cup_{j=1}^n E_j$. Then, $m^*(A) = m^*(A \cap F_n) + m^*(A \cap F_n^c)$. Then, $F_n^c \supset E^c \implies m^*(A \cap F_n^c) \geq m^*(A \cap E^c)$. Since we have established from the previous lemma that outer measure on disjoint measurable sets satisfies finite additivity, we can write

$$m^*(A \cap F_n) = m^* \bigg(\bigcup_{j=1}^n (A \cap E_j) \bigg) = \sum_{j=1}^n m^*(A \cap E_j) \tag{72}$$

Then,

$$m^*(A) \geq \sum_{j=1}^n m^*(A \cap E_j) + m^*(A \cap E^c) \tag{73}$$

for every n, therefore also with ∞. But by countable subadditivity of the outer measure,

$$\sum_{j=1}^\infty m^*(A \cap E_j) \geq m^*(A \cap E) \tag{74}$$

If follows that $m^*(A) \geq m^*(A \cap E) + m^*(A \cap E^c)$.

[a] TBD: why is setminus measurable?

Corollary 2.1 (Measurable Sets form a σ-Algebra)

The set of all Lebesgue measurable sets of $\mathbb{R}^n$ form a σ-algebra.

Proof. Listed.

1. $\emptyset$ is measurable since it has outer measure 0.
2. $\mathbb{R}^n$ is measurable since for any set $A \subset \mathbb{R}^n$,

$$m^*(A) = m^*(A \cap \mathbb{R}^n) + m^*(A \cap (\mathbb{R}^n)^c) = m^*(A) \tag{75}$$

3. We proved closure under complements.
4. We just proved closure under countable union.

Now that we have established that Lebesgue measurable sets form a σ-algebra, let's give some examples.

Theorem 2.4 (Rays are Measurable)

Every interval $(a, +\infty)$ is measurable.

Proof. We simply prove using Carathéodory,[a] so we wish to show that for any set $A \subset \mathbb{R}$,

$$m^*(A) \geq m^*(A \cap (a, +\infty)) + m^*(A \cap (-\infty, a]) = m^*(A') + m^*(A'') \tag{76}$$

Again, the fact that we have to prove this nontrivial part of the inequality reminds us of using the ϵ of room trick. Let us $\{I_k\}_{k=1}^{\infty}$ is a countable cover of A s.t.

$$m^*(A) + \epsilon > \sum_{k=1}^{\infty} \ell(I_k) \tag{77}$$

We can take this cover and create two smaller covers, covering A' and A''.

$$\{I'_k \coloneqq I_k \cap A'\}_k, \qquad \{I''_k \coloneqq I_k \cap A''\}_k \tag{78}$$

Since these are valid covers, by definition it must be bounded below by the outer measures.

$$\sum_{k=1}^{\infty} \ell(I'_k) \geq m^*(A'), \qquad \sum_{k=1}^{\infty} \ell(I''_k) \geq m^*(A'') \tag{79}$$

Now all that remains is to connect the sums together. For each k, we have $\ell(I_k) = \ell(I'_k) + \ell(I''_k)$, and since both series converge in $[0, +\infty]$, we can indeed sum them up as limits to get

$$\sum_{k=1}^{\infty} \ell(I'_k) + \sum_{k=1}^{\infty} \ell(I''_k) = \sum_{k=1}^{\infty} \ell(I'_k) + \ell(I''_k) = \sum_{k=1}^{\infty} \ell(I_k) \tag{80}$$

Putting the bounds together gives

$$m^*(A') + m^*(A'') \leq \sum_{k=1}^{\infty} \ell(I'_k) + \sum_{k=1}^{\infty} \ell(I''_k) = \sum_{k=1}^{\infty} \ell(I_k) \leq m^*(A) + \epsilon \tag{81}$$

Since this is true for every $\epsilon > 0$, we are done.

[a] WLOG $a \notin A$ (since we can take the point out without affecting outer measure). TBD: Do we need this assumption really?

Note again that this ϵ of room trick is used so that we can *fix* some open cover that acts as a middle ground between the inequalities that we are trying to prove. Then as we let $\epsilon \to 0$, we are done.

Corollary 2.2 (All Intervals are Measurable)

All intervals $I \subset \mathbb{R}$ are measurable.

Example 2.5 (Cantor Set is Measurable)

Let us define

$$C_0 = [0, 1], \qquad C_1 = [0, 1/3] \cup [2/3, 1], \ldots \tag{82}$$

Basically, we take our every middle third of each subinterval. So C_k is the union of 2^k intervals of size 3^{-k}. Note that $C_k \subset C_{k-1}$. Now define the **Cantor set** as

$$C \coloneqq \bigcap_{k=1}^{\infty} C_k \tag{83}$$

The Cantor set is measurable since it is a countable intersection of closed sets, which are measurable.

Theorem 2.5 (Translations of Sets are Measurable)

If $E \subset \mathbb{R}^n$ is measurable, then for any $a \in \mathbb{R}^n$, $E + a := \{x + a \in \mathbb{R}^n \mid x \in E\}$ is measurable.

Proof. We again use Carathéodory. Let $A \subset \mathbb{R}^n$ by any set. Then by translation invariance of outer measure, we have

$$\begin{aligned} m^*(A) &= m^*(A - a) && (84) \\ &= m^*((A - a) \cap E) + m^*((A - a) \cap E^c) && (85) \\ &= m^*(A \cap (E + a)) + m^*(A \cap (E + x)^c) && (86) \end{aligned}$$

and so $E + a$ is measurable.

This next theorem is a different flavor of Littlewood's first principle. It tells us that we can use a finite union of intervals that "symmetrically" approximates measurable sets on the real line.

Theorem 2.6 (Finite Union of Intervals are Good Symmetric Approximations)

Suppose E is measurable, with $m^*(E) < +\infty$. Fix $\epsilon > 0$. Then there exists a finite number of intervals $\{I_k\}_{k=1}^n$ s.t. if $O = \cup_{k=1}^n I_k$, then

$$m^*(O \setminus E) + m^*(E \setminus O) < \epsilon \tag{87}$$

Proof. In here, we use the outer approximately open definition of measurable sets. Since every open set can be written as a countable union of open intervals[a], we can find a collection of open intervals $\{I_k\}_{k=1}^\infty$ s.t.

$$E \subset U := \bigcup_{k=1}^\infty I_k, \qquad m^*(U \setminus E) \leq \frac{\epsilon}{2} \tag{88}$$

The major thing to do now is to reduce the countable union into a finite union. Note that we can just take any subcollection of the I_k's, and we are guaranteed that their union O will satisfy

$$m^*(O \setminus E) \leq m(U \setminus E) \leq \frac{\epsilon}{2} \tag{89}$$

The problem is that we don't want to take too small of a collection so that the other difference is too big. To do this, we can just select the tail of the series: Find n s.t. $\sum_{k=n+1}^\infty \ell(I_k) \leq \epsilon/2$ where WLOG, I_k are disjoint. Define $O = \cup_{k=1}^n I_k$. Then, we have

$$\begin{aligned} m^*(O \setminus E) &\leq m(U \setminus E) \leq \frac{\epsilon}{2} && (90) \\ m^*(E \setminus O) &\leq m(U \setminus O) \leq \sum_{k=n+1}^\infty \ell(I_k) \leq \frac{\epsilon}{2} && (91) \end{aligned}$$

[a]since $\mathbb{R}^n$ is second countable

This symmetry in difference induced me to use the inner approximately closed definition in addition. My idea was to just find a closed set F s.t. $F \subset E \subset U$, but there is no straightforward way of finding one finite collection of intervals O.

Example 2.6 (idk where to put this)

One should note that in particular, if E is m^*-measurable and A is any set disjoint from E, then we must have

$$m^*(A \cup E) = m^*((A \cup E) \cap E) + m^*((A \cup E) \cap E^c) \quad (92)$$
$$= m^*(E) + m^*(A) \quad (93)$$

which solves a bit of the theorem on measures.

2.3 Measures

Now by restricting our outer measure to measurable sets, we get our measure.

Definition 2.4 (Lebesgue Measure)

The restriction the Lebesgue outer measure m^* to the set of all measurable sets $\mathcal{A}$, is called the **Lebesgue measure**

$$m = m^*\big|_{\mathcal{A}} \quad (94)$$

Note that for outer measures, they satisfy both countable subadditivity and finite additivity. With measures, we get the best of both worlds: countable subadditivity.

Lemma 2.7 (Axiomatic Properties of Lebesgue Measure)

The Lebesgue measure satisfies the following.

1. *Null Empty Set.* $m(\emptyset) = 0$.
2. *Countable Additivity.* For all countable collections $\{A_k\}_{k=1}^\infty$ of pairwise disjoint subsets $A_k \in \mathcal{A}$[a],

$$m\bigg(\bigsqcup_{k=1}^{\infty} A_k\bigg) = \sum_{k=1}^{\infty} m(A_k) \quad (95)$$

[a] Remember that we are allowed to take countable unions inside our σ-algebra, so this makes sense.

Proof. Listed.

1. *Null Empty Set.* Since this is true for outer measure m^*.
2. *Countable Additivity.* $m(\cup E_j) = \sum_j m(E_j)$. $\leq$ is trivial by countable subadditivity of the outer measure. For $\geq$, note that for every $n \in \mathbb{N}$,

$$m\bigg(\bigcup_{j=1}^{\infty} E_j\bigg) \geq m\bigg(\bigcup_{j=1}^{n} E_j\bigg) = \sum_{j=1}^{n} m(E_j) \quad (96)$$

   where the inequality comes from monotonicity and the equality comes from finite additivity of the outer measure. Now take $n \to \infty$.

Unlike the outer measure, monotonicity is not an axiomatic property because of two independent reasons, both sufficient. First, the Lebesgue outer measure suffices. Second, it is actually a direct consequence of the two axiomatic properties.

Lemma 2.8 (Translation Invariance)

The Lebesgue measure is translation invariant.

Proof. We know that translations of Lebesgue measurable sets are also Lebesgue measurable, and the Lebesgue outer measure is translation invariant.

Now we provide some "continuity" properties of the Lebesgue measure.

Theorem 2.7 (Continuity From Below)

If $A_1 \subset A_2 \subset A_3 \subset \ldots$, then

$$m\bigg(\bigcup_{k=1}^\infty A_k\bigg) = \lim_{k\to\infty} m(A_k) \tag{97}$$

Proof. First, note that the limit on the RHS is defined, since $m(A_k)$ is nondecreasing and so must converge in $[0, +\infty]$. But why does this limit equal to the left hand side? The only property that makes sense to work with is countable additivity, so we should define the disjoint collection

$$B_1 = A_1, \quad B_k = A_k \setminus A_{k-1} \tag{98}$$

Then, it becomes straightforward

$$\begin{aligned}
m\bigg(\bigcup_{k=1}^\infty A_k\bigg) &= m\bigg(\bigcup_{k=1}^\infty B_k\bigg) && \text{(Construction)} \\
&= \sum_{k=1}^\infty m(B_k) && \text{(Countable Additivity)} \\
&= \lim_{n\to\infty} \sum_{k=1}^n m(B_k) && \text{(Definition of Series)} \\
&= \lim_{n\to\infty} m\bigg(\bigcup_{k=1}^\infty B_k\bigg) && \text{(Finite Additivity)} \\
&= \lim_{n\to\infty} m(A_k) && (99)
\end{aligned}$$

Proof. Old proof. We can see that

$$\begin{aligned}
m\bigg(\bigcup_{k=1}^\infty A_k\bigg) &= m(A_1) + \sum_{k=2}^\infty m(B_k) && (100) \\
&= m(A_1) + \lim_{k\to\infty} \sum_{k=2}^\infty m(B_k) && (101) \\
&= \lim_{k\to\infty} m(A_1 \cup B_2 \cup \ldots B_k) = \lim_{k\to\infty} m(A_k) && (102)
\end{aligned}$$

Now a similar theorem, but with a little twist to it.

Theorem 2.8 (Continuity from Above)

If $A_1 \supset A_2 \supset A_3 \supset \ldots$, then

$$m\bigg(\bigcap_{k=1}^{\infty} A_k\bigg) = \lim_{k\to\infty} m(A_k) \tag{103}$$

if $m(A_1) < \infty$.

Proof. First, note that the $m(A_1) < +\infty$ is a necessary condition, since if we take $A_k = [k, \infty)$ on the real number line, then we have $\cap_{k=1}^{\infty} A_k = \emptyset$, but the limit of the measure is ∞. We did not have this problem for continuity from below.
Well we can define $B_k = A_k \setminus A_{k+1}$ and write $\cap_{k=1}^{\infty} A_k = A_1 \setminus \cup_{k=1}^{\infty} B_k$, which means that

$$\begin{aligned}
m\bigg(\bigcap_{k=1}^{\infty} A_k\bigg) &= m\bigg(A_1 \setminus \bigcup_{k=1}^{\infty} B_k\bigg) && (104)\\
&= m(A_1) - m\bigg(\bigcup_{k=1}^{\infty} B_k\bigg) && \text{(Excision)}\\
&= m(A_1) - \sum_{k=1}^{\infty} m(B_k) && \text{(Countable Additivity)}\\
&= m(A_1) - \lim_{n\to\infty} \sum_{k=1}^{n} m(B_k) && \text{(Definition of Series)}\\
&= \lim_{n\to\infty} \bigg(m(A_1) - \sum_{k=1}^{n} m(B_k)\bigg) && (105)\\
&= \lim_{n\to\infty} m\bigg(A_1 \setminus \bigcup_{k=1}^{n} B_k\bigg) && \text{(Excision)}\\
&= \lim_{n\to\infty} m(A_n) && (106)
\end{aligned}$$

Now the first line uses the fact that if $A \subset B$, then $m(B \setminus A) + m(A) = m(B)$, and with the further assumption that $m(A) < \infty$, we can subtract on both sides like we do with regular arithmetic.

We will see two applications of continuity from above.

Example 2.7 (Cantor Set has Measure 0)

The Cantor set has measure 0. We can see that it is the intersection of all C_k's which are nested $C_k \supset C_{k+1}$ and $m(C_0) = m([0, 1]) = 1$. Therefore, by continuity from above,

$$m\bigg(\bigcap_{k=1}^{\infty} C_k\bigg) = \lim_{k\to\infty} m(C_k) = \lim_{k\to\infty} \frac{2^k}{3^k} = 0 \tag{107}$$

It is also closed as an intersection of closed sets. It is also uncountable, since we can just do it using a tradic system and see that the Cantor set are all reals with infinite triadic representation of digits 0 and 2. Then create a bijection with binary represenation of reals. Here's a new way I learned. Suppose C is countable, so enumerate it: $c_1, c_2, \ldots$. Pick one interval I_1 in C_1 that doesn't contain c_1. Then, pick $I_2 \subset I_1 \cap C_2$ s.t. it doesn't contain c_2. Keep going, and we get

$$I_1 \supset I_2 \supset I_3 \supset \ldots \tag{108}$$

By nested intervals lemma, these are closed, bounded, and nested, which is nonempty. So we've found a point not in the Cantor set, contradicting the fact that we have enumerated it.

Lemma 2.9 (Borel-Cantelli)

Suppose $\{E_k\}_{k=1}^\infty$ are measurable, with $\sum_k m(E_k) < +\infty$. Then,

$$m\big(\limsup E_k\big) \coloneqq m\bigg(\bigcap_{n=1}^\infty \bigcup_{k=n}^\infty E_k\bigg) = 0 \tag{109}$$

That is, all $x \in \mathbb{R}$ belonging to only a finite number of E_k, has measure 0.

Proof. This is a direct application of continuity from above, where we can set

$$B_n \coloneqq \bigcup_{k=n}^\infty E_k \tag{110}$$

Notice that $B_n \supset B_{k+1}$ and B_1 has finite measure since by countable subadditivity[a],

$$m(B_1) = m\bigg(\bigcup_{k=1}^\infty E_k\bigg) \leq \sum_{k=1}^\infty m(E_k) < +\infty \tag{111}$$

Therefore, we can derive

$$\begin{aligned} m\bigg(\bigcap_{n=1}^\infty \bigcup_{k=n}^\infty E_k\bigg) &= \lim_{n\to\infty} \bigg(\bigcup_{k=n}^\infty E_k\bigg) && \text{(Continuity from Above)} \\ &= \lim_{n\to\infty} \sum_{k=n}^\infty m(E_k) && \text{(Countable Additivity)} \\ &= 0 && (112) \end{aligned}$$

because the tail of series converging to a finite value must tend to 0. Note that for the last step, we could have used countable subadditivity as well.

[a]not countable additivity!

Definition 2.5 (Almost Everywhere)

Given a measure space $(X, \mathcal{A}, m)$, a subset $A \in \mathcal{A}$ is said to be a m-null set if $m(A) = 0$. If some property holds for all points $x \in X$ except on a null set, then we say that the property holds **almost everywhere**.

Example 2.8 (Rational Function)

The function $f(x) = \frac{1}{\sqrt{|x|}}$ is less than ∞ almost everywhere.

2.4 Nonmeasurable Sets

Lemma 2.10 (Quotienting over Countable Set Implies Lebesgue Measure 0)

Suppose E is measurable, bounded, and there exists a countably infinite, bounded set Λ s.t. $\{E+m\}_{m\in\Lambda}$ are disjoint. Then, $m(E) = 0$.

Proof. Consider $\cup_{m\in\Lambda}\{E+m\}$. It is bounded, so its measure is finite. Also, by countable addivitiy and translation invariance, we get

$$+\infty > m\bigg(\bigcup_{m\in\Lambda}\{E+m\}\bigg) = \sum_{m\in\Lambda} m(E+m) = \sum_{m\in\Lambda} m(E) \tag{113}$$

Since Λ is infinite, we must have $m(E) = 0$.

Recall that we say x is *rationally equivalent* to y if $x - y \in \mathbb{Q}$. This is an equivalence relation on $\mathbb{R}$, giving us a quotient set of an uncountable number of classes. A *choice set* for this equivalence relation is a set containing exactly one element from each class.[3] We can do this on any $E \subset \mathbb{R}$.

Lemma 2.11 ()

If C_E is any choice set on such E, then
1. $\forall x, y \in C_E$, if $x - y \notin \mathbb{Q}$, then for all $\Lambda \in \mathbb{Q}$, $\{m + C_E\}_{m\in\Lambda}$ are disjoint.
2. $\forall x \in E$, there exists $y \in C_E$ s.t. $x - y \in \mathbb{Q}$.

Theorem 2.9 (Every Set of Positive Outer Measure Contains Nonmeasurable Set)

Any set $E \subset \mathbb{R}^n$ of positive outer measure contains a nonmeasurable set.

Proof. WLOG, let E be bounded.[a] Let C_E be any choice set. Suppose C_E is measurable. Let b be such that $E \subset [-b, b]$. Let $\Lambda = \mathbb{Q} \cap [-2b, 2b]$. Consider the disjoint family of sets $\{C_E + \lambda\}_{\lambda\in\Lambda}$. Also,

$$E \subset \bigcup_{\lambda\in\Lambda}\{C_E + \lambda\} \tag{114}$$

Indeed, $\forall x \in E$, there exists $y \in C_E$ s.t. $x - y \in \mathbb{Q}$ and in Λ by definition of Λ and $E \subset [-b, b]$. By the lemma, $m(C_E) = 0$. But also,

$$m^*(E) \leq \sum_{\lambda\in\Lambda} \underbrace{m(C_E + \lambda)}_{=m(C_E)} = 0 \tag{115}$$

which is a contradiction, so C_E is not measurable.

[a] Otherwise, just take a bounded subset.

Now let's talk about fractals. The Cantor set is the simplest case of a fractal. Fractals were of interest when people tried to find the length of the coast of England. We would like to measure the "size" of fractals, and the measure is not a very good method since you always just get measure 0. The Hausdorff dimension is a nicer way.

[3] This assumes axiom of choice.

Definition 2.6 (Hausdorff Content)

The α**-Hausdorff content** of an arbitrary set S is defined[a]

$$h^\alpha(S) \coloneqq \lim_{\delta \to 0} \inf_{S \subset \bigcup_{r_i \leq \delta} B_{r_i}(x_i)} \sum_i r_i^\alpha \tag{116}$$

[a]To clarify, the term in the limit is an inf over a countable cover of open balls, with centers x_i and radius r_i, each $\leq \delta$.

Definition 2.7 (Hausdorff Dimension)

The **Hausdorff dimension** of an arbitrary set S is

$$d^H \coloneqq \inf_\alpha \{\alpha \mid h^\alpha(S) = 0\} \tag{117}$$

Example 2.9 ()

Take a look at a straight line $S = [0, 1]$. Intuitively, you need (on the order of) $1/r$ balls of radius r. So,

$$h^\alpha(S) \sim r^\alpha \cdot \frac{1}{r} = r^{\alpha - a} \to \begin{cases} +\infty & \text{if } \alpha < 1 \\ 0 & \text{if } \alpha > 1 \end{cases} \tag{118}$$

So the Hausdorff dimension is 0.

Example 2.10 ()

For the Cantor set, the Hausdorff dimension is $\log 2 / \log 3$.

Definition 2.8 (Cantor-Lebesgue Function, Devil's Staircase)

The **Cantor-Lebesgue function** $\phi : [0, 1] \to \mathbb{R}$ is defined as such.

1. Let us define $O_k = [0, 1] \setminus C_k{}^a$, which is an open set. So O_k consists of $2^k - 1$ open intervals I_j (j indexed from left to right). For O_k, we define $\phi(x) = j/2^k$, where j is the number of the interval I_j, indexed left to right. This defines ϕ on $O = \cup_{k=1}^\infty O_k = [0, 1] \setminus C$.
2. On C, let us define $\phi(x) \coloneqq \inf_{y \geq x, y \in O} \phi(x)$.

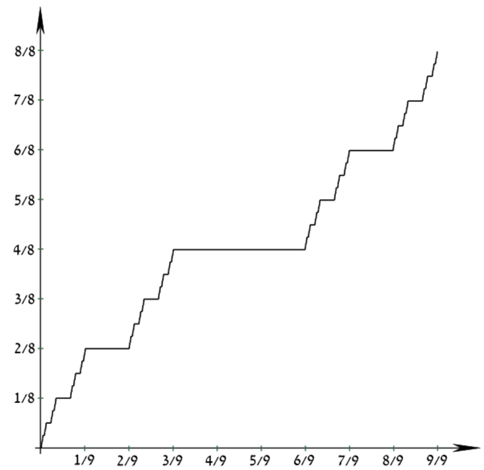


Figure 1: Plot

[a] C_k defined as before when constructing the Cantor set.

Theorem 2.10 ()

ϕ is a nondecreasing, continuous function s.t. $\phi'(x) = 0$ for all $x \in O$ and $m(O) = 1$.

Proof. Listed.

1. *Increasing.* ϕ is increasing on each O_k, and so on O. Then, it is also increasing on C by definition.
2. *Continuity.* If $x \in C$, it lies between 2 intervals of O_k for any k. The difference in function values between 2 neighboring intervals of O_k is 2^{-k}, so ϕ is continuous.
3. *Derivative.* The derivative is 0 because it is constant around an interval.

Theorem 2.11 (Pathological Strictly Increasing Devil's Staircase)

Define $\psi(x) = \phi(x) + x$. Then,

1. ψ is continuous and strictly increasing.
2. ψ maps C into a set of positive measure.
3. ψ maps some subset of C into a nonmeasurable set.

Proof. Listed.

1. Continuity is from sum of continuous functions, and strictly increasing since ϕ is nondecreasing and x is strictly increasing.
2. We know that $\psi([0,1]) = [0,2]$ and $m(\psi(O)) = 1$, where for each interval $I_j \subset O$, $m(\psi(I_j)) = \ell(I_j)$. Therefore, $m(\psi(C)) = 1$.
3. Since ψ is strictly increasing, there exists a continuous inverse ψ^{-1}. Find $Z \subset \psi(C)$ that is nonmeasurable, which we can do from the previous theorem. Then, there exists $E \subset C$ s.t. $\psi(E) = Z$, and E is not Borel since if it were, then Z would be Borel, too.

3 Measurable Functions

So far, we have defined measurable sets, constructed the Lebesgue measure, and shown that Lebesgue measurable sets can be approximated by nice open sets. Now, let's talk about measurable functions. Just like for measurable sets, there is a general sense in which we can define them and there is the more "Euclidean" way of defining them.

Definition 3.1 (Measurable Function)

Given a measurable space $(X, \mathcal{A})$, $f : (X, \mathcal{A}) \longrightarrow \mathbb{R}$ is **measurable** if

$$f^{-1}(A) \in \mathcal{A} \text{ for all } A \text{ open} \tag{119}$$

Note that measurable functions are always defined on measurable sets, so we don't have to state that its domain is always measurable.

Theorem 3.1 (Measurable Functions on Real Line)

Let $f : E \subset \mathbb{R} \to \mathbb{R} \cup \{\pm\infty\}$ and E be measurable. Then, TFAE

1. $\forall c \in \mathbb{R}$, $\{x \in E \mid f(x) > c\}$ is measurable.
2. $\forall c \in \mathbb{R}$, $\{x \in E \mid f(x) \geq c\}$ is measurable.
3. $\forall c \in \mathbb{R}$, $\{x \in E \mid f(x) < c\}$ is measurable.
4. $\forall c \in \mathbb{R}$, $\{x \in E \mid f(x) \leq c\}$ is measurable.
5. f is Lebesgue measurable.

Furthermore, if any of these hold, then also $\{x \in E \mid f(x) = c\}$ is measurable for all c (but not the converse!).

Proof. We know that (1) $\iff$ (4) and (2) $\iff$ (3) by taking complements. We prove (1) $\iff$ (2).

1. (1) $\implies$ (2).
$$\{x \in E \mid f(x) \geq c\} = \bigcap_{k=1}^{\infty} \{x \in E \mid f(x) > c - \frac{1}{k}\} \tag{120}$$
2. (2) $\implies$ (1).
$$\{x \in E \mid f(x) > c\} = \bigcup_{k=1}^{\infty} \{x \in E \mid f(x) \geq c + \frac{1}{k}\} \tag{121}$$

For (5), we know that

1. (5) $\implies$ (1) is trivial, since open intervals are open sets.
2. (1) $\implies$ (5). Any open set is a countable union of disjoint open intervals, and so let
$$U = \bigcup_{k=1}^{\infty} I_k, \qquad I_k = (a_k, b_k) = \underbrace{(-\infty, b_k)}_{B_k} \cap \underbrace{(a_k, +\infty)}_{A_k} \tag{122}$$
Therefore,
$$f^{-1}(U) = \bigcup_{k=1}^{\infty} f^{-1}(B_k \cap A_k) = \bigcup_{k=1}^{\infty} \{f^{-1}(B_k) \cap f^{-1}(A_k)\} \tag{123}$$
which is measurable since countable union/intersections are measurable (by definition of σ-algebra).

For the final implication, we can use (2) and (4) to get

$$\{x \in E \mid f(x) = c\} = \{x \in E \mid f(x) \leq c\} \cup \{x \in E \mid f(x) \geq c\} \tag{124}$$

The first question is how you would relate this to continuity.

Theorem 3.2 (Continuous Functions are Measurable)

If $f : X \to \mathbb{R}$ is continuous, then it is measurable.

Proof. If f is continuous, then $f^{-1}(O) = U \cap X$ for every open O with open U.

Theorem 3.3 (Monotonic Functions are Measurable)

Let I be an interval. If $f : I \subset \mathbb{R}$ is monotone, then f is measurable.

Proof. You can probably see that it is more advantageous to prove using the definition of measurability using rays. We wish to show that for all c, $E_c \coloneqq \{x \in I \mid f(x) > c\}$ is measurable. We wish to show that E_c is an interval, though there seems to be some complications with potential discontinuities. Therefore, we use an equivalent definition of an interval: I is an interval if for every $x, y \in I$, $x < t < y \implies t \in I$. Therefore, we can see that if $x, y \in E_c$, then $f(x) > c, f(y) > c$. Therefore, if t is in between them, $f(t) > f(\min\{x, y\}) > c$, and so $t \in E_c$. Since intervals are measurable, we are done.

There is also some notion of robustness.

Theorem 3.4 (Function Difference on Measure 0 Set Doesn't Affect Measurability)

Suppose $f : E \subset \mathbb{R} \to \mathbb{R} \cup \{\pm\infty\}$ with E measurable, and let g be some other function. If f is measurable on E and $g(x) = f(x)$ a.e. for $x \in E$, then g is measurable on E.

Proof. We wish to show that for any open $O \subset \mathbb{R}$, $g^{-1}(O)$ is measurable. We might start with Carathéodory and try to show that for all $A \subset E$,

$$m^*(A) = m^*(A \cap g^{-1}(O)) + m^*(A \cap g^{-1}(O)^c) \tag{125}$$

But this turns out to be overkill. Since this is about 0 measure sets, you should be thinking about how 0-measure sets do not affect measurability and try to use this. In $g^{-1}(O)$, there is a portion of it that overlaps with f—call it $A \subset E$—and a portion that doesn't. We know that $m^*(E \setminus A) = 0$[a] and a measure 0 set difference doesn't affect measurability, so A is measurable. So let's decompose it.

$$\begin{aligned} g^{-1}(O) &= \big(g^{-1}(O) \cap A\big) \cup \big(g^{-1}(O) \cap (E \setminus A)\big) &(126)\\ &= \big(f^{-1}(O) \cap A\big) \cup \big(g^{-1}(O) \cap (E \setminus A)\big) &(127) \end{aligned}$$

If we try to take the measure of this, the first term is the union of measurable sets $f^{-1}(O)$ and A. The second term is also measurable since the outer measure is 0, by subadditivity compared to $m^*(E \setminus A) = 0$. Therefore $g^{-1}(O)$ is measurable.

[a] TBD: Can we write m?

Proof. In class. Consider $S = \{x \in E \mid g(x) < c\}$. Let $A \subset E$ be the set where $g(x) = f(x)$, with $m(E \setminus A) = 0$. Then,

$$S = \big(\{x \in E \mid g(x) < c\} \cap (E \setminus A)\big) \cup \big(\{x \in E \mid f(x) < c\} \cap A\big) \tag{128}$$

where the first term is measure 0 by monotonicity with $E \setminus A$, $m(E \setminus A) = 0$, and the second term is measurable since $A = E \setminus (E \setminus A)$. So, S is measurable.

You preserve measurability if you split the domain in a "measurable way."

Theorem 3.5 (Measurable Partition Induces Measurable Restrictions of Functions)

Take a measurable subset $D \subset E$ and let $f : E \to \mathbb{R} \cup \{\pm\infty\}$ be a function. Then, the following are equivalent.

1. f is measurable on E
2. f is measurable on D and on $E \setminus D$.

Proof. We prove bidirectionally.

1. $(\to)$. Let's prove measurability on D. We can see that
$$\{x \in D \mid f(x) \in O\} = \{x \in E \mid f(x) \in O\} \cap D \tag{129}$$
as the intersection of measurable sets, is measurable. Then we can just take the complement of both sides to get.
$$\begin{aligned} \{x \in E \setminus D \mid f(x) \in O\} &= E \setminus \{x \in D \mid f(x) \in O\} && (130)\\ &= E \setminus \big(\{x \in E \mid f(x) \in O\} \cap D\big) && (131)\\ &= \underbrace{\big(E \setminus \{x \in E \mid f(x) \in O\}\big)}_{\text{measurable}} \cup \underbrace{\big(E \setminus D\big)}_{\text{measurable}} && (132) \end{aligned}$$
which is also measurable.
2. $(\leftarrow)$. Take some open $O \subset \mathbb{R} \cup \{\pm\infty\}$ and take its preimage. Then,
$$f^{-1}(O) = \{x \in D \mid f(x) \in O\} \cup \{x \in E \setminus D \mid f(x) \in O\} \tag{133}$$
as the finite union and intersection of measurable sets, is measurable.

3.1 Arithmetic and Composition of Measurable Functions

The following theorem is useful, since we don't want to manually check measurability of every single new function we create.

Theorem 3.6 (Arithmetic on Measurable Functions)

Given measurable functions $f, g : E \subset \mathbb{R} \to \mathbb{R}$, the following standard operations on them create new measurable functions:

1. αf is measurable for all $\alpha \in \mathbb{R}$.
2. $f + g$ is measurable
3. $f \cdot g$ is measurable
4. f/g is measurable on $\{x \mid g(x) \neq 0\}$

Proof. WLOG, we can assume f, g are finite everywhere since changing these values to finite values over a set of measure 0 doesn't affect measurability.

1. If $\alpha = 0$, this is trivially true. If not, then
$$\{x \in E \mid (\alpha f)(x) < c\} = \{x \in E \mid f(x) < \frac{c}{\alpha}\} \tag{134}$$
2. Suppose $f(x) + g(x) < c \iff f(x) < c - g(x) \iff \exists q \in \mathbb{Q}$ s.t. $f(x) < q < c - g(x)$.[a] Then,
$$\{x \in E \mid f(x) + g(x) < c\} = \bigcup_{q \in \mathbb{Q}} \big(\{x \in E \mid f(x) < q\} \cap \{x \in E \mid g(x) < c - q\}\big) \tag{135}$$

which is a countable union of measurable sets, and is measurable.

3. We use a nice trick from analysis.

$$fg = \frac{1}{4}\big((f+g)^2 - (f-g)^2\big) \tag{136}$$

and so it suffices to prove that h measurable implies h^2 measurable. For $c \geq 0$[b], we have

$$\{x \in E \mid h^2(x) > c\} = \{x \in E \mid h(x) > \sqrt{c}\} \cup \{x \in E \mid h(x) < -\sqrt{c}\} \tag{137}$$

4.

[a]The reason we want to introduce rationals is that we want to take advantage of countability.
[b]We only need to consider this case since h^2 is always nonnegative and so $c < 0$ would mean preimage is empty set.

Theorem 3.7 (Finite Min/Max of Measurable Functions are Measurable)

If $f_1, \ldots, f_n : E \subset \mathbb{R} \to \mathbb{R}$ are measurable, then so are $\max_k f_k$ and $\min_k f_k$.

Proof. We can prove by induction, but this is still a one-liner. For maximum,

$$\{x \in E \mid (\max_k f_k)(x) > c\} = \bigcup_{k=1}^n \{x \in E \mid f_k(x) > c\} \tag{138}$$

and for the minimum,

$$\{x \in E \mid (\min_k f_k)(x) > c\} = \bigcap_{k=1}^n \{x \in E \mid f_k(x) > c\} \tag{139}$$

Example 3.1 (Composition of Two Functions need not be Measurable)

Recall from 2.4 that we built a function $\psi(x)$ that maps some measurable A to nonmeasurable $\psi(A)$. Let's extend ψ to all $\mathbb{R}$ and keep it strictly increasing. Let χ_A be the characteristic function of A. Consider $f = \chi_A \circ \psi^{-1}$, and take the preimage of $(1/2, +\infty)$ under ψ.

$$f^{-1}((\frac{1}{2}, +\infty)) = \{x \mid \psi^{-1}(x) \in A\} = \{x \in \psi(A)\} = \psi(A) \tag{140}$$

which we have proven that there exists some A s.t. $\psi(A)$ is not measurable.

So this is bad news, but we have a compromise.

Theorem 3.8 (Composition of Measurable then Continuous is Measurable)

Suppose g is measurable on E, f is continuous on $\mathbb{R}$. Then, $f \circ g$ is measurable.

Proof. Take any open O. Then,

$$(f \circ g)^{-1}(O) \iff g(x) \in f^{-1}(O) \tag{141}$$

where $f^{-1}(O)$ is open, which implies measurable.

So we get much more results, like that $|f|$ or $|f|^p$ is measurable if f is measurable.

3.2 Sequences of Measurable Functions

Let's compare continuous functions and measurable functions. In terms of composition, continuity is a little more robust since we can compose continuous functions to get continuous functions. Meanwhile, we know that measurable functions don't necessarily compose to measurable functions. The relation is reversed when we talk about convergence. Recall from analysis the definitions of pointwise convergence and uniform convergence of a sequence of functions. First, we present an analogous measure-theoretic definition of pointwise convergence.

Definition 3.2 (Almost Sure Convergence)

A sequence of functions $(f_n : E \to \mathbb{R})_n$ is said to **converge almost surely to** f or **converge to** f **almost everywhere**—denoted $f_n \to f$—if $f_n(x) \to f(x)$ for all $x \in A \subset E$ where $m(E \setminus A) = 0$.

If you have uniform convergence, this is great since the uniform limit of continuous (Riemann integrable) functions is continuous (Riemann integrable). However, the pointwise limit of continuous (Riemann integrable) functions may fail to be continuous (Riemann integrable). It turns out that measurability is preserved through almost sure convergence.

Theorem 3.9 (Almost Sure Convergence of Measurable Functions are Measurable)

Suppose f_n are measurable on E and $f_n \to f$ a.e. on E. Then, f is measurable.

Proof. We can assume $f_n \to f$ everywhere (since behavior on measure 0 sets don't affect measurability). Now, we wish to show that $E_0 := \{x \in E \mid f(x) < c\}$ is a measurable set. Carathéodory would be a bad idea since there's so many moving parts, and so the next best thing to do is to try and construct it as a countable union/intersection of measurable sets.
Now if $f(x) < c$ for some fixed x, then this is equivalent to saying $\forall \epsilon > 0$, $\exists N \in \mathbb{N}$ s.t.

$$k \geq N \implies f_k(x) < c - \epsilon \tag{142}$$

Therefore, if $x \in E_0$, then we can construct this set in multiple stages.

1. For fixed $\epsilon > 0$ and $N \in \mathbb{N}$, it better be the case that $f_k(x) < c - \epsilon$ for all $k \geq N$, meaning that we must take the intersections of the set for each such k.

$$\bigcap_{k=N}^{\infty} \{x \in E \mid f_k(x) < c - \epsilon\} \tag{143}$$

2. Now we unfix N. There must exist just one N, so we want to take the union.

$$\bigcup_{N \in \mathbb{N}} \bigcap_{k=N}^{\infty} \{x \in E \mid f_k(x) < c - \epsilon\} \tag{144}$$

3. Now we unfix ϵ. This must be true for all ϵ, so this indicates that this must be a union. However, we don't want an uncountable union, so using the Archimidean principle, we might as well let $\epsilon = \frac{1}{m}$ vary over a countable set.

$$\bigcup_{m=1}^{\infty} \bigcup_{N \in \mathbb{N}} \bigcap_{k=N}^{\infty} \{x \in E \mid f_k(x) < c - \frac{1}{m}\} \tag{145}$$

Since we have only taken countable unions and intersections of each measurable $\{x \in E \mid f_k(x) < x - \epsilon\}$, we are done.

Note in the previous proof, that when we are constructing sets, the "for all" usually means an intersection, and "there exists" usually means a union.

So though continuous functions are more robust w.r.t. composition, measurable functions are more robust w.r.t. convergence. Consider a generalization of step functions called *simple functions*.

Definition 3.3 (Simple Functions)

For $A \subset X$ (any subset, not just in some σ-algebra), the **characteristic**, or **indicator function** of A is the function $\chi_A : X \longrightarrow \mathbb{R}$ defined

$$\chi_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if else} \end{cases} \tag{146}$$

A function $\phi : \mathbb{R} \longrightarrow \mathbb{R}$ is called a **simple function** if it is a finite linear combination of characteristic functions.

$$\phi = \sum_{i=1}^n a_i \chi_{A_i} \tag{147}$$

Lemma 3.1 (Simple Approximations Lemma)

Assume f is bounded on $E \subset \mathbb{R}$, measurable. For every $\epsilon > 0$, there exists simple functions $\phi_\epsilon, \psi_\epsilon$ s.t.

$$\phi_\epsilon \leq f \leq \psi_\epsilon, \qquad \psi_\epsilon - \phi_\epsilon \leq \epsilon \tag{148}$$

for all $x \in E$.

Proof. Suppose $|f(x)| \leq M$. Consider a "partition of $[-M, M]$ into intervals of size ϵ.

$$y_0 = -M < y_1 < y_2 < \ldots < y_{n-1} < y_n = M \tag{149}$$

where $y_k - y_{k-1} = h < \epsilon$ for all k. Define $E_k = f^{-1}([y_{k-1}, y_k])$, which are measurable. Then, we define

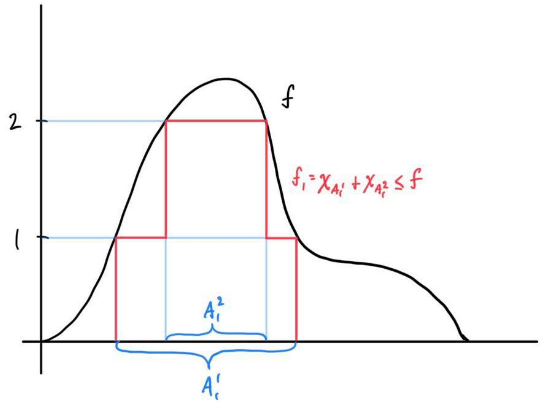
$$\phi_\epsilon(x) = \sum_{k=1}^n y_{k-1} \chi_{E_k}(x), \qquad \psi_\epsilon(x) = \sum_{k=1}^n y_k \chi_{E_k}(x) \tag{150}$$

and can show that this satisfies the properties.

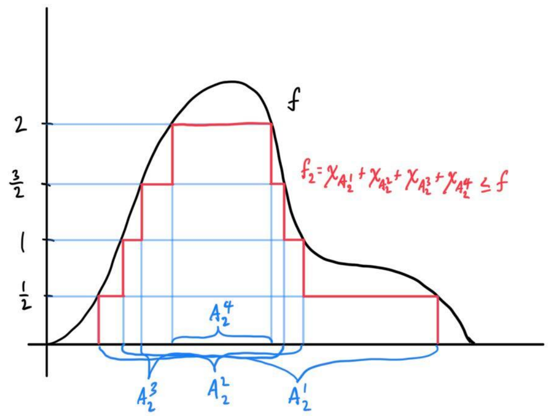
Theorem 3.10 (Simple Approximation Theorem)

Let $f : E \subset \mathbb{R} \cup \{\pm\infty\}$ be measurable. Then, there is a sequence of simple functions ϕ_n s.t. $\phi_n \to f$ for all $x \in E$, and $\|\phi_n(x)\| \leq \|f(x)\|$ for all $x \in E$.

Proof. We give a general picture of this proof for a function $f : \mathbb{R} \longrightarrow [0, \infty]$. We can first divide the codomain of the graph below into segments of $t = 1, 2, \ldots$, and take the preimage of all these units under f to get f_1. More specifically, $A_1^t = f^{-1}([t, \infty])$ for all t. By measurability of f, A_1^t is measurable, and we can assign $f_1 = \chi_{A_1^1} + \chi_{A_1^2} \leq f$.



Doing this again with finer subintervals of the codomain gives us, with $f_2 = \chi_{A_2^1} + \chi_{A_2^2} + \chi_{A_2^3} + \chi_{A_2^4} \leq f$.



and in general, we have $f_k = \sum_{j=1}^{\infty} \frac{1}{2^{k-1}} \chi_{A_k^j}$. But we said a simple function is a *finite* sum, and if ∞ is in the range of f, then this becomes a problem. We can quickly fix this by just truncating the summation at a certain point in the codomain (f_1 only considers intervals up to 1, f_2 up to 2 and so on), ultimately giving us

$$f_k = \sum_{j=1}^{k2^{k-1}} \frac{1}{2^{k-1}} \chi_{A_k^j} \tag{151}$$

3.3 Nearly Uniform Convergence of Measurable Functions

We have proved one of the first big theorems on measurable functions: that limit of measurable functions is also measurable. But we would like to find ways to extract *uniform* convergence. This is where another major idea of measure theory—Littlewood's third principle—comes in.[4]

Lemma 3.2 ()

Let $(f_n : E \subset \mathbb{R} \to \mathbb{R})$ be a sequence of measurable functions with $m(E) < +\infty$ that converges pointwise to f. Then, for each $\eta > 0$ and $\delta > 0$, there exists a measurable subset $A \subset E$ and index N such that

$$|f_n - f| < \eta \text{ on } A \text{ for all } n \geq N, \text{ and } m(E \setminus A) < \delta \tag{152}$$

Proof.

[4]The first principle was that every measurable set is nearly an open set. The second principle, ironically, will come last.

The next theorem is one we will use all the time. It basically tells us a way to turn a sequence of pointwise convergent functions into a sequence of uniformly convergent functions. It seems similar to Dini's theorem in that it gives conditions of uniform integrability for a sequence of pointwise convergence functions.

Theorem 3.11 (Egorov)

Let $(f_n : E \subset \mathbb{R} \to \mathbb{R})$ be a sequence of measurable functions with $m(E) < +\infty$ that converges pointwise to f. Then, for each $\epsilon > 0$, there exists a closed set $F \subset E$ s.t.

$$f_n \to f \text{ uniformly on } F \text{ and } m(E \setminus F) < \epsilon \tag{153}$$

Proof.

3.4 Continuous Approximations of Measurable Functions

So if we throw out a small set, we can approximate measurable functions with continuous functions.

Lemma 3.3 (Luzin, for Simple Functions)

Suppose $f : E \subset \mathbb{R} \to \mathbb{R}$ is simple. Then $\forall \epsilon > 0$, $\exists$ a closed $F \subset E$, $g \in C(\mathbb{R})$, s.t. $f|_F = g|_F$ and $m(E \setminus F) < \epsilon$.

Proof.

Theorem 3.12 (Luzin)

Suppose $f : E \subset \mathbb{R} \to \mathbb{R}$ is measurable. Then $\forall \epsilon > 0$, $\exists$ a closed $F \subset E$, $g \in C(\mathbb{R})$, s.t. $f|_F = g|_F$ and $m(E \setminus F) < \epsilon$.

Proof. Assume $m(E) < \infty$. If it is ∞, then we can divide the set into countable sets, each with finite measure, and we can do it for $\epsilon/2^n$. So it suffices for finite measure. Suppose f_n are simple and $f_n \to f$ pointwise on E. By the lemma, we can find closed $F_n \subset E$ s.t. $m(E \setminus F_n) < \epsilon/2^{n+1}$ and $g_n \in C(\mathbb{R})$, with $g_n|_{F_n} = f_n|_{F_n}$. Also, by Egorov, we can find F_0 s.t. f_n converges uniformly on F_0, and $m(E \setminus F_0) < \epsilon/2$. Define

$$F = \bigcap_{n=0}^{\infty} F_n, \qquad E \setminus F = \bigcup_{n=0}^{\infty} E \setminus F_n \tag{154}$$

Then by subadditivity,

$$m(E \setminus F) \leq \sum_{n=0}^{\infty} m(E \setminus F_n) < \epsilon \tag{155}$$

Finally, f_n converges uniformly on F and $f_n|_F = g_n|_F$, so f_n are continuous on F. Since uniform limit of continuous functions is continuous, the limit f is continuous on F.

This is an argument for the interval, but this can be generalized to more general sets.

4 Lebesgue Integration

Remember that Riemann integration is characterized by the approximation of step functions, which are the "building blocks" of Riemann integrable functions. To define the Lebesgue integral, we will start off with simple functions—which are a generalization of step functions. A function will be Lebesgue integrable if it can be approximated by these simple functions in some appropriate way. There are parallels between how we construct the Riemann and Lebesgue integral, namely that we can define the upper and lower integrals as the infimums and supremums of some set. However—while we were able to do this in "one shot" for the Riemann integral, the Lebesgue integral requires us to take intermediate steps. There are two ways that we can construct the Lebesgue integral.

1. We define the integral for simple functions ϕ over finite measure. Since simple functions are made up of a linear combination of indicators, which are themselves bounded, the finite measure locks in the property that $\int \phi$ will be bounded. This allows us to take the integral of simple functions that have real-valued coefficients—both positive and negative—but is inherently limiting as we can't define the integral over infinite measures. With this, we can define the integral of bounded measurable functions using the lower and upper integrals. This is similar to the construction of the Riemann integral. Furthermore, if we want to define the general integral, we have to now restrict what we have built up into the nonnegative case—a bit unnatural.
2. We first define the integral for *nonnegative* simple functions ϕ over any measure. We are sacrificing the ability to integrate negative simple functions early on, but this doesn't matter since we will split functions into a positive and negative part anyways. The true advantage of doing this is that we gain the ability to define integrals for infinite measures. This allows us to define for positive (possibly unbounded) measurable functions, and finally when we define the general Lebesgue integral, we deal with the $\infty - \infty$ problem by defining it only if at least one of the positive or negative integrals are finite.

I personally think the second way is superior, since in the end we can define integrals over infinite measure. Furthermore, since we are only working with positive functions, we only need to define the lower integral rather than checking to see if the lower and upper integrals coincide.

As we redefine these integrals over and over (sort of like method overriding), we really want to preserve three properties of the integral.

1. Linearity.
$$\int_E (\alpha f + \beta g) = \alpha \int_E f + \beta \int_E g \tag{156}$$
2. Monotonicity
$$f \leq g \implies \int_E f \leq \int_E g \tag{157}$$
3. Additivity.
$$A, B \text{ disjoint, measurable, and } E = A \cup B \implies \int_E f = \int_A f + \int_B f \tag{158}$$

4.1 Simple Functions

We will use ϕ, φ, ψ to denote simple functions.

Definition 4.1 (Lebesgue Integral of Simple Functions over Finite Measure)

Suppose $\phi = \sum_{k=1}^n a_k \chi_{E_k} : E \subset \mathbb{R} \to \mathbb{R}$ with $a_k \in \mathbb{R}$.

1. If $m(E) < +\infty$, then we can define the integral which is guaranteed to be a finite number.

$$\int_E \phi := \sum_{k=1}^n a_k m(E_k) \tag{159}$$

2. If $a_k \geq 0$, we define the integral which lives in $\mathbb{R} \cup \{+\infty\}$.[a]

$$\int_E \phi := \sum_{k=1}^n a_k m(E_k) \tag{160}$$

This is well defined for any representation of ϕ[b]

[a] As we have stated before, we could also define the Lebesgue integral of simple functions by letting a_k takes values in $\mathbb{R}$. But then, we might have a case where $E = A \sqcup B$, with $m(A) = +\infty, m(B) = +\infty$, and $\phi = \chi_A - 2\chi_B \implies \int_E \phi = \infty - \infty$. To prevent this from happening, some authors add the assumption that $m(E) < +\infty$, and I cover this case to make it as comprehensive as possible.

[b] We need this since the coefficients need not be unique. For example, we can write $1 \cdot \chi_{[0,1]} + 1 \cdot \chi_{[0.5,1]} = 1 \cdot \chi_{[0,0.5]} + 2 \cdot \chi_{[0.5,1]}$. If the E_i's are disjoint, then this decomposition is unique and is called the **standard representation** of ϕ.

Proof. It is clear that the two definitions coincide if $m(E) < +\infty$ and $a_k \geq 0$ is true; it is the same formula.

For bounded functions, we will temporarily focus on the first definition, and for general functions, we rely on the second definition.

Theorem 4.1 (Integral Properties for Simple Functions over Finite Measure)

Suppose $\phi, \psi : E \subset \mathbb{R} \to \mathbb{R}$ are simple with $m(E) < +\infty$. Then, the following properties hold.

1. Linearity. For $\alpha, \beta \in \mathbb{R}$,

$$\int_E (\alpha\phi + \beta\psi) = \alpha \int_E \phi + \beta \int_E \psi \tag{161}$$

2. Monotonicity.

$$\phi \leq \psi \implies \int_E \phi \leq \int_E \psi \tag{162}$$

3. Additivity. If A, B are disjoint, measurable, and $E = A \cup B$, then

$$\int_E \phi = \int_A \phi + \int_B \phi \tag{163}$$

Proof. Listed.

1. We can subdivide sets s.t. ϕ and ψ can be rewritten using the same finite family of sets A_k.
2. We can use (1) to rewrite

$$\int_E \psi - \int_E \phi = \int \underbrace{(\psi - \phi)}_{\geq 0 \forall x} \geq 0 \tag{164}$$

3. Trivial by definition of simple functions.

Example 4.1 (Step Function as Simple Function)

For $a, b \in \mathbb{R}$, with $a < b$, let $f : [a, b] \longrightarrow \mathbb{R}$ be a step function. That is, there exists a partition $a = x_0 < x_1 < \ldots < x_n = b$ and constants $c_1, c_2, \ldots, c_n \in \mathbb{R}$ s.t. $f(x) = c_i$ for all $x \in (x_{i-1}, x_i)$ and each $i = 1, \ldots, n$. Then, f is equal to the following simple function, taken over all open intervals and the points x_j at the boundary of each interval.

$$f = \sum_{i=1}^n c_i \chi_{(x_{i-1}, x_i)} + \sum_{j=0}^n f(x_j) \chi_{\{x_j\}} \tag{165}$$

If we ignore the behavior of f on the partition points x_j's, then f agrees almost everywhere with the simple function

$$\sum_{i=1}^n c_i \chi_{(x_{i-1}, x_i)} \tag{166}$$

4.2 Bounded Measurable Functions over Finite Measure

This isn't the most popular way to define Lebesgue integrability, but I'd like to compare it to Riemann integral. First, note that a Riemann integral

Definition 4.2 (Lower, Upper Lebesgue Integral)

Let $f : E \subset \mathbb{R} \to \mathbb{R}$ be bounded, with $m(E) < +\infty$. Then, the **upper and lower Lebesgue integrals** are defined

$$\overline{L}f := \inf_\phi \bigg\{ \int \phi \,\bigg|\, f \leq \phi, \phi \text{ simple} \bigg\}, \qquad \underline{L}f := \sup_\phi \bigg\{ \int \phi \,\bigg|\, \phi \leq f, \phi \text{ simple} \bigg\} \tag{167}$$

If the upper and lower Lebesgue integrals are equal, then f is said to be **Lebesgue integrable**.

This is exactly the same form as that of Riemann integration, notably that f must be bounded and we define integrability as the equivalence of the upper and lower integrals. The fact that $m(E)$ is finite is realized implicitly with partitions, though in this sense it's more similar to the Riemann-Stieltjes integral. From this, it's pretty easy to intuit that the Lebesgue integral agrees with the Riemann integral for step functions. Let $c_1, \ldots, c_n \in [0, \infty)$ and $a = x_0 < x_1 < \ldots < x_n = b$ be a partition. Let $f : [a, b] \longrightarrow [0, \infty]$ be a step function taking the value c_i on the interval (x_{i-1}, x_i) for $i = 1, \ldots, n$. Then the Riemann integral of f is simply

$$\int f(x)\, dx = \sum_{i=1}^n c_k |x_i - x_{i-1}| \tag{168}$$

The Lebesgue integral is

$$\int f\, d\mu = \sum_{i=1}^n c_i \mu((x_{i-1}, x_i)) + \sum_{j=0}^n f(x_j) \mu(\{x_j\}) = \sum_{i=1}^n c_k |x_i - x_{i-1}| \tag{169}$$

which agrees with the Riemann integral. In the Riemann integral, we write dx to indicate the variable that is being integrated over, but in the Lebesgue integral, we write $d\mu$, the measure which we are integrating over. We make this more rigorous in the following theorem.

Theorem 4.2 (Riemann Integrability Implies Lebesgue Integrability)

If f is Riemann integrable, then it is Lebesgue integrable.

Proof. Recall that f is Riemann integrable if

$$\sup_P L(P, f) = \inf_P U(P, f) \tag{170}$$

for a partition P. But for any $L(P, f)$ (or $U(P, f)$), there exist simple ϕ (or ψ) s.t.

$$\int_E \phi = L(P, f), \qquad \left(\int_E \psi = U(P, f) \right) \tag{171}$$

So,

$$\sup_P L(P, f) \leq \underline{L}f \leq \overline{L}f \leq \inf_P U(P, f) \tag{172}$$

where the first and third inequalities we just showed, and the middle inequality $\underline{L}f \leq \overline{L}f$ comes from monotonicity. So if the inf = sup, then $\underline{L}f = \overline{L}f$ has nowhere to go.

However, the converse is not true.

Example 4.2 (Lebesgue Integrable but Not Riemann Integrable)

Consider the simple function (consisting of one characteristic function) $\chi_{\mathbb{Q} \cap [0,1]}$. $\mathbb{Q} \cap [0, 1]$ is a Lebesgue measurable set of $\mathbb{R}$, and we have $\chi_{\mathbb{Q} \cap [0,1]} \geq 0$, so its Lebesgue integral is given by the above definition:

$$\int_{\mathbb{R}} \chi_{\mathbb{Q} \cap [0,1]} \, d\lambda = 1 \cdot \lambda(\mathbb{Q} \cap [0, 1]) = 0 \tag{173}$$

Remember that continuous functions are Riemann integrable. There is indeed an analogous result between measurable functions and Lebesgue-integrable functions.

Theorem 4.3 (Measurable Functions are Lebesgue Integrable)

Let $f : E \subset \mathbb{R} \to \mathbb{R}$ be measurable, bounded with $m(E) < +\infty$. Then, f is Lebesgue integrable.

Proof. We prove that $\forall n$, $\exists$ a simple ϕ_n, ψ_n s.t. $\phi_n \leq f \leq \psi_n$ and $\psi_n - \phi_n < 1/n$. Then,

$$\overline{L}f - \underline{L}f \leq \int \psi_n - \int \phi_n \leq \frac{1}{n} m(E) \tag{174}$$

So take $n \to \infty$.

Now, we have a much larger class of functions we can integrate. It turns out that there is an iff condition for Lebesgue integrability.

Theorem 4.4 (Lebesgue Integrable iff Set of Discontinuities have Measure 0)

A function $f : E \subset \mathbb{R}$ is Lebesgue integrable iff f is continuous almost everywhere.

Proof.

To prove additivity, we'll need a specific lemma.

Lemma 4.1 ()

Let $f : E \subset \mathbb{R} \to \mathbb{R}$ be measurable, bounded with $m(E) < +\infty$, and $A \subset E$ measurable. Then,

$$\int_E f \cdot \chi_A = \int_A f \tag{175}$$

Proof.

Theorem 4.5 (Integral Properties for Bounded Measurable Functions over Finite Measure)

Suppose $f, g : E \subset \mathbb{R} \to \mathbb{R}$ are bounded and measurable with $m(E) < +\infty$. Then, the following properties hold.

1. Linearity. For $\alpha, \beta \in \mathbb{R}$,
$$\int_E (\alpha f + \beta g) = \alpha \int_E f + \beta \int_E g \tag{176}$$
2. Monotonicity.
$$f \leq g \implies \int_E f \leq \int_E g \tag{177}$$
3. Additivity. If $A \subset E$ is measurable $B = A \setminus E$, then
$$\int_E f = \int_A f + \int_B f \tag{178}$$

Proof. Listed.

1. For scalar multiplication, we just use $\alpha\phi, \alpha\psi$ in the simple function approximation. Now for the sums of functions, we can see that
$$f \leq \psi_1, g \leq \psi_2 \implies f + g \leq \psi_1 + \psi_2 \implies \int (f+g) \leq \int f + \int g \tag{179}$$
$$f \geq \psi_1, g \geq \psi_2 \implies f + g \geq \psi_1 + \psi_2 \implies \int (f+g) \geq \int f + \int g \tag{180}$$
$$\tag{181}$$
and by taking the limit as the simple functions approach f, g, we get equality.
2.
3. We can define $f_1 = f \cdot \chi_A, f_2 = f \cdot \chi_B$. f_1, f_2 are measurable and bounded (as the product of measurable functions) implying that
$$\int_E f = \int_E (f_1 + f_2) = \int_E f_1 + \int_E f_2 = \int_E f \cdot \chi_A + \int_E f \cdot \chi_B = \int_A f + \int_B f \tag{182}$$
where the final equality follows from the lemma.

Corollary 4.1 ()

Suppose $f : E \subset \mathbb{R} \to \mathbb{R}$ is bounded and measurable. Then

$$\left| \int_E f \right| \leq \int_E |f| \tag{183}$$

Proof. We know $-|f| \leq f \leq |f|$. By monotonicity, we have

$$\int -|f| \leq \int f \leq \int |f| \tag{184}$$

The next theorem is sort of like transferring convergence of functions to that of integrals. This results shouldn't be surprisingly since uniform convergence is so strong.

Theorem 4.6 (Uniform Convergence Implies Convergence of Integrals)

Suppose f_n are bounded, measurable, and $f_n \to f$ uniformly on E with $m(E) < +\infty$. Then,

$$\int_E f_n \to \int_E f \tag{185}$$

Proof. Since f is bounded, there exists $N \in \mathbb{N}$ s.t. $\|f_N - f\|_\infty \leq 1$. Then, by reverse triangle inequality (?), $\|f\|_\infty \leq \|f_N\|_\infty + 1$. Also, f is measurable as the limit of f_N. Given $\epsilon > 0$, find N_1 s.t. $\forall n \geq N_1$,

$$\|f_n - f\|_\infty \leq \frac{\epsilon}{m(E)} \implies \left| \int f - \int f_n \right| = \left| \int_E (f - f_n) \right| \leq \int_E |f - f_n| \leq \frac{\epsilon}{m(E)} \cdot m(E) = \epsilon \tag{186}$$

Naturally we might see if this results holds with weaker assumptions. Unfortunately, this is not true for pointwise convergence.

Example 4.3 (Integrals Don't Converge Under Pointwise Convergence)

Let $f_n(x) = n \cdot \chi_{[0,1/n]}(x)$, so $f_n(x) \to 0$ a.e. in $[0,1]$ but $\int f_n(x) = 1$ for all $n \in \mathbb{N}$.

But not all hope is lost. This is where the key theorems of measure theory comes in. The following theorem is more general, since uniform convergence implies uniformly bounded, but it isn't used in practice as much (according to Kiselev).

Theorem 4.7 (Bounded Convergence Theorem)

Suppose f_n are measurable, uniformly bounded.[a] Suppose $f_n \to f$ pointwise on E, with $m(E) < \infty$. Then,

$$\int_E f_n \to \int_E f \tag{187}$$

[a] $\|f_n\|_\infty \leq M < \infty$

Proof. The idea is to use Egorov to split the domain into F and $E \setminus F$. Over F, we can then use the fact that $f_n \to f$ uniformly, and over $E \setminus F$, we can control the non-uniformness with a small measure $m(E \setminus F)$.
f is bounded because f_n are uniformly bounded, and it is measurable since it's a limit of measurable functions. Fix $\epsilon > 0$. by Egoroff, $\exists$ closed $F \subset E$ s.t. $f_n \to f$ uniformly on F, and $m(E \setminus F) \leq \frac{\epsilon}{4M}$. Then,

$$\left| \int_E f - \int_E f_n \right| \leq \int_E |f - f_n| = \int_F |f - f_n| + \int_{E \setminus F} |f - f_n| \tag{188}$$

For the first term, $\exists N \in \mathbb{N}$ s.t. if $n \geq N$, then $|f_n(x) - f(x)| \leq \frac{\epsilon}{2m(F)} \leq \epsilon 2$ for all $x \in F$. For the second term, we can bound this by $2M \cdot \frac{\epsilon}{4M} \leq \frac{\epsilon}{2}$. Therefore, the whole expression $\leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$.

However, this assumption is still too strong for convergence and is therefore not used much in practice unlike other theorems (e.g. monotone convergence theorem and dominated convergence theorem). One nice part is that it works for unbounded functions, which is good since working with bounded functions is not the most natural assumption to have in practice. This is why we will step back and reconstruct the Lebesgue integral using positive—and possibly unbounded—simple functions (1st definition).

4.3 Positive Measurable Functions

The next natural step to generalize Lebesgue integral is to look at unbounded functions and/or infinite measures. In order to do this, we must start off by stepping back into only nonnegative functions first.

Unlike Riemann integration and Lebesgue integration of signed bounded functions, which looks at both the supremum and infimum of integrals of simple functions, Lebesgue integration of positive only looks at the supremum, given that f is nonnegative, so for all these f, the Lebesgue integral always exists.

Definition 4.3 (Lebesgue Integral on Positive Measurable Functions)

Let $f : E \subset \mathbb{R} \to \mathbb{R}$ satisfy $f \geq 0$, with E measurable (but not necessarily that $m(E) < +\infty$). Then, the **Lebesgue integral** is defined

$$\int_E f := \sup\left\{ \int_E h \;\middle|\; 0 \leq h \leq f, h \text{ measurable, bounded} \right\} \tag{189}$$

$$= \sup\left\{ \int_E \phi \;\middle|\; 0 \leq \phi \leq f, \phi \text{ simple} \right\} \tag{190}$$

A positive measurable function f is **Lebesgue integrable** if $\int f < +\infty$.

Proof. The only thing to show is that it suffices to check for simple functions only, so all we have to do is check the latter form.

Theorem 4.8 (Integral Properties for Nonnegative Measurable Functions)

Suppose nonnegative $f, g : E \subset \mathbb{R} \to \mathbb{R}$ are measurable. Then,

1. Linearity. For all $\alpha, \beta \geq 0$,[a]
$$\int_E (\alpha f + \beta g) = \alpha \int f + \beta \int g \tag{191}$$
2. Monotonicity.
$$f \leq g \implies \int f \leq \int g \tag{192}$$
3. Additivity. If A, B are disjoint with $E = A \cup B$, then
$$\int_E f = \int_A f + \int_B f \tag{193}$$

[a] Note that we have ≥ 0 since we are dealing with positive functions!

Proof. Listed.

1. We can easily show $\int \alpha f = \alpha \int f$ simply by multiplying the simple functions ϕ by α. For integrals of sums of functions, we show the following.

(a) $\int f + \int g \leq \int (f+g)$. Since given simple ϕ_1, ϕ_2 with $\phi_1 \leq f, \phi_2 \leq g$, we have $\phi_1 + \phi_2 \leq f+g$, and so by taking the supremum, we have the bound.

(b) $\int f + \int g \geq \int (f+g)$. Suppose h is simple s.t. $0 \leq h \leq f+g$. Define[a]

$$\ell \coloneqq \min\{h, f\}, \qquad k \coloneqq h - \ell \tag{194}$$

Note that ℓ, k are both measurable. Furthermore, ℓ is at most h (which is simple) and so is bounded, and k is also bounded since it is bounded h minus some nonnegative ℓ that is at most h. Therefore, we invoke our previous definition of the Lebesgue integral for bounded functions to get

$$\int k + \int \ell = \int h \tag{195}$$

Since $\ell \leq f$ and $k = h - \ell \leq g$, by monotonicity we have $\int k \leq \int g$ and $\int \ell \leq \int f$. Substituting this in gives

$$\int h \leq \int f + \int g \implies \int (f+g) = \sup_h \int h \leq \int f + \int g \tag{196}$$

2. Direct.
3.

[a]Intuitively, we are trying to split the $f+g$ into ℓ and k, such that $\ell + k = h \leq f+g$.

Lemma 4.2 (Chebyshev)

Suppose $f : E \subset \mathbb{R} \to \mathbb{R}$, $f \geq 0$, and f is measurable. Then for all $\alpha > 0$,[a]

$$m(\{x \in E \mid f(x) \geq \alpha\}) \leq \frac{1}{\alpha} \int_E f \tag{197}$$

[a]In probability, this refers to a bound on the variance of a random variable, but in analysis, it seems to be a more general result of the bounds of form $\int |f|^p$.

Proof. Let us call the set on the LHS E_α. Then, E_α is measurable and define $\phi(x) = \alpha \chi_{E_\alpha}(x)$. Then, $\phi(x) \leq f(x)$, and so

$$\int f \geq \int \phi = \alpha m(E_\alpha) \tag{198}$$

Theorem 4.9 (Vanishing Integral iff a.e. Vanishing Nonnegative Function)

Suppose $f \geq 0$ on E. Then,

$$\int_E f = 0 \iff f = 0 \text{ a.e. on } E \tag{199}$$

Proof. We prove bidirectionally.

1. $(\rightarrow)$. By Chebyshev, we see that

$$m(\underbrace{\{x \in E \mid f(x) \geq 1/n\}}_{E_n}) \leq n \int f = 0 \tag{200}$$

for all $n \in \mathbb{N}$. But $\{x \in E \mid f(x) > 0\} = \cup_{n=1}^\infty E_n$. By countable additivity, $m(\{x \in E \mid f(x) > 0\}) = 0$.

2. $(\leftarrow)$. Since $f = 0$ a.e., any $0 \leq \phi \leq f$—ϕ simple—will satisfy $m(\{x \in E \mid \phi(x) = 0\}) = 0$, but it

will never get off 0. Therefore, $\int_E \phi = 0$.

Theorem 4.10 (Integrable Functions can be Infinite on at Most 0 Measure Set)

If f is integrable on E, then $f(x) < +\infty$ a.e. on E.

Proof. By Chebyshev,

$$m(\{x \in E \mid f(x) \geq n\}) \leq \frac{1}{n} \int_E f \tag{201}$$

As $n \to \infty$, the LHS is bounded above by all $\epsilon > 0$.

$$m(\{x \in E \mid f(x) = \infty\}) = m\bigg(\bigcap_{n=1}^\infty E_n\bigg) \tag{202}$$

Lemma 4.3 (Fatou's Lemma)

Let $f_n \geq 0$ be measurable, and it converges to f. Then,

$$\int_E \liminf f_n \leq \liminf \int_E f_n, \qquad \int_E f \leq \liminf \int_E f_n \tag{203}$$

Proof. Here is a proof that doesn't require MCT. The idea is that we want to use bounded convergence theorem to do the work for us. The f_n's are all measurable already, but we are missing two things. First, the measure of the total space may not be finite. Second we don't have a uniformly bounded sequence of functions. We can solve both of these by appealing to the definition of the integral.
It suffices to show that if h is any bounded measurable function of finite support s.t. $0 \leq h \leq f$, then

$$\int_E h \leq \liminf_{n \to \infty} \int_E f_n \tag{204}$$

since taking the supremum w.r.t. h gives the result. The h being bounded gives us a start, so let it be bounded by M. As for the finite measure, we know that h has finite support, so define

$$E_0 = \{x \in E \mid h(x) \neq 0\} \tag{205}$$

and we have $m(E_0) < +\infty$. Now we must construct a sequence of uniformly bounded functions. Define the measurable functions

$$h_n \coloneqq \min\{h, f_n\} \tag{206}$$

Since $0 \leq h_n \leq M$ on E_0 and $h_n = 0$ on $E \setminus E_0$, we know that h_n is uniformly bounded on E_0. For each $x \in E$, we know that

$$h(x) \leq f(x), f_n(x) \to f(x) \implies h_n(x) \to h(x) \tag{207}$$

So we actually know what the (h_n) converges to. Therefore, we finally invoke BCT

$$\lim_{n \to \infty} \int_E h_n = \lim_{n \to \infty} \int_{E_0} h_n = \int_{E_0} h = \int_E h \tag{208}$$

Now we use the fact that $h_n \leq f_n$ on E to show that $\int_E h_n \leq \int_E f_n$, and by taking the liminf,

$$\int_E h = \lim_{n \to \infty} \int_E h_n = \liminf_{n \to \infty} \int_E h_n \leq \liminf_{n \to \infty} \int_E f_n \tag{209}$$

Proof. Define $g_k(x) = \int_{j \geq k} f_j(x)$. Note that $g_k(x) \leq f_k(x)$ by definition, and $g_k(x)$ is increasing. Since by definition $\lim_{k\to\infty} g_k(x) = \liminf f_k(x)$, by MCT,

$$\lim_{k\to\infty} \int_E g_k = \int \liminf f_k \tag{210}$$

But since $\int g_k \leq \int f_k$ for all k,

$$\liminf \int f_k \geq \int \liminf f_k \tag{211}$$

Note that equality does not have to hold. For example, take $f_n(x) = \chi_{[n,n+1]}(x)$.

We know that the integral doesn't behave well under weak kinds of convergence, e.g. pointwise or a.e. The following theorem gives us a slightly stronger assumption.

Theorem 4.11 (Monotone Convergene Theorem (MCT))

Given a nondecreasing sequence of nonnegative measurable functions $f_1 \leq f_2 \leq f_3 \leq \ldots : E \subset \mathbb{R} \longrightarrow [0, +\infty]$, its limit $\lim_{n\to\infty} f_n$ always exists[a], is measurable, and

$$\int_E \lim_{k\to\infty} f_k = \lim_{k\to\infty} \int_E f_k \tag{212}$$

[a]since for every $x \in E$, $f_n(x)$ is a monotonic sequence in $[0, +\infty]$, which is guaranteed to converge.

Proof. Set the RHS $\lim_{n\to\infty} \int f_n = \alpha$ (could be infinity) and let $\lim_{n\to\infty} f_n = f$. $f(x) \geq f_n(x)$ for all x, n, so $\int f \geq \alpha$. Consider $0 \leq \phi \leq f$, ϕ simple. Take $0 < c < 1$[a] and define

$$E_n = \{x \mid f_n(x) \geq c\phi(x)\} \tag{213}$$

Then, the E_n are increasing and $\cup_{n=1}^\infty E_n = E$.[b] Observe that

$$\int_{E_n} f_n \geq c \int_{E_n} \phi \tag{214}$$

Suppose that $\phi(x) = \sum_{k=1}^M a_k \chi_{F_k}(x)$. Then,

$$\int_{E_n} \phi = \sum_{k=1}^M a_k m(F_k \cap E_n) \to \sum_{k=1}^M a_k m(F_k) = \int_E \phi \text{ as } n \to +\infty \tag{215}$$

Therefore, by taking the limit of 214 as $n \to \infty$, we get

$$\lim_{n\to\infty} \int_E f_n \geq c \int_E \phi \quad \forall c < 1 \implies \text{ also true for } c = 1 \tag{216}$$

Since $\phi \leq f$ is arbitrary, just take the supremum over all ϕ and we get $\lim_{n\to\infty} \int f_n \geq \int f$.

[a]We introduce the c because we can then claim that the union of E_n is E.
[b]This may not be true is $c = 1$.

Here is a variant of MCT.

Lemma 4.4 (Levi's Lemma)

Suppose $f_n \geq 0$ is increasing and $\int_E f_n \leq M$ for all n. Then,
1. $f(x) = \lim_{n\to\infty} f_n(x)$ is integrable, and
2. $\int f \leq M$.

Proof. By MCT, $\int_E f_n \to \int_E f$.

The huge problem with Riemann integrals is that this theorem doesn't hold, but it is the case for Lebesgue integration. In practice, the MCT is not as useful, but the following is.

4.4 Lebesgue Integral for Signed Functions

Definition 4.4 (Lebesgue Integral for Signed Functions)

Given $f : E \subset \mathbb{R} \to \mathbb{R} \cup \{\pm\infty\}$, let us write

$$f = f^+ - f^-, \qquad f^+ := \max\{0, f\}, f^- := \max\{-f, 0\} \tag{217}$$

f is **Lebesgue integrable** iff f^+ and f^- and we define the **Lebesgue integral** of f as

$$\int f := \int f^+ - \int f^- \tag{218}$$

given that at least one of these integrals is finite. If one is infinite and the other is finite, then we can call it infinite. If we have *both* infinite integrals, then the integral doesn't exist.

If f^+, f^- are Lebesgue integrable, then it can only be $\pm\infty$ on a set of measure 0, so it generally won't affect anything. Also, there is a difference between the *existence* of the integral and a function *being* Riemann-integrable.

Since $|f| = f^+ + f^-$, f is also Lebesgue integrable if

$$\int |f| \, d\mu < \infty \tag{219}$$

since by triangle inequality, we have

$$\left| \int f \, d\mu \right| \leq \int |f| \, d\mu \tag{220}$$

Theorem 4.12 ()

$f^\pm$ integrable iff $|f|$ integrable.

Proof. $|f| = f^+ + f^-$.

Theorem 4.13 (Integral Properties of Signed Measurable Functions)

Suppose $f, g : E \subset \mathbb{R} \to \mathbb{R}$ are measurable. Then,
1. Linearity. For all $\alpha, \beta \in \mathbb{R}$,

$$\int_E (\alpha f + \beta g) = \alpha \int f + \beta \int g \tag{221}$$

2. Monotonicity.

$$f \leq g \implies \int f \leq \int g \tag{222}$$

3. Additivity. If A, B are disjoint with $E = A \cup B$, then

$$\int_E f = \int_A f + \int_B f \tag{223}$$

Proof. Listed. As always, linearity is nontrivial.

1. For scalar multiplication, we divide into 2 cases.
   (a) $\alpha > 0$. Then, $(\alpha f)^+ = \alpha f^+$ and $(\alpha f)^- = \alpha f^-$, so

   $$\int (\alpha f) = \int \alpha f^+ - \int \alpha f^- = \alpha \int f \tag{224}$$

   (b) $\alpha < 0$. Then, $(\alpha f)^+ = -\alpha f^+$ and $(\alpha f)^- = \alpha f^+$, so

   $$\int (\alpha f) = \int -\alpha f^- + \int \alpha f^+ = \alpha \bigg(\int f^+ - \int f^- \bigg) = \alpha \int f \tag{225}$$

2. For addition, note that $|f + g| \leq |f| + |g|$, so it is integrable. So

   $$\int (f + g) = \int (f + g)^+ - \int (f + g)^- \tag{226}$$

   Now observe that

   $$(f + g)^+ - (f + g)^- = f + g = f^+ + g^+ - f^- - g^- \tag{227}$$
   $$(f + g)^+ + f^- + g^- = (f + g)^- + f^+ + g^+ \tag{228}$$

   Since it is nonnegative, our previous properties of linearity holds, and then we can just rearrange.
3.
4.

Theorem 4.14 (Dominated Convergence Theorem)

Let (f_n) be a sequence of measurable functions on E. Suppose that there is a function g that is integrable over E and dominates (f_n) on E in the sense that

$$|f_n| \leq g \text{ on } E \quad \forall n \tag{229}$$

Then, if $f_n \to f$ pointwise a.e. on E, then f is integrable over E and

$$\lim_{n \to \infty} \int_E f_n = \int_E f \tag{230}$$

Proof.

Definition 4.5 (Normed Vector Space of Lebesgue Integrable Functions)

The set of all functions $f : (X, \mathcal{A}, \mu) \longrightarrow \mathbb{R}$ that are Lebesgue integrable is denoted $\mathcal{L}^1(X, \mathcal{A}, \mu; \mathbb{R})$, or for short $\mathcal{L}^1(X, \mathcal{A}, \mu)$.

Theorem 4.15 ()

Suppose $f : (\mathbb{R}, \mathcal{A}, \mu) \longrightarrow \mathbb{R}$ is 0 almost everywhere. Then f is Lebesgue integrable with

$$\int_{\mathbb{R}} f \, d\mu = 0 \tag{231}$$

If $g : \mathbb{R} \longrightarrow \mathbb{R}$ is such that $f = g$ μ-almost everywhere, then

$$\int_{\mathbb{R}} f \, d\mu = \int_{\mathbb{R}} g \, d\mu \tag{232}$$

4.5 Uniformly Integrable Functions

The following theorem gives an *almost* equivalent condition of integrability.

Theorem 4.16 (Conditions for Integrability)

Let f be measurable on E.

1. If f is integrable over E, then $\forall \epsilon > 0$, $\exists \delta > 0$ s.t. if $m(A) < \delta$, $A \subset E$, then $\int_A |f| < \epsilon$.
2. Conversely, if $m(E) < +\infty$, then the converse holds.

Proof. We prove bidirectionally.

1. $(\rightarrow)$. The idea is that first, if f is bounded, then the result is trivial since we can choose $\delta = \frac{\epsilon}{M}$, where M is bounded. Otherwise, we want to choose a nice approximation f_ϵ s.t.
$$0 \leq f_\epsilon \leq |f|, \qquad \int_E |f| - \int f_\epsilon \leq \frac{\epsilon}{2} \tag{233}$$
Maybe take $f_\epsilon = \min\{f, n\}$, which is bounded, and by MCT, we can prove it. Or choose a simple function.
2. $(\leftarrow)$. There are two steps. If E is bounded, then say it is in $[-m, m]$, and divide it into subintervals each $< \delta$. If not, then take increasing intervals s.t. $m(E \setminus [-n, n]) < \delta$ (?).

Now we state the family analogue of integrability.

Definition 4.6 (Uniformly Integrable)

A family $\mathscr{F}$ of measurable functions is **uniformly integrable (u.i.) over** E—also called **equi-integrable**—iff $\forall \epsilon > 0$, $\exists \delta > 0$ s.t.

$$A \subseteq E, m(A) < \delta \implies \int_A f < \epsilon \quad \forall f \in \mathscr{F} \tag{234}$$

Example 4.4 (Dominated Family of Integrable Function is UI)

Fix some integrable g, and consider the uniformly integrable family

$$\mathscr{F} := \{f \mid |f| \leq g\} \tag{235}$$

But not every function in a uniformly integrable family is integrable due to the extra awkward $m(E) < +\infty$ assumption.

Example 4.5 (Finite Family of Integrable Functions is UI)

A finite family of integrable functions $\{f_1, \ldots, f_n\}$ is always uniformly integrable, since we can take $\delta = \min_{1 \leq i \leq n}\{\delta_i\}$, which will satisfy the first theorem.

Example 4.6 (Family of Integrable Functions May not be UI)

Even if all functions are integrable, the family may not be.

1. Consider
$$\{f_n \coloneqq n \cdot \chi_{[0,1/n]}\}_{n \in \mathbb{N}} \tag{236}$$
Take $\epsilon > 1/2$ over $E = [-1, 1]$.
2. Consider
$$\{f_n \coloneqq \min\{|x|^{-1}, n\}\}_n \tag{237}$$
which is not UI.

This is useful for showing convergence of integrals.

Lemma 4.5 ()

Assume $m(E) < +\infty$.[a] Let $\{f_n : E \to \mathbb{R}\}$ be u.i. over E and $f_n \to f$ a.e. on E. Then, f is integrable.

[a] This is to avoid the awkward assumption above.

Proof. It is similar to the proof for ϵ-δ criterion. Fix $\epsilon \geq 1$. By definition of u.i., we can find corresponding δ. Now split E into finite union of disjoint E_j.

$$E = \bigsqcup_{j=1}^{N} E_j \qquad m(E_j) < \delta \tag{238}$$

Then, $\int_E |f_n| \leq N$, so N is independent of n. By Fatou,

$$\int_E |f| \leq \liminf_{n\to\infty} \int_E |f_n| \leq N \tag{239}$$

The following convergence theorem gives the same conclusion in MCT or DCT, but under weaker assumptions (though now we must assue that the measure of the whole set is finite).

Theorem 4.17 (Vitali Convergence Theorem)

Assume $m(E) < +\infty$. Let $\{f_n : E \to \mathbb{R}\}$ be u.i. over E and $f_n \to f$ a.e. on E. Then,

$$\lim_{n\to\infty} \int_E f_n = \int f, \qquad \lim_{n\to\infty} \int_E |f_n - f| = 0 \tag{240}$$

Proof. Note that the proof of the previous lemma tells us that f is integrable only if $\int |f_n| \leq C$ and $f_n \to f$ a.e. Equality does not hold under weaker conditions, since

$$f_n(x) = n\chi_{[0,1/n]}(x), f_n \to 0 \text{ a.e.} \tag{241}$$

but $\int f_n = 1$ and $\int f = 0$. So this is why we need integrability.
Fix $\epsilon > 0$ (by definition of u.i.) s.t. if $m(A) < \delta$, $\int_A |f_n| < \frac{\epsilon}{3}$. By Fatou, $\int_A |f| < \frac{\epsilon}{3}$. Now by Egorov,

since $m(E) < +\infty$, $f_n \to f$ a.e. We can find some set E_0 s.t. $m(E_0) < \delta$ and $f_n \to f$ uniformly on $E \setminus E_0$.
Choose N s.t. if $n > N$,

$$|f_n(x) - f(x)| < \frac{\epsilon}{3m(E)} \quad \forall x \in E \setminus E_0 \tag{242}$$

Then, if $n > N$,

$$\int_E |f_n - f| = \int_{E_0} |f_n - f| + \int_{E \setminus E_0} |f_n - f| \tag{243}$$

We know

1. For the first term, we know that the measure is small, so we use uniform integrability.

$$\int_{E_0} |f_n - f| \leq \int_{E_0} |f_n| + \int_{E_0} |f| \leq \frac{\epsilon}{3} + \frac{\epsilon}{3} \tag{244}$$

2. For the second term, we know that at each point, $|f_n - f|$ is bounded by $\frac{\epsilon}{3m(E)}$. We use uniform convergence to bound it.

$$\int_{E \setminus E_0} |f_n - f| \leq \frac{\epsilon}{3m(E)} \cdot m(E) \tag{245}$$

What is great about Vitali is that this conclusion is essentially “sharp.” Let’s make this more rigorous.

Theorem 4.18 ()

If $m(E) < +\infty$, $h_n \geq 0$ is integrable, and $h_n \to 0$ a.e. on E, then

$$\lim_{n\to\infty} \int_E h_n = 0 \iff (h_n) \text{ is uniformly integrable over } E \tag{246}$$

Proof. For the backwards implication, this is Vitali. For the forward implication, fix $\epsilon \geq 0$. Choose $N \in \mathbb{N}$ large s.t.

$$\int_E h_n < \epsilon \quad \forall n \geq N \tag{247}$$

Note that $\{h_1, \ldots, h_n\}$ is a finite family, which we showed was u.i. So by definition $\exists \delta > 0$ s.t. if $m(A) < \delta$,

$$\int_A h_n < \epsilon \quad \forall n \in \{1, \ldots, N\} \tag{248}$$

But $\int_A h_n < \epsilon$ is also true for $n \geq N + 1$.

So we *need* the uniformly integrability condition for Vitali. We might ask if $m(E) < +\infty$ can be taken out, but in general it cannot. It is necessary because you can consider $f_n(x) = \chi_{[n,n+1]}(x)$ with $E = \mathbb{R}$. So a natural question is to ask: what should we add in order to replace $m(E) < +\infty$?

Lemma 4.6 ()

Suppose f is integrable over set E. Then $\forall \epsilon > 0$, $\exists E_0 \subset E$ with $m(E_0) < +\infty$ s.t.

$$\int_{E \setminus E_0} |f| < \epsilon \tag{249}$$

Colloquially, the integral should be “concentrated” over a finite measure set, even if the space is infinite.

Proof. We again approximate by bounded functions. By definition of integrability, $\exists g$ bounded, with compact support (since integral is finite), s.t.

$$0 \leq g \leq |f|, \qquad \int_E |f| - \int_E g < \epsilon \tag{250}$$

Then, just take $E_0 = \text{supp}(g)$, and split the integral to write the inequality as

$$\int_{E \setminus E_0} |f| + \underbrace{\int_{E_0} |f| - \int_E g|}_{\geq 0 \text{ since } 0 \leq g \leq |f|} < \epsilon \implies \int_{E \setminus E_0} |f| < \epsilon \tag{251}$$

This lemma tells you that if you have even 1 integrable function, you can find that the integral is concentrated around a set of finite measure. But this is for 1 function only, and we want to do it for a *sequence* of functions.

Definition 4.7 (Tight)

A family $\mathscr{F}$ of measurable functions on E is **tight over** E if $\forall \epsilon > 0$, $\exists E_0 \subset E$ with $m(E_0) < +\infty$ s.t.

$$\int_{E \setminus E_0} |f| < \epsilon \quad \forall f \in \mathscr{F} \tag{252}$$

This is a more general condition since if $m(E) < +\infty$, then we can just take $E_0 = E$ and it is automatically tight. Therefore, we can replace the finite measure assumption with tightness.

Theorem 4.19 (Generalized Vitali Convergence Theorem)

Let (f_n) be a sequence of functions on E that is u.i. and tight, with $f_n \to f$ a.e. on E.

$$\lim_{n \to \infty} \int_E f_n = \int f, \qquad \lim_{n \to \infty} \int_E |f_n - f| = 0 \tag{253}$$

Proof. Fix $\epsilon > 0$. By tightness, $\exists E_0 \subset E$ with $m(E_0) < +\infty$ s.t.

$$\int_{E \setminus E_0} |f_n| < \frac{\epsilon}{4} \quad \forall n \in \mathbb{N} \tag{254}$$

By Fatou, $\int_{E \setminus E_0} |f| < \frac{\epsilon}{4}$. Now combine both inequalities and triangle inequality to get

$$\int_{E \setminus E_0} |f_n - f| < \int_{E \setminus E_0} (|f_n| + |f|) < \frac{\epsilon}{2} \tag{255}$$

Now $m(E_0) < +\infty$ and (f_n) is u.i. on E_0. By VCT, $\exists N \in \mathbb{N}$ s.t.

$$n \geq N \implies \int_{E_0} |f_n - f| < \frac{\epsilon}{2} \tag{256}$$

Now combine 255 and 256 to get

$$\int_E |f_n - f| = \int_{E \setminus E_0} |f_n - f| + \int_{E_0} |f_n - f| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon \tag{257}$$

Note that the first term $\int_{E \setminus E_0} |f_n - f|$ has an infinite measure, though the integral itself is small. The second is over a finite measure set, which we can use VCT.

But this is not a complete generalization, since in VCT we don't assume f_n is integrable, but generalized VCT assumes uniform integrability.

5 Convergence

Definition 5.1 (Convergence in Measure)

Let (f_n) be a sequence of measurable and finite[a] a.e. $f_n \to f$ **in measure** if for every $\eta > 0$,

$$\lim_{n\to\infty} m\big(\{x \mid |f_n(x) - f(x)| > \eta\}\big) = 0 \tag{258}$$

Colloquially, the set over which f_n and f differ too much is small.

[a]We add finite condition since we avoid dealing with $+\infty - \infty$.

So we have 3 types of convergence: uniform convergence, a.e. convergence, and now convergence in measure. Now we want to relate this convergence to the ones we already have.

Theorem 5.1 ()

Suppose E is measurable, $m(E) < +\infty$, and $f_n \to f$ a.e. in E (assume f_n all measurable). Then , $f_n \to f$ in measure.

Proof. Observe that if $f_n \to f$ uniformly, then it converges in measure, because given some $\eta > 0$, $\exists N$ s.t.

$$\{x \mid |f_n(x) - f(x)| > \eta\} = \emptyset \tag{259}$$

by definition. It doesn't go to 0; it is 0. You can guess why we started with this, because now we can directly use Egorov's theorem. Fix any $\epsilon > 0$. Find $E_0 \subset E$ s.t. $m(E \setminus E_0) < \epsilon$, and $f_n \to f$ uniformly on E_0. It follows that for all $\eta > 0$,

$$m(\{x \mid |f_n(x) - f(x)| > \eta\}) \leq \epsilon \tag{260}$$

for all $n \geq N(\eta)$. Since this is true for every $\epsilon > 0$, so this implies

$$\lim_{n\to\infty} m(\{x \mid |f_n(x) - f(x)| > \eta\}) = 0 \tag{261}$$

for every $\eta > 0$.

A few remarks. First, if the measure of E is infinite, this need not be true. Consider $f_n(x) = \chi_{[n,n+1]}(x)$. Then, this converges to 0 pointwise, but it does not converge to 0 in measure. There is always a measure 1 set where f is 1. Where the proof breaks down is in Egorov's theorem, since it does not work when $m(E) = +\infty$.

The second remark is that the converse is not true. Consider $[0, 1]$ and the sequence of functions

$$\chi_{[0,1/2]}, \chi_{[1/2,1]}, \chi_{[0,1/4]}, \chi_{[1/4,1/2]}, \chi_{[1/2,3/4]}, \ldots \tag{262}$$

Then $f_n \to 0$ in measure since the size shrinks at the rate of 2^{-n}. However, it doesn't converge a.e. since for any point $x \in [0, 1]$, the function will be 1 eventually as we hit the subinterval containing x, like "waves." So $f_n(x)$ diverges for all $x \in [0, 1]$. So indeed, convergence in measure is the weakest type of convergence.

Here is a sort-of converse.

Theorem 5.2 (Riesz)

Suppose $f_n \to f$ in measure. Then, there exists a subsequence $f_{n_k} \to f$ a.e.

Proof. For every k, find n_k s.t. for all $n \geq n_k$,

$$m(\underbrace{\{x \mid |f_n(x) - f(x)| > 1/k\}}_{E_k}) < 2^{-k} \tag{263}$$

Then,

$$\sum_{k=1}^{\infty} m(E_k) < +\infty \tag{264}$$

By Borel-Cantelli, the set of all x's that are in infinitely many E_k have measure 0. So, almost everywhere, x is only in a finite number of E_k. So for a.e., x, there exists $N(x)$ s.t. $x \notin E_k$ for all $k \geq N(x)$. This means

$$|f_{n_k}(x) - f(x)| < 1/k \tag{265}$$

for all $k \geq N(x)$. Therefore, $f_{n_k}(x) \to f(x)$ for a.e. x.

In the example above, we can just skip the functions that evaluate x to 1.

Practically, proving convergence in measure is still pretty good since we can pass in a subsequence that converges a.e. Here is a corollary.

Corollary 5.1 ()

Let $f_n \geq 0$, integrable on E. Then,

$$\lim_{n \to +\infty} \int_E f_n \, dx = 0 \iff f_n \to 0 \text{ in measure} \tag{266}$$

f_n are tight and uniformly integrable.

Proof. We prove bidirectionally.

1. $(\rightarrow)$. Tight, uniformly integrable is true be definition. Also, $f_n \to 0$ in measure by Chebyshev.

$$m(\{x \mid f_n(x) > \eta\}) \leq \frac{1}{\eta} \int_E f_n \, dx \tag{267}$$

2. $(\leftarrow)$ For the opposite, we use the previous theorem. Find f_{n_k} s.t. that it converges to 0 a.e., and then use Vitali's convergence theorem.

In general, if $f_n \to 0$ in measure, it doesn't mean that the integral will go to 0 since you can take larger and larger bumps. So we need extra assumptions.

Lemma 5.1 ()

Suppose f is bounded, and there exists measurable sequences of functions ϕ_n, ψ_n s.t.

$$\psi_n(x) \leq f(x) \leq \psi_n(x) \quad \forall x \in E \tag{268}$$

and

$$\lim_{n \to +\infty} \int_E (\psi_n - \phi_n) = 0 \tag{269}$$

Then, there exists $\tilde{phi}_n \to f$ and $\tilde{psi}_n \to f$ a.e.

Proof. Define

$$\tilde{\phi}_n(x) = \max\{\phi_1(x), \ldots, \phi_n(x)\}, \quad \tilde{\psi}_n(x) = \min\{\psi_1(x), \ldots, \psi_n(x)\} \tag{270}$$

We still have $\tilde{\phi}_n(x) \leq f(x) \leq \tilde{\phi}_n(x)$ for all n and for all x. Also, $\tilde{\phi}_n(x)$ is increasing, $\tilde{\psi}_n(x)$ is decreasing. Now define

$$\phi^*(x) \coloneqq \lim_{n\to\infty} \tilde{\phi}_n(x), \qquad \psi^*(x) \coloneqq \lim_{n\to\infty} \tilde{\psi}_n(x) \tag{271}$$

Observe that

$$\int (\tilde{\psi}_n - \tilde{\phi}_n) \leq \int (\psi_n - \phi_n) \implies \int (\tilde{\psi}_n - \tilde{\phi}_n) \to 0 \text{ as } n \to \infty \tag{272}$$

Also,

$$\int (\underbrace{\psi^*(x) - \phi^*(x)}_{\geq 0})\, dx \leq \int (\tilde{\psi}^* - \tilde{\phi}^*) \tag{273}$$

for all n. Therefore,

$$\int (\psi^*(x) - \phi^*(x)) = 0 \implies \psi^*(x) = \phi^*(x) \text{ a.e.} \tag{274}$$

And so $f(x)$, which is sandwiched between ψ^* and ϕ^*, must be equal a.e. We didn't assume that f was measurable, but these ψ_n, ϕ_n is measurable by assumption.

Now, we can prove this master theorem.

Theorem 5.3 (Characterization of Lebesgue Integrability)

Let f be bounded on measurable set E of finite measure. Then f is Lebesgue integrable iff f is measurable.

Proof. The backward implication is true in general. We want to show that f is measurable. Recall that for bounded functions, we defined Lebesgue integrals with $\underline{L}f$ and $\overline{L}f$. Therefore, we can find simple ϕ_n, ψ_n s.t. $\phi_n \leq f \leq \psi_n$, and $\int \psi_n - \int \phi_n \leq 1/n$. Now we are exactly in the setting of the lemma, and so by the lemma, we can find measurable $\tilde{\psi}_n(x) \to f$ a.e. (in fact, $\tilde{\psi}$ will be simple). Since the limit of measurable functions is measurable, f is measurable.

This is a very reasonable criterion, and you can't really hope for more then Lebesgue measurability. This following theorem on Riemmann integrability is much more restrictive, while for above, measurable functions can be very wild.

Theorem 5.4 (Characterization of Riemann Integrability)

f is Riemann integrable on $[a, b]$ if the set of its discontinuities has measure 0.

Proof. Not stated. In book.

Oct 8. Now we will build the theory of differentiation on monotone functions.

Definition 5.2 ()

Given a set E, a collection $\mathcal{F}$ **covers** E **in Vitali sense** if $\forall x \in E, \forall \epsilon > 0$, there exists $I \in \mathcal{F}$ s.t. $x \in I$, $\ell(I) < \epsilon$.

Note that we don't assume that E is measurable. It is easy to see that a Vitali set can be uncountable (all subintervals of $[0,1]$) and even countable (all subintervals with rational endpoints). Nevertheless, we can still select a finite set of intervals that almost covers E.

Lemma 5.2 (Vitali Covering Lemma)

Suppose $m^*(E) < +\infty$ and $\mathcal{F}$ covers E in Vitali sense. Then, $\forall \epsilon > 0$, $\exists$ a disjoint finite collection $I_1, I_2, \ldots, I_n$ of intervals from $\mathcal{F}$ s.t.

$$m^*\left(E \setminus \bigcup_{k=1}^{n} I_k\right) < \epsilon \tag{275}$$

Proof. Since $m^*(E) < +\infty$, by definition $\exists$ open O s.t. $E \subset O$, $m(O) < +\infty$. WLOG we can assume that all intervals in $\mathcal{F}$ lie in O.[a]
Note two things.

1. If $I_1, I_2, \ldots, I_n$ are disjoint and belong to O, then $\sum_{k=1}^{n} \ell(I_k) < +\infty$ since it is less than the measure of O which is finite.
2. Second, if we have finite collection $\{I_k\}_{k=1}^{n} \in \mathcal{F}$, define

$$\mathcal{F}_n \coloneqq \{I \in \mathcal{F} \mid I \cap \bigcup_{k=1}^{n} I_k = \emptyset\} \tag{276}$$

Then every $x \in E \setminus \bigcup_{k=1}^{n} I_k$ lies in some $I \in \mathcal{F}_n$.

The ideal is to define $I_1, \ldots, I_n \in \mathcal{F}$ s.t. they are disjiont and

$$E \setminus \bigcup_{k=1}^{n} I_k \subset \bigcup_{k=n+1}^{\infty} 5I_k \quad \forall n \tag{277}$$

where $5I$ means that we keep the center of the interval fixed and scale it up by 5 times. If we do that, then $\forall \epsilon > 0$, find n s.t. $\sum_{k=n+1}^{\infty} \ell(I_k) < \epsilon/5$. Take $I_1, \ldots, I_n$ as our intervals

$$m^*\left(E \setminus \bigcup_{k=1}^{n} I_k\right) \leq \sum_{k=n+1}^{\infty} \ell(5I_k) < \epsilon \tag{278}$$

So it remains to select these intervals $I_1, \ldots, I_n$. We will do this inductively.

1. I_1 be any interval in $\mathcal{F}$ s.t.

$$\ell(I_1) \geq \frac{1}{2} \sup_{I \in \mathcal{F}} \ell(I) \tag{279}$$

2. Once $I_1, \ldots, I_n$ have been selected, we select I_{n+1} from $\mathcal{F}_n$ s.t.

$$\ell(I_{n+1}) \geq \frac{1}{2} \sup_{I \in \mathcal{F}_n} \ell(I) \tag{280}$$

So these intervals are clearly disjoint from the ones that we have selected earlier. So it remains to show 277. Suppose $x \in E \setminus \cup_{k=1}^{n} I_k$. Then, $\exists I \in \mathcal{F}_n$ s.t. $x \in I$.

Suppose $I \in \mathcal{F}_m$ for all $m \geq n$. This is impossible since by construction, $\ell(I_m) \geq \frac{1}{2}\ell(I)$. This contradicts $\sum \ell(I_m)$ is finite. Therefore, $\exists m$ s.t. $I \in \mathcal{F}_{m-1}$ but $I \notin \mathcal{F}_m$. This implies that $I \cap I_m \neq$ (while the intersection with the previous ones were empty). But then, $I \subset 5I_m$, since $\ell(I_m) \geq \frac{1}{2}\ell(I)$.[b]

[a] We can just discard any interval that it not Vitali in O and keep only those intervals in O such that it would still be in a Vitali cover. Indeed, we can discard all $I \subset \mathcal{F}$ s.t. $I \not\subset O$. Given $x \in E, x \in O$, so $d(x, O^c) > 0$ for all $\epsilon > 0$, $\exists I \in F$ s.t. $\ell(I) < \epsilon$, $x \in I, I \subset O$ (just take $\ell(I), < \min(\epsilon, d(x, O^c))$). So even remaining intervals cover E in Vitali sense.

[b] The 5 is needed since we have $1/2$. So we are taking the midpoint $3/4$ of the interval $[1/2, 1] \subset [0, 1]$, which should be blown up by 5.

This may be a heavy proof, but this lemma seems to be very convenient.

Definition 5.3 (Derivative)

Given any f and x in the interior of its domain, we can define the **upper and lower derivative** as

$$\overline{D}f(x) := \lim_{h \to 0} \sup_{0<|t|<h} \frac{f(x+t) - f(x)}{t}, \qquad \underline{D}f(x) := \lim_{h \to 0} \inf_{0<|t|<h} \frac{f(x+t) - f(x)}{t} \tag{281}$$

If they are equal, then we can define the **derivative** as either one, and we say f is **differentiable** at x.

Note that as h goes to 0, the first is nondecreasing and the second is nonincreasing, and clearly

$$\underline{D}f(x) \leq \overline{D}f(x) \tag{282}$$

Lemma 5.3 ()

Suppose f is increasing on $[a, b]$. Then $\forall \alpha > 0$, then,

$$m^*\{x \mid \overline{D}f(x) \geq \alpha\} \leq \frac{1}{\alpha}\big(f(b) - f(a)\big) \tag{283}$$

and

$$m^*\{x \mid \overline{D}f(x) = \infty\} = 0 \tag{284}$$

Proof. Fix $\alpha > 0$, define $E_\alpha = \{x \mid \overline{D}f(x) \geq \alpha\}$. Take any $\alpha' < \alpha$, any $\epsilon > 0$. Consider all intervals $[c, d] \subset [a, b]$ s.t. $f(d) - f(c) > \alpha'(d - c)$. This collection covers E_α in Vitali sense. Since no matter how small h is, we can find t so that this ratio term is bigger than α'.
Now, we can use the covering lemma to find a finite disjoint collection $\{[c_k, d_k]\}_{k=1}^n$ s.t. $m^*(E \setminus \cup_{k=1}^n [c_k, d_k]) < \epsilon$. Then,

$$m^*(E) \leq \sum_{k=1}^n (d_k - c_k) + \epsilon \tag{285}$$

by subadditivity of outer measure. Using the inequality,

$$\leq \frac{1}{\alpha'} \sum_{k=1}^n \big(f(d_k) - f(c_k)\big) + \epsilon \tag{286}$$

But f is monotone, so

$$\leq \frac{1}{\alpha'}\big(f(b) - f(a)\big) + \epsilon \tag{287}$$

This is true for all $\alpha' < \alpha$ for all $\epsilon > 0$, proving the first claim. The second part follows since it is an intersection of all sets for $\alpha = n$ for all $n \in \mathbb{N}$, which go to 0.

Since we are using outer measure, we don't have to worry nor rely on about measurability. Also, there can be uncountable set at which f is infinite, but it just guarantees outer measure 0.

Theorem 5.5 (Lebesgue)

Suppose f is increasing on (a, b). Then, it is differentiable a.e. on (a, b).

Proof. WLOG, (a, b) is bounded.[a] Consider the countable family of sets

$$E_{\alpha,\beta} = \{x \mid \overline{D}f(x) > \alpha > \beta > \underline{D}f(x), \alpha, \beta \in \mathbb{Q}\} \tag{288}$$

Note that if the derivatives aren't equal, we can always squeeze 2 rationals in, so

$$\{x \mid \overline{D}f(x) > \underline{D}f(x)\} \subset \bigcup_{\alpha,\beta\in\mathbb{Q}} E_{\alpha,\beta} \tag{289}$$

We want to prove that $m^*(E_{\alpha,\beta}) = 0 \quad \forall \alpha, \beta$. Let's find O open s.t. $E_{\alpha,\beta} \subset O$ and $m(O) < m^*(E) + \epsilon$, where we will denote $E = E_{\alpha,\beta}$.
Consider all intervals $[c, d] \subset O$ s.t. $f(d) - f(c) < \beta(d - c)$. Since we know $\underline{D}f(x) < \beta$, these intervals cover E in Vitali sense. So you find a disjoint subcollections $[c_k, d_k]$ for $k = 1, \ldots, n$ s.t.

$$m^*\left(E \setminus \bigcup_{k=1}^{n} [c_k, d_k]\right) < \epsilon \tag{290}$$

Observe that

$$\begin{aligned} \sum_{k=1}^{n} \big(f(d_k) - f(c_k)\big) &< \beta \sum_{k=1}^{n} (d_k - c_k) && (291) \\ &\leq \beta\big(m^*(E) + \epsilon\big) && (292) \end{aligned}$$

On the other hand, we can apply the previous lemma to $E \cap [c_k, d_k]$ to get

$$m^*(E \cap [c_k, d_k]) \leq \frac{1}{\alpha}\big(f(d_k) - f(c_k)\big) \tag{293}$$

and so

$$\begin{aligned} m^*(E) &\leq \frac{1}{\alpha} \sum_{k=1}^{n} \big(f(d_k) - f(c_k)\big) + \epsilon && (294) \\ &\leq \frac{\beta}{\alpha}\big(m^*(E) + \epsilon\big) + \epsilon, \quad \forall \epsilon > 0 && (295) \end{aligned}$$

So, $m^*(E) \leq \frac{\beta}{\alpha} m^*(E)$, where $\frac{\beta}{\alpha} < 1$. Therefore $m^*(E) = 0$.

[a] Otherwise, we can always split it into a countable union of bounded intervals.

But being differentiable doesn't imply that fundamental theorem of calculus holds. So integrating the derivative won't get you back these functions (think of step functions). So we will have to specify a class of functions such that this holds.

Corollary 5.2 ()

Suppose f is increasing on $[a, b]$. Then f' is integrable, and

$$\int_a^b f' \leq f(b) - f(a) \tag{296}$$

Proof. Define

$$D_n f(x) = \frac{f(x + 1/n) - f(x)}{1/n} \tag{297}$$

by Lebesgue theorem, $D_n f \to f'$ a.e. By Fatau,

$$\int_a^b \leq \liminf \int_a^b D_n f = \tag{298}$$

where

$$\begin{aligned} \int_a^b D_n f &= n \int_{a+1/n}^{b+1/n} f(x) - n \int_a^b f(x) && (299) \\ &= n\bigg(\int_b^{b+1/n} f(b) - \int_a^{a+1/n} f(x) \bigg) && (300) \\ &\leq f(b) - f(a) && (301) \end{aligned}$$

Here we extended $f(x)$ by $f(b)$ for $x \in [b, b + \frac{1}{n}]$.

Example 5.1 ()

If f is not monotone but continuous, f' doesn't have to be integrable. Consider

$$f(x) = x^2 \sin\left(\frac{1}{x^2}\right) \tag{302}$$

on $[0, 1]$. Then

$$f'(x) 2x \sin\left(\frac{1}{x^2}\right) - \frac{2}{x} \cos\left(\frac{1}{x^2}\right) \tag{303}$$

6 Differentiation

Now, we will establish differentiation and culminate in the fundamental theorem of calculus. Monotone functions are a nice class of functions to study for differentiation and for constructing more general measures.

Theorem 6.1 ()

Suppose f is monotone, increasing on $[a, b]$. Then, the set of discontinuities of f at most countable.

Proof. Let x_k be any point of discontinuity. Note that

$$\lim_{x \to x_k^-} f(x), \qquad \lim_{x \to x_k^+} f(x) \tag{304}$$

both exist by monotonicity, but since there is a discontinuity, we have

$$L_k^- = \lim_{x \to x_k^-} f(x) < \lim_{x \to x_k^+} f(x) = L_k^+ \tag{305}$$

Then, $L_k^+ - L_k^-$ is a jump of f at x_k. These intervals $[L_k^-, L_k^+]$ are disjoint due to monotonicity, and each interval contains a rational number. So there can only be at most countable intervals.

One piece of info from this trick.

Definition 6.1 ()

A point x is a discontinuity of the first kind of $f(x)$ if both one-sided limits exist.

Now here's a generalization for not necessarily monotone functions.

Theorem 6.2 (Detour)

The set of discontinuities of the first kind is countable.

Proof. Idea of the proof. Look at some jump discontinuity and record the jump $\eta > 0$. Then, find $\delta > 0$ s.t. if $0 < y - x < \delta$, then

$$\left| f(x) - \lim_{y \to x^+} f(y) \right| < \frac{\eta}{10} \tag{306}$$

Then look at the rectangle on the graph associated with each jump. Because the limits exist, you can pick the rectangles so small that they are completely disjoint. Look at picture.

Now back to monotone functions.

Theorem 6.3 ()

For any countable set $C \subset (a, b)$ (where the interval doesn't need to be bounded), there exists monotonically increasing f with a jump at each $x \in C$ and continuous at every $x \notin C$.

Proof. Let $x_1, x_2, \ldots$ be C, and define

$$f(x) = \sum_{x_k \leq x, x \in C} 2^{-k} \tag{307}$$

The sum is increasing and convergent (since it's dominated by geometric series). f also has a jump of 2^{-k} at every x_k.
Now we prove continuity. Suppose $x \notin C$. Take $N \in \mathbb{N}$. Find $\delta_N > 0$ s.t.

$$x_1, x_2, \ldots, x_N \notin (x - \delta_N, x + \delta_N) \tag{308}$$

which is possible since this is a finite set. The remaining sum can only add up to 2^{-N}, and so $f(x + \delta_N) - f(x - \delta_N) \leq 2^{-N}$.

7 Measure (TBD)

The introduction of the σ-algebra seemed quite arbitrary, but bear with me as it will make sense very soon. In general, we want to define a measure $\mu : 2^X \to [0, +\infty]$ that satisfies two properties.

1. *Null empty set.* $\mu(\emptyset) = 0$.
2. *Countable Additivity.* For all countable collections $\{A_k\}_{k=1}^\infty$ of pairwise disjoint subsets $A_k \subset 2^X$,

$$\mu\bigg(\bigsqcup_{k=1}^\infty A_k\bigg) = \sum_{k=1}^\infty \mu(A_k) \tag{309}$$

The first condition is important because it allows us to take finite disjoint unions. That is, since $\mu(A_1 \cup A_2) = \mu(A_1 \cup A_2 \cup \emptyset \cup \ldots)$, we have

$$\sum_{k=1}^\infty = \mu(A_1) + \mu(A_2) \tag{310}$$

Disjointness is clearly important since if it wasn't, then $\mu(A) = \mu(A \cup A) = 2\mu(A)$, which is absurd.

It turns out that this second property is highly restrictive, and in fact some measures cannot be even defined—in the sense that we can create partitions of weird sets and rearrange them to get paradoxes (the most famous being the Banach-Tarski paradox). Therefore, we need to find a certain subset $\mathcal{A} \subset 2^X$ that is consistent with this definition of measure.

1. We want to define a function $\mu^* : 2^X \to [0, +\infty]$ that has a slightly less restrictive form of property 2.[5] We should always be able to construct such a function, which we will call the *outer measure.*
2. Then, we want to use this outer measure to define sets that should like in $\mathcal{A}$. We call these *measurable sets.* It will turn out that $\mathcal{A}$ must be a σ-algebra.
3. Finally, we take the restriction of the outer measure to only measurable sets, and this defines our measure: $\mu = \mu^*\big|_{\mathcal{A}}$.

[6]

Definition 7.1 (Measure)

Given a measurable space $(X, \mathcal{A})$, a **measure** is a function $\mu : \mathcal{A} \longrightarrow [0, +\infty]$[a] satisfying

1. Null empty set $\mu(\emptyset) = 0$.
2. Countable additivity: For all countable collections $\{A_k\}_{k=1}^\infty$ of pairwise disjoint subsets $A_k \in \mathcal{A}$,

$$\mu\bigg(\bigsqcup_{k=1}^\infty A_k\bigg) = \sum_{k=1}^\infty \mu(A_k) \tag{311}$$

Remember that we are allowed to take countable unions inside our σ-algebra, so this makes sense. This immediately implies that given $A, B \in \mathcal{A}$, then $A \subset B \implies \mu(A) \leq \mu(B)$. The triplet $(X, \mathcal{A}, \mu)$ is called a **measure space**.

[a] We usually introduce this by taking the codomain to be either $[0, +\infty]$ or $(-\infty, +\infty)$, which is the signed measure.

[5] How we implement such a function is a different question, though.

[6] Old but good explanation: Now let's try to construct a measure λ on the Borel σ-algebra $\mathcal{B}(\mathbb{R})$ that assigns length, i.e. $\lambda([a, b]) = b - a$. We will do so by constructing outer measures $\lambda^* : 2^{\mathbb{R}} \longrightarrow \mathbb{R}$ that acts on the power set of $\mathbb{R}$ s.t. $\lambda^*([a, b]) = b - a$. But this turns out to have its own problems and contradictions, so once we construct such a λ^*, we will "throw away" all the sets that don't behave nicely under λ^* and just use its restriction on the Borel algebra. It turns out that the sets that do behave well under λ^* is bigger than the Borel algebra, call it $\mathcal{M}_{\lambda^*}$. So, we have $\mathcal{B}(\mathbb{R}) \subset \mathcal{M}_{\lambda^*} \subset 2^{\mathbb{R}}$. We will do this in full generality in the following way. We take any space X and construct an outer measure μ^* on its power set 2^X. Then, we construct the σ-algebra of well-behaved sets $\mathcal{M}_{\mu^*} \subset 2^X$, and define our measure μ on $\mathcal{M}_{\mu^*}$. When defining our outer measure, the condition that the outer measure of a disjoint union of subsets is equal to the sum of the outer measure of the subsets is a bit too restricting, so we use a softer condition.

Let's go through each of these three steps in detail.

7.1 Outer Measure

Definition 7.2 (Outer Measure)

Given a space X, an **outer measure** is a function $\mu^* : 2^X \to [0, +\infty]$ satisfying either the two properties.

1. *Null Empty Set.* $\mu^*(\emptyset) = 0$.
2. *Countable Subadditivity.* For arbitrary subset $A, B_1, B_2, \ldots$,

$$A \subset \bigcup_{k=1}^{\infty} B_k \implies \mu(A) \leq \sum_{k=1}^{\infty} \mu(B_k) \tag{312}$$

or equivalently, the three properties.

1. *Null Empty Set.* $\mu^*(\emptyset) = 0$.
2. *Monotonicity.* If $A, B \subset X$, then

$$A \subset B \implies \mu^*(A) \leq \mu^*(B) \tag{313}$$

3. *Countable Subadditivity.* For any countable collection of subsets $\{A_k\}$ of X,

$$\mu^*\bigg(\bigcup_k A_k\bigg) \leq \sum_k \mu^*(A_k) \tag{314}$$

Proof. Prove that the two definitions are equal.

It's a hard definition, but a natural one, since we're taking all these intervals and trying to make them as snug as possible to define the outer measure of an arbitrary set. As always, let's begin with the simplest case in the real line. The following definition suffices.

Lemma 7.1 (Lebesgue Outer Measure is an Outer Measure)

The Lebesgue outer measure λ^* on $\mathbb{R}$ is indeed an outer measure.

We can also generalize this further by introducing a increasing, continuous function $F : \mathbb{R} \to \mathbb{R}$ and defining the outer measure to be

$$\lambda^*(A) = \inf_{C_A} \sum_{j=1}^{\infty} \big(F(b_j) - F(a_j)\big) \tag{315}$$

7.2 Measurable Sets

Definition 7.3 (Carathéodory's criterion)

Given outer measure μ^* on X, a set $E \subset X$ is called μ^***-measurable** if for every set $A \subset X$,

$$\mu^*(A \cap E) + \mu^*(A \cap E^c) = m^*(A) \tag{316}$$

In general it says that no matter how nasty a subset A is, E should be nice enough that we can cut E into two pieces C and D. Due to the definition of the outer measure, we are guaranteed to have $\mu^*(C \cup D) \leq \mu^*(C) + \mu^*(D)$. The sets with which this inequality is strict is not measurable, and the measurable sets specifically satisfy

1. equality
2. for countable sets.

One should note that in particular, if E is μ^*-measurable and A is any set disjoint from E, then we must have

$$\mu^*(A \cup E) = \mu^*((A \cup E) \cap E) + \mu^*((A \cup E) \cap E^c) \tag{317}$$
$$= \mu^*(E) + \mu^*(A) \tag{318}$$

which solves a bit of the theorem on measures. In practice, we will often prove that $\mu^*(A \cap E) + \mu^*(A \cap E^c) \leq m^*(A)$, since the properties of outer measure implies $\geq$.

Example 7.1 ()

Take $X = \mathbb{R}$ and have $B = (-\infty, b]$. Then $B^c = (b, \infty)$, and B divides $\mathbb{R}$ into a right side and a left side. If we take any subset $A \subset \mathbb{R}$, then B is nice enough to divide A into a left and a right side.

Now we want to establish some nice properties.

Theorem 7.1 (Outer Measure 0 Sets are Measurable)

For any outer measure μ^* on X, $E \subset X$ with $\mu^*(E) = 0$ implies that E is μ^*-measurable.

Proof. Take any A. Then $(A \cap E) \subset E$ and $(A \cap E^c) \subset A$. So by monotonicity,

$$\mu^*(A \cap E) + \mu^*(A \cap E^c) \leq \mu^*(E) + \mu^*(A) = \mu^*(A) \tag{319}$$

and this by definition means that E is measurable.

Now let's talk about constructing measurable sets.

Theorem 7.2 (Finite Unions are Measurable)

A finite union of μ^*-measurable sets is μ^*-measurable.

Proof. It suffices to prove for E_1, E_2, and the rest follows by induction. Fix any A. Then

$$\mu^*(A) = \mu^*(A \cap E_1) + \mu^*(A \cap E_1^c) \tag{320}$$
$$= \mu^*(A \cap E_1) + \mu^*\big((A \cap E_1^c) \cap E_2\big) + \mu^*\big((A \cap E_1^c) \cap E_2^c\big) \tag{321}$$

But

$$(A \cap E_1^c) \cap E_2^c = A \cap (E_1 \cup E_2)^c \tag{322}$$
$$(A \cap E_1^c) \cap E_2 = (A \setminus E_1) \setminus E_2 \tag{323}$$

So, $(A \cap E_1) \cup \big((A \setminus E_1) \cap E_2\big) = A \cap \big(A \cap (E_1 \cup E_2)^c\big)$. Therefore, we get

$$\mu^*(A \cap E_1) + \mu^*\big((A \cap E_1^c) \cap E_2\big) + \mu^*\big((A \cap E_1^c) \cap E_2^c\big) \geq \mu^*(A \cap (E_1 \cup E_2)) + \mu^*(A \cap (E_1 \cup E_2)^c) \tag{324}$$

So we have proved that the set of all measurable sets is closed under finite unions. By definition it works for finite intersections. This makes it into an *algebra*, but we want to upgrade this to a σ-algebra by proving closure under *countable* unions. We will need the lemma.

Lemma 7.2 ()

Suppose $E_1, \ldots, E_n$ are disjoint. Then,

$$\mu^*\bigg(\bigcup_{j=1}^n E_j\bigg) = \sum_{j=1}^n \mu^*(E_j) \tag{325}$$

Proof. We already did this for 2 sets, and just use induction.

Now we prove lemma, which is more general (arbitrary intersections than finite?).

Lemma 7.3 ()

Suppose A is any set, E_j disjoint and measurable. Then,

$$\mu^*\bigg(A \cap \Big(\bigcup_{j=1}^n E_j\Big)\bigg) = \sum_{j=1}^n \mu^*(A \cap E_j) \tag{326}$$

Proof. By induction, $n = 1$ is true. Then,

$$\begin{aligned}
\mu^*\bigg(A \cap \Big(\bigcup_{j=1}^n E_j\Big)\bigg) &= \mu^*\Bigg(\bigg(A \cap \Big(\bigcup_{j=1}^n E_j\Big)\bigg) \cap E_n\Bigg) + \mu^*\Bigg(\bigg(A \cap \Big(\bigcup_{j=1}^n E_j\Big)\bigg) \cap E_n^c\Bigg) && (327)\\
&= \mu^*(A \cap E_n) + \mu^*\bigg(A \cap \Big(\bigcup_{j=1}^{n-1} E_j\Big)\bigg) && (328)\\
&= \sum_{j=1}^n \mu^*(A \cap E_j) && (329)
\end{aligned}$$

by the induction hypothesis.

Theorem 7.3 (Countable Unions are Outer Measurable)

Suppose $E_1, E_2, \ldots$ are a countable collection of measurable sets. Then, $E = \cup_{j=1}^\infty E_j$ is measurable.

Proof. They key is to look at disjoint sets. WLOG, one can assume E_j are disjoint. Indeed, we can define new sets

$$E_n' \coloneqq E_n \setminus \bigg(\bigcup_{j=1}^{n-1} E_j\bigg) \tag{330}$$

that are measurable, with $\cup E_n' = \cup E_n$. Now, fix any set A. Define sets $F_n = \cup_{j=1}^n E_j$. Then, $\mu^*(A) = \mu^*(A \cap F_n) + \mu^*(A \cap F_n^c)$. Then, $F_n^c \supset E^c \implies \mu^*(A \cap F_n^c) \geq \mu^*(A \cap E^c)$. Through the previous lemma, we have

$$\mu^*(A \cap F_n) = \mu^*\bigg(\bigcup_{j=1}^n (A \cap E_j)\bigg) = \sum_{j=1}^n \mu^*(A \cap E_j) \tag{331}$$

Then,

$$\mu^*(A) \geq \sum_{j=1}^n \mu^*(A \cap E_j) + \mu^*(A \cap E^c) \tag{332}$$

for every n, therefore also with ∞. But

$$\sum_{j=1}^{\infty} \mu^*(A \cap E_j) \geq \mu^*(A \cap E) \tag{333}$$

If follows that $\mu^*(A) \geq \mu^*(A \cap E) + \mu^*(A \cap E^c)$.

Corollary 7.1 (Measurable Sets form a σ-Algebra)

The set of all μ^*-measurable sets of X form a σ-algebra.

With this, we can construct a lot of measurable sets.

Lemma 7.4 (Sets of Measure 0 have no Effect)

Suppose $\mu^*(E) = 0$ and A is any set. Then, $\mu^*(A \cup E) = \mu^*(A)$.

Proof. We have

$$\mu^*(A \cup E) = \underbrace{\mu^*\big((A \cup E) \cap E\big)}_{=0} + \mu^* \underbrace{\big((A \cup E) \cap E^c\big)}_{\subset A} \leq \mu^*(A) \leq \mu^*(A) \tag{334}$$

But $A \cup E \supset A$, so $\mu^*(A \cup E) = \mu^*(A)$.

So we can always drop an outer-measure 0 set and it won't affect the outer measure of the original set.

Theorem 7.4 ()

Every interval $(a, +\infty)$ is measurable.

Proof. Take any set A, and WLOG $a \notin A$ (since we can take the point out without affecting outer measure). Suppose $\{I_k\}_{k=1}^{\infty}$ is a cover of A s.t.

$$\mu^* > \left(\sum_{k=1}^{\infty} \ell(I_k) \right) - \epsilon \tag{335}$$

Then,

1. $I_k' \coloneqq I_k \cap (a, +\infty)$ will cover $A_1 = A \cap (a, +\infty)$, and
2. $I_k'' \coloneqq I_k \cap (-\infty, a)$ will cover $A_2 = A \cap (-\infty, a)$.

Therefore, $\mu^*(A_1) \leq \sum_k \ell(I_k')$, $\mu^*(A_2) \leq \sum_k \ell(I_k'')$. Also,

$$\ell(I_k) = \ell(I_k') + \ell(I_k'') \implies \mu^*(A_1) + \mu^*(A_2) \leq \sum_k \ell(I_k) \leq \mu^*(A) + \epsilon \tag{336}$$

for every $\epsilon > 0$. Since this is true for every $\epsilon > 0$, we are done.

The next theorem shows that we can construct measurable sets with "nice" sets on the real line.

Theorem 7.5 (λ^*-measurable Sets)

TFAE in $\mathbb{R}$ with the Lebesgue outer measure.
1. E is measurable.
2. $\forall \epsilon > 0$, $\exists$ open set $O \supset E$ s.t. $\mu(O \setminus E) \leq \epsilon$.
3. $\forall \epsilon > 0$, $\exists$ closed set $F \subset E$ s.t. $\mu^*(E \setminus F) < \epsilon$.
4. $\exists$ a G_δ set G s.t. $E \subset G$ and $\mu^*(G \setminus E) = 0$.
5. $\exists$ a F_σ set F s.t. $F \subset E$ and $\mu^*(E \setminus F) = 0$.

Proof. Listed.
1. (2) $\implies$ (1). Then for every $k \in \mathbb{N}$, we can find $O_k \supset E$ s.t. $m^*(O_k \setminus E) \leq 1/k$. Define the G_δ set $G = \cap_{k=1}^\infty O_k$. Then, $(G \setminus E) \subset (O_k \setminus E)$ for all $k \implies m^*(G \setminus E) \leq 1/k$ for all k. Therefore $m^*(G \setminus E) = 0$, and $E = G \setminus (G \setminus E)$ is measurable.
2. (1) $\implies$ (2). Assume $m^*(E) < +\infty$. Find a cover $\{I_k\}_{k=1}^\infty$ s.t. $\sum_{k=1}^\infty \ell(I_k) \leq m^*(E) + \epsilon$. Call $O = \cup_k I_k$. Since E is measurable, $m^*(O \setminus E) = m^*(O) - m^*(E) \leq \sum_{k=1}^\infty \ell(I_k) - m^*(E) \leq \epsilon$
3. (1) $\iff$ (3). Straightforward with argument above.
4. (1) $\iff$ (4). Generally, we use the fact that E measurable iff E^c measurable. Find $O \supset E^c$ open, with $m^*(O \setminus E^c) \leq \epsilon$. Then $F = O^c$ is closed, $F \subset E$, and $m^*(E \setminus F) \leq \epsilon$.
5. (1) $\iff$ (5). Same argument as (1) $\iff$ (4).

The following theorem is a major one, showing that measurable sets can be approximated well by Borel sets.

Theorem 7.6 ()

Suppose E is measurable, with $m^*(E) < +\infty$. Fix $\epsilon > 0$. Then there exists a finite number of intervals $\{I_k\}_{k=1}^n$ s.t. if $O = \cup_{k=1}^n I_k$, then

$$m^*(O \setminus E) + m^*(E \setminus O) < \epsilon \tag{337}$$

Proof. Find $\{I_k\}_{k=1}^\infty$ s.t. $U = \cup_{k=1}^\infty I_k$ satisfies $E \subset U$, $m^*(U \setminus E) \leq \epsilon/2$. Find n s.t. $\sum_{k=n+1}^\infty \ell(I_k) \leq \epsilon/2$ where WLOG, I_k are disjoint. Define $O = \cup_{k=1}^n I_k$. Then, we have

$$m^*(O \setminus E) \leq m(U \setminus E) \leq \frac{\epsilon}{2} \tag{338}$$

$$m^*(E \setminus O) \leq m(U \setminus O) \leq \sum_{k=n+1}^\infty \ell(I_k) \leq \frac{\epsilon}{2} \tag{339}$$

For $\mathbb{R}$, we can create our Lebesgue outer measure λ^* on it, which generates the Lebesgue σ-algebra $\mathcal{M}_{\lambda^*}$. This turns out to be bigger than the Borel σ-algebra $\mathcal{B}(\mathbb{R})$, but there is little difference in which one we choose when we actually integrate.

Theorem 7.7 ()

A set $E \subset \mathbb{R}$ is Lebesgue measurable implies that it is also Borel measurable.

$$\mathcal{B}(\mathbb{R}) \subset \mathcal{M}_{\lambda^*} \subset 2^{\mathbb{R}} \tag{340}$$

Lemma 7.5 ()

If $E \subset \mathbb{R}$ and $\lambda^*(E) = 0$, then $E \in \mathcal{M}_{\lambda^*}$, i.e. E is Lebesgue outer-measurable.

Proof. We must prove that E satisfies the Carathéodory's criterion. For all $E \subset \mathbb{R}$, we know that $\lambda^*(A) \leq \lambda^*(A \cap E) + \lambda^*(A \cap E^c)$ by definition of outer measure. Now, since $\lambda^*(E) = 0$ and $A \cap E \subset E$, this means that $\lambda^*(A \cap E) = 0$ also. Furthermore, $A \cap E^c \subset A$, meaning that $\lambda^*(A) \geq \lambda^*(A \cap E^c)$, and we get

$$\lambda^*(A) \geq \lambda^*(A \cap E) + \lambda^*(A \cap E^c) \tag{341}$$

which proves equality.