Measure Theory

Muchang Bahng

Spring 2024

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 Sigma Algebras

In here, we will develop a deeper formalism of set theory and topology, now that we have the tools of analysis.

1.1 Set-Theroetic Limits

Let's talk about sequences of sets $(A_n)_n$.

Definition 1.1 (Monotone Sequence)

A sequence of sets $(A_n)_n$ is called
1. **(strictly) increasing** if $A_n \subsetneq A_{n+1}$.
2. **nondecreasing** if $A_n \subseteq A_{n+1}$.
3. **(strictly) decreasing** if $A_n \supsetneq A_{n+1}$.
4. **nonincreasing** if $A_n \supseteq A_{n+1}$.

Definition 1.2 (Limsup and Liminf of Sets)

Given a sequence of sets $(A_n)_n$, the **limsup** and **liminf** of them can be defined in the equivalent ways.
1. The **liminf** is the set of points that are missing in only a finite number of sets, and the **limsup** is the set of points that are in an infinite number of sets.

$$\liminf_{n\to\infty} A_n := \bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} A_m \tag{9}$$

$$\limsup_{n\to\infty} A_n := \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} A_m \tag{10}$$

2. The **liminf** and **limsup** are the set of points x where the liminf and limsup of the indicator function function evaluated at x equals 1.

$$\liminf_{n\to\infty} A_n := \{x \in X \mid \liminf_{n\to\infty} \mathbb{1}_{A_n}(x) = 1\} \tag{11}$$

$$\limsup_{n\to\infty} A_n := \{x \in X \mid \limsup_{n\to\infty} \mathbb{1}_{A_n}(x) = 1\} \tag{12}$$

Both liminf and limsup always exist for any sequence of sets.

Proof.

DeMorgan's law.

Lemma 1.1 (Monotonicity)

For any sequence of sets

$$\liminf_{n\to\infty} A_n \subseteq \limsup_{n\to\infty} A_n \tag{13}$$

Lemma 1.2 (Complements)

$$\liminf_{n\to\infty} A_n = \left(\limsup_{n\to\infty} A_n^c\right)^c \tag{14}$$

Proof.

Definition 1.3 (Limit of Sets)

1.2 Borel Hierarchy

Definition 1.4 (F_σ Sets)

A F_σ**-set** is a subset of a topological space that is a countable union of closed sets.

Definition 1.5 (G_δ Sets)

A G_δ**-set** is a subset of a topological space that is a countable intersection of open sets.

Lemma 1.3 ()

The complement of a F_σ set is a G_δ set.

1.3 Sigma Algebra

Now, given any set X, we can construct its power set 2^X. But we can't naively just give a measure to every $A \in 2^X$, since for certain spaces, this causes nasty contradictions shown through the Banach-Tarski Paradox.[1] A nice set of subsets of X to work with is the σ-algebra of X.

Definition 1.6 (σ-Algebra)

A **σ-algebra** on a set X is a collection of subsets of X, denoted $\mathcal{A} \subset 2^X$ that contains $\emptyset$, X itself, is stable under a countable union, and is stable under complementation. This pair $(X, \mathcal{A})$ is called a **measurable space**.

Lemma 1.4 (Additional Property of σ-Algebras)

A commonly known property of any σ-algebra $\mathcal{A}$ is that it is stable under countable intersections, too.

$$A_1, A_2, \ldots, \in \mathcal{A} \implies \bigcap_{k=1}^\infty A_k \in \mathcal{A} \tag{15}$$

Proof.

We can utilize the fact that

$$\bigcap_{k=1}^\infty A_k = X \setminus \bigcup_{k=1}^\infty A_k^c \tag{16}$$

A σ-algebra is similar to the topology τ of topological space. Both $\mathcal{A}$ and τ require $\emptyset$ and X to be in it. The three differences are that (i) τ does not allow compelmentation, (ii) τ allows any (even uncountable) union

[1] Given any two bounded subsets A and B of $\mathbb{R}^n$ where $n \geq 3$, both of which have a nonempty interior, there are partitions of A and B into a finite number of disjoint subsets, $A = A_1 \cup \ldots \cup A_k$, $B = B_1 \cup \ldots \cup B_k$, such that A_i and B_i are congruent for every $i \in [k]$.

of sets (condition is strengthened), and (iii) τ allows only finite intersection of sets (condition is weakened). Now in order to construct σ-algebras, the following theorems are useful since they allow us to construct σ-algebras from other σ-algebras. It turns out that the intersection of σ-algebras is a σ-algebra, but not for unions.

Theorem 1.1 (Intersection of Sigma Algebras is a Sigma Algebra)

Let $\{\mathcal{A}_k\}$ be a family of σ-algebras of X. Then, $\cap \mathcal{A}_k$ is also a σ-algebra of X.

Proof.

Clearly, $\emptyset, X$ is in $\cap \mathcal{A}_k$. To prove complementation,

$$A \in \bigcap \mathcal{A}_k \implies A \in \mathcal{A}_k \; \forall k \implies A^c \in A_k \; \forall k \implies A^c \in \bigcap \mathcal{A}_k \tag{17}$$

To prove countable union, let $\{A_j\}_{j \in J}$ be some countable family of subsets in $\cap \mathcal{A}_k$. Then,

$$A_j \in \bigcap \mathcal{A}_k \; \forall j \in J \implies A_j \in \mathcal{A}_k \; \forall k \forall j \implies \bigcup A_j \in \mathcal{A}_k \; \forall k \implies \bigcup A_j \in \bigcap \mathcal{A}_k \tag{18}$$

This allows us to easily prove the following theorem, which just establishes the existence of σ-algebras.

Theorem 1.2 (Unique Smallest Sigma Algebra)

Let $F \subset 2^X$. Then there exists a unique smallest σ-algebra $\sigma(F)$ containing F, called the σ-algebra **generated** by F.

Proof.

Let us denote $\mathcal{M}$ as the set of all possible σ-algebras $\mathcal{B}$ of X. $\mathcal{M}$ is nonempty since it contains 2^X. Then, the intersection

$$\bigcap_{\mathcal{B} \in \mathcal{M}} \mathcal{B} \tag{19}$$

is the unique smallest σ-algebra.

With this guarantee, we can now define what it means for a set of subsets to *generate* a σ-algebra.

Definition 1.7 (σ-Algebra Generated by a Set)

Given a collection of sets $\mathscr{C}$, the σ-algebra **generated** by $\mathscr{C}$ is the unique smallest σ-algebra containing $\mathscr{C}$, denoted $\sigma(\mathscr{C})$.

This gives us a convenient way to construct σ-algebras. The general method is to identify a collection of "important" subsets that we would like to be included in the σ-algebra, and then just generate it.

Definition 1.8 (Borel σ-algebra)

The **Borel σ-algebra** of a topological space $(X, \mathscr{T})$ is the σ-algebra generated by the topology $\mathscr{T}$, denoted $\mathcal{B}(X) := \sigma(\mathscr{T})$. An element of the Borel algebra is called a **Borel set**.

Note that the Borel algebra contains:

1. all open sets,
2. all closed sets due to closure under complements,

3. all G_δ sets due to closure under countable unions,
4. all F_σ sets due to closure under countable intersection.

Definition 1.9 (Measure Space)

A **measure set** is a tuple $(X, \mathcal{A})$, where X is an arbitrary space and $\mathcal{A}$ a σ-algebra.

2 Measures

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 $A_k \subset 2^X$,

$$\mu\bigg(\bigsqcup_{k=1}^\infty A_k\bigg) = \sum_{k=1}^\infty \mu(A_k) \tag{20}$$

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

$$\sum_{k=1}^\infty = \mu(A_1) + \mu(A_2) \tag{21}$$

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. But this is self-contradictory, as it turns out that 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.[2] 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: $\mu = \mu^*\big|_{\mathcal{A}}$.

[3]

Definition 2.1 (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 $A_k \in \mathcal{A}$,

$$\mu\bigg(\bigsqcup_{k=1}^\infty A_k\bigg) = \sum_{k=1}^\infty \mu(A_k) \tag{22}$$

   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**.

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

[3] Old but good explanataion: 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.

2.1 Outer Measure

Definition 2.2 (Outer Measure)

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

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

$$A \subset \bigcup_{k=1}^{\infty} B_k \implies \mu(A) \leq \sum_{k=1}^{\infty} \mu(B_k) \tag{23}$$

or equivalently, the three properties.

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

$$A \subset B \implies \mu^*(A) \leq \mu^*(B) \tag{24}$$

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

$$\mu^*\bigg(\bigcup_k A_k\bigg) \leq \sum_k \mu^*(A_k) \tag{25}$$

Proof.

Prove that the two definitions are equal.

Okay, so what are some examples of outer measures that we can define on $\mathbb{R}$, or in general $\mathbb{R}^n$? Well one approach would be to generalize the concepts length/area/volume. 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 2.3 (Elementary Set)

An **elementary set** $E \subset \mathbb{R}^n$ 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 elementary set $E \subset \mathbb{R}^n$ is $E = I_1 \times \ldots \times I_n$ for intervals $I_1, \ldots, I_k$.

Definition 2.4 (Size)

The **size** of an elementary set $E = I_1 \times \ldots \times I_n \subset \mathbb{R}^n$ is defined recursively as

1. The **length** of an interval I is $\ell(I) = b - a$.
2. The **size** of E is $s(E) = \prod_{i=1}^{n} (b_i - a_i)$.

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

Definition 2.5 (Lebesgue Outer Measure)

Given any set $A \subset \mathbb{R}$, the **Lebesgue outer measure** is defined

$$\lambda^*(A) = \inf \left\{ \sum_{k=1}^\infty \ell(I_k) \;\middle|\; A \subset \bigcup_{k=1}^\infty I_k \right\} \tag{26}$$

Intuitively, it is just the infimum of the sums of lengths of the intervals that cover A.[a]

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 2.1 (Lebesgue Outer Measure is an Outer Measure)

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

Proof.

We prove the three properties. The first two are trivial. For the third, we wish to show that $\lambda^*(\cup A_n) \leq \sum \lambda^*(A_n)$. For each n, find a specific cover $\{I_{n_k}\}_{k=1}^\infty$ of A_n such that it "just covers" enough (this is possible since λ^* is an infimum) such that for any $\epsilon > 0$,

$$\sum_k \ell(I_{n_k}) - \frac{\epsilon}{2^k} \leq \lambda^*(A_k) \tag{27}$$

Then,

$$\bigcup A_n \subset \bigcup_{n,k=1} I_{n_k} \implies \sum_{n,k=1} \ell(I_{n_k}) \leq \sum_{n=1}^\infty \lambda^*(A_n) + \epsilon \tag{28}$$

and since ϵ is arbitrary, we are done.
The first condition is trivial. As for 2, if I have $A \subset B \subset \mathbb{R}$ and have a covering of B, then I also have a covering of A, and so the infimum corresponding to the covering of B must be greater than or equal to the infimum of that corresponding to the covering of A. For 3, we want to prove that the outer measure of the union of A_k's is less than or equal to the sum of the outer measures of the A_k's. We pick $\epsilon > 0$ and have some covering $\{(a_j^k, b_j^k)\}_{j=1}^\infty \in C_{A^k}$. So we have

$$\lambda^*(A_k) \leq \sum_{j=1}^\infty b_j^k - a_j^k \tag{29}$$

We want the inequality to go the other way around, but we can't do that. But note that $\lambda^*(A_k)$ is the infimum of all coverings $\{(a_j^k, b_j^k)\}_{j=1}^\infty$ of A_k, and so we can choose a covering that is as close to $\lambda^*(A_k)$, and then add a term of ϵ to $\lambda^*(A_k)$ to make it greater than this covering. This is an important step of the proof that is used often!

$$\frac{\epsilon}{2^k} + \lambda^*(A_k) \geq \sum_{j=1}^\infty b_j^k - a_j^k \tag{30}$$

Now,

$$A = \bigcup_{k=1}^\infty A_k \subset \bigcup_{k=1}^\infty \bigcup_{j=1}^\infty (a_j^k, b_j^k) \tag{31}$$

[a] I use the notation μ^* to represent general outer measures, and λ^* to represent specifically the Lebesgue outer measure.

and we can see that $\{(a_j^k, b_j^k)\}_{j,k=1}^\infty \in C_A$ is a countable covering of A (since the countable union of a countable union is countable), implying that

$$\lambda^*(A) \leq \sum_{k=1}^\infty \sum_{j=1}^\infty (b_j^k - a_j^k) \leq \sum_{k=1}^\infty \bigg(\lambda^*(A_k) + \frac{\epsilon}{2^k} \bigg) = \epsilon + \sum_{k=1}^\infty \lambda^*(A_k) \tag{32}$$

and so setting ϵ arbitrarily small we have $\lambda^*(A) \leq \sum_{k=1}^\infty \lambda^*(A_k)$.

Not only is it an outer measure; it also satisfies the property that we wanted, along with the bonus property of translation invariance!

Theorem 2.1 (Lebesgue Outer Measure Coincides with Interval Length)

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

Proof.

Let $I = [a, b]$. Take $I_1 = [a - \epsilon, b + \epsilon]$.

1. For an upper bound, we see $\lambda^*(I) \leq b - a + 2\epsilon$, where $2\epsilon \to 0$.
2. For a lower bound, suppose

$$\bigcup_{n=1}^\infty I_n \supset [a, b] \tag{33}$$

   By Heine-Borel, we can extract a finite subcollection $I_1, \ldots, I_n$ that still covers $[a, b]$. Under (a_k, b_k) so that a_k is increasing.
   (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

$$b_2 - a_2 + b_1 - a_1 \geq b_2 - a_1 > b - a \tag{34}$$

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

$$b_n - a_n + b_{n-1} - a_{n-1} + \ldots + b_1 - a_1 \geq b_n - a_1 > b - a \tag{35}$$

Corollary 2.1 (Translation Invariance)

λ^* is translation invariant. That is, for any $A \subset \mathbb{R}$,

$$\lambda^*(A) = \lambda^*(A + x) \tag{36}$$

where $A + x \coloneqq \{a + x \in \mathbb{R} \mid a \in A\}$.

Theorem 2.2 (Countable Sets have Outer Measure 0)

Any countable set of $\mathbb{R}$ has Lebesgue outer measure 0.

Proof.

Just enumerate $A = \{x_1, \ldots\}$. Then, we set $I_k = \big(x_k - \frac{\epsilon}{2^k}, x_k + \frac{\epsilon}{2^k}\big)$. Then,

$$\sum_{k=1}^\infty \ell(I_k) = \epsilon \tag{37}$$

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

$$\lambda^*(A) = \inf_{C_A} \sum_{j=1}^{\infty} \big(F(b_j) - F(a_j)\big) \tag{38}$$

In $\mathbb{R}^n$, this construction is exactly the same, since we can take rectangular prisms, which we know the area/volume of, make a countable covering of some arbitrary set $A \subset \mathbb{R}^n$, and then find the infimum of the volume of this set. But we can't apply the outer measure on power sets since there exists some sets that do not behave like how we want it to behave under a measure. For example, there exists disjoint $A, B \subset (0, 1)$ s.t. $A \cup B = (0, 1)$, but $\lambda^*(A) + \lambda^*(B) > 1$.

2.2 Measurable Sets

Definition 2.6 (Carathéodory's criterion)

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

$$\mu^*(A \cap E) + \mu^*(A \cap E^c) = m^*(A) \tag{39}$$

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

$$\begin{aligned} \mu^*(A \cup E) &= \mu^*((A \cup E) \cap E) + \mu^*((A \cup E) \cap E^c) && (40) \\ &= \mu^*(E) + \mu^*(A) && (41) \end{aligned}$$

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 2.1 ()

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 2.3 (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,

$$\mu^*(A \cap E) + \mu^*(A \cap E^c) \leq \mu^*(E) + \mu^*(A) = \mu^*(A) \tag{42}$$

and this by definition means that E is measurable.

Now let's talk about constructing measurable sets.

Theorem 2.4 (Finite Unions are Outer 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

$$\begin{aligned} \mu^*(A) &= \mu^*(A \cap E_1) + \mu^*(A \cap E_1^c) && (43) \\ &= \mu^*(A \cap E_1) + \mu^*\big((A \cap E_1^c) \cap E_2\big) + \mu^*\big((A \cap E_1^c) \cap E_2^c\big) && (44) \end{aligned}$$

But

$$\begin{aligned} (A \cap E_1^c) \cap E_2^c &= A \cap (E_1 \cup E_2)^c && (45) \\ (A \cap E_1^c) \cap E_2 &= (A \setminus E_1) \setminus E_2 && (46) \end{aligned}$$

So, $(A \cap E_1) \cup \big((A \setminus E_1) \cap E_2\big) = A \cap \big(A \cap (E_1 \cup E_2)^c\big)$.

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 2.2 ()

Suppose $E_1, \ldots, E_n$ are disjoint. Then,

$$\mu^*\bigg(\bigcup_{j=1}^n E_j\bigg) = \sum_{j=1}^n \mu^*(E_j) \tag{47}$$

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 2.3 ()

Suppose A is any set, E_j disjoint and measurable. Then,

$$\mu^*\bigg(A \cap \Big(\bigcup_{j=1}^n E_j\Big)\bigg) = \sum_{j=1}^n \mu^*(A \cap E_j) \tag{48}$$

Proof.

By induction, $n = 1$ is true. Then,

$$\mu^*\Big(A \cap \Big(\bigcup_{j=1}^n E_j\Big)\Big) = \mu^*\bigg(\Big(A \cap \Big(\bigcup_{j=1}^n E_j\Big)\Big) \cap E_n\bigg) + \mu^*\bigg(\Big(A \cap \Big(\bigcup_{j=1}^n E_j\Big)\Big) \cap E_n^c\bigg) \tag{49}$$

$$= \mu^*(A \cap E_n) + \mu^*\Big(A \cap \Big(\bigcup_{j=1}^{n-1} E_j\Big)\Big) \tag{50}$$

$$= \sum_{j=1}^n \mu^*(A \cap E_j) \tag{51}$$

by the induction hypothesis.

Theorem 2.5 (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

$$E_n' \coloneqq E_n \setminus \Big(\bigcup_{j=1}^{n-1} E_j\Big) \tag{52}$$

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

$$\mu^*(A \cap F_n) = \mu^*\Big(\bigcup_{j=1}^n (A \cap E_j)\Big) = \sum_{j=1}^n \mu^*(A \cap E_j) \tag{53}$$

Then,

$$\mu^*(A) \geq \sum_{j=1}^n \mu^*(A \cap E_j) + \mu^*(A \cap E^c) \tag{54}$$

for every n, therefore also with ∞. But

$$\sum_{j=1}^\infty \mu^*(A \cap E_j) \geq \mu^*(A \cap E) \tag{55}$$

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

Corollary 2.2 (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 2.4 (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

$$\mu^*(A \cup E) = \underbrace{\mu^*\big((A \cup E) \cap E\big)}_{=0} + \mu^* \underbrace{\big((A \cup E) \cap E^c\big)}_{\subset A} \leq \mu^*(A) \leq \mu^*(A) \tag{56}$$

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 2.6 ()

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.

$$\mu^* > \bigg(\sum_{k=1}^\infty \ell(I_k)\bigg) - \epsilon \tag{57}$$

Then,

1. $I_k' := I_k \cap (a, +\infty)$ will cover $A_1 = A \cap (a, +\infty)$, and
2. $I_k'' := I_k \cap (-\infty, a)$ 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,

$$\ell(I_k) = \ell(I_k') + \ell(I_k'') \implies \mu^*(A_1) + \mu^*(A_2) \leq \sum_k \ell(I_k) \leq \mu^*(A) + \epsilon \tag{58}$$

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

Theorem 2.7 (λ^*-measurable Sets)

TFAE in $\mathbb{R}$ with the Lebesgue outer measure. E is measurable.

1. $\forall \epsilon > 0$, $\exists$ open set $O \supset E$ s.t. $\mu(O \setminus E) \leq \epsilon$.
2. $\forall \epsilon > 0$, $\exists$ closed set $F \subset E$ s.t. $\mu^*(E \setminus F) < \epsilon$.
3. $\exists$ a G_δ set G s.t. $E \subset G$ and $\mu^*(G \setminus E) = 0$.
4. $\exists$ a F_σ set F s.t. $F \subset E$ and $\mu^*(E \setminus F) = 0$.

So essentially, we can construct measurable sets with "nice" sets.

Proof.

Listed.

1.

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 2.8 ()

A set $E \subset \mathbb{R}$ is Lebesgue measurable implies that it is also Borel measurable.

$$\mathcal{B}(\mathbb{R}) \subset \mathcal{M}_{\lambda^*} \subset 2^{\mathbb{R}} \tag{59}$$

Lemma 2.5 ()

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

$$\lambda^*(A) \geq \lambda^*(A \cap E) + \lambda^*(A \cap E^c) \tag{60}$$

which proves equality.

2.3 Measures

Theorem 2.9 ()

The restriction of an outer measure μ^* to the set of all μ^*-measurable sets $\mathcal{A}$, denoted $\mu = \mu^*\big|_{\mathcal{A}}$, is measurable.

Now there are nice properties that we want Lebesgue measures to have: completeness, regularity, and translation invariance.

1. Completeness: Given sets $A \subset B \subset C$ with $\mu(A) = \mu(C)$ and $A, C \in \mathcal{A}$, this implies that $B \in \mathcal{A}$. 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.
2. Regularity: Given sets $A \subset B \subset C$, regularity talks about whether I can approximate B well. Most nice measures have this property.

$$\sup_{A \text{ compact}} \mu(A) = \mu(B) = \inf_{C \text{ open}} \mu(C) \tag{61}$$

3. Translation invariance: Lebesgue measure is translation invariant. $\mu(x + A) = \mu(A)$ for all $x \in \mathbb{R}^n$ on $\mathcal{B}(\mathbb{R}^n)$.

Definition 2.7 (Almost Everywhere)

Given a measure space $(X, \mathcal{A}, \mu)$, a subset $A \in \mathcal{A}$ is said to be a μ-null set if $\mu(A) = 0$. 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 2.2 (Rational Function)

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

Let us first look into some properties of measures, which all seem natural.

Theorem 2.10 ()

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

$$\mu\bigg(\bigcup_{k=1}^\infty A_k\bigg) = \lim_{k\rightarrow\infty} \mu(A_k) \tag{62}$$

Proof.

This is the first time we introduce limits. With the fact that $\mu(A_k)$ must be nondecreasing, we can use real analysis and see that it is bounded by ∞, meaning that it must have a limit. But why does this limit equal to the left hand side? We can see that

$$\mu\bigg(\bigcup_{k=1}^\infty A_k\bigg) = \mu(A_1) + \sum_{k=2}^\infty \mu(B_k) \tag{63}$$

$$= \mu(A_1) + \lim_{k\rightarrow\infty} \sum_{k=2}^\infty \mu(B_k) \tag{64}$$

$$= \lim_{k\rightarrow\infty} \mu(A_1 \cup B_2 \cup \ldots B_k) = \lim_{k\rightarrow\infty} \mu(A_k) \tag{65}$$

where $B_k = A_k \setminus A_{k-1}$.

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

Theorem 2.11 ()

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

$$\mu\bigg(\bigcap_{k=1}^\infty A_k\bigg) = \lim_{k\rightarrow\infty} \mu(A_k) \tag{66}$$

if $\mu(A_1) < \infty$.

Proof.

The $\mu(A_1) < \infty$ 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 ∞. 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

$$\begin{aligned}
\mu\bigg(\bigcap_{k=1}^\infty A_k\bigg) &= \mu\bigg(A_1 \setminus \bigcup_{k=1}^\infty B_k\bigg) \\
&= \mu(A_1) - \mu\bigg(\bigcup_{k=1}^\infty B_k\bigg) \\
&= \mu(A_1) - \sum_{k=1}^\infty \mu(B_k) \\
&= \mu(A_1) - \lim_{K\to\infty} \sum_{k=1}^K \mu(B_k) \\
&= \lim_{K\to\infty} \bigg(\mu(A_1) - \sum_{k=1}^K \mu(B_k)\bigg) \\
&= \lim_{K\to\infty} \mu\bigg(A_1 \setminus \bigcup_{k=1}^K B_k\bigg) = \lim_{K\to\infty} \mu(A_K)
\end{aligned}$$

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

3 Measurable Functions and Integration

Now that we've discussed measurability of sets, we need to talk about measurability of functions, and then we can integrate over them.

3.1 Measurable Functions

Definition 3.1 (Measurable Function)

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

$$f^{-1}(A) \in \mathcal{A} \text{ for all } A \text{ open} \tag{67}$$

where $f^{-1}(A)$ denotes the preimage of A.

Note that if we take $\mathbb{R}^n$, it can have either its Borel σ-algebra $\mathcal{B}(\mathbb{R}^n)$ or its Lebesgue σ-algebra $\mathcal{M}_{\lambda^*}$. Therefore, a function $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ is said to be **Lebesgue measurable** (**Borel measurable**) if for every $E \in \mathcal{B}(\mathbb{R})$, $f^{-1}(E) \in \mathcal{M}_{\lambda^*}$ ($f^{-1}(E) \in \mathcal{B}(\mathbb{R}^n)$). Since $\mathcal{B}(\mathbb{R}^n) \subset \mathcal{M}_{\lambda^*}$, all Borel measurable functions are Lebesgue measurable. It follows that any continuous function $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ is Borel (and hence Lebesgue measurable).

There are many ways to prove measurability, which we will list below.

Theorem 3.1 (TFAE)

The following are equivalent.

1. f is measurable
2. $f^{-1}(U) \in \mathcal{A}$ for all $U \in \mathcal{B}(\mathbb{R}$
3. $f^{-1}((-\infty, t)) \in \mathcal{A} \; \forall t \in \mathbb{R}$.

This immediately implies that monotonic functions on $\mathbb{R}$ are measurable. For example, take $f : [a, b] \longrightarrow \mathbb{R}$ that is nondecreasing. Then, we would like to show that the preimage of every half-interval $(-\infty, t)$ under f is in $\mathcal{B}(\mathbb{R})$. Well if we assume $f(a) \geq t$, then $f(x) > t \; \forall t \in [a, b]$, and so its preimage is $\emptyset$. If $f(a) < t$, having $f(b) < t$ also leads to the preimage being $[a, b]$ (which is the entire space and is in $\mathcal{B}(\mathbb{R})$), and having $f(b) > t$ implies that the preimage is $[a, f^{-1}(t)]$.

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

Theorem 3.2 (Sloppy Version)

Given measurable functions f, g, the following standard operations on them create new measurable functions:

1. $f + g$ is measurable
2. $f \cdot g$ is measurable
3. αf is measurable
4. f/g is measurable on $\{x \mid g(x) \neq 0\}$
5. $f \vee g \coloneqq \max(f, g)$ is measurable
6. $f \wedge g \coloneqq \min(f, g)$ is measurable

Theorem 3.3 ()

Given a sequence of measurable functions $f_1, f_2, \ldots$, we have

$$\lim_{k \to \infty} f_k \tag{68}$$

is measurable where it exists.

3.2 Simple Functions

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 consider a generalization of step functions called *simple functions*. A function will be Lebesgue integrable if it can be approximated by these simple functions in some appropriate way.

Definition 3.2 (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

$$\chi_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if else} \end{cases} \tag{69}$$

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

$$\phi = \sum_{i=1}^{n} a_i \chi_{A_i} \tag{70}$$

Lemma 3.1 (Measurability on Simple Functions)

Now, let $(X, \mathcal{A})$ be a measurable space. Then,

$$\phi = \sum_{i=1}^{n} a_i \chi_{A_i} : (X, \mathcal{A}) \longrightarrow \mathbb{R} \tag{71}$$

is measurable if all A_i are measurable, i.e. $A_i \in \mathcal{A}$ for all i.

Proof.

Let T be an open set in $\mathbb{R}$. Then, for characteristic function χ_A,

$$\chi_A^{-1}(T) = \begin{cases} \emptyset & \text{if } 0, 1 \not\in T \\ A & \text{if } 1 \in T, 0 \not\in T \\ X \setminus A & \text{if } 0 \in T, 1 \not\in T \\ X & \text{if } 0, 1 \in T \end{cases} \tag{72}$$

and so χ_A must be measurable if $A \in \mathcal{A}$ (which also by definition implies that $A^c = X \setminus A \in \mathcal{A}$). If χ_{A_i} is measurable, then the linear combination of measurable functions is also measurable.

Also observe that the coefficients need not be unique, since we can write

$$1 \cdot \chi_{[0,1]} + 1 \cdot \chi_{[0.5,1]} = 1 \cdot \chi_{[0,0.5]} + 2 \cdot \chi_{[0.5,1]} \tag{73}$$

If the E_i's are disjoint, then this decomposition is unique and is called the **standard representation** of ϕ.

Example 3.1 (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.

$$f = \sum_{i=1}^{n} c_i \chi_{(x_{i-1}, x_i)} + \sum_{j=0}^{n} f(x_j) \chi_{\{x_j\}} \tag{74}$$

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

$$\sum_{i=1}^{n} c_i \chi_{(x_{i-1}, x_i)} \tag{75}$$

If the A_i's above are just intervals in $\mathbb{R}$, then ϕ reduces to a step function. But the entire problem with intervals is that they are too coarse. We can't work with them, so we generalize them to all measurable sets in $(X, \mathcal{A})$. The Riemann integral is built on an approximation scheme of a function, which we usually want to be continuous to satisfy this approximation, and so, if we want to build an approximation scheme for Lebesgue integrals, we want a similar scheme, i.e. if we take a sequence of simple measurable functions, I can get arbitrarily close to any measurable function f. This is exactly what we show below.

Theorem 3.4 ()

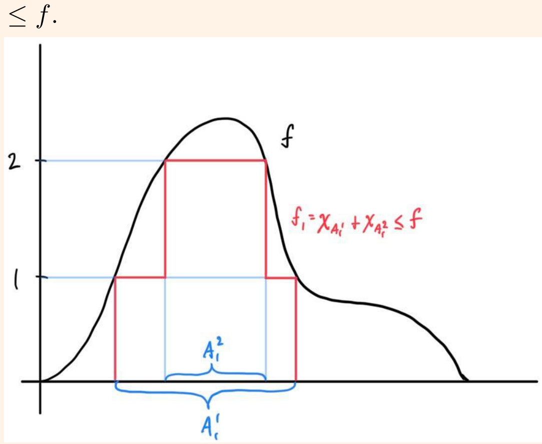
If $f : (X, \mathcal{A}) \longrightarrow [0, \infty]$ is measurable, there are simple measurable functions $f_k : (X, \mathcal{A}) \longrightarrow [0, \infty)$ s.t.

$$f_k \leq f_{k+1} \text{ and } f = \lim_{k \to \infty} f_k \tag{76}$$

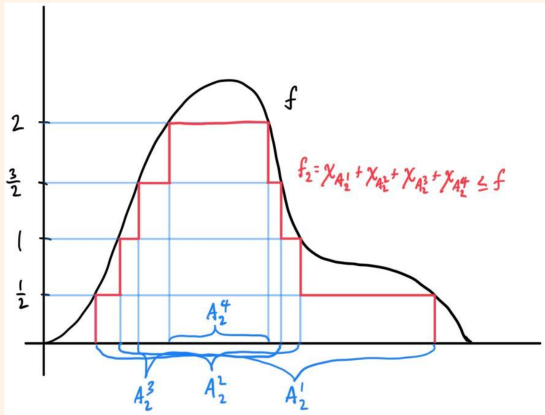
where the inequalities and limits are pointwise.

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

$$f_k = \sum_{j=1}^{k2^{k-1}} \frac{1}{2^{k-1}} \chi_{A_k^j} \tag{77}$$

3.3 Lebesgue Integral

Finally, we can learn how to integrate. We require the positiveness condition on f below because our previous theorem on approximating arbitrary functions with simple measurable functions f_k requires that it be positive, too.

Definition 3.3 (Lebesgue Integral of Positive Simple Functions)

If $f = \sum_{k=1}^{n} c_k \chi_{A_k}$ is a positive simple Lebesgue measurable function on measure space $(X, \mathcal{A}, \mu)$, then the **Lebesgue integral** of f is

$$\int f \, d\mu = \sum_{k=1}^{n} c_k \mu(A_k) \tag{78}$$

This 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

$$\int f(x) \, dx = \sum_{i=1}^{n} c_k |x_i - x_{i-1}| \tag{79}$$

The Lebesgue integral is

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

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.

Therefore, there are many possible values that can come out of a Lebesgue integral of a certain function, while a Riemann integral outputs only one value if exists.

Example 3.2 ()

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:

$$\int_{\mathbb{R}} \chi_{\mathbb{Q}\cap[0,1]} \, d\lambda = 1 \cdot \lambda(\mathbb{Q} \cap [0,1]) = 0 \tag{80}$$

Definition 3.4 (Lebesgue Integral on Positive Measurable Functions)

If $f : (X, \mathcal{A}, \mu) \longrightarrow [0, \infty]$ is measurable, then

$$\int_X f \, d\mu = \sup \bigg\{ \int g \, d\mu \, \bigg| \, g \text{ simple }, g \leq f \bigg\} \tag{81}$$

Unlike Riemann integration, which looks at both the supremum and infimum of integrals of simple functions, Lebesgue integration only looks at the supremum, given that f is nonnegative, so for all these f, the Lebesgue integral always exists. Defining Lebesgue integration for all real-valued functions, requires a simple extension.

Definition 3.5 (Lebesgue Integral)

Given a function $f : (X, \mathcal{A}, \mu) \longrightarrow \mathbb{R}$, we can split f into a positive and negative part:

$$f = f^+ - f^- \tag{82}$$

where $f^+ = \max(f, 0)$ and $f^- = \max(-f, 0)$. Then, the Lebesgue integral of f is

$$\int f \, d\mu = \int f^+ \, d\mu - \int f^- \, d\mu \tag{83}$$

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. It has the properties:

1. Monotonicity:
$$g \leq f \implies \int g \, d\mu \leq \int f \, d\mu \tag{84}$$
2. Scalar Multiplication:
$$\int cf \, d\mu = c \int f \, d\mu \tag{85}$$
3. Addition:
$$\int f + g \, d\mu = \int f \, d\mu + \int g \, d\mu \tag{86}$$

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

$$\int |f| \, d\mu < \infty \tag{87}$$

since by triangle inequality, we have

$$\left| \int f \, d\mu \right| \leq \int |f| \, d\mu \tag{88}$$

Definition 3.6 ()

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 3.5 ()

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

$$\int_{\mathbb{R}} f \, d\mu = 0 \tag{89}$$

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

$$\int_{\mathbb{R}} f \, d\mu = \int_{\mathbb{R}} g \, d\mu \tag{90}$$

3.4 Monotone Convergence Theory

From now on, we will assume that all spaces X are measure spaces $(X, \mathcal{A}, \mu)$ and all functions f are measurable functions. The huge problem with Riemann integrals is that this theorem doesn't hold, but it is the case for Lebesgue integration.

Theorem 3.6 (Monotone Convergene Theorem (MCT))

Given a nondecreasing sequence of measurable functions $f_1 \leq f_2 \leq f_3 \leq \ldots : X \longrightarrow [0, \infty]$, its limit $\lim_{k\to\infty} f_k$ always exists (since f_k is nondecreasing), is measurable, and

$$\int \lim_{k\to\infty} f_k \, d\mu = \lim_{k\to\infty} \int f_k \, d\mu \tag{91}$$

This allows us to integrate the limit of nice functions f_k by integrating these f_k first and then finding what the values converge to.

3.5 Riemann vs Lebesgue Integral

Theorem 3.7 ()

$f : \mathbb{R} \longrightarrow \mathbb{R}$ is Riemann integrable iff it is continuous λ almost everywhere. If so, then f is Lebesgue measurable and

$$\int_{[a,b]} f \, d\lambda = \int_a^b f \, dx \tag{92}$$

for all $a < b \in \mathbb{R}$.