Real Analysis

Hai Le **Professor**
LaTeX by Rui Zhou
University of Michigan

Winter 2024

Contents

Chapter 1

Riemann Integration

1.1 Introduction: Fourier Series

$$F(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} [a_n \cos nx + b_n \sin nx]$$

$$\begin{cases} a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} F(x) \cos nx \mathrm{d}x \\ b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} F(x) \sin nx \mathrm{d}x \end{cases}$$

In order to validate such results, the operations of $\int$ and $\sum$ for infinite terms should be interchangeable in computation order.

Justifying Fourier Analysis/Series under Riemann Integral is crazy. Alternatively, we would define as Lebesgue's Integration:

$$\int \sum^{\infty} = \sum^{\infty} \int$$

1.2 Riemann Integral

Definition 1.2.1: Partition

Suppose $a, b \in \mathbb{R}$ with $a < b$. A partition of $[a, b]$ is a finite list of the form $x_0, x_1, \cdots, x_n$, where $a = x_0 < x_1 < \cdots < x_n = b$.

A partition do not necessarily need to be an even partition.

Definition 1.2.2: Infimum and Supremum of a Function

If f is a real-valued function, A is a subset of the domain of f, then

$$\inf_A f := \inf \{f(x) : x \in A\}$$
$$\sup_A f := \sup \{f(x) : x \in A\}$$

Definition 1.2.3: Lower and Upper Riemann Sums

Suppose $f : [a, b] \to \mathbb{R}$ is a bounded function and P is a partition $x_0, x_1, \cdots, x_n$ of $[a, b]$.

- The lower Riemann sum $L(f, P, [a, b])$ is defined as

$$L(f, P, [a, b]) := \sum_{j=1}^{n} (x_j - x_{j-1}) \cdot \inf_{[x_{j-1}, x_j]} f$$

- The upper Riemann sum $U(f, P, [a, b])$ is defined as

$$U(f, P, [a, b]) := \sum_{j=1}^{n} (x_j - x_{j-1}) \cdot \sup_{[x_{j-1}, x_j]} f$$

The next result states that adjoining more points to a partition increases the lower Riemann sum and decreases the upper Riemann sum.

Proposition 1.2.4

Suppose $f : [a, b] \to \mathbb{R}$ is a bounded function and P, P' are partitions of $[a, b]$ such that the listing defining P is a sublist of the list defining P'. Then

$$L(f, P, [a, b]) \leq L(f, P', [a, b]) \leq U(f, P', [a, b]) \leq U(f, P, [a, b]).$$

The following result states that if the function is fixed, then each lower Riemann sum is less than or equal to each upper Riemann sum.

Proposition 1.2.5: Lower Riemann Sums No More Than Upper Riemann Sums

Suppose $f : [a, b] \to \mathbb{R}$ is a bounded function and P, P' are partitions of $[a, b]$. Then

$$L(f, P, [a, b]) \leq U(f, P', [a, b])$$

Now that we have been working with lower and upper Riemann sums, here we define the lower and upper Riemann integrals.

Definition 1.2.6: Lower and Upper Riemann Integrals

Suppose $f : [a, b] \to \mathbb{R}$ is a bounded function. The *lower Riemann integral* $L(f, [a, b])$ and the *upper Riemann integral* $U(f, [a, b])$ of f are defined by

$$L(f, [a, b]) = \sup_{P} L(f, P, [a, b])$$

$$U(f, [a, b]) = \inf_{P} U(f, P, [a, b])$$

In the definition above, we take the supremum (over all partitions) of the lower Riemann sums because adjoining more points to a partition increases the lower Riemann sum and should provide a more accurate estimate of the area under the graph. Similarly, in the definition above, we take the infimum (over all partitions) of the upper Riemann sums because adjoining more points to a partition decreases the upper Riemann sum and should provide a more accurate estimate of the area under the graph.

Our intuition suggests that for a partition with only a small gap between consecutive points, the lower Riemann sum should be a bit less than the area under the graph, and the upper Riemann sum should be a bit more than the area under the graph. Intuitively, the lower and upper Riemann Integrals will converge to each other, or converge to a definite number, as long as the partition goes more and more finely. Both of them are approximations of the area under f on $[a, b]$.

Remark.

> By definition we immediately have $L(f, [a, b]) \leq U(f, [a, b])$. Instead of choosing between the lower Riemann integral and the upper Riemann integral, the standard procedure in Riemann integration is to consider only functions for which those two quantities are equal. When we take equality to this relationship, that is, $L(f, [a, b]) = U(f, [a, b])$, we call f (a bounded function on a closed bounded interval) *Riemann integrable*.

Definition 1.2.7: Riemann Integrable, Riemann Integral

A bounded function on a closed bounded interval is called *Riemann integrable* if its lower Riemann integral equals its upper Riemann integral. If $f : [a, b] \to \mathbb{R}$ is Riemann integrable, then the *Riemann integral* $\int_a^b f$ is defined by

$$\int_a^b f = L(f, [a, b]) = U(f, [a, b])$$

Proposition 1.2.8: Continuity and Riemann Integrability

Every continuous real-valued function on each closed bounded interval is Riemann integrable.

Proposition 1.2.9: Bounds on Riemann Integral

Suppose $f : [a, b] \to \mathbb{R}$ is Riemann integrable. Then

$$(b-a) \inf_{[a,b]} f \leq \int_a^b f \leq (b-a) \sup_{[a,b]} f$$

1.3 Why is Riemann Integral NOT so good?

1. **Riemann integration does NOT handle functions with (too many) discontinuities.**

   Example.

   > Consider Dirichlet's function $f : [0,1] \to \mathbb{R}$:
   >
   > **Definition 1.3.1: Dirichlet Function**
   >
   > Dirichlet Function $f : [0,1] \to \mathbb{R}$ is defined as
   >
   > $$f(x) = \begin{cases} 1, & \text{if } x \text{ is rational} \\ 0, & \text{if } x \text{ is irrational} \end{cases}$$
   >
   > By definition, for $[a,b] \subseteq [0,1]$ with $a < b$, we have
   >
   > $$\begin{cases} \inf_{[a,b]} f = 0 \\ \sup_{[a,b]} f = 1 \end{cases} \implies \begin{cases} L(f,[0,1]) = 0 \\ U(f,[0,1]) = 1 \end{cases} \implies L(f,[0,1]) \neq U(f,[0,1])$$
   >
   > So we conclude that f is not Riemann integrable.
   >
   > Intuitively, since the set of rational numbers is countable and the set of irrational numbers is uncountable, $f(x)$ should be "not too" different from the function 0. In this way of reasoning, f should, in some sense, have integral 0. However, the Riemann integral of f is not defined.

2. **Riemann integration does not work well with unbounded functions.**

   Example.

   > Consider the following integration:
   >
   > $$\int_0^1 \frac{1}{\sqrt{x}} \mathrm{d}x$$
   >
   > Using the idea of improper integral, this gives an outcome of 2. However, say $x_0, x_1, \cdots, x_n$ is a partition of $[0,1]$, then $\sup_{[x_0,x_1]} f = \infty$. Thus if we tried to apply the definition of the upper Riemann sum to f, we would have $U(f,P,(0,1)) = \infty$ for every partition P of $[0,1]$.

3. **Riemann integration does NOT work well with pointwise limits.**

   Example.

   > Consider the Dirichlet's function. $\mathbb{Q}$ is countable. Let $r_1, r_2, \cdots$ be the sequence that indexes the rational numbers. For each positive integer k, we define f_k :

> $[0,1] \to \mathbb{R}$ such that
>
> $$f_k(x) = \begin{cases} 1, & \text{if } x \in \{r_1, \cdots, r_k\} \\ 0, & \text{otherwise} \end{cases}$$
>
> From Real analysis, we see that f_k is Riemann integrable, and $\int f_k = 0$ (which is a Riemann integral). However, when k goes to ∞, f_k converges to the Dirichlet's function, which is in contrast not Riemann integral.

Because analysis relies heavily upon limits, a good theory of integration should allow for interchange of limits and integrals, at least when the functions are appropri- ately bounded. Based on the three reasons, we hope to develop a new kind of integration that works well with the issues unsolved by the Riemann integration. We have to develop a new way to measure subsets of $\mathbb{R}$.

Chapter 2

Measure

Throughout this semester we will work with ”infinity” a lot. It's more convenient to think of ∞ as a number. Therefore, to begin with, we base our analysis on the extended real line:

$$[-\infty.\infty] = \mathbb{R} \cup \{-\infty, \infty\}$$

The underlying intuition is that, ∞ is recognized as a number instead of a never-reaching limit here.

The computations of ∞ follows that

$$\infty + \infty = \infty$$
$$\infty \cdot \infty = \infty$$

However, notice that $\infty - \infty$, ∞/∞ and $0 \cdot \infty$ is generally undefined.

2.1 Outer Measure

Our goal is to have a ”measure function” of subsets of $\mathbb{R}$. The primary question is that, what are the desirable properties?

Suppose $\mu : \mathcal{P}(\mathbb{R}) = 2^{\mathbb{R}} \to [-\infty, \infty]$, with $\mu(\mathbb{R}) = \infty$ (where $\mathcal{P}(\mathbb{R})$ is the set of all subsets of $\mathbb{R}$, notated as $2^{\mathbb{R}}$, or the power set of $\mathbb{R}$). We would wish μ to have the following properties:

Definition 2.1.1: Desired Properties of Outer Measure

1. $\mu(\varnothing) = 0$, $\mu(\{x\}) = 0$.
2. $\mu([a, b]) = b - a$.
3. "Linearity"
   - $\mu(A) = \mu(A + x)$, where $x \in \mathbb{R}$.
   - $\mu(\alpha A) = |\alpha| \cdot \mu(A)$.
4. $\mu(A) \leq \mu(B)$ if $A \subseteq B$. That is, measure preserves order.
5. $$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i) \text{ if } A_i\text{'s are pairwise disjoint.}$$

Remark.

1. The desired property of linearity guarantees the measure to be translation invariant, in which translation is defined as

   ### Definition 2.1.2: Translation

   If $t \in \mathbb{R}$ and $A \subseteq \mathbb{R}$, then the translation $t + A$ is defined by

   $$t + A := \{t + a : a \in A\}$$

2. Countable additivity is defined as follows:

   ### Definition 2.1.3: Countable Additivity

   Measure μ is said to have countable additivity if

   $$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i) \text{ if } A_i\text{'s are pairwise disjoint.}$$

   If we alternatively define additivity only in finite meanings, such limit operation may not be applicable.

Given all those desired properties of a measure, we would depart with "outer measure", which is at first glance intuitive and then proves to be a useful tool. We would develop the theoretical works by checking its properties.

Definition 2.1.4: Length of Open Interval

The length $l(I)$ of an interval I is defined by

$$l(I)=\begin{cases} b-a, \text{ if } I=(a,b) \text{ for some } a,b\in\mathbb{R} \text{ with } a<b \\ 0, \text{ if } I=\varnothing \\ \infty, \text{ if } I=(-\infty,a) \text{ or } I=(a,\infty) \text{ for some } a\in\mathbb{R} \\ \infty, \text{ if } I=(-\infty,\infty) \end{cases}$$

Suppose $A\subseteq\mathbb{R}$. The size of A should be at most the sum of the lengths of a sequence of open intervals whose union contains A. Taking the infimum of all such sums gives a reasonable definition of the size of A, denoted $\mu^*(A)$ and called the outer measure of A.

Definition 2.1.5: Outer Measure

The *outer measure* $\mu^*(A)$ of a set $A\subseteq\mathbb{R}$ is defined by

$$\mu^*(A)=\inf\left\{\sum_{k=1}^{\infty} l(I_k): I_1,I_2,\cdots \text{ are open intervals such that } A\subseteq\bigcup_{k=1}^{\infty} I_k\right\}$$

Intuitively, $\mu^*(A)$ corresponds to the length of the smallest open covering of A.

The definition of outer measure is well-defined, that is, for all $A\subset\mathbb{R}$ the outer measure returns a number. One simple argument is $A\subset\bigcup_{n=1}^{\infty}(-n,n), \forall A\subset\mathbb{R}$, so there are at least one element in the operation of inf, which makes sense.

Proposition 2.1.6

Finite Sets Have Zero Outer Measure

Proof for Proposition.

Suppose $A=\{a_1,\cdots,a_n\}$ is a finite set of real numbers. Suppose $\varepsilon>0$. Define a sequence $I_1,I_2,\cdots$ of open intervals by

$$I_k=\begin{cases}(a_k-\varepsilon,a+\varepsilon) & \text{if } k\le n\\ \varnothing & \text{if } k>n\end{cases}$$

Then $I_1,I_2,\cdots$ is a sequence of open intervals whose union contains A. Clearly $\sum_{k=1}^{\infty} l(I_k)=2\varepsilon n$. Hence $|A|\le 2\varepsilon n$. Because ε is an arbitrary positive number, this implies that $|A|=0$. ∎

Proposition 2.1.7

Every countable subset of $\mathbb{R}$ has outer measure 0.

Proof for Proposition.

By definition, $\mu^*(A) = \inf\{\sum l(I_n),\ \text{where } A \subset (\cup I_n)\}$. Construct $I_n = \left(a_n - \frac{\varepsilon}{2^n}, a_n + \frac{\varepsilon}{2^n}\right)$. Clearly, $A \subseteq \bigcup_{n=1}^{\infty} I_n$. According to the definition,

$$\begin{aligned}\mu^*(A) &\leq \sum_{n=1}^{\infty} l(I_n) \\ &= \sum_{n=1}^{\infty} 2 \cdot \frac{\varepsilon}{2^n} = \varepsilon \sum_{n=1}^{\infty} \frac{1}{2^{n-1}} = 2\varepsilon \to 0\end{aligned}$$

Therefore, for any $\varepsilon > 0$ small, we have $0 \leq \mu^*(A) \leq 0$; immediately we have $\mu^*(A) = 0$.

Since finite sets are countable, and every countable subset of $\mathbb{R}$ has zero outer measure, it follows that finite sets have outer measure 0. The former is the immediate corollary of the more general latter.

Proposition 2.1.8: Outer Measure Preserves Order

Suppose A and B are the subsets of $\mathbb{R}$ with $A \subseteq B$. Then $|A| \leq |B|$.

Proof for Proposition.

Suppose $I_1, I_2, \cdots$ is a sequence of open intervals whose union contains B. Then the union of this sequence of open intervals also contains A. Hence,

$$\mu^*(A) \leq \sum_{k=1}^{\infty} l(I_k)$$

Taking the infimum over all sequences of open intervals whose union contains B, we have $\mu^*(A) \leq \mu^*(B)$.

Moreover, we expect that the size of a subset of $\mathbb{R}$ should not change if the set is shifted to the right or to the left. The next definition allows it to be more precise.

Definition 2.1.9: Translation

If $t \in \mathbb{R}$ and $A \subseteq \mathbb{R}$, then the *translation* $t + A$ is defined by

$$t + A = \{t + a : a \in A\}$$

Obviously, translation does not change the length of an open interval. The next result states that translation invariance carries over to outer measure.

Proposition 2.1.10: Outer Measure is Translation Invariant

Suppose $t \in \mathbb{R}$ and $A \subseteq \mathbb{R}$. Then

$$\mu^*(t + A) = \mu^*(A).$$

Proof for Proposition.

Suppose $I_1, I_2, \cdots$ is a sequence of open intervals whose union contains A. Then $t + I_1, t + I_2, \cdots$ is a sequence of open intervals whose union contains $t + A$. Thus

$$\mu^*(t + A) \leq \sum_{k=1}^{\infty} l(t + I_k) = \sum_{k=1}^{\infty} l(I_k)$$

Taking the infimum of the last term over all sequences $I_1, I_2, \cdots$ of open intervals whose union contains A, we have $\mu^*(t + A) \leq \mu^*(A)$.

To get the inequality in the other direction, note that $A = -t + (t + A)$, so we have $\mu^*(A) = \mu^*(-t + (t + A)) \leq \mu^*(t + A)$. Hence $\mu^*(t + A) = \mu^*(A)$.

Proposition 2.1.11: Countable Subadditivity of Outer Measure

Suppose $A_1, A_2, \cdots$ is a sequence of subsets of $\mathbb{R}$, then

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

Remark.

Countable subadditivity implies finite subadditivity, meaning that

$$\mu^*(A_1 \cup \cdots \cup A_n) \leq \mu^*(A_1) + \cdots + \mu^*(A_n)$$

for all $A_1, \cdots, A_n \in \mathbb{R}$, because we can take $A_k = \varnothing$ for $k > n$ in the inequality of countable subadditivity.

Proposition 2.1.12: Outer Measure of Closed Bounded Interval

Suppose $a, b \in \mathbb{R}$ with $a < b$, then

$$\mu^*([a, b]) = b - a.$$

Proof for Proposition.

If $\varepsilon > 0$, then $(a - \varepsilon, b + \varepsilon), \varnothing, \varnothing, \cdots$ is a sequence of open intervals whose union contains $[a, b]$. Thus $|[a, b]| \leq b - a + 2\varepsilon$. Because this inequality holds for all $\varepsilon > 0$, we conclude that $|[a, b]| \leq b - a$.

A proof of the inequality in the opposite direction requires that the completeness of $\mathbb{R}$ is used in some form. For example, suppose $\mathbb{R}$ was a countable set. Then we would have $|[a, b]| = 0$ since the outer measure of any countable set is 0. Thus something deeper is going on with the ingredients needed to prove that $|[a, b]| \geq b - a$. The following definition and Heine-Borel theorem would be useful.

Definition 2.1.13: Open Cover; Finite Subcover

Suppose $A \subseteq \mathbb{R}$.

- A collection of $\mathcal{C}$ of open subsets of $\mathbb{R}$ is called an *open cover* of A if A is contained in the union of all the sets in $\mathcal{C}$.
- An open cover of $\mathcal{C}$ of A is said to have a *finite subcover* if A is contained in the union of some *finite* list of sets in $\mathcal{C}$.

Example.

The collection $\{(k, k+2) : k \in \mathbb{Z}^+\}$ is an open cover of $[2, +\infty)$ because $[2, +\infty) \subseteq \bigcup_{k=1}^{\infty} (k, k+2)$. This open cover does not have a finite subcover because there do not exist finitely many sets of the form $(k, k+2)$ whose union contains $[2, +\infty)$.

Theorem 2.1.14: Heine-Borel Theorem

Every open cover of a **closed bounded** subset of $\mathbb{R}$ has a finite subcover.

Proof for Proposition. (Cont.)

To prove the inequality in the other direction, suppose $I_1, I_2, \cdots$ is a sequence of open intervals such that $[a, b] \subseteq \bigcup_{k=1}^{\infty} I_k$. By the Heine-Borel theorem, there exists $n \in \mathbb{Z}^+$ such that

$$[a, b] \subseteq I_1 \cup \cdots \cup I_n$$

We will now prove by induction that the inclusion above implies that

$$\sum_{k=1}^{n} l(I_k) \geq b - a$$

This will then imply that $\sum_{k=1}^{\infty} l(I_k) \geq \sum_{k=1}^{n} l(I_k) \geq b - a$, completing the proof that $|\,[a, b]\,| \geq b - a$.

To get started with our induction, note that in the case of $n = 1$ such relationship stands. Now for the induction step: Suppose $n > 1$ and $[a, b] \subseteq I_1 \cup \cdots \cup I_n$ implies $\sum_{k=1}^{n} l(I_k) \geq b - a$ for all choices of $a, b \in \mathbb{R}$ with $a < b$. Suppose $I_1, \cdots, I_n, I_{n+1}$ are open intervals such that

$$[a, b] \subseteq I_1 \cup \cdots \cup I_n \cup I_{n+1}$$

Thus b is in at least one of the intervals $I_1, \cdots, I_n, I_{n+1}$. By rebelling, we can assume that $b \in I_{n+1}$. Suppose $I_{n+1} = (c, d)$. If $c < a$, then $l(I_{n+1}) \geq b - a$ and there is nothing further to prove; thus we can assume that $a < c < b < d$. Hence, $[a, c] \subseteq I_1 \cup \cdots \cup I_n$.

By our induction hypothesis, we have $\sum_{k=1}^{n} l(I_k) \geq c - a$. Thus,

$$\begin{aligned}\sum_{k=1}^{n+1} l(I_k) &= \sum_{k=1}^{n} l(I_k) + l(I_{n+1}) \\ &\geq (c-a) + (d-c) \\ &= d - a \\ &\geq b - a\end{aligned}$$

which immediately completes the proof. So we conclude that $|[a,b]| \geq b-a$. ∎

Corollary 2.1.15: Nontrivial Intervals are Uncountable

Every interval in $\mathbb{R}$ that contains at least two distinct elements is uncountable.

Proof for Corollary.

Suppose I is an interval that contains $a, b \in \mathbb{R}$ with $a < b$. Then

$$\mu^*(I) \geq \mu^*([a,b]) = b - a > 0$$

where the first inequality above holds because outer measure preserves order, and the equality above comes from the property that the outer measure of each open interval equals its length. Because every countable subset of $\mathbb{R}$ has outer measure 0, we can conclude that I is uncountable. ∎

2.2 Measurable Spaces and Functions

Proposition 2.2.1: Nonexistence of Extension of Length to All Subsets of $\mathbb{R}$

There does not exist a function μ with all the following properties:

1. μ is a function from the set of subsets of $\mathbb{R}$ to $[0,\infty]$ ($\mu : \mathcal{P}(\mathbb{R}) = 2^{\mathbb{R}} \to [0,\infty]$).
2. $\mu(I) = l(I)$ for every open interval I of $\mathbb{R}$.
3. $\mu\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} \mu(A_k)$ for every disjoint sequence $A_1, A_2, \cdots$ of subsets of $\mathbb{R}$.
4. $\mu(t + A) = \mu(A)$ for every $A \subseteq \mathbb{R}$ and every $t \in \mathbb{R}$.

The last result shows that we need to give up one of the desirable properties in our goal of extending the notion of size from intervals to more general subsets of $\mathbb{R}$. We cannot give up (2) because the size of an interval needs to be its length. We cannot give up (3) because countable additivity is needed to prove theorems about limits. We cannot give up (4) because a size that is not translation invariant does not satisfy our intuitive notion of size as a generalization of length. Thus, we are forced to relax the requirement in (1) that the size is defined for all subsets of $\mathbb{R}$. Experience shows that to have a viable theory that allows for taking limits, the collection of subsets for which the size is defined should be *closed under complementation and closed under countable unions*. Thus we make the following definition of σ-algebra.

2.2.1 σ-Algebra

Definition 2.2.2: σ-Algebra

Suppose X is a set and $\mathcal{S}$ is a set of subsets of X. Then $\mathcal{S}$ is called a σ*-algebra* on X if the following three statements are satisfied:

1. $\varnothing \in \mathcal{S}$. (Notice the notation of $\in$, an element inclusion instead of a set inclusion.)
2. If $E \in \mathcal{S}$, then the complement $X \backslash E \in \mathcal{S}$.
3. If $E_1, E_2, \cdots$ is a sequence of elements in $\mathcal{S}$, then $\bigcup_{k=1}^{\infty} E_k \in \mathcal{S}$.

To verify $\mathcal{S}$ is indeed a σ-algebra, it is obvious for the first two bullet points. For the third one, you need to use the result that the countable union of countable sets is countable.

Example.

1. Trivial σ-algebra: X is any set, $S = \{\varnothing, X\}$.
2. Discrete/Full σ-algebra: X is any set, $S = \mathcal{P}(X) = 2^X$ (the power set of X, set of all subsets of X).
3. Suppose X is a set. Then the set of all subsets E of X such that E is countable or $X \backslash E$ is a σ-algebra on X.
4. $X = \{a, b, c, d\}$, $S = \{\varnothing, \{a, b, c, d\}, \{a, b\}, \{c, d\}\}$.

Proposition 2.2.3: σ-Algebras are Closed Under Countable Intersection

Suppose $\mathcal{S}$ is a σ-algebra on a set X. Then

- $X \in \mathcal{S}$;
- If $D, E \in \mathcal{S}$, then $D \cup E \in \mathcal{S}$ and $D \cap E \in \mathcal{S}$ and $D \backslash E \in \mathcal{S}$;
- If $E_1, E_2, \cdots$ is a sequence of elements of $\mathcal{S}$, then $\bigcap_{k=1}^{\infty} E_k \in \mathcal{S}$.

Informally, we think of σ-algebra as a collection of sets that are ***closed under intersection, union and complement.***

2.2.2 Measurable Spaces

The word *measurable* is used in the terminology below because in the next section we introduce a size function, called a *measure*, defined on *measurable sets*.

Definition 2.2.4: Measurable Space; Measurable Set

A *measurable space* is an ordered pair (X, S) where X is a set and $\mathcal{S}$ is a σ-algebra on X. An element of $\mathcal{S}$ is called an $\mathcal{S}$*-measurable* set, or just a *measurable set* if $\mathcal{S}$ is clear from the context.

In every discussion you should point out which σ-algebra you are currently working on.

Example.

If $X = \mathbb{R}$ and $\mathcal{S}$ is the set of all subsets of $\mathbb{R}$ that are countable or have a countable complement, then the set of rational numbers is $\mathcal{S}$-measurable but the set of positive real numbers is not $\mathcal{S}$-measurable.

Now that the full σ-algebra is not a "good" collection of $\mathbb{R}$, what is a "good" σ-algebra? The next result guarantees that there is a smallest σ-algebra on a set X containing a given set $\mathcal{A}$ of subsets of X, which will then pave way for the introduction of Borel set.

Proposition 2.2.5: Smallest σ-Algebra Containing a Collection of Subsets

Suppose X is a set and $\mathcal{A}$ is a set of subsets of X. Then the intersection of all σ-algebras on X that contain $\mathcal{A}$ is a σ-algebra on X.

Example.

- Suppose X is a set and $\mathcal{A}$ is the set of subsets of X that consist of exactly one element:

$$\mathcal{A} = \{\{x\} : x \in X\}$$

  Then the smallest σ-algebra on X containing $\mathcal{A}$ is the set of all subsets E of X such that E is countable or $X \backslash E$ is countable.

- Suppose $\mathcal{A} = \{(0,1), (0,\infty)\}$. Then the smallest σ-algebra on $\mathbb{R}$ containing $\mathcal{A}$ is

$$\{\varnothing, (0,1), (0,\infty), (-\infty,0] \cup [1,\infty), [1,\infty), (-\infty,1), \mathbb{R}\}$$

Definition 2.2.6: Borel Set

The smallest σ-algebra on $\mathbb{R}$ containing all *open subsets* of $\mathbb{R}$ is called the collection of *Borel subsets* of $\mathbb{R}$. An element of this σ-algebra is called a *Borel set*.

We have defined the collection of Borel subsets of $\mathbb{R}$ to be the smallest σ-algebra on $\mathbb{R}$ containing all the open subsets *of $\mathbb{R}$. We could have defined the collection of Borel subsets of $\mathbb{R}$ to be the smallest σ-algebra on $\mathbb{R}$ containing all the* open intervals*, because every open subset of $\mathbb{R}$ is the union of a sequence of open intervals.*

Example.

- Every closed subset of $\mathbb{R}$ is a Borel set because every closed subset of $\mathbb{R}$ is the complement of an open subset of $\mathbb{R}$.
- Every countable subset of $\mathbb{R}$ is a Borel set because if $B = \{x_1, x_2, \cdots\}$. then $B = \bigcup_{k=1}^{\infty} \{x_k\}$, which is a Borel set because each $\{x_k\}$ is a closed set.
- Every half-open interval $[a, b)$ (where $a, b \in \mathbb{R}$) is a Borel set because $[a, b) = \bigcap_{k=1}^{\infty} \left(a - \frac{1}{k}, b\right)$.
- If $f : \mathbb{R} \to \mathbb{R}$ is a function, then the set of points at which f is continuous is the intersection of a sequence of open sets, and thus is a Borel set.

We may informally say that any subset of $\mathbb{R}$ that you can think of is (highly-probably) a Borel set.

2.2.3 Inverse Images

Definition 2.2.7: Inverse Image

If $f : X \to Y$ is a function and $A \subseteq Y$, then the inverse image of A (which is a set), denoted as $f^{-1}(A)$, is defined by

$$f^{-1}(A) = \{x \in X : f(x) \in A\}$$

Example.

Suppose $f : \mathbb{R} \to \mathbb{R},\ f(x) = x^2$.

- $f^{-1}((1, 4)) = (-2, -1) \cup (1, 2)$.
- $f^{-1}((-1, -1)) = (-1, 1)$, since $f^{-1}((-1, -1)) = f^{-1}([0, 1))$.

Inverse images work very well with set operations.

Proposition 2.2.8: Algebra of Inverse Images

Suppose $f : X \to Y$ is a function. Then

1. $f^{-1}(Y \backslash A) = X \backslash f^{-1}(A)$, for $A \subseteq Y$. That is, $f^{-1}\left(A^C\right) = \left(f^{-1}(A)\right)^C$.
2. $f^{-1}\left(\bigcup_{A \in \mathcal{A}} A\right) = \bigcup_{A \in \mathcal{A}} f^{-1}(A)$, for every set $\mathcal{A}$ of subsets of Y.
3. $f^{-1}\left(\bigcap_{A \in \mathcal{A}} A\right) = \bigcap_{A \in \mathcal{A}} f^{-1}(A)$, for every set $\mathcal{A}$ of subsets of Y.

Remark.

To prove $A = B$, a "trick" is usually employed:

$$\begin{aligned} & A = B \\ \iff & A \subseteq B \text{ and } B \subseteq A \\ \iff & \begin{cases} x \in A \implies x \in B \\ x \in B \implies x \in A \end{cases} \end{aligned}$$

Similarly, to prove $a = b$ is equivalent to prove $a \leq b$ and $b \leq a$.

Proposition 2.2.9: Inverse Image of a Composition

Suppose $f : X \to Y$ and $g : Y \to W$. Then

$$(g \circ f)^{-1}(A) = f^{-1}\left(g^{-1}(A)\right)$$

for every $A \subseteq W$.

2.2.4 Measurable Functions

The next definition tells us which real-valued functions behave reasonably with respect to a σ-algebra on their domain.

Definition 2.2.10: Measurable Function

Suppose $(X, \mathcal{S})$ is a measurable space. A function $f : X \to \mathbb{R}$ is called $\mathcal{S}$-measurable if $f^{-1}(B) \in \mathcal{S}$ for all Borel set $B \subseteq \mathbb{R}$.

Remark.

- This definition is a special case of a general class of measurable function. In general, suppose $(X, \mathcal{S})$ is a measurable space, and a function $f : X \to Y$ is called $\mathcal{T}$-measurable if $f^{-1}(B) \in \mathcal{S}$ for all $B \in \mathcal{T}$, where $\mathcal{T}$ is a σ-algebra of Y.
- In fact, most of the time we will work with real-valued functions, but the works can be extended to more general settings. Real-valued functions are the simplest to study

with.

Example.

Consider a set X.

- If $\mathcal{S} = \{\varnothing, X\}$, then the only $\mathcal{S}$-measurable functions $f : X \to \mathbb{R}$ are constant functions.
- If $S = \mathcal{P}(X)$, then every function is $\mathcal{S}$-measurable.
- $X = \mathbb{R}$. If $\mathcal{S} = \{\varnothing, (-\infty, 0), [0, \infty), \mathbb{R}\}$, then $f : \mathbb{R} \to \mathbb{R}$ is $\mathcal{S}$-measurable if and only if f is constant on $(-\infty, 0)$ and f is constant on $[0, +\infty)$.
- Suppose E is a subset of a set X. The *characteristic function* of E is the function $\chi_E : X \to \mathbb{R}$ defined by $\chi_E(x) = \begin{cases} 1 & \text{if } x \in E \\ 0 & \text{if } x \notin E \end{cases}$. Suppose $(X, \mathcal{S})$ is a measurable space, $E \subseteq X$ and $B \subseteq \mathbb{R}$. Then

$$\chi_E{}^{-1}(B) = \begin{cases} E & \text{if } 0 \notin B \text{ and } 1 \in B \\ X \backslash E & \text{if } 0 \in B \text{ and } 1 \notin B \\ X & \text{if } 0 \in B \text{ and } 1 \in B \\ \varnothing & \text{if } 0 \notin B \text{ and } 1 \notin B \end{cases}$$

  Thus we see that χ_E is an $\mathcal{S}$-measurable function if and only if $E \in \mathcal{S}$.

Next, we are delving into the conditions for a function to be measurable. Roughly speaking, we have the following claim:

Claim

- Every ”well-behaved” function is measurable.
- Algebraic operations preserve measurability.
- $\lim_{k\to\infty} f_k = f$ is measurable.

Weaker Condition for Measurability The definition of an $\mathcal{S}$-measurable function requires the inverse image of every Borel subset of $\mathbb{R}$ to be in $\mathcal{S}$. The following useful result shows that to verify that a function is $\mathcal{S}$-measurable, we can check the inverse images of a *much smaller* collection of subsets of $\mathbb{R}$.

Proposition 2.2.11: Condition for Measurable Function

Suppose $(X, \mathcal{S})$ is a measurable space, and $f : X \to \mathbb{R}$ is a function that satisfies:

$$f^{-1}((a, \infty)) = \{x \in X : f(x) > a\} \in \mathcal{S}$$

for all $a \in \mathbb{R}$. Then f is an $\mathcal{S}$-measurable function.

Proof for Proposition.

Recall the properties of f^{-1}:

- $f^{-1}(\varnothing) = \varnothing$.
- If $A \subseteq \mathbb{R}$, then $f^{-1}(\mathbb{R}\backslash A) = X \backslash f^{-1}(A)$.
- If $A_1, A_2, \cdots \subseteq \mathbb{R}$, then $f^{-1}\left(\bigcup_{i=1}^{\infty} A_i\right) = \bigcup_{i=1}^{\infty} f^{-1}(A_i)$.

In some sense, the three properties correspond to those of σ-algebra. Define

$$\mathcal{T} = \left\{A \subseteq \mathbb{R} : f^{-1}(A) \in \mathcal{S}\right\}$$

We hope to show that every Borel subset of $\mathbb{R}$ is in $\mathcal{T}$. To achieve this, we will first show that $\mathcal{T}$ is a σ-algebra on $\mathbb{R}$.

First because $f^{-1}(\varnothing) = \varnothing \in \mathcal{S}$, we have $\varnothing \in \mathcal{T}$.

If $A \in \mathcal{T}$, then $f^{-1}(A) \in \mathcal{S}$; hence

$$f^{-1}(\mathbb{R}\backslash A) = X \backslash f^{-1}(A) \in \mathcal{S}$$

The last element inclusion relationship stands since $\mathcal{S}$ is a σ-algebra. Therefore, $\mathbb{R}\backslash A \in \mathcal{T}$. In other words, $\mathcal{T}$ is closed under complementation.

If $A_1, A_2, \cdots \in \mathcal{T}$, then $f^{-1}(A_1), f^{-1}(A_2), \cdots \in \mathcal{S}$; hence

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

The last element inclusion relationship stands since $\mathcal{S}$ is a σ-algebra. Therefore, $\bigcup_{k=1}^{\infty} A_k \in \mathcal{T}$. In other words, $\mathcal{T}$ is closed under countable unions.

Based on the above three properties, $\mathcal{T}$ is a σ-algebra on $\mathbb{R}$.

We note that

- $(a, b] = (a, \infty) \setminus (b, \infty)$, for all $a < b$.
- $(a, b) = \bigcup_{k=1}^{\infty} (a, b - \frac{1}{k}]$.

By hypothesis, $\mathcal{T}$ contains $\{(a, \infty) : a \in \mathbb{R}\}$. Because $\mathcal{T}$ is closed under complementation, $\mathcal{T}$ also contains $\{(-\infty, b] : b \in \mathbb{R}\}$. Because the σ-algebra $\mathcal{T}$ is closed under finite intersections, we see that $\mathcal{T}$ contains $\{(a, b] : a, b \in \mathbb{R}\}$. Because $(a, b) = \bigcup_{k=1}^{\infty} (a, b - \frac{1}{k}]$ and $(-\infty, b) = \bigcup_{k=1}^{\infty} (-k, b - \frac{1}{k}]$ and $\mathcal{T}$ is closed under countable unions, we can conclude that $\mathcal{T}$ contains every open subset of $\mathbb{R}$. Thus, the σ-algebra $\mathcal{T}$ contains the smallest

σ-algebra on $\mathbb{R}$ that contains all open subsets of $\mathbb{R}$. In other words, $\mathcal{T}$ contains every Borel subset of $\mathbb{R}$. Thus f is an $\mathcal{S}$-measurable function.

By construction we see that, in the result above, we could replace the collection of sets $\{(a,\infty) : a \in \mathbb{R}\}$ by any collection of subsets of $\mathbb{R}$ such that the smallest σ-algebra containing that collection contains the Borel subsets of $\mathbb{R}$. ■

Borel Measurable Functions We have been dealing with $\mathcal{S}$-measurable functions from X to $\mathbb{R}$ in the context of an arbitrary set X and a σ-algebra $\mathcal{S}$ on X. An important special case of this setup is when X is a Borel subset of $\mathbb{R}$ and $\mathcal{S}$ is the set of Borel subsets of $\mathbb{R}$ that are contained in X. In this special case, the $\mathcal{S}$-measurable functions are called *Borel measurable.*

Definition 2.2.12: Borel Measurable Function

Suppose $X \subseteq \mathbb{R}$ is a Borel set. A function $f : X \to \mathbb{R}$ is called *Borel measurable* if $f^{-1}(B)$ is a Borel set for every Borel set $B \subseteq \mathbb{R}$.

Remark.

- If $X \subseteq \mathbb{R}$ and there exists a Borel measurable function $f : X \to \mathbb{R}$, then X must be a Borel set since $X = f^{-1}(\mathbb{R})$.
- If $X \subseteq \mathbb{R}$ and $f : X \to \mathbb{R}$ is a function, then f is a Borel measurable function if and only if $f^{-1}((a,\infty))$ is a Borel set for every $a \in \mathbb{R}$.

Continuity and Borel Measurability The next result states that continuity interacts well with the notion of Borel measurability.

Proposition 2.2.13: Every Continuous Function is Borel Measurable

Every continuous real-valued function defined on a Borel subset of $\mathbb{R}$ is a Borel measurable function.

Proof for Proposition.

Suppose $X \subseteq \mathbb{R}$ is a Borel set and $f : X \to \mathbb{R}$ is continuous. To prove that f is Borel measurable, fix $a \in \mathbb{R}$. If $x \in X$ and $f(x) > a$, then by the continuity of f, there exists some $\delta_x > 0$ such that $f(y) > a$ for all $y \in (x-\delta_x, x+\delta_x) \cap X$. Thus

$$f^{-1}((a,\infty)) = \left(\bigcup_{x \in f^{-1}((a,\infty))} (x-\delta_x, x+\delta_x) \right) \cap X$$

The union inside the large parentheses above is an open subset of $\mathbb{R}$; hence its intersection with X is a Borel set. Thus we can conclude that $f^{-1}((a,\infty))$ is a Borel set. From the previous result we conclude that f is a Borel measurable function. ■

Another way to see this result is using the theorem that, a function f is continuous if and only if the inverse image of an open set is an open set. An immediate corollary is that (combined with the properties of inverse image), the inverse image of a Borel set is a Borel set, which completes our proof.

Increasing Function and Borel Measurability Next we come to another class of Borel measurable functions. A similar definition could be made for decreasing functions, with a corresponding similar result.

Definition 2.2.14: Increasing Function

Definition (Increasing) Suppose $X \subseteq \mathbb{R}$ and $f : X \to \mathbb{R}$ is a function.

- f is called *increasing* if $f(x) \leq f(y)$ for all $x, y \in X$ with $x < y$.
- f is called *strictly increasing* if $f(x) < f(y)$ for all $x, y \in X$ with $x < y$.

Proposition 2.2.15: Every Increasing Function is Borel Measurable

Every increasing function defined on a Borel subset of $\mathbb{R}$ is a Borel measurable function.

Proof for Proposition.

Suppose $X \subseteq \mathbb{R}$ is a Borel set and $f : X \to \mathbb{R}$ is increasing. To prove that f is Borel measurable, fix $a \in \mathbb{R}$. Let $b = \inf f^{-1}((a, \infty))$, then $f^{-1}((a, \infty)) = (b, \infty) \cap X$, or $f^{-1}((a, \infty)) = [b, \infty) \cap X$. Either way, we can conclude that $f^{-1}((a, \infty))$ is a Borel set, which implies that f is a Borel measurable function. ∎

We use the result that, since f is an increasing function, $b = \inf f^{-1}((a, \infty))$ is well-defined; that is, the infimum always exists.

Composition and Measurability The next result shows that measurability interacts well with composition. (Notice that here we are back discussing the general case of measurability.)

Proposition 2.2.16: Composition of Measurable Functions is Measurable

Suppose $(X, \mathcal{S})$ is a measurable space, and $f : X \to \mathbb{R}$ is an $\mathcal{S}$-measurable function. Suppose g is a real-valued Borel measurable function defined on a subset of $\mathbb{R}$ that includes the range of f. Then $g \circ f : X \to \mathbb{R}$ is $\mathcal{S}$-measurable.

Proof for Proposition.

Suppose $B \in \mathbb{R}$ is a Borel set. Then using the property of inverse image of composition:

$$(g \circ f)^{-1}(B) = f^{-1}\left(g^{-1}(B)\right)$$

Because g is a Borel measurable function, $g^{-1}(B)$ is a Borel subset of $\mathbb{R}$. Because f is an $\mathcal{S}$-measurable function, $f^{-1}\left(g^{-1}(B)\right) \in \mathcal{S}$. Thus the equation above implies that $(g \circ f)^{-1}(B) \in \mathcal{S}$. Thus $g \circ f$ is an $\mathcal{S}$-measurable function. ∎

Corollary 2.2.17

If f and g are continuous, then $g \circ f$ is measurable.

Example.

If f is measurable, then so are $-f, \frac{1}{2}f, |f|, f^2, \exp(f)$, etc. (Because each of these functions can be written as the composition of f with a continuous and thus Borel measurable function g.)

Algebraic Operations and Measurability Measurability also interacts well with algebraic operations.

Proposition 2.2.18: Algebraic Operations with Measurable Functions

Suppose $(X, \mathcal{S})$ is a measurable space and $f, g : X \to \mathbb{R}$ are $\mathcal{S}$-measurable. Then

- $f+g$, $f-g$, and fg are $\mathcal{S}$-measurable functions;
- If $g(x) \neq 0$ for all $x \in X$, then $\dfrac{f}{g}$ is an $\mathcal{S}$-measurable function.

Proof for Proposition.

In order to prove $f+g$ is an $\mathcal{S}$-measurable function, by definition we are to show for any $a \in \mathbb{R}$:

$$(f+g)^{-1}((a, \infty)) \in \mathcal{S}$$

The intuition begins with the observation that, since both f and g are $\mathcal{S}$-measurable, if we can find a way (of countable unions) to represent the inverse image of $f+g$, then we are done.

$$(f+g)^{-1}((a, \infty)) = \bigcup_{r \in} \left(f^{-1}((r, \infty)) \cup g^{-1}((a-r, \infty))\right)$$

In order to prove the equivalence of sets, we use the trick: $A = B \iff A \subseteq B$ and $B \subseteq A$.

If $x \in \bigcup_{r \in} \left(f^{-1}((r, \infty)) \cup g^{-1}((a-r, \infty))\right)$, this implies that there exists an $r \in$ such that $x \in f^{-1}((r, \infty)) \cup g^{-1}((a-r, \infty))$, since x must belong to at least one of the intervals. The immediate result is that $x \in f^{-1}((r, \infty))$ and $x \in g^{-1}((a-r, \infty))$. By definition of inverse images, we have $f(x) \in (r, \infty)$ and $g(x) \in (a, r, \infty)$. Thus, it follows that $f(x) + g(x) \in (a, \infty)$.

Suppose $x \in (f+g)^{-1}((a,\infty))$; that is, $f(x)+g(x) > a \iff f(x) > a - g(x)$. By the completeness of $\mathbb{R}$, there must exist some $r \in$ such that $f(x) > r > a - g(x)$. Therefore, x belongs to the set in the right-hand side.

Based on the two directions of intersection relationships we have proved, the equivalence of the two sets is then proved. From the intuition we have introduced, the remainder of the proof is trivial.

- By property of composition and measurability, $-g$ is an $\mathcal{S}$-measurable function. Thus $f-g = f+(-g)$ is an $\mathcal{S}$-measurable function.
- Since $fg = \dfrac{(f+g)^2 - f^2 - g^2}{2}$, combined with all the properties that we have proved above, the proof is trivial.
- Suppose $g(x) \neq 0, \forall x \in X$. The function defined on $\mathbb{R}\setminus\{0\}$ (which is a Borel subset of $\mathbb{R}$) that takes x to $\frac{1}{x}$ is continuous and thus is a Borel measurable function. It follows that $\frac{1}{g}$ is an $\mathcal{S}$-measurable function. Combining this result with what we have already proved about the product of $\mathcal{S}$-measurable functions, we conclude that $\frac{f}{g}$ is an $\mathcal{S}$-measurable function.

Limit and Measurability The next result shows that the pointwise limit of a sequence of $\mathcal{S}$-measurable functions is $\mathcal{S}$-measurable. This is a highly desirable property (recall that the set of Riemann integrable functions on some interval is not closed under taking pointwise limits).

Proposition 2.2.19: Limit of $\mathcal{S}$-Measurable Functions

Suppose $(X,\mathcal{S})$ is a measurable space and $f_1, f_2, \cdots$ is a sequence of $\mathcal{S}$-measurable functions from X to $\mathbb{R}$. Suppose $\lim_{k\to\infty} f_k(x)$ exists for each $x \in X$. Define $f : X \to \mathbb{R}$ by

$$f(x) := \lim_{k\to\infty} f_k(x)$$

Then f is an $\mathcal{S}$-measurable function.

Proof for Proposition.

In order to prove $f^{-1}((a,\infty)) \in \mathcal{S}$, we follow the same strategy as before; that is, to represent $f^{-1}((a,\infty))$ as a result of countable unions or intersections.

$$f^{-1}((a,\infty)) = \bigcup_{j=1}^{\infty} \bigcup_{m=1}^{\infty} \bigcap_{k=m}^{\infty} f_k^{-1}\left(\left(a+\frac{1}{j},\infty\right)\right)$$

Obviously, $f_k^{-1}\left(\left(a+\frac{1}{j},\infty\right)\right)$ is an element of $\mathcal{S}$. If such equivalence stands, then $f^{-1}((a,\infty))$ is also an element of $\mathcal{S}$, which completes our proof.

2.2.5 Working on Extended Real Line

Sometimes it is more convenient for us to work with the *extended real line*. Hence, we need extension of Borel set for the case of extended real line.

Definition 2.2.20: Borel Subsets of $[-\infty, \infty]$

A subset of $[-\infty, \infty]$ is called a *Borel set* if its intersection with $\mathbb{R}$ is a Borel set.

Remark.

- In other words, a set $C \subseteq [-\infty, \infty]$ is a Borel set if and only if there exists a Borel set $B \subseteq \mathbb{R}$ such that $C = B$ or $C = B \cup \{\infty\}$ or $C = B \cup \{-\infty\}$ or $C = B \cup \{-\infty, \infty\}$.
- It can be verified that the set of Borel subsets of $[-\infty, \infty]$ is a σ-algebra on $[-\infty, \infty]$.

Proposition 2.2.21: Weaker Condition for Measurability

Suppose $(X, \mathcal{S})$ is a measurable space. A function $f : X \to [-\infty, \infty]$ is such that $f^{-1}((a, \infty]) \in \mathcal{S}$ for all $a \in \mathbb{R}$. Then f is an $\mathcal{S}$-measurable function.

Such proposition tells us that everything is still good even when we extend the definition to the extended real line.

Proposition 2.2.22: Infimum and Supremum of a Sequence of $\mathcal{S}$-Measurable Functions are Measurable

Suppose $(X, \mathcal{S})$ is a measurable space. $f_1, f_2, \cdots$ is a sequence of $\mathcal{S}$-measurable functions defined on $X \to [-\infty, \infty]$. Define $g, h : X \to [-\infty, \infty]$ by

$$g(x) := \inf \{f_k(x) : k \in\}$$
$$h(x) := \sup \{f_k(x) : k \in\}$$

Then g and h are $\mathcal{S}$-measurable functions.

Proof for Proposition.

Let $a \in \mathbb{R}$. The definition of supremum implies that

$$h^{-1}((a, \infty]) = \bigcup_{k=1}^{\infty} f_k^{-1}((a, \infty])$$

The equation above, along with the condition for measurability, implies that h is an $\mathcal{S}$-measurable function. Then, note that

$$g(x) = -\sup \left\{-f_k(x) : k \in \mathbb{Z}^+\right\}, \forall x \in X$$

Thus the result about the supremum implies that g is an $\mathcal{S}$-measurable function.

One of the reason that we hope to work with the extended real line is that, the $\inf$ *or the* $\sup$ *might be* $\pm\infty$.

2.3 Measures and Their Properties

The original motivation for the next definition came from trying to extend the notion of the length of an interval. However, the definition below allows us to discuss size in many more contexts.

Definition 2.3.1: Measure

Let $(X, \mathcal{S})$ be a measurable space. A measure on $(X, \mathcal{S})$ is a function $\mu : \mathcal{S} \to [0, \infty]$ such that

- $\mu(\varnothing) = 0$.
- $\mu\left(\bigcup_{k=1}^{\infty} E_k\right) = \sum_{k=1}^{\infty} \mu(E_k)$, for every pairwise disjoint sequence $E_1, E_2, \cdots \in \mathcal{S}$.

Example.

- Counting measure: X is a set, and $\mathcal{S} = \mathcal{P}(X)$.

$$\mu(E) = \begin{cases} n \text{ if } E \text{ is finite and } E \text{ has } n \text{ elements.} \\ \infty \text{ if } E \text{ is infinite.} \end{cases}$$

- *Dirac* measure: $(X, \mathcal{S})$ is a measure space. Fix an element $c \in X$.

$$\delta_c(E) = \begin{cases} 1 \text{ if } c \in E \\ 0 \text{ if } c \notin E \end{cases}$$

- Lebesque measure: $X = \mathbb{R}, \mathcal{S} = \mathcal{B}(\mathbb{R})$. The outer measure is a measure.

  Though we have shown that the outer measure is not a measure on the power set of $\mathbb{R}$, since it does not satisfy countable additivity. However, it is a measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, as we would show later.

- Simple probability measure on a coin toss: $X = \{H, T\}, S = \mathcal{P}(X) = \{\varnothing, \{H\}, \{T\}, \{H, T\}\}$. Define μ as:

$$\begin{cases} \mu(\varnothing) = 0 \\ \mu(\{H\}) = \frac{1}{2} \\ \mu(\{T\}) = \frac{1}{2} \\ \mu(\{H, T\}) = 1 \end{cases}$$

  It can be verified that μ is indeed a measure.

The following terminology is frequently useful.

Definition 2.3.2: Measure Space

$(X, \mathcal{S}, \mu)$ is called a measure space, where X is a set, $\mathcal{S}$ is a σ-algebra on X, and μ is a measure on the measurable space $(X, \mathcal{S})$.

Based on the definition of measure, a series of related properties can be derived.

Proposition 2.3.3: Measure Preserves Order

Suppose $(X, \mathcal{S}, \mu)$ is a measure space. Suppose $D, E \in \mathcal{S}$ and $D \subseteq E$. Then

$$\mu(D) \leq \mu(E).$$

Proposition 2.3.4: Measure of a Set Difference

Suppose $(X, \mathcal{S}, \mu)$ is a measure space. Suppose $D, E \in \mathcal{S}$ and $D \subseteq E$. Then

$$\mu(E \backslash D) = \mu(E) - \mu(D), \text{ given that } \mu(D) < \infty.$$

Proposition 2.3.5: Countable Additivity

For any $E_1, E_2, \cdots \in \mathcal{S}$, we have

$$\mu\left(\bigcup_{k=1}^{\infty} E_k\right) \leq \sum_{k=1}^{\infty} \mu(E_k).$$

Proof for Proposition.

Following the definition of measure, we construct a sequence of disjoint sets whose union is $\bigcup_{k=1}^{\infty} E_k$. Let $D_1 = \varnothing$ and $D_k = E_1 \cup \cdots \cup E_{k-1}$ for $k \geq 2$. Then $E_1 \backslash D_1, E_2 \backslash D_2, \cdots$ is a disjoint sequence of subsets of X whose union equals $\bigcup_{k=1}^{\infty} E_k$. Thus,

$$\begin{aligned}
\mu\left(\bigcup_{k=1}^{\infty} E_k\right) &= \mu\left(\bigcup_{k=1}^{\infty} (E_k \backslash D_k)\right) \\
&= \sum_{k=1}^{\infty} \mu(E_k \backslash D_k) \\
&\leq \sum_{k=1}^{\infty} \mu(E_k)
\end{aligned}$$

Such countable subadditivity does not require $E_1, E_2, \cdots \in \mathcal{S}$ to be pairwise disjoint, which is a general and useful result.

Below is the two important properties of the ”*continuity of measure*”.

Proposition 2.3.6: Measure of an Increasing Union

Suppose $(X, \mathcal{S}, \mu)$ is a measure space. Let $E_1 \subseteq E_2 \subseteq E_3 \subseteq \cdots$ be an increasing sequence of $\mathcal{S}$ measurable sets. Then

$$\mu\left(\bigcup_{k=1}^{\infty} E_k\right) = \lim_{k\to\infty} \mu(E_k).$$

Proof for Proposition.

Denote $E_0 = \varnothing$ to make notations consistent. By the construction of an increasing sequence of $E_1 \subseteq E_2 \subseteq E_3 \subseteq \cdots$, we can partition $\bigcup_{k=1}^{\infty} E_k$ as:

$$\bigcup_{k=1}^{\infty} E_k = \bigcup_{k=1}^{\infty} (E_k \backslash E_{k-1})$$

Hence,

$$\begin{aligned}
\mu\left(\bigcup_{k=1}^{\infty} E_k\right) &= \mu\left(\bigcup_{k=1}^{\infty} (E_k \backslash E_{k-1})\right) \\
&= \sum_{k=1}^{\infty} \mu(E_k \backslash E_{k-1}) \\
&= \sum_{k=1}^{\infty} (\mu(E_k) - \mu(E_{k-1})) \\
&= \lim_{k\to\infty} \mu(E_k) - \mu(E_0) \\
&= \lim_{k\to\infty} \mu(E_k)
\end{aligned}$$

The second equality holds because of countable additivity, and the third one stands because of the proposition we derived above.

Proposition 2.3.7: Measure of an Decreasing Intersection

Suppose $(X, \mathcal{S}, \mu)$ is a measure space. Let E_1 be a decreasing sequence of $\mathcal{S}$ measurable sets with $\mu(E_1) < \infty$. Then

$$\mu\left(\bigcap_{k=1}^{\infty} E_k\right) = \lim_{k\to\infty} \mu(E_k).$$

Proof for Proposition.

Proposition (Measure of an Decreasing Intersection)
Proof By De Morgan's Law,

$$E_1 \backslash \left(\bigcap_{k=1}^{\infty} E_k\right) = \bigcup_{k=1}^{\infty} (E_1 \backslash E_k)$$

Hence,

$$\begin{aligned}\mu\left(E_1\backslash\left(\bigcap_{k=1}^{\infty}E_k\right)\right)&=\mu\left(\bigcup_{k=1}^{\infty}(E_1\backslash E_k)\right)\\&=\lim_{k\to\infty}\mu\left(E_1\backslash E_k\right)\end{aligned}$$

Meanwhile, for the left-hand side:

$$\mu\left(E_1\backslash\left(\bigcap_{k=1}^{\infty}E_k\right)\right)=\mu\left(E_1\right)-\mu\left(\bigcap_{k=1}^{\infty}E_k\right)$$

For the right-hand side:

$$\lim_{k\to\infty}\mu\left(E_1\backslash E_k\right)=\lim_{k\to\infty}\left(\mu\left(E_1\right)-\mu\left(E_k\right)\right)=\mu\left(E_1\right)-\lim_{k\to\infty}\mu\left(E_k\right)$$

It is noteworthy that we can subtract $\mu(E_1)$ from the both sides of equality only if E_1 is finite. Jointly we have $\mu\left(\bigcap_{k=1}^{\infty}E_k\right)=\lim_{k\to\infty}\mu\left(E_k\right)$, which completes the proof.

The condition of $\mu(E_1)<\infty$ can be relaxed to $N\in$: $\mu(E_N)<\infty$.

Proposition 2.3.8: Measure of a Union

Suppose $(X,\mathcal{S},\mu)$ is a measure space. Let $D,E\in\mathcal{S}$ and $\mu(D\cap E)<\infty$. Then

$$\mu(D\cup E)=\mu(D)+\mu(E)-\mu(D\cap E).$$

Proof for Proposition.

From graphical intuition we separate $D\cup E$ into three disjoint parts:

$$D\cup E=(D\backslash(D\cap E))\cup(E\backslash(D\cap E))\cup(D\cap E).$$

By property of measure, we have

$$\begin{aligned}\mu(D\cup E)&=\mu(D\backslash(D\cap E))+\mu(E\backslash(D\cap E))+\mu(D\cap E)\\&=\mu(D)-\mu(D\cap E)+\mu(E)-\mu(D\cap E)+\mu(D\cap E)\\&=\mu(D)+\mu(E)-\mu(D\cap E)\end{aligned}$$

More generally, the *Exclusion - Inclusion Principle* is:

$$\mu\left(\bigcup_{k=1}^{n}E_k\right)=\sum_{k=1}^{\infty}(-1)^{k+1}\left(\sum_{1\le i_1<\cdots<i_k\le n}\mu\left(E_{i_1}\cap\cdots\cap E_{i_k}\right)\right)$$

2.4 Lebesgue Measure

2.4.1 Outer Measure is a Measure

The main goal is to show that the "outer measure" μ^* is a measure on $(\mathbb{R},\mathcal{B}(\mathbb{R}))$. (Note that "outer measure" is not itself a measure; the name is given artificially.) The core task is to answer whether the "outer measure" satisfies *countable additivity*. The procedure unfolds in the following steps:

- μ^* is countably additive with open sets.
- μ^* is countably additive with closed sets.
- μ^* is countably additive with Borel sets.

Proposition 2.4.1: Additivity of Outer Measure if One of the Sets is Open

Suppose A and G are disjoint subsets of $\mathbb{R}$, and G is open. Then

$$\mu^*(A \cup G) = \mu^*(A) + \mu^*(G).$$

Proof for Proposition.

The trick we will deploy is the same as before; that is, to prove $x = y$ is equivalent to prove both $x \leq y$ and $x \geq y$.

Previously, we have proved that the outer measure has countable subadditivity, so the direction of $\leq$ holds. We will only need to focus on the opposite direction of $\geq$. Because of the fact that any open set can be expressed as a countable union of open intervals, we first start from the simplest case of G being an open interval and then progress to the case where G is a general open set.

- $G = (a, b)$ is an open interval with $a, b \in \mathbb{R}$. We also assume that $a, b \notin A$. Let $I_1, I_2, \cdots$ be any open intervals whose union covers $A \cup G$. For $n > 0$, let

$$\begin{aligned} J_n &= I_n \cap (-\infty, a) \\ K_n &= I_n \cap (a, b) \\ L_n &= I_n \cap (b, \infty) \end{aligned}$$

Note that J_n, K_n, L_n are themselves open finite intervals. Hence,

$$l(I_n) = l(J_n) + l(K_n) + l(L_n)$$

Moreover, we have

$$\begin{aligned} A &\subseteq \left(\bigcup_{n=1}^{\infty} J_n\right) \cup \left(\bigcup_{n=1}^{\infty} L_n\right) \\ G &\subseteq \bigcup_{n=1}^{\infty} K_n \end{aligned}$$

By definition of outer measure,

$$\begin{cases} \mu^*(A) \leq \sum_{n=1}^{\infty} \left(l(J_n) + l(L_n)\right) \\ \mu^*(G) \leq \sum_{k=1}^{\infty} l(K_n) \end{cases}$$

$$\Longrightarrow \mu^*(A) + \mu^*(G) \leq \sum_{n=1}^{\infty} \left(l(J_n) + l(K_n) + l(L_n)\right) = \sum_{n=1}^{\infty} l(I_n)$$

Taking infimum to both sides (over all open coverings of $A \cup G$), we have

$$\mu^*(A) + \mu^*(G) \leq \mu^*(A \cup G)$$

Here G is assumed to be open to ensure that K_n's are open intervals, otherwise some of the properties of outer measure are not applicable.

- G is any open set. (In fact, any open set can always be decomposed into the union of disjoint intervals. Equivalently, $G = \bigcup_{n=1}^{\infty} I_n$.) Since $\left(\bigcup_{n=1}^{N} I_n\right) \subseteq G$, we have

$$\begin{aligned} \mu^*(A \cup G) &\geq \mu^*\left(A \cup \left(\bigcup_{n=1}^{N} I_n\right)\right) \\ &= \mu^*(A) + \sum_{n=1}^{N} \mu^*(I_n) \\ &= \mu^*(A) + \sum_{n=1}^{N} l(I_n) \end{aligned}$$

Note that the second equality is true because of result in the first case we have just discussed.

Taking limit of $N \to \infty$ of both sides, we have

$$\begin{aligned} \mu^*(A \cup G) &\geq \mu^*(A) + \sum_{n=1}^{\infty} l(I_n) \\ &\geq \mu^*(A) + \mu^*(G) \end{aligned}$$

Proposition 2.4.2: Additivity of Outer Measure if One of the Sets is Closed

Suppose A and F are disjoint subsets of $\mathbb{R}$, and F is closed. Then

$$\mu^*(A \cup F) = \mu^*(A) + \mu^*(F).$$

Proof for Proposition.

Suppose $I_1, I_2, \cdots$ are open intervals whose union covers $A \cup F$. Let $G = \bigcup_{k=1}^{\infty} I_k$. Since $A \cap F = \varnothing$, we have $A \subseteq (G \backslash F)$, which implies $\mu^*(A) \leq \mu^*(G \backslash F)$. Note that

$G\backslash F = G \cap F^c$, and both G and F^c are open, so $G\backslash F$ is open.

$$\begin{aligned}
\mu^*(A) &\leq \mu^*(G\backslash F) \\
\implies \mu^*(A) + \mu^*(F) &\leq \mu^*(G\backslash F) + \mu^*(F) \\
&= \mu^*((G\backslash F) \cup F) \\
&= \mu^*(G) \\
&\leq \sum_{k=1}^{\infty} l(I_k)
\end{aligned}$$

Note that the second equality holds because of the proposition we have just proved above.

Taking infimum of both sides, we get the final result that

$$\mu^*(A) + \mu^*(F) \leq \mu^*(A \cup F).$$

For F to be closed, the complement of F is open.

Before moving on to the proof of additivity when one of the set is Borel, we need the following lemma.

Lemma 2.4.3: Approximation of Borel Sets by Some Closed Set

Suppose $B \subseteq \mathbb{R}$ is a Borel set. Then for every $\varepsilon > 0$, there exists a closed set $F \subseteq B$ such that

$$\mu^*(B\backslash F) \leq \varepsilon.$$

Proof for Lemma

To prove something is true for Borel set, you need to construct a collection of sets that satisfies such property, and show such collection is

- Closed under intersection/union.
- Closed under complement.
- Contains open sets.

Let

$$\mathcal{L} = \{D \subseteq \mathbb{R} \mid \text{ For } \varepsilon > 0, \text{ there exists a closed set } F \subseteq D \text{ such that } \mu^*(D\backslash F) \leq \varepsilon\}$$

We hope to show:

- $\mathcal{L}$ is itself a σ-algebra.
- $\mathcal{L}$ contains open sets/open intervals.

Then based on the two points, $\mathcal{L}$ is a σ-algebra that contains the Borel σ-algebra.

It is in fact trivial to see the second point. Let $D = (a, b)$ with $a < b$ and $a, b \in \mathbb{R}$. Take $F = \left[a + \frac{\varepsilon}{2}, b - \frac{\varepsilon}{2}\right]$, and then we are done.

The main contents leave for proving $\mathcal{L}$ being a σ-algebra.

- $\mathcal{L}$ contains $\varnothing$: If $D = \varnothing$, then take $F = \varnothing$, which is open and then valid.
- $\mathcal{L}$ is closed under countable intersection and union. However, note that it is equivalent for being closed under countable intersection and being closed under countable union.

  Suppose $D_1, D_2, \cdots \in \mathcal{L}$. Fix $\varepsilon > 0$. For each $k > 0$, there exists a closed set $F_k \subseteq D_k$ such that

  $$\mu^* \left(D_k \backslash F_k\right) \leq \frac{\varepsilon}{2^k}$$

  A fact from set theory, that the intersection of closed sets is closed, regardless of whether or not such operation is finite or infinite. Then $\bigcap_{k=1}^{\infty} F_k$ is a closed set. By definition, since each F_k is a subset of D_k, we have

  $$\bigcap_{k=1}^{\infty} F_k \subseteq \bigcap_{k=1}^{\infty} D_k$$

  Moreover, you can check it for yourself that

  $$\left(\bigcap_{k=1}^{\infty} D_k\right) \backslash \left(\bigcap_{k=1}^{\infty} F_k\right) \subseteq \bigcup_{k=1}^{\infty} \left(D_k \backslash F_k\right)$$

  Then we have

  $$\mu^* \left(\left(\bigcap_{k=1}^{\infty} D_k\right) \backslash \left(\bigcap_{k=1}^{\infty} F_k\right)\right) \leq \sum_{k=1}^{\infty} \mu^* \left(D_k \backslash F_k\right) \leq \sum_{k=1}^{\infty} \frac{\varepsilon}{2^k} = \varepsilon$$

  The result shows that $\bigcap_{k=1}^{\infty} D_k \in \mathcal{L}$.
- $\mathcal{L}$ is closed under complements.

  Suppose $D \in \mathcal{L}$. We had to show $D^c \in \mathcal{L}$.

  - Suppose $\mu^* (D) < \infty$. Fix $\varepsilon > 0$. Let $F \subseteq D$ such that F is closed and $\mu^* (D \backslash F) \leq \frac{\varepsilon}{2}$. From definition of $\mu^* (D)$, there is an open set G (which can be imagined as the union of an open covering) with $G \subseteq D$ such that

    $$\mu^* (G) \leq \mu^* (D) + \frac{\varepsilon}{2}$$

    First because G is open, G^c is closed. Because $G \subseteq D$, $G^c \subseteq D^c$. And

    $$D^c \backslash G^c = G \backslash D \subseteq G \backslash F$$

Then

$$\begin{aligned}\mu^*\left(D^c \backslash G^c\right) &= \mu^*\left(G \backslash D\right)\\ &\leq \mu^*\left(G \backslash F\right)\\ &= \mu^*\left(G\right) - \mu^*\left(F\right)\\ &= \left(\mu^*\left(G\right) - \mu^*\left(D\right)\right) + \left(\mu^*\left(D\right) - \mu^*\left(F\right)\right)\\ &\leq \frac{\varepsilon}{2} + \mu^*\left(D \backslash F\right)\\ &\leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon\end{aligned}$$

Note that the equality in the second line comes from additivity of outer measure when one of the sets is closed, and the fourth line above uses subadditivity applied to the union $D = (D \backslash F) \cup F$.

– When $\mu^*\left(D\right) = \infty$, for any $D \in \mathcal{L}$, consider

$$D_k = D \cap [-k, k], \forall k \in$$

Because $\mathcal{L}$ is closed under intersection, and both D and $[-k, k]$ are in $\mathcal{L}$, we have $D_k \in \mathcal{L}$. From construction, we see that D_k has a finite measure, so $D_k^c \in \mathcal{L}$. It follows that

$$\begin{aligned} D &= \bigcup_{k=1}^{\infty} D_k\\ \implies D^c &= \bigcap_{k=1}^{\infty} D_k^c\end{aligned}$$

Again, since each $D_k^c \in \mathcal{L}$, and $\mathcal{L}$ is closed under intersection, so $D^c \in \mathcal{L}$, which completes our proof.

■

Utilizing the lemma, we are allowed to prove the additivity of outer measure if one of the sets is a Borel set.

Proposition 2.4.4: Additivity of Outer Measure if One of the Sets is a Borel Set

Suppose A and B are two disjoint sets, and B is a Borel set. Then

$$\mu^*\left(A \cup B\right) = \mu^*\left(A\right) + \mu^*\left(B\right).$$

Proof for Proposition.

From the previous proposition, since B is a Borel set, there exists a closed set $F \subseteq B$

such that $\mu^*(B\backslash F) \le \varepsilon$. Since F is a closed set, again from the previous results,

$$\begin{aligned}\mu^*(A \cup B) &\ge \mu^*(A \cup F)\\ &= \mu^*(A) + \mu^*(F)\\ &\ge \mu^*(A) + \mu^*(B) - \mu^*(B\backslash F)\\ &\ge \mu^*(A) + \mu^*(B) - \varepsilon\end{aligned}$$

The first inequality holds because of the fact that $F \subseteq B \implies (A \cup F) \subseteq (A \cup B)$. The inequality in the third line is true by countable subadditivity of the outer measure. The result above holds for any $\varepsilon > 0$. Take limit $\varepsilon \to 0$, we end up with

$$\mu^*(A \cup B) \ge \mu^*(A) + \mu^*(B).$$

Remark.

We should never write $\mu^*(B\backslash F) = \mu^*(B) - \mu^*(F)$ for $F \subseteq B$, because we are not sure whether B and F are infinite or not. What we always have is countable subadditivity of the outer measure,

$$\begin{aligned}&B = F \cup (B\backslash F) \implies \mu^*(B) \le \mu^*(F) + \mu^*(B\backslash F)\\ &F = B \cup (F\backslash B) \implies \mu^*(F) \le \mu^*(B) + \mu^*(F\backslash B)\\ \implies &\begin{cases}\mu^*(B\backslash F) \ge \mu^*(B) - \mu^*(F)\\ \mu^*(F\backslash B) \ge \mu^*(F) - \mu^*(B)\end{cases}\end{aligned}$$

Corollary 2.4.5: Existence of a Subset of $\mathbb{R}$ that is Not a Borel Set

There exists a set $B \subseteq \mathbb{R}$ such that $\mu^*(B) < \infty$ and B is not a Borel set.

Proof for Corollary.

In the section of countable subadditivity of the outer measure we showed that there exist disjoint sets $A, B \subseteq \mathbb{R}$ such that $\mu^*(A \cup B) \neq \mu^*(A) + \mu^*(B)$. For any such sets, we must have $\mu^*(B) < \infty$ because otherwise both $\mu^*(A \cup B)$ and $\mu^*(A) + \mu^*(B)$ equal ∞ (as follows from the inequality $\mu^*(B) \le \mu^*(A \cup B)$). Now the additivity of outer measure if one of the sets is a Borel sets implies that B is not a Borel set.

The tools we have constructed now allow us to prove that outer measure, when restricted to the Borel sets, is a measure.

Proposition 2.4.6: Outer Measure is a Measure on Borel Sets

The outer measure μ^* is a measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

Proof for Proposition.

Consider $A_1, A_2, \cdots \in \mathcal{B}(R)$ are disjoint. To prove the countable additivity, because we have had countable subadditivity of the outer measure, we only need to prove

$$\mu^*\left(\bigcup_{n=1}^{\infty} A_n\right) \geq \sum_{n=1}^{\infty} \mu^*(A_n)$$

To see this, we first focus on finite cases to employ the preceding result:

$$\begin{aligned}\mu^*\left(\bigcup_{n=1}^{\infty} A_n\right) &\geq \mu^*\left(\bigcup_{n=1}^{N} A_n\right)\\ &= \sum_{n=1}^{N} \mu^*(A_n) \overset{N\to\infty}{\longrightarrow} \sum_{k=1}^{\infty} \mu^*(A_n)\end{aligned}$$

The second equality holds because each of A_n is a Borel set and we can then apply the previous proposition. Note that the limit here is valid itself just from our definition of infinite series.

2.4.2 Definitions and Properties

Definition 2.4.7: Lebesgue Measure

"The" Lebesgue measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ is $\mu : \mathcal{B}(\mathbb{R}) \to [0, \infty]$ such that

$$\mu(B) := \mu^*(B), \forall B \in \mathcal{B}(\mathbb{R}).$$

The actual Lebesgue measure is defined on a "slightly" bigger σ-algebra. There are some Lebesgue measurable sets that are not Borel sets; but that definition is enough to serve our purpose.

As we will see in this section, outer measure is actually a measure on a somewhat larger class of sets called the Lebesgue measurable sets. The approach chosen here has the advantage of emphasizing that a Lebesgue measurable set differs from a Borel set by a set with outer measure 0. The attitude here is that sets with outer measure 0 should be considered sets small enough that do not matter much.

Definition 2.4.8: Lebesgue Measurable Set

A set $A \subseteq \mathbb{R}$ is called *Lebesgue measurable* if there exists a Borel set $B \subseteq A$ such that the outer measure

$$\mu^*(A \backslash B) = 0.$$

Intuitively, Lebesgue just added some sets that differ from Borel sets by a set with outer measure of 0 into the Borel sets and call them together the *Lebesgue measurable* sets. Each Borel set is Lebesgue measurable because if $A \subseteq \mathbb{R}$ is a Borel set, then we can take $B = A$ in the definition above.

Proposition 2.4.9: Different Definitions for Lebesgue Measurability

Suppose $A \subseteq \mathbb{R}$. Then all of the following statements are equivalent:

- A is Lebesgue measurable.
- **There exists a Borel set** $B \subseteq A$ **such that** $\mu^* (A \backslash B) = 0$**.**
- For each $\varepsilon > 0$, there exists a **closed** set $F \subseteq A$ such that

$$\mu^* (A \backslash F) < \varepsilon$$

- There exists a (countable) sequence of closed sets $F_1, F_2, \cdots$ such that

$$\mu^* \left(A \backslash \left(\bigcup_{k=1}^{\infty} F_k \right) \right) = 0$$

- For each $\varepsilon > 0$ there exists an **open** set $G \supseteq A$ such that

$$\mu^* (G \backslash A) < \varepsilon$$

- There exists a (countable) sequence of closed sets $G_1, G_2, \cdots$ such that

$$\mu^* \left(\left(\bigcap_{k=1}^{\infty} G_k \right) \backslash A \right) = 0$$

Remark.

- The fourth condition implies that every Borel set is almost a countable union of closed sets. The last condition implies that every Borel set is almost a countable intersection of open sets.
- The reason why we have developed such list of definitions is that, they would allow us to work on some more general settings (such as beyond the $\mathbb{R}$; e.g., the set of functions, vectors, etc.)
- The infinite union of closed sets is not necessarily a closed set. That is why we developed the second and the third definitions (instead of regard them as being equivalent). For example, $\bigcup_{n=1}^{\infty} \left[\frac{1}{n}, 1 - \frac{1}{n}\right] = (0, 1)$.
- Carethedory's criterion: A is Lebesgue measurable if and only if, for all $E \subseteq \mathbb{R}$, we have

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

  The intuition is that, if A is Lebesgue measurable, it should separate each set $E \in \mathbb{R}$ well. Compared with the previous definitions, this definition is even more general; that is, it imposes on restriction on A's being closed or open.

Proposition 2.4.10: Outer Measure is a Measure on Lebesgue Measurable Sets

The collection $\mathcal{L}$ of Lebesgue measurable sets is a σ-algebra on $\mathbb{R}$. The outer measure is a measure on $(\mathbb{R}, \mathcal{L}(\mathbb{R}))$, which is called the Lebesgue measure.

Proof for Proposition.

$\mathcal{L}$ indeed contains both the Borel sets and every set with outer measure 0, which is a σ-algebra on $\mathbb{R}$. To prove the countable additivity, suppose $A_1, A_2, \cdots$ is a disjoint sequence of Lebesgue measurable sets. By the definition of Lebesgue measurable set, for each $k \in \mathbb{Z}^+$ there exists a Borel set $B_k \subseteq A_k$ such that $\mu^*(A_k \backslash B_k) = 0$. Now

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

Note that the second line above holds because $B_1, B_2, \cdots$ is a disjoint sequence of Borel sets and outer measure is a measure on Borel sets; the last line above holds because

$$\left.\begin{aligned} B_k \subseteq A_k &\Longrightarrow \mu^*(B_k) \leq \mu^*(A_k) \\ A_k = B_k \cup (A_k \backslash B_k) &\Longrightarrow \mu^*(A_k) \leq \mu^*(B_k) + \mu^*(A_k \backslash B_k) = \mu^*(B_k) \end{aligned}\right. \Longrightarrow \mu^*(A_k) = \mu^*(B_k)$$

The inequality above, combined with countable subadditivity of outer measure, implies that $\mu^*\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} \mu^*(A_k)$, completing the proof of the outer measure being a measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. ∎

From now on, when we say "measure" we mean $\mu^* : \mathcal{B}(\mathbb{R}) \to \mathbb{R}$, and elements in $\mathcal{B}(\mathbb{R})$ are called "measurable sets".

2.5 Convergence of Measurable Functions

Claim: Littlewood's Three Principles of Measure Theory

1. Every measurable set is "nearly" a finite union of intervals.
2. (Luzin's Theorem) Every measurable function is "nearly" continuous.
3. (Egorov's Theorem) Every convergent sequence of functions is "nearly" uniformly convergent.

Definition 2.5.1: Pointwise Covergence and Uniform Convergence

Suppose X is a set, $f_1, f_2, \cdots : X \to \mathbb{R}$ is a sequence of functions and $f : X \to \mathbb{R}$ is a function.

- Pointwise convergence: $f_1, f_2, \cdots$ converges pointwise to f on X if for both $x \in X$ and $\varepsilon > 0$, there exists $n \in$ such that $|f_k(x) - f(x)| < \varepsilon$ for all $k \geq n$. $(\lim_{k\to\infty} f_k(x) - f(x), \forall x \in X)$
- Uniform convergence: $f_1, f_2, \cdots$ converges uniformly to f on X if for every $\varepsilon > 0$, there exists $n \in$ such that $|f_k(x) - f(x)| < \varepsilon$ for all $k \geq n$ and for all $x \in X$.

Theorem 2.5.2: Uniform Limit of Continuous Functions is Continuous

Suppose $B \subseteq \mathbb{R}$ and $f_1, f_2, \cdots$ is a sequence of functions from B to $\mathbb{R}$ that converges uniformly on B to a function $f : B \to \mathbb{R}$. Suppose $b \in B$ and f_k is continuous at b for each $k \in \mathbb{Z}^+$. Then f is continuous at b.

2.5.1 Egorov's Theorem

A sequence of functions that converges pointwise need not converge uniformly. However, the next result says that a pointwise convergent sequence of functions on a measure space with finite total measure almost converges uniformly, in the sense that it converges uniformly except on a set that can have arbitrarily small measure.

Theorem 2.5.3: Egorov's Theorem

Suppose $(X, \mathcal{S}, \mu)$ is a measure space with $\mu(X) < \infty$. Suppose $f_1, f_2, \cdots$ is a sequence of $\mathcal{S}$-measurable function from X to $\mathbb{R}$ that converges pointwise to a function $f : X \to \mathbb{R}$. Then for every $\varepsilon > 0$, there exists an $\mathcal{S}$-measurable set E such that $\mu(X \backslash E) < \varepsilon$ and $f_1, f_2, \cdots$ converges uniformly to f on E.

Proof for Theorem

Let $\varepsilon > 0$. Fix $n \in$. The fact that $f_n \to f$ pointwise on X implies that

$$X = \bigcup_{m=1}^{\infty} \bigcap_{k=m}^{\infty} \left\{ x \in X : |f_K(x) - f(x)| < \frac{1}{n} \right\}$$

For $m \in$, let

$$A_{m,n} = \bigcap_{k=m}^{\infty} \left\{ x \in X : |f_K(x) - f(x)| < \frac{1}{n} \right\}$$

Note that each $A_{m,n} \in \mathcal{S}$ because each $f_k - f$ is an $\mathcal{S}$-measurable function. Then $A_{1.n} \subseteq A_{2,n} \subseteq A_{3,n} \subseteq \cdots$ is an increasing sequence of sets. Hence, by continuity of measure, we have

$$\lim_{m\to\infty} \mu\left(A_{m,n}\right) = \mu\left(\bigcup_{m=1}^{\infty} A_{m,n}\right) = \mu\left(X\right)$$

The equality means there exists $m_n \in$ (an integer dependent of n) such that

$$\mu\left(X\right) - \mu\left(A_{m_n,n}\right) < \frac{\varepsilon}{2^n}$$

Let

$$E = \bigcap_{n=1}^{\infty} A_{m_n,n}$$

Then

$$\begin{aligned}\mu\left(X\backslash E\right) &= \mu\left(X\backslash\left(\bigcap_{n=1}^{\infty} A_{m_n,n}\right)\right)\\ &= \mu\left(\bigcup_{n=1}^{\infty}\left(X\backslash A_{m_n,n}\right)\right)\\ &\leq \sum_{n=1}^{\infty} \mu\left(X\backslash A_{m_n,n}\right)\\ &\leq \sum_{n=1}^{\infty} \frac{\varepsilon}{2^n} = \varepsilon\end{aligned}$$

where the equality in the second line is by De Morgan's law, and the inequality in the third line is by subadditivity of μ. (Note that countable additivity is only applicable to the case of disjoint sets, while subadditivity is true for all interested sets.)

To complete the proof, we must verify that $f_1, f_2, \cdots$ converges uniformly to f on E. To do this, suppose $\varepsilon' > 0$. Let $n \in \mathbb{Z}^+$ be such that $\frac{1}{n} < \varepsilon'$. Then $E \subseteq A_{m_n,n}$, which implies that

$$\left|f_k\left(x\right) - f\left(x\right)\right| < \frac{1}{n} < \varepsilon'$$

for all $k \geq m_n$ and all $x \in E$. Hence $f_1, f_2, \cdots$ does indeed converge uniformly to f on E. ■

Remark.

- Since $f_1, f_2, \cdots$ are measurable functions, it follows that $A_{m_n,n}$'s and E are measurable.
- To prove a set is measurable, try to describe it as a union or intersection of sets.

2.5.2 Luzin's Theorem

Here we are going to introduce the concept of simple functions, and show that any function can be approximated by a sequence of simple functions that is pointwise and uniformly convergent. The results of simple function would be quite helpful for further works.

Definition 2.5.4: Characteristic/Indicator Function

Let $E \subseteq X$ be any subset of X. The characteristic function (or, indicator function) on E, $\chi_E : X \to \mathbb{R}$ is defined as

$$\chi_E(x) = \begin{cases} 1 \text{ if } x \in E \\ 0 \text{ if } x \in E \end{cases}$$

Definition 2.5.5: Simple Function

A function is called *simple* if it takes on only finitely many values. Let $(X, \mathcal{S})$ be a measurable space. An $\mathcal{S}$-measurable simple function $f : X \to \mathbb{R}$ has the form

$$f(x) = c_1\chi_{E_1} + c_2\chi_{E_2} + \cdots + c_n\chi_{E_n}$$

for *distinct* nonzero $c_1, c_2, \cdots, c_n$ and $E_1, E_2, \cdots, E_n \in \mathcal{S}$.

Step functions are in analog to simple functions. However, different from step functions that are defined on intervals, simple functions are defined on measurable sets.

Theorem 2.5.6: Approximation by Simple Functions

Suppose $(X, \mathcal{S})$ is a measurable space and $f : [-\infty, \infty]$ is $\mathcal{S}$-measurable. Then there exists a sequence $f_1, f_2, \cdots$ of functions from X to $\mathbb{R}$ such that

- each f_k is a simple $\mathcal{S}$-measurable function;
- $|f_k(x)| \leq |f_{k+1}(x)| \leq |f(x)|$ for all $k \in$ and all $x \in X$;
- $\lim_{k\to\infty} f_k(x) = f(x)$ for every $x \in X$;
- $f_1, f_2, \cdots$ converges *uniformly* on X to f if f is bounded.

Proof for Theorem

Instead of directly defining f on its intervals (like the Riemann integrals do), we define f on the inverse image of evenly-cut intervals of the codomain of f:

For $k > 0$ and $n \in \mathbb{Z}$, we divide $[n, n+1)$ into 2^k equally sized half-open intervals, and define f as:

$$f_k(x) = \begin{cases} \frac{m}{2^k} & \text{if } 0 \leq f(x) \leq k \text{ and } f(x) \in \left[\frac{m}{2^k}, \frac{m+1}{2^k}\right) \\ \frac{m+1}{2^k} & \text{if } -k \leq f(x) \leq 0 \text{ and } f(x) \in \left[\frac{m}{2^k}, \frac{m+1}{2^k}\right) \\ k & \text{if } f(x) > k \\ -k & \text{if } f(x) < -k \end{cases}$$

The first two conditions can easily be checked by construction of f_k's. For any $f(x) \in [-k, k]$, we have $|f_k(x) - f(x)| \leq \frac{1}{2^k}$, for every $x \in X$ such that $f(x) \in [-k, k]$ and for every k. This implies pointwise convergence that $\lim_{k\to\infty} f_k(x) = f(x)$. Furthermore, if f is bounded, then we can remove the condition in the previous statement that $f(x) \in [-k, k]$, which means that $|f_k(x) - f(x)| \leq \frac{1}{2^k}$ for all $x \in X$ and all k. Therefore, $f_1, f_2, \cdots$ converges *uniformly* on X to f if f is bounded.

Be careful about the interpretation of the conclusion of Luzin's theorem that $f|_B$ is a continuous function on B. This is not the same as saying that f (on its original domain) is continuous at each point of B. For example, χ is discontinuous at every point of $\mathbb{R}$. However, $\chi|_{\mathbb{R}\backslash}$ is a continuous function on $\mathbb{R}\backslash$ (because this function is identically 0 on its domain).

Theorem 2.5.7: Luzin's Theorem

Suppose $g : \mathbb{R} \to \mathbb{R}$ is a Borel measurable function. Then for every $\varepsilon > 0$, there exists a closed set $F \subseteq \mathbb{R}$ such that $\mu(\mathbb{R}\backslash F) < \varepsilon$ and $g|_F$ is a continuous function on F ($g|_F : F \to \mathbb{R}$).

Proof for Theorem

The proof goes with the following sketch:

- Prove Luzin's theorem for simple function g_k, and $g_k \to g$. (pointwise convergence)
- Use Egorov's theorem to make the previous pointwise convergence to uniform convergence.
- Uniform limit of continuous functions is continuous.

First prove Luzin's theorem for simple functions. Assume g to be a simple function that

$$g = d_1\chi_{D_1} + \cdots d_n\chi_{D_n}$$

for some $d_1, d_2, \cdots, d_n \neq 0$ and some disjoint $D_1, D_2, \cdots, D_n \subseteq \mathbb{R}$.

Let $\varepsilon > 0$. For each k, there exists a closed set $F_k \subseteq D_k$, and an open set $G_k \supseteq D_k$ such that

$$\mu(D_k\backslash F_k) \leq \frac{\varepsilon}{2n}\mu(G_k\backslash D_k) \leq \frac{\varepsilon}{2n}$$

Because $G_k\backslash F_k = (G_k\backslash D_k) \cup (D_k\backslash F_k)$, by subadditivity we have

$$\mu(G_k\backslash F_k) \leq \frac{\varepsilon}{2n} + \frac{\varepsilon}{2n} = \frac{\varepsilon}{n}, \forall k \in \{1, 2, \cdots, n\}$$

Note that simple function must be continuous. Let

$$F = \left(\bigcup_{k=1}^{n} F_k\right) \cup \left(\bigcap_{k=1}^{n} (\mathbb{R}\backslash G_k)\right)$$

Because $F_k \subseteq D_k$, g is identically d_k on F_k. Then $g|_{F_k}$ is constant and then continuous for each $k \in \{1, 2, \cdots, n\}$, and $g|_{\mathbb{R}\backslash G_k}$ is 0 and then continuous for each k. In conclusion, $g|_F$ is continuous.

By theorem of approximation by simple functions, there exists a sequence of simple functions $g_1, g_2, \cdots : \mathbb{R} \to \mathbb{R}$ such that the sequence converges pointwise to g. Let $\varepsilon > 0$.

By the first case above, for each $k > 0$, there exists a closed set $C_k \subseteq \mathbb{R}$ such that

- $\mu(\mathbb{R}\backslash C_k) \leq \frac{\varepsilon}{2^{k+1}}$.
- $g_k|_{C_k}$ is continuous.

Let $C = \bigcap_{k=1}^{\infty} C_k$. Then C is a closed set and $C \subseteq C_k$. From this we can see that $g_k|_C$ is continuous for every k. By De Morgan's law:

$$\mu(\mathbb{R}\backslash C) \leq \sum_{k=1}^{\infty} \mu(\mathbb{R}\backslash C_k) \leq \sum_{k=1}^{\infty} \frac{\varepsilon}{2^{k+1}} = \varepsilon$$

For each $m \in \mathbb{Z}$, consider $(m, m+1)$. The sequence $g_1|_{(m,m+1)}, g_2|_{(m,m+1)}, \cdots$ converges pointwise to $g|_{(m,m+1)}$. This is obviously true because g_k's converge pointwise to g on $\mathbb{R}$, so is the case when restricted to $(m, m+1)$. By Egorov's theorem, there exists a Borel set E_m such that

$$\mu((m, m+1)\backslash E_n) \leq \frac{\varepsilon}{2^{|m|+3}}$$

and such that the sequence $g_1, g_2, \cdots$ converges uniformly to g on E_m. Then we have $g_1, g_2, \cdots$ converges uniformly to g on $C \cap E_m$. Because each g_k is continuous, $g|_{C \cap E_m}$ is continuous. Define

$$D = \bigcup_{m \in \mathbb{Z}} (C \cap E_m)$$

Then $g|_D$ is continuous. The last step is to show that

$$\mu(\mathbb{R}\backslash D) \leq \cdots \leq \varepsilon$$

There exists a closed set $F \subseteq D$ such that

$$\mu(D\backslash F) \leq \varepsilon - \mu(\mathbb{R}\backslash D)$$

Then

$$\mu(\mathbb{R}\backslash F) \leq \mu(D\backslash F) + \mu(\mathbb{R}\backslash D) \leq \varepsilon$$

$g|_F$ is continuous because $g|_D$ is continuous. ■

Remark.

- Every time you come across the case of $A \subseteq B$, you may not write $\mu(B\backslash A) = \mu(A) - \mu(B)$.
- Any intersection of closed set, regardless of finitely or infinitely many, must be closed. However this is not true for union operation.

Luzin's theorem also has a second version.

Theorem 2.5.8: Continuous Extensions of Continuous Functions

Every continuous function on a closed subset of $\mathbb{R}$ can be extended to a continuous function on all of $\mathbb{R}$. More precisely, if $F \subseteq \mathbb{R}$ is closed and $g : F \to \mathbb{R}$ is continuous, then there exists a continuous function $h : \mathbb{R} \to \mathbb{R}$ such that $h|_F = g$.

Theorem 2.5.9: Luzin's Theorem, Second Version

Suppose $E \subseteq \mathbb{R}$ and $g : E \to \mathbb{R}$ is a Borel measurable function. Then for every $\varepsilon > 0$, there exists a closed set $F \subseteq E$ and a continuous function $h : \mathbb{R} \to \mathbb{R}$ such that $\mu^* (E \backslash F) < \varepsilon$ and $h|_F = g|_F$.

Proof for Theorem

Suppose $\varepsilon > 0$. Extend g to a function $\tilde{g} : \mathbb{R} \to \mathbb{R}$ by defining

$$\tilde{g}(x) = \begin{cases} g(x), & \text{if } x \in E \\ 0, & \text{if } x \in \mathbb{R} \backslash E \end{cases}$$

By the first version of Luzin's theorem, there is a closed set $C \subseteq \mathbb{R}$ such that $\mu^* (\mathbb{R} \backslash C) < \varepsilon$ and $\tilde{g}|_C$ is a continuous function on C. There exists a closed set $F \subseteq (C \cap E)$ such that

$$\mu^* ((\mathcal{C} \cap E) \backslash F) < \varepsilon - \mu^* (\mathbb{R} \backslash C)$$

Thus,

$$\mu^* (E \backslash F) \leq \mu^* (((C \cap E) \backslash F) \cup (\mathbb{R} \backslash C)) \leq \mu^* ((C \backslash E) \backslash F) + \mu^* (\mathbb{R} \backslash C) < \varepsilon$$

Now $\tilde{g}|_F$ is a continuous function on F. Also, $\tilde{g}|_F = g|_F$ (because $F \subseteq E$). Use the theorem of continuous extensions of continuous functions to extend $\tilde{g}|_F$ to a continuous function $h : \mathbb{R} \to \mathbb{R}$. ■