A Primer on Measure Theory

Pramook Khungurn

January 23, 2022

This is a primer on measure theory and Lebesgue integration. The materials are taken from Bartle's "The Elements Of Integration And Lebesgue Measure" [Bartle, 1995], Billingsley's "Probability and Measure" [Billingsley, 1995], and Hunter's note on measure theory [Hunter, 2011].

1 Introduction

- Why do we care about measure theory and Lebesgue integration?
  - They expand the class of functions for which integrations are defined compared to what can be achieved by Riemann integration.
  - Theorems relating to the intechange of limits and integrals are valid under less stringent conditions (again, compared to Reimann integration).
  - In particular, the dominated convergence theorem [1] is a very powerful tool. For examples, it can be use to easily show that

    $$\lim_{n \to \infty} \int_0^\infty \frac{e^{-nx}}{\sqrt{x}}\, \mathrm{d}x = 0$$

    and

    $$\frac{\mathrm{d}}{\mathrm{d}x} \int_0^\infty x^2 e^{-tx}\, \mathrm{d}x = -\int_0^\infty x^3 e^{-tx}\, \mathrm{d}x.$$

  - Measure theory is the foundation of modern probability theory, and the dominated convergence theorem shows up everywhere in it.
- How is Lebesgue integration different from Reimann integration?
  - Riemann integrals are defined in terms of approximating a function with constant functions over intevals.
  - An **interval** is a subset of the real line which is of one of the following forms:

    $$\begin{aligned} [a,b] &= \{x \in \mathbb{R} : a \le x \le b\}, \\ [a,b) &= \{x \in \mathbb{R} : a \le x < b\}, \\ (a,b] &= \{x \in \mathbb{R} : a < x \le b\}, \\ (a,b) &= \{x \in \mathbb{R} : a < x < b\}. \end{aligned}$$

    The real number a and b are said to be the **endpoints** of the interval, and $b - a$ is the **length** of the interval.

[1] https://en.wikipedia.org/wiki/Dominated_convergence_theorem

- A **step function** φ is a linear combination of a finite number of characteristic functions of intevals.

$$\varphi(x) = \sum_{i=1}^{n} c_j \chi_{E_j}(x)$$

where $c_j \in \mathbb{R}$ and each E_j is an interval with endpoints a_j and b_j.

- The integral of a step function φ is defined to be

$$\int \varphi = \sum_{i=1}^{n} c_j (b_j - a_j).$$

- If f is a bounded function on $[a, b]$, then the **Reimann integral** is defined to the limit of the integrals of step functions that approximate f.
- The **lower Rieman integral** is defined to be the supremum of integrals of all step functions ϕ such that $\phi(x) \leq f(x)$ for all $x \in [a, b]$ and $\phi(x) = 0$ for all $x \notin [a, b]$.

• The Lebesgue integral is defined similarly, with some differences.

  - Intervals are replaced by a larger collection of sets (called **measurable sets**).
  - The notion of "length" is generalize to the notation of **measure**.
    * Here, the measure is a function μ that maps a set of a non-negative real number.
  - The step function is replaced by the **simple function**, which is a finite linear combination of characteristic functions of measurable sets.

$$\varphi(x) = \sum_{j=1}^{n} c_j \chi_{E_j}(x)$$

where each E_j is a measurable set. The integral of ϕ is defined to be

$$\int \varphi = \sum_{j=1}^{n} c_j \mu(E_j).$$

  - If f is a non-negative function defined on $\mathbb{R}$, then the **Lebesgue integral** of f is the supremum of all simple functions ϕ such that $\phi(x) \leq f(x)$ for all $x \in \mathbb{R}$.
    * This notation can later be generalized to functions taking both signs.

• When studying integration, it is convenient to work with the **extended real number system** $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$.

  - For any $x \in \mathbb{R}$, we have that $-\infty < x < \infty$.
  - We say that the length of the real line is ∞.
  - We define the supremum of non-empty set of real numbers which does not have an upper bound to be ∞, and the infemum of the a non-empty set of real numbers which does not have a lower bound to be $-\infty$.
    * In this way, all non-empty sets of real numbers (or subsets of $\overline{\mathbb{R}}$) have unique supremums and infemums.

- The arithematic operations between the infiniites and real numbers are as follows:

$$
\begin{aligned}
(\pm\infty) + (\pm\infty) = x + (\pm\infty) = (\pm\infty) + x &= \pm\infty \\
(\pm\infty)(\pm\infty) &= +\infty \\
(\pm\infty)(\mp\infty) &= -\infty \\
(\pm\infty)x = x(\pm\infty) &= \begin{cases} \pm\infty, & \text{if } x > 0, \\ 0, & \text{if } x = 0, \\ \mp\infty, & \text{if } x < 0 \end{cases}
\end{aligned}
$$

  for any (finite) real number x.

- Note that we do not define $(\pm\infty) - (\pm\infty)$. We also do not define quotients when the denominators are $\pm\infty$.

- If (x_n) is a sequence of extended real numbers, define the **limit superior** and **limit inferior** by

$$
\begin{aligned}
\limsup_{n\to\infty} x_n &= \lim_{n\to\infty}\Big(\sup_{m\geq n} x_m\Big) = \inf_n \Big(\sup_{m\geq n} x_m\Big) \\
\liminf_{n\to\infty} x_n &= \lim_{n\to\infty}\Big(\inf_{m\geq n} x_m\Big) = \sup_n \Big(\inf_{m\geq n} x_m\Big).
\end{aligned}
$$

  If the limit superior and limit inferere of a sequence both exist and are equal, then the **limit** of the sequence exists and is equal to that value.

2 Sigma-Algebras and Measures

2.1 Basic Definitions

- Let us denote the power set of set X with $\mathcal{P}(X)$.

- A σ-algebra is the domain upon which we define measures. It is a collection of sets with some nice properties.

  Definition 1. *A* **σ-algebra** *(or a* **σ-field***) on a set X is a collection $\mathcal{X} \in \mathcal{P}(X)$ of subsets of X, called* **measurable sets***, such that the following properties hold.*

  1. *$\emptyset, X \in \mathcal{X}$.*
  2. *It is closed under complementation: if $A \in \mathcal{X}$, then $A^c = X - A \in \mathcal{X}$.*
  3. *It is closed under countable unions: if (A_n) is a sequence of sets in $\mathcal{X}$, then $\bigcup_{n=1}^{\infty} A_n \in \mathcal{X}$.*

- We can show that a σ-algebra is closed under countable intersections as well. To see this, we note that $A_n^c \in \mathcal{X}$ for all $n \in \mathbb{N}$, and so $\bigcup_{i=1}^{\infty} A_n^c \in \mathcal{X}$ As result, $\big(\bigcup_{i=1}^{\infty} A_n^c\big)^c \in \mathcal{X}$. Applying de Morgan's law, we have that $\bigcap_{i=1}^{\infty} A_n = \big(\bigcup_{i=1}^{\infty} A_n^c\big)^c \in \mathcal{X}$.

- **Definition 2.** *A measurable space $(X, \mathcal{X})$ is a non-empty set X equipped with a σ-algebra $\mathcal{X}$ on X.*

- **Definition 3.** *Let $\mathcal{A}$ be a non-empty collection of subsets of X. The* **σ-algebra generated by** *$\mathcal{A}$, denoted by $\sigma(\mathcal{A})$ is the smallest σ-algebra that contains $\mathcal{A}$. In other words,*

$$\sigma(\mathcal{A}) = \bigcap \Big\{ \tilde{\mathcal{A}} \subseteq \mathcal{P}(X) : \mathcal{A} \subseteq \tilde{\mathcal{A}} \text{ and } \tilde{\mathcal{A}} \text{ is a } \sigma\text{-algebra} \Big\}.$$

- **Definition 4.** *The* **Borel algebra** *is the σ-algebra $\mathcal{B}$ generated by all the open intervals (a, b) in $\mathbb{R}$. Any set in $\mathcal{B}$ is called a* **Borel set***.*

- Observe that we can write any open interval (a, b) as a countable unions of closed intervals:

$$(a, b) = \bigcup_{n \geq N}^{\infty} \left[a + \frac{1}{n}, b - \frac{1}{n} \right]$$

where N is an integer such that $b - a - \frac{2}{N} > 0$. As a result, $\mathcal{B}$ is also generated by the collection of close intervals $[a, b]$ in $\mathbb{R}$. The same is also true for half-open intervals of the form $(a, b]$ and $[a, b)$.

- Let X be the set $\overline{\mathbb{R}}$ of extended real numbers. If E is a Borel set, then define

$$\begin{aligned} E_1 &= E \cup \{\infty\} \\ E_2 &= E \cup \{-\infty\} \\ E_3 &= E \cup \{-\infty, \infty\} \end{aligned}$$

Let $\overline{\mathcal{B}}$ the collection of all sets E, E_1, E_2, and E_3 as E varies over $\mathcal{B}$. We have that $\overline{B}$ is a σ-algebra, and it is called the **extended Borel algebra**.

- A "measure" encapsulates the notion of length, area, volume, mass, etc. of a set.

Definition 5. *Let $(X, \mathcal{X})$ be a measurable space. A* **measure** *is a function $\mu : \mathcal{X} \to [0, \infty]$ with the following properties.*

1. $\mu(\emptyset) = 0$.
2. *μ is* **countably additive**. *That is, for a sequence (E_n) of disjoint sets, it holds that*

$$\mu\left(\bigcup_{n=1}^{\infty} E_n \right) = \sum_{n=1}^{\infty} \mu(E_n).$$

- If $\mu(E) < \infty$ for all $E \in \mathcal{X}$, we say that μ is **finite**.

- A **probability measure** is a finite measure with $\mu(X) = 1$.

- If there exists a sequence (E_n) of sets in $\mathcal{X}$ with $\bigcup_{i=1}^{\infty} E_n = X$ and such that $\mu(E_n) < \infty$ for all n, then we say that μ is **σ-finite**.

- Here is an example of a measure that is σ-finite but not finite. Let $X = \mathbb{N}$, and $\mathcal{X} = \mathcal{P}(\mathbb{N})$. Define $\mu(E)$ to be the number of elements in E with the convention that $\mu(E) = \infty$ when E is infinite. Obviously, $\mu(\mathbb{N}) = \infty$. However, $\mathbb{N} = \{1\} \cup \{2\} \cup \cdots$, and $\mu(\{n\}) = 1$ for all $n \in \mathbb{N}$. The measure μ is called the **counting measure** on $\mathbb{N}$.

- **Lemma 6.** *Let μ be a measure defined on a σ-algebra $\mathcal{X}$. Let $E, F \in \mathcal{X}$ be such that $E \subseteq F$, then $\mu(E) \leq \mu(F)$. If $\mu(E) < \infty$, then $\mu(F - E) = \mu(F) - \mu(E)$.*

Proof. Since $F = E \cup (F - E)$ and $E \cap (F - E) = \emptyset$, it follows that

$$\mu(F) = \mu(E) + \mu(F - E).$$

Because $\mu(F - E) \geq 0$, it follows that $\mu(F) \geq \mu(E)$. If $\mu(E) < \infty$, we can subtract from both sides of the equation. $\square$

- A sequence of sets (E_n) is **increasing** if $E_n \subseteq E_{n+1}$ for all n.

- A sequence of sets (E_n) is **decreasing** if $E_n \supseteq E_{n+1}$ for all n.

- **Lemma 7.** *If (E_n) is an increasing sequence of measurable sets, then*

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

If (E_n) is a decreasing sequence of measurable sets and $\mu(E_1) < \infty$, then

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

Proof. Let (E_n) be increasing. Set $F_0 = E_1$, and $F_n = E_{n+1} - E_n$ for all $n \geq 1$. We have that (F_n) is a sequence of disjoint sets. So,

$$\mu\left(\bigcup_{n=1}^{\infty} E_n\right) = \mu\left(\bigcup_{n=0}^{\infty} F_n\right) = \sum_{n=0}^{\infty} \mu(F_n)$$

Also because $E_n = \bigcup_{i=0}^{n} F_i$, we have that

$$\mu(E_n) = \sum_{i=0}^{n} \mu(F_i).$$

So,

$$\mu\left(\bigcup_{n=1}^{\infty} E_n\right) = \sum_{n=0}^{\infty} \mu(F_n) = \lim_{n\to\infty} \sum_{i=0}^{n} \mu(F_i) = \lim_{n\to\infty} \mu(E_n).$$

Next, let (E_n) be decreasing and $\mu(E_1) < \infty$. Let $F_n = E_1 - E_n$. We have that (F_n) is increasing and $\mu(F_n) = \mu(E_1) - \mu(E_n)$. If follows that

$$\mu\left(\bigcup_{n=1}^{\infty} F_n\right) = \lim_{n\to\infty} \mu(F_n) = \lim_{n\to\infty} \mu(E_1) - \mu(E_n) = \mu(E_1) - \lim_{n\to\infty} \mu(E_n)$$

Now,

$$\bigcap_{n=1}^{\infty} E_n = E_1 - \bigcup_{n=1}^{\infty} F_n$$

So,

$$\mu\left(\bigcap_{n=1}^{\infty} E_n\right) = \mu(E_1) - \mu\left(\bigcup_{n=1}^{\infty} F_n\right) = \mu(E_1) - \left(\mu(E_1) - \lim_{n\to\infty} \mu(E_n)\right) = \lim_{n\to\infty} \mu(E_n)$$

are required. □

- **Definition 8.** *A* **measure space** *is a triple $(X, \mathcal{X}, \mu)$ where X is a non-empty set, $\mathcal{X}$ is a σ-algebra on X, and μ is a measure on $\mathcal{X}$.*

- **Definition 9.** *In a measure space $(X, \mathcal{X}, \mu)$, a set $N \in \mathcal{X}$ is set to be of* **measure zero** *or a* **null set** *if $\mu(N) = 0$. A property that holds on N^c is said to hold* μ**-almost everywhere**. *In the context where μ is clear, we says that a property holds just* **almost everywhere**.

- For examples, we say that two functions f and g are equal almost everywhere if $f(x) = g(x)$ for all $x \notin N$ where N is a set of measure zero. We also say that a sequence of functions (f_n) converges almost everywhere in X if $\lim_{n\to\infty} f_n(x)$ exists for all $x \notin N$.

- **Definition 10.** *A measure space* $(X, \mathcal{X}, \mu)$ *is* **complete** *if every subset of a set of measure zero is measureable.*

- **Theorem 11.** *Let* $(X, \mathcal{X}, \mu)$ *be a measure space. Define* $(X, \overline{\mathcal{X}}, \overline{\mu})$ *by*

$$\overline{\mathcal{X}} = \{A \cup M : A \in \mathcal{X}, M \subseteq N \textit{ where } N \in \mathcal{X} \textit{ and } \mu(N) = 0\}$$

and

$$\overline{\mu}(A \cup M) = \mu(A).$$

Then, $(X, \overline{\mathcal{X}}, \overline{\mu})$ *is a complete measure space such that* $\mathcal{X} \subseteq \overline{X}$ *and* $\overline{\mu}$ *is the unique extension of* μ *to* $\overline{X}$.

Proof (sketch). The hardest bit of the proof is to show that $\overline{\mathcal{X}}$ is close under complementation. Let $A \in \mathcal{X}$, $N \in \mathcal{X}$ be a set of measure zero, and $M \subseteq N$. We have that $(A \cup M)^c = A^c \cap M^c$. Because $M^c = N^c \cup (N - M)$, we have that

$$(A \cup M)^c = A^c \cap M^c = A^c \cap (N^c \cup (N - M)) = (A^c \cap N^c) \cup (A^c \cap (N - M)).$$

We note that $A^c \cap N^c \in \mathcal{X}$ and $A^c \cap (N - M) \subseteq N$, so $(A \cup M)^C \in \overline{\mathcal{X}}$. The other parts of the proof seems straightforward, and we refer to [Hunter, 2011] for a longer proof sketch. □

2.2 Why Do We Need the Sigma-Algebra?

- Why is the σ-algebra necessary? In other words, why do we care to limit ourselves to only some subsets of $\mathcal{P}(X)$? It turns out that we can arrive at contradictions if we want to define measures on all subsets.

- Suppose that we want to define a probability measure μ on the set $[0, 1)$ such that we define μ for all subsets of X. Then, there exists a set A where the measure cannot consistenly be defined.

  To construct this set, we first partition the elements of $[0, 1)$ into equivalent classes. Let $x \in [0, 1)$, we let $[x] = \{y \in [0, 1) : (y - x) \bmod 1 \in \mathbb{Q} \cap [0, 1)\}$ be the equivalence class containing x. Consider the collection $\{[x] : x \in [0, 1)\}$ of equivalent classes. By the axiom of choice, we can construct the set A that contains exactly one element from each equivalent class.

  Let $x \in [0, 1)$. Then, $x \in [y]$ for some $y \in [0, 1)$, and $y - x \in \mathbb{Q}$. Let $T_q(A) = \{(x + q) \bmod 1 : x \in A\}$ be a circlically translated copy of A by q. We have that there exists one $q \in \mathbb{Q} \cap [0, 1)$ such that $x \in T_q(A)$. As a result, $[0, 1) = \bigcup_{q \in \mathbb{Q} \cap [0,1)} T_q(A)$.

  Moreover, we have that $T_q(A) \cap T_r(A) = \emptyset$ for any $q \neq r \in \mathbb{Q} \cap [0, 1)$. To see this, let there be x such that $x \in T_q(A)$ and $x \in T_r(A)$. Then, $(x + q) \bmod 1 \in A$ and $(x + r) \bmod 1 \in A$. Note that $(x + q) \bmod 1$ and $(x + r) \bmod 1$ must belong in the same equivalent class, say $[y]$ where $y \in A$. Since both numbers are in A, it must be that $(x + q) \bmod 1 = (x + r) \bmod 1 = y$. So, $q = r$.

  A sensible probability measure should be translation invariant. In other words, it is natural to require that $\mu(T_q(A)) = A$ for all q.

  Then, we run into a problem. What value should we assign to $\mu(A)$? If $\mu(A) = 0$, we have that $1 = \mu([0, 1)) = \sum_{q \in \mathbb{Q} \cap [0,1)} \mu(T_q(A)) = 0$. However, if $\mu(A) > 0$, then $\mu([0, 1)) = \infty$. In both cases, we cannot make $\mu([0, 1)) = 1$.

- In general, this problem crops up a lot as a consequence of the Banach–Tarski paradox.[2]

- To avoid this problem, we have to forego one of the nice things: the axiom of choice, the requirement that the measure is invariant to translations, or the fact that we can define measures on all subsets. The approach taken by the math world is to choose the last by limitting the sets we can define measures on to those in a σ-algebra.

[2] `https://en.wikipedia.org/wiki/Banach%E2%80%93Tarski_paradox`

2.3 Pi-Systems and Lambda-Systems

- We take a detour to explore structures that are related to the σ-algebra that will be useful in further studies.

- **Definition 12.** *A collection of sets $\mathcal{A} \subseteq \mathcal{P}(X)$ is a π-system if it is closed under finite intersections. In other words, if $A, B \in \mathcal{A}$, then $A \cap B \in \mathcal{A}$.*

- **Definition 13.** *A collection of sets $\mathcal{A} \subseteq \mathcal{P}(X)$ is a λ-system it satisfies the following properties.*

  1. *$X \in \mathcal{A}$.*
  2. *It is closed under complementation. That is, $A \in \mathcal{A}$ implies $A^c \in \mathcal{A}$.*
  3. *It is closed under countable disjoint unions. That is, if (A_n) is a sequence of disjoint sets in $\mathcal{A}$, then $\bigcup_{i=n}^{\infty} A_n \in \mathcal{A}$.*

- Note that the π-system and the σ-system both have weaker conditions than that of the σ-algebra.

- **Lemma 14.** *A λ-system is closed under proper set differences. In other words, If $\mathcal{A}$ is a λ-system, $A, B \in \mathcal{A}$ and $A \subseteq B$, then $B - A \in \mathcal{A}$.*

  Proof. We have that $B^c \in \mathcal{A}$, and $A \cap B^c = \emptyset$. As a result, $A \cup B^c \in \mathcal{A}$ and so is $(A \cup B^c)^c = B - A$. □

- **Lemma 15.** *A collection $\mathcal{A} \subseteq \mathcal{P}(X)$ that is both a π-system and a λ-system is a σ-algebra.*

  Proof. We only need to show that $\mathcal{A}$ is closed under countable (general) unions. Let (A_n) be a sequence of sets in $\mathcal{A}$. Let $B_n = A_n \cap A_{n-1}^c \cap A_{n-2}^c \cap \cdots A_1^c$. We have that $B_n \in \mathcal{A}$ because it is formed by finite intersections. Moreover, the B_n's are disjoint from one other, so $\bigcup_{n=1}^{\infty} B_n = \bigcup_{n=1}^{\infty} A_n$ is a member of $\mathcal{A}$ as well. □

- **Definition 16.** *Let $\mathcal{A} \subseteq \mathcal{P}(X)$ be a collection of sets. The* **λ-system generated by** *$\mathcal{A}$, denoted by $\lambda(\mathcal{A})$ is the smallest λ-system that contains $\mathcal{A}$. In other words,*

$$\lambda(\mathcal{A}) = \bigcap \left\{ \mathcal{L} \subseteq \mathcal{P}(X) : \mathcal{L} \text{ is a } \lambda\text{-system and } \mathcal{A} \subseteq \mathcal{L} \right\}.$$

  .

- Note that since a σ-algebra is a λ-system, we have that $\lambda(\mathcal{A}) \subseteq \sigma(\mathcal{A})$ for all $\mathcal{A} \subseteq \mathcal{P}(X)$.

- The following theorem goes in the opposite direction and is useful in proving many uniqueness theorems.

  Theorem 17 (Dynkin's π–λ theorem). *If $\mathcal{A}$ is a π-system, $\mathcal{B}$ is a λ-system, and $\mathcal{A} \subseteq B$, then $\sigma(\mathcal{A}) \subseteq \mathcal{B}$.*

  Proof. We will show that $\lambda(\mathcal{A})$ is a σ-algebra. If this is the case, then $\sigma(\mathcal{A}) \subseteq \lambda(\mathcal{A}) \subseteq \mathcal{B}$, and the theorem holds.

  By the last lemma, it is sufficient to show that $\lambda(\mathcal{A})$ is closed under intersection. Doing this is rather convoluted. First, for any $A \subseteq X$, define

$$\mathcal{G}_A = \{B \subseteq X : A \cap B \in \lambda(\mathcal{A})\}.$$

  We will show that, if $A \in \lambda(\mathcal{A})$, then $\mathcal{G}_A$ is a λ-system.

  To see this, we first note that $X \in \mathcal{G}_A$ because $A \cap X = A \in \lambda(\mathcal{A})$.

  Next, we show that $\mathcal{G}_A$ is closed under complementation. Let $B \in \mathcal{G}_A$. We have that $A \cap B \in \lambda(\mathcal{A})$ Because $A \cap B \subseteq A$, it follows from Lemma 14 that $A - (A \cap B) = A \cap B^c \in \lambda(\mathcal{A})$. Hence, $B^c \in \mathcal{G}_A$.

Next, we show that $\mathcal{G}_A$ is closed under countable disjoin unions. Let $B = \bigcup_{n=1}^{\infty} B_n$ where each $B_n \in \mathcal{G}_A$. We have that $A \cap B_n \in \lambda(\mathcal{A})$ for all n, and so is $\bigcup_{n=1}^{\infty}(A \cap B_n) = A \cap \cup_{n=1}^{\infty} B_n = A \cap B$. Hence, $B = \bigcup_{n=1}^{\infty} B_n$.

We have now established that $A \in \lambda(\mathcal{A}) \implies \mathcal{G}_A$ is a λ-system.

We shall now show that $A \in \mathcal{A} \implies \lambda(\mathcal{A}) \subseteq \mathcal{G}_A$. To see this, let $B \in \mathcal{A}$. It follows that $A \cap B \in \mathcal{A}$ because $\mathcal{A}$ is a π-system. Because $\mathcal{A} \subseteq \lambda(\mathcal{A})$, it follows that $A \cap B \in \lambda(\mathcal{A})$. In other words, $B \in \mathcal{G}_A$, and so $\mathcal{A} \subseteq \mathcal{G}_A$. Because $A \in \mathcal{A} \subseteq \lambda(\mathcal{A})$, it follows that $\mathcal{G}_A$ is a λ-system that contains $\mathcal{A}$. Thus, $\lambda(\mathcal{A}) \subseteq \mathcal{G}_A$.

Now,

$$\begin{aligned} A \in \mathcal{A} &\implies \lambda(\mathcal{A}) \subseteq \mathcal{G}_A \\ A \in \mathcal{A} &\implies (B \in \lambda(\mathcal{A}) \implies B \in \mathcal{G}_A) \\ (A \in \mathcal{A}) \wedge (B \in \lambda(\mathcal{A})) &\implies B \in \mathcal{G}_A. \end{aligned}$$

Because $B \in \mathcal{G}_A$ iff $A \in \mathcal{G}_B$, it follows that:

$$\begin{aligned} (A \in \mathcal{A}) \wedge (B \in \lambda(\mathcal{A})) &\implies A \in \mathcal{G}_B \\ B \in \lambda(\mathcal{A}) &\implies (A \in \mathcal{A} \implies A \in \mathcal{G}_B) \\ B \in \lambda(\mathcal{A}) &\implies \mathcal{A} \subseteq \mathcal{G}_B. \end{aligned}$$

Because $\mathcal{G}_B$ is a λ-system, it follows that $\lambda(\mathcal{A}) \subseteq \mathcal{G}_B$. So,

$$B \in \lambda(\mathcal{A}) \implies \lambda(\mathcal{A}) \subseteq \mathcal{G}_B.$$

Hence,

$$\begin{aligned} A \in \lambda(\mathcal{A}) \wedge B \in \lambda(\mathcal{A}) &\implies A \in \lambda(\mathcal{A}) \wedge \lambda(\mathcal{A}) \subseteq \mathcal{G}_B \\ A \in \lambda(\mathcal{A}) \wedge B \in \lambda(\mathcal{A}) &\implies A \in \mathcal{G}_B \\ A \in \lambda(\mathcal{A}) \wedge B \in \lambda(\mathcal{A}) &\implies A \cap B \in \lambda(\mathcal{A}) \\ A, B \in \lambda(\mathcal{A}) &\implies A \cap B \in \lambda(\mathcal{A}). \end{aligned}$$

It follows that $\lambda(A)$ is a π-system, and thus a σ-algebra. □

- It follows that, if $\mathcal{A}$ is a π-system, then $\lambda(\mathcal{A}) = \sigma(\mathcal{A})$.
- Here's an application of the Dynkin's theorem.

Theorem 18. *Let μ_1 and μ_2 be two finite measures on $\sigma(\mathcal{A})$ where $\mathcal{A} \subseteq \mathcal{P}(X)$ is a π-system that contains X. If they agree on $\mathcal{A}$, then they also agree on $\sigma(\mathcal{A})$.*

Proof. Let $\mathcal{E} = \{E \in \sigma(\mathcal{A}) : \mu_1(E) = \mu_2(E)\}$. We will show that $\mathcal{E}$ is a λ-system. If so, it follows that $\sigma(\mathcal{A}) \subseteq \mathcal{E}$, and we would be done.

By the assumption of the theorem, $X \in \mathcal{A}$, so $\mu_1(X) = \mu_2(X)$. Thus, $X \in \mathcal{E}$.

Next, let $E \in \mathcal{E}$. We have that

$$\mu_1(E^c) = \mu_1(X - E) = \mu_1(X) - \mu_1(E) = \mu_2(X) - \mu_2(E) = \mu_2(X - E) = \mu_2(E^c).$$

As a result, $E^c \in \mathcal{E}$. (Note that we can do this because the measure is finite, so we can apply Lemma 6.)

Lastly, let (E_n) be a sequence of disjoint sets in $\mathcal{E}$, and let $E = \bigcup_{n=1}^{\infty} E_n$. It follows that

$$\mu_1(E) = \mu_1\bigg(\bigcup_{n=1}^{n} E_n\bigg) = \sum_{n=1}^{\infty} \mu_1(E_n) = \sum_{n=1}^{\infty} \mu_2(E_n) = \mu_2\bigg(\bigcup_{n=1}^{n} E_n\bigg) = \mu_2$$

So, $E \in \mathcal{E}$ as well. □

2.4 Algebras and Monotone Classes

- In this section, we take another detour on a structure that is similar to the σ-algebra and a theorem about it that is similar to the Dynkin's π-λ theorem.

- **Definition 19.** *A family $\mathcal{A}$ of subsets of a set X is said to be an* **algebra** *or a* **field** *on X if the following properties are satisfied.*

  1. $\emptyset, X \in \mathcal{A}$.
  2. *If $E \in \mathcal{A}$, then $E^c = X - E \in \mathcal{A}$.*
  3. *If $E_1, E_2, \dots, E_n \in \mathcal{A}$, then $\bigcup_{i=1}^n E_i \in \mathcal{A}$.*

- Note that an algebra is a π-system because $A \cap B = (A^c \cup B^c)^c$. However, it is not a λ-system.

- **Definition 20.** *A class $\mathcal{M}$ of subsets of X is called a* **monotone class** *if it is closed under the formation of monotone unions and intersections. In other words,*

  – *If $A_1, A_2, \dots \in \mathcal{M}$ and $A_i \subseteq A_{i+1}$ for all i, then $\bigcup_{i=1}^\infty A_i \in \mathcal{M}$, amd*
  – *If $A_1, A_2, \dots \in \mathcal{M}$ and $A_i \supseteq A_{i+1}$ for all i, then $\bigcap_{i=1}^\infty A_i \in \mathcal{M}$.*

- Note that a σ-algebra is a monotone class because it is already closed under arbitrary countable unions and intersections.

- **Lemma 21.** *Let $\mathcal{A}$ be a subset of $\mathcal{P}(X)$. If $\mathcal{A}$ is both an algebra and a monotone class, then it is a σ-algebra.*

  Proof. We only have to show that $\mathcal{A}$ is closed under arbitrary countable unions. Let (A_n) be a sequence of sets in $\mathcal{A}$. Define $B_n = \bigcup_{i=1}^n A_i$. We have that $B_n \in \mathcal{A}$ for all n because $\mathcal{A}$ is an algebra. Moreover, because $B_n \subseteq B_{n+1}$ for all n, it follows that $\bigcup_{n=1}^\infty B_n \in \mathcal{A}$ because $\mathcal{A}$ is a monotone class. Because $\bigcup_{n=1}^\infty A_n = \bigcup_{n=1}^\infty B_n$, we have that $\bigcup_{n=1}^\infty A_n \in \mathcal{A}$ too, so $\mathcal{A}$ is a σ-algebra. □

- **Definition 22.** *Let $\mathcal{A}$ be a subset of $\mathcal{P}(X)$. The* **monotone class generated by** *$\mathcal{A}$, denoted by $m(\mathcal{A})$, is intersection of all monotone classes that contains $\mathcal{A}$. That is,*

$$m(\mathcal{A}) = \bigcap \{\mathcal{M} : \mathcal{M} \text{ is a monotone class and } \mathcal{A} \subseteq \mathcal{M}\}.$$

- **Theorem 23 (Halmos's monotone class lemma).** *Let $\mathcal{A}$ be an algebra and $\mathcal{M}$ be a monotone class. Then, $\mathcal{A} \subseteq \mathcal{M}$ implies $\sigma(\mathcal{A}) \subseteq \mathcal{M}$. In particular, $\sigma(\mathcal{A}) = m(\mathcal{A})$.*

  Proof. Since $\mathcal{A} \subseteq \sigma(\mathcal{A})$ and $\sigma(\mathcal{A})$ is a monotone class that contains $\mathcal{A}$, it follows that $m(\mathcal{A}) \subseteq \sigma(\mathcal{A})$. As a result, it suffices to show that $\sigma(\mathcal{A}) \subseteq m(\mathcal{A})$. For this, it suffices to show that $m(\mathcal{A})$ is an algebra (which will imply that it is a σ-algebra according to Lemma 21).

  (Two basic sets) Because $\mathcal{A} \subseteq m(\mathcal{A})$, it follows that $\emptyset$ and $\mathcal{X}$ are both in $m(\mathcal{A})$.

  (Closure under complementation) Consider $\mathcal{G} = \{A : A^c \in m(\mathcal{A})\}$. Because $m(\mathcal{A})$ is a monotone class, it follows that $\mathcal{G}$ is a monotone class as well. It follows that $m(\mathcal{A}) \subseteq \mathcal{G}$. This means that, for any $A \in m(\mathcal{A})$, it means that $A \in \mathcal{G}$, which means that $A^c \in m(\mathcal{A})$. As a result, $m(\mathcal{A})$ is closed under complementation.

  (Closure under finite union) Define $\mathcal{G}_1 = \{A : A \cup B \in m(\mathcal{A}) \text{ for all } B \in \mathcal{A}\}$. We have that $\mathcal{G}_1$ is a monotone class. To see this, suppose that (A_n) is an increasing sequence of sets in $\mathcal{G}_1$. It follows from definition that $(A_n \cup B)$ is also an increasing sequence of sets in $m(\mathcal{A})$. As a result,

$$\bigcup_{n=1}^\infty (A_n \cup B) = \Big(\bigcup_{n=1}^\infty A_n \Big) \cup B \in m(\mathcal{A}),$$

which implies that $\bigcup_{n=1}^{\infty} A_n \in \mathcal{G}_1$. The condition involving countable monotone intersections can be shown similarly.

Because $\mathcal{A}$ is an algebra, it follows that

$$A \in \mathcal{A} \land B \in \mathcal{A} \implies A \cup B \in \mathcal{A}.$$

Because $\mathcal{A} \subseteq m(\mathcal{A})$, we can also say that

$$\begin{aligned} A \in \mathcal{A} \land B \in \mathcal{A} &\implies A \cup B \in m(\mathcal{A}) \\ A \in \mathcal{A} &\implies \Big(B \in \mathcal{A} \implies A \cup B \in m(\mathcal{A})\Big) \\ A \in \mathcal{A} &\implies \forall_{B \in \mathcal{A}}[A \cup B \in m(\mathcal{A})], \\ A \in \mathcal{A} &\implies A \in \mathcal{G}_1. \end{aligned}$$

In other words, $\mathcal{A} \subseteq \mathcal{G}_1$. Because $\mathcal{G}_1$ is a monotone class and $\mathcal{G}_1$ contains $\mathcal{A}$, it follows that $m(\mathcal{A}) \subseteq \mathcal{G}_1$. Now,

$$\begin{aligned} m(\mathcal{A}) \subseteq \mathcal{G}_1 &\equiv A \in m(\mathcal{A}) \implies A \in \mathcal{G}_1 \\ &\equiv A \in m(\mathcal{A}) \implies \forall_{B \in \mathcal{A}}[A \cup B \in m(\mathcal{A})] \\ &\equiv A \in m(\mathcal{A}) \implies \Big(B \in \mathcal{A} \implies A \cup B \in m(\mathcal{A})\Big) \\ &\equiv A \in m(\mathcal{A}) \land B \in \mathcal{A} \implies A \cup B \in m(\mathcal{A}). \end{aligned}$$

Define $\mathcal{G}_2 = \{B : A \cup B \in m(\mathcal{A}) \text{ for all } A \in m(\mathcal{A})\}$. It follows from the above statement that $\mathcal{A} \subseteq \mathcal{G}_2$. Moreover, we can show again that $\mathcal{G}_2$ is a monotone class by repeating the same argument we used for $\mathcal{G}_1$. As a result, $m(\mathcal{A}) \subseteq \mathcal{G}_2$. In other words,

$$\begin{aligned} B \in m(\mathcal{A}) &\implies B \in \mathcal{G}_2 \\ B \in m(\mathcal{A}) &\implies \forall_{A \in m(\mathcal{A})}[A \cup B \in m(\mathcal{A})] \\ B \in m(\mathcal{A}) &\implies \Big(A \in m(\mathcal{A}) \implies A \cup B \in m(\mathcal{A})\Big) \\ A \in m(\mathcal{A}) \land B \in m(\mathcal{A}) &\implies A \cup B \in m(\mathcal{A}), \end{aligned}$$

which means that $m(\mathcal{A})$ is closed under finite unions. So, $m(\mathcal{A})$ is an algebra. □

3 Lebesgue Measures

The Lebesgue measure on $\mathbb{R}$ is a measure that corresponds to the notion of "length" on the real line. We will construct it in this section.

3.1 Length as a Premeasure

- The natural notion of **length** can be defined as follows.
  - Let ℓ denote the length function.
  - The length of the half-open interval $(a, b]$ is defined to be $b - a$.
  - The lengths of $(-\infty, b]$, $(a, +\infty)$, and $(-\infty, \infty)$ are defined to be ∞.
  - The length of the union of a finite number of disjoint sets of intervals of these forms is defined to be the sum of the corresponding lengths.

$$\ell\bigg(\bigcup_{i=1}^{n}(a_i, b_n]\bigg) = \sum_{i=1}^{n}(b_i - a_i).$$

- Let $\mathcal{F}$ be the collection of subsets of $\mathbb{R}$ that contains all intervals of the forms

$$(a, b], (-\infty, b], (a, \infty), (\infty, \infty), \tag{1}$$

and all the finite unions of such intervals.

- By the notion above, we have that ℓ is a function of signature $\mathcal{F} \to [0, \infty]$. However, we cannot claim that it is a measure (1) we have not defined how to deal with countable unions of such intervals yet, and (2) we have not identified the σ-algebra upon which it acts.

- Furthermore, we cannot claim that $\mathcal{F}$ is a σ-algebra. However, $\mathcal{F}$ is an algebra.

- **Lemma 24.** *$\mathcal{F}$ is an algebra on $\mathbb{R}$.*

  Proof. We have that:

  - $\emptyset \in \mathcal{F}$ because $\emptyset = \{x : 1 < x \leq 1\} = (1, 1]$.
  - $\mathbb{R} = (-\infty, \infty) \in \mathcal{F}$ by definition.
  - $\mathcal{F}$ is closed under complementation.
    - $(-\infty, b]^c = (b, \infty) \in \mathcal{F}$.
    - $(a, \infty)^c = (-\infty, a] \in \mathcal{F}$.
    - $(a, b]^c = (\infty, a] \cup (b, \infty) \in \mathcal{F}$.
    - $\emptyset$ and $(-\infty, \infty)$ are in $\mathcal{F}$.
  - $\mathcal{F}$ is closed under finite unions by definition. □

- We now turn back to the length function ℓ. It turns out that we can show that it has more nice properties that are worthy of being encapsulated with an abstraction.

- **Definition 25.** *Let $\mathcal{A}$ be an algebra on X. A* **premeasure** *on $\mathcal{A}$ is an extended real valued function μ defined on $\mathcal{A}$ that satisfies the following properties.*

  1. $\mu(\emptyset) = 0$.
  2. $\mu(E) \geq 0$ *for all* $E \in \mathcal{A}$.
  3. *If (E_n) is any disjoint sequence of sets in $\mathcal{A}$ such that $\bigcup_{i=n}^{\infty} E_n$ belongs to $\mathcal{A}$, then*

$$\mu\left(\bigcup_{n=1}^{\infty} E_n\right) = \sum_{n=1}^{\infty} \mu(E_n).$$

- **Lemma 26.** *The length function ℓ is a premeasure on $\mathcal{F}$.*

  Proof. First, $\ell(\emptyset) = \ell((1, 1]) = 1 - 1 = 0$.

  Second, if $E \in \mathcal{F}$, then $\mathcal{E}$ is a finite union of intervals of the forms in (1). We can subdivide these intervals into disjoint pieces, and the length of the union is the sum of the length of the individual pieces, each of which is non-negative. So, we have that $\mu(E) \geq 0$.

  Lastly, suppose that (E_n) is a sequence of disjoint sets in $\mathcal{F}$ such that $\bigcup_{i=n}^{\infty} E_n$ is also in $\mathcal{F}$. Note that $\bigcup_{i=n}^{\infty} E_n$ is a finite union of intervals, so we can again partition them into disjoint pieces, each of which is of a form in (1). Each piece is now a union of a countable collection of disjoint elements of $\mathcal{F}$. WLOG, we may treat each of the E_n's as an interval which is disjoint from any other. Our goal now would be to show that the lengths of the constituent intervals add up to the length of the piece.

A piece can be any of the 4 types. We will only deal with the $(a, b]$ type in this proof as the proof of other types are similar. Suppose, then, that

$$(a, b] = \bigcup_{j=1}^{\infty} (a_j, b_j]$$

where the intervals are disjoint. Consider the first n intervals. We may assume that

$$a \le a_1 < b_1 \le a_2 < b_2 \le \cdots \le a_n < b_n \le b.$$

We have that

$$\sum_{i=1}^{n} \ell((a_i, b_i]) = \sum_{i=1}^{n} (b_i - a_i) \le b_n - a_1 \le b - a = \ell((a, b]).$$

Because n is arbitrary, we have that

$$\sum_{i=1}^{\infty} \ell((a_i, b_i]) \le \ell((a, b]).$$

For the other direction, let $\varepsilon > 0$ be arbitrary. Let (ε_j) be a sequence of positive numbers with $\sum \varepsilon_j < \varepsilon/2$. Consider the interval $I_j = (a_j - \varepsilon_j, b_k + \varepsilon_j)$. The collection $\{I_j\}$ of open sets is a cover of the interval $[a, b]$. Since $[a, b]$ is compact, it has a finite subcover, say, $I_1, I_2, \dots, I_m$. By reordering and discarding some intervals, we may assume that

$$\begin{aligned} a_1 - \varepsilon_1 &< a \\ b &< b_m + \varepsilon_m \\ a_j - \varepsilon_j &< b_{j-1} + \varepsilon_{j-1}. \end{aligned}$$

If follows that

$$b - a \le (b_m + \varepsilon_m) - (a_1 - \varepsilon_1) \le \sum_{j=1}^{m} [(b_j + \varepsilon_j) - (a_j - \varepsilon_j)] \le \varepsilon + \sum_{j=1}^{m} (b_j - a_j) \le \varepsilon + \sum_{j=1}^{\infty} (b_j - a_j).$$

Since $\varepsilon > 0$ is arbitrary, it follows that

$$\ell((a, b]) \le \sum_{i=1}^{\infty} \ell((a_i, b_i]).$$

As a result, $\ell((a, b]) = \sum_{i=1}^{\infty} \ell((a_i, b_i])$.

Combining the results of all cases, we can conclude that ℓ is countably additive in $\mathcal{F}$. $\square$

3.2 Length as an Outer Measure

- The good news is that, given a premeasure μ on an algebra $\mathcal{A}$, we can show that μ can be extended to a measure on a σ-algebra.
  - In other words, there exist a σ-algebra $\mathcal{A}^*$ containing $\mathcal{A}$ and a measure μ^* defined on $\mathcal{A}^*$ such that $\mu^*(E) = \mu(E)$ for all $E \in \mathcal{A}$.

  As a result, we can extend our ℓ so that it becomes a measure on a σ-algebra.
- The way to extend a premeasure μ is as follows.

Definition 27. *Given a premeasure μ defined on an algebra $\mathcal{A}$ on set X, define $\mu^* : \mathcal{P}(\mathbf{X}) \to [0, \infty]$ to be*

$$\mu^*(B) = \inf \left\{ \sum_{j=1}^{\infty} \mu(E_j) : (E_j) \in \mathfrak{C}(B) \right\}$$

where $\mathfrak{C}(B)$ is the set of all sequences (E_j) of sets in $\mathcal{A}$ such that $B \subseteq \bigcup_{j=1}^{\infty} E_j$.

- The following lemma is useful for proving other results

  Lemma 28. *For any set $B \subseteq X$ and any real number $\varepsilon > 0$, there exists a sequence $(E_j) \in \mathfrak{C}(B)$ such that $\sum_{j=1}^{\infty} \mu(E_j) \leq \mu^*(B) + \varepsilon$.*

  Proof. Suppose that the lemma is false. Then, there exists $\varepsilon > 0$ such that, for all $(E_j) \in \mathfrak{C}(B)$, it must be the case that $\sum_{j=1}^{\infty} \mu(E_j) > \mu^*(B) + \varepsilon$. It follows that $\mu^*(B) + \varepsilon$ is a lower bound of $\{\sum \mu(E_j) : E_j \in \mathfrak{C}(B)\}$. So,

$$\mu^*(B) = \inf \left\{ \sum_{j=1}^{\infty} \mu(E_j) : (E_j) \in \mathfrak{C}(B) \right\} \geq \mu^*(B) + \varepsilon,$$

  and we have arrived at a contradiction. □

- **Lemma 29.** *The function μ^* has the following properties.*

  (a) $\mu^*(\emptyset) = 0$.
  (b) $\mu^*(B) \geq 0$ *for any* $B \subseteq X$.
  (c) If $A \subseteq B$, *then* $\mu^*(A) \leq \mu^*(B)$.
  (d) If $A \in \mathcal{A}$, *then* $\mu^*(A) = \mu(A)$.
  (e) If (B_n) *is a sequence of subsets of* X, *then*

$$\mu^*\left(\bigcup_{n=1}^{\infty} B_n \right) \leq \sum_{i=1}^{\infty} \mu^*(B_n).$$

  Proof. Let us call a sequence $(E_j) \in \mathfrak{C}(B)$ a *cover* of B.

  For (b), recall that μ is a measure, so $\mu(E_j) \geq 0$ for all j. Hence, $\sum \mu(E_j) \geq 0$ for any sequence (E_j). As a result, $\mu^*(B) = \inf\{\sum \mu(E_j) : (E_j) \in \mathfrak{C}(\emptyset)\} \geq 0$.

  For (a), note that the sequence $(\emptyset, \emptyset, \dots)$ is a cover of $\emptyset$, so $\mu^*(\emptyset) = \inf\{\sum \mu(E_j) : (E_j) \in \mathfrak{C}(\emptyset)\} \leq 0$. However, from (b), we have that $\mu^*(\emptyset) \geq 0$, so $\mu^*(\emptyset) = 0$.

  For (c), let $A \subseteq B$. Let $(E_j) \in \mathfrak{C}(B)$. We have that $A \subseteq B \subseteq \bigcup E_j$, so $(E_j) \in \mathfrak{C}(A)$ too. It follows that $\mathfrak{C}(A) \supseteq \mathfrak{C}(B)$, which implies that

$$\left\{ \sum \mu(E_j) : (E_j) \in \mathfrak{C}(A) \right\} \supseteq \left\{ \sum \mu(E_j) : (E_j) \in \mathfrak{C}(B) \right\},$$

  and so

$$\mu^*(A) = \inf \left\{ \sum \mu(E_j) : (E_j) \in \mathfrak{C}(A) \right\} \leq \inf \left\{ \sum \mu(E_j) : (E_j) \in \mathfrak{C}(B) \right\} = \mu^*(B).$$

  For (d), let $A \in \mathcal{A}$. We have that $\{A, \emptyset, \emptyset, \dots)$ is a cover of A. As a result,

$$\mu^*(A) \leq \mu(A) + \mu(\emptyset) + \mu(\emptyset) + \cdots = \mu(A).$$

Let (E_j) be a cover of A. We have that $A = \bigcup(A \cap E_j)$. Because μ is a measure,

$$\mu(A) \leq \sum_{j=1}^{\infty} \mu(A \cap E_j) \leq \sum_{j=1}^{\infty} \mu(E_j).$$

It follows that $\mu(A) \leq \inf\{\sum \mu(E_j) : (E_j) \in \mathfrak{C}(A)\} = \mu^*(A)$. Hence, $\mu(A) = \mu^*(A)$.

For (e), let $\varepsilon > 0$ be arbitrary. For each n, apply Lemma 28 to choose a sequence (E_{nk}) of sets in $\mathcal{A}$ such that (E_{nk}) covers B_n and

$$\sum_{k=1}^{\infty} \mu(E_{nk}) \leq \mu^*(B_n) + \frac{\varepsilon}{2^n}.$$

Since $\{E_{nk} : n, k \in \mathbb{N}\}$ is a countable collection from $\mathcal{A}$ whose union contains $\bigcup B_n$, it follows that

$$\mu^*\left(\bigcup_{n=1}^{\infty} B_n\right) \leq \sum_{n=1}^{\infty}\sum_{k=1}^{\infty} \mu(E_{nk}) \leq \varepsilon + \sum_{k=1}^{\infty} \mu^*(B_n).$$

Since ε can be arbitrarily small, we have that Property (e) holds. □

- **Definition 30.** *An* **outer measure** μ^* *on a set* X *is a function* $\mu^* : \mathcal{P}(X) \to [0, \infty]$ *such that the following properties hold.*

  1. $\mu^*(\emptyset) = 0$.
  2. *If* $E \subseteq F \subseteq X$, *then* $\mu^*(E) \leq \mu^*(F)$.
  3. μ^* *is* **countably subadditive**. *In other words, if* $\{E_i \subseteq X : i \in \mathbb{N}\}$ *is a countable collection of subsets of* X, *then*

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

- Lemma 29 shows that μ^* is an outer measure on X if μ is a premeasure on an algebra $\mathcal{A}$ on X. The outer measure μ^* is called the **outer measured generated by** μ.

3.3 Measure from Outer Measure

- μ^* is defined for arbitrary subsets of X, so it is also defined on countable unions of subsets of X as well. However, it is not yet a fully fledged measure because we cannot yet find a σ-algebra on which it is countably additive on.

- The following criterion is used to classify members of such a σ-algebra.

  Definition 31. *A subset* E *of* X *is said to be* μ^***-measurable** *if*

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

  for every subset A *of* X.

- A μ^*-measurable set E splits any set A into pieces whose output measures add up to the outer measure of A. In other words, a set is μ^*-measurable if it splits other sets in a “nice” way.

- **Theorem 32 (Carathéodory's extension theorem).** *Let* μ^* *be an outer measure on* X. *The collection of* μ^**-measurable sets is a* σ*-algebra on* X. *Moreover,* μ^* *is a measure on this collection.*

Proof. Let $\mathcal{X}^*$ denote the set of μ^*-measurable sets on X. We shall show that $\mathcal{X}^*$ is a σ-algebra, and μ^* is a measure on it.

($\infty, X \in \mathcal{X}^*$) Because μ^* is a premeasure, we have that $\mu^*(\emptyset) = 0$. For any set $A \subseteq X$, we have that

$$\mu^*(A) = \mu^*(\emptyset) + \mu^*(A) = \mu^*(A \cap \emptyset) + \mu^*(A \cap X).$$

Because $\emptyset$ and X are complements of each other, it follows that they are μ^*-measurable and so belong to $\mathcal{X}^*$.

(Closure under complementation) Let E be a μ^*-measurable set. It follows that

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

This implies that E^c is also μ^*-measurable.

(Closure under finite unions) This step is required to show closure under countable union. Suppose that E and F are μ^*-measurable. Let $A \subseteq X$. We need to show that

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

Because $A = (A \cap (E \cup F)) \cup (A \cap (E \cup F)^c)$ and μ^* is subadditive, we have that

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

Thus, it remains to show that $\mu^*(A) \geq \mu^*(A \cap (E \cup F)) + \mu^*(A \cap (E \cup F)^c)$. Because E is μ^*-measurable, we have that

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

Since both $A \cap E$ and $A \cap E^c$ are subsets of X and B is μ^*-measurable, we have

$$\begin{aligned} \mu^*(A \cap E) &= \mu^*(A \cap E \cap F) + \mu^*(A \cap E \cap F^c) \\ \mu^*(A \cap E^c) &= \mu^*(A \cap E^c \cap F) + \mu^*(A \cap E^c \cap F^c). \end{aligned}$$

As a result,

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

Because $E \cup F = (E \cap F) \cup (E \cap F^c) \cup (E^c \cap F)$, we have that

$$A \cap (E \cup F) = (A \cap E \cap F) \cup (A \cap E \cap F^c) \cup (A \cap E^c \cap F).$$

So,

$$\mu^*(A \cap (E \cup F)) \leq \mu^*(A \cap E \cap F) \cup \mu^*(A \cap E \cap F^c) \cup \mu^*(A \cap E^c \cap F).$$

Thus,

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

It follows that $\mu^*(A) = \mu^*(A \cap (E \cup F)) + \mu^*(A \cap (E \cup F)^c)$, and so $E \cup F$ is also measurable.

(Closure under finite intersections) We briefly mention that closure under complementation and closure under finite union implies closure under complementation. This is a consequence of de Morgan's law: $E \cap F = (E^c \cup F^c)^c$.

(Closure under disjoint countable unions implies closure under countable unions) We need to show closure under countable union. However, it suffices to only show that $\bigcup E_j$ is measurable for all sequences (E_j) where the sets are measurable and disjoint. To see this, let F_j denote the union of the first j sets:

$$F_j = \bigcup_{i=1}^{j} E_j,$$

and let $G_0 = E_1$ and $G_j = F_{j+1} - F_j = F_{j+1} \cap F_j^c$ for $j \geq 1$. It follows that, G_j is measurable for all j because $\mathcal{X}^*$ is closed under finite unions and intersections. Moreover, the G_j's are disjoint, and

$$\bigcup_{j=1}^{\infty} E_j = \bigcup_{j=0}^{\infty} G_j.$$

Hence, the measurability of $\bigcup_{j=0}^{\infty} G_j$ imples the measurability of $\bigcup_{j=1}^{\infty} E_j$.

(Finite additivity) In order to establish closure under countable unions, it is useful to show that μ^* is additive for finite disjoint unions. Let E and F be any disjoint measurable sets. Because E is measurable, we have that

$$\mu^*(E \cup F) = \mu^*((E \cup F) \cap E) + \mu^*((E \cup F) \cap E^c) = \mu^*(E) + \mu^*((E \cup F) - E) = \mu^*(E) + \mu^*(F).$$

(Closure under disjoint countable unions) Let (E_j) be a sequence of disjoint measurable sets. Let

$$F_j = \bigcup_{i=1}^{j} E_j, \qquad\qquad F = \bigcup_{i=1}^{\infty} E_j.$$

We need to show show that, for any set $A \subseteq X$, it is true that $\mu^*(A) = \mu^*(A \cap F) + \mu^*(A \cap F^c)$. Because μ^* is subadditive, we already know that $\mu^*(A) \leq \mu^*(A \cap F) + \mu^*(A \cap F^c)$. So, we only need to show that $\mu^*(A) \geq \mu^*(A \cap F) + \mu^*(A \cap F^c)$.

For any j, we have that F_j is measurable because of closure under finite unions. So

$$\begin{aligned}
\mu^*(A) &= \mu^*(A \cap F_j) + \mu^*(A \cap F_j^c) \\
&= \mu^*\left(A \cap \bigcup_{i=1}^{j} E_i\right) + \mu^*(A \cap F_j^c) \\
&= \mu^*\left(\bigcup_{i=1}^{n} (A \cap E_i)\right) + \mu^*(A \cap F_j^c).
\end{aligned}$$

Because $A \cap E_1$, $A \cap E_2$, ..., and $A \cap E_j$ are mutually disjoint, we have that

$$\mu^*(A) = \sum_{i=1}^{j} \mu^*(A \cap E_j) + \mu^*(A \cap F_j^c).$$

Moreover, because $F_j \subseteq F$, it follows that $A \cap F_j^c \supseteq A \cap F^c$. Hence, $\mu^*(A \cap F_j^c) \geq \mu^*(A \cap F^c)$. Thus,

$$\mu^*(A) \geq \sum_{i=1}^{j} \mu^*(A \cap E_i) + \mu^*(A \cap F^c).$$

Taking the limit as $j \to \infty$, we have that

$$
\begin{aligned}
\mu^*(A) &\geq \sum_{i=1}^{\infty} \mu^*(A \cap E_i) + \mu^*(A \cap F^c) \\
&\geq \mu^*\left(\bigcup_{i=1}^{\infty} (A \cap E_i)\right) + \mu^*(A \cap F^c) \\
&= \mu^*\left(A \cap \bigcup_{i=1}^{\infty} E_i\right) + \mu^*(A \cap F^c) \\
&= \mu^*(A \cap F) + \mu^*(A \cap F^c),
\end{aligned}
$$

which implies that F is μ^*-measurable.

(Countable additivity) In the proof of closure under countable union, we established that

$$
\begin{aligned}
\mu^*(A) &\geq \sum_{i=1}^{\infty} \mu^*(A \cap E_i) + \mu^*(A \cap F^c) \\
&\geq \mu^*\left(\bigcup_{i=1}^{\infty} (A \cap E_i)\right) + \mu^*(A \cap F^c) \\
&= \mu^*(A \cap F) + \mu^*(A \cap F^c) \\
&\geq \mu^*(A).
\end{aligned}
$$

Therefore, it must be the case that

$$
\sum_{i=1}^{\infty} \mu^*(A \cap E_i) + \mu^*(A \cap F^c) = \mu^*\left(\bigcup_{i=1}^{\infty} (A \cap E_i)\right) + \mu^*(A \cap F^c),
$$

and this is true for any set $A \subseteq X$. Taking $A = F$, we have that

$$
\begin{aligned}
\sum_{i=1}^{\infty} \mu^*(F \cap E_i) + \mu^*(F \cap F^c) &= \mu^*\left(\bigcup_{i=1}^{\infty} (F \cap E_i)\right) + \mu^*(F \cap F^c) \\
\sum_{i=1}^{\infty} \mu^*(E_i) + \mu^*(\emptyset) &= \mu^*\left(\bigcup_{i=1}^{\infty} E_i\right) + \mu^*(\emptyset) \\
\sum_{i=1}^{\infty} \mu^*(E_i) &= \mu^*\left(\bigcup_{i=1}^{\infty} E_i\right),
\end{aligned}
$$

which shows that μ^* is countably additive. □

- In the proof of Theorem 32, there is a useful fact that we will use later, so let us extract it into a lemma here.

  Lemma 33. *For any sequence of sets (E_j) where each set belongs to an algebra $\mathcal{A}$, there exists a sequence of disjoint sets (H_j) such that, for each j, we have that $H_j \in \mathcal{A}$, and $H_j \subseteq E_j$. Moreover,*

  $$\bigcup_{j=1}^{\infty} H_j = \bigcup_{j=1}^{\infty} E_j$$

  Proof. Choose H_j to be G_{j-1} as defined in the proof of Theorem 32. We have that all the properties we want have already been proven except for $H_j \subseteq E_j$. However, this should be clear because

  $$H_j = G_{j-1} = F_j \cap F_{j-1}^c = (E_j \cup F_{j-1}) \cap F_{j-1}^c = E_j \cap F_{j-1}^c \subseteq E_j.$$

  We are done. □

- **Proposition 34.** *The measure space $(X, \mathcal{X}^*, \mu^*)$ in Theorem 32 is complete.*

  Proof. Let E be a μ^*-measurable set of measure zero. Let $B \subseteq E$. We need to show that B is also μ^*-measurable, and that $\mu^*(B) = 0$. The latter statement is immediate because μ^* is subadditive.

  Let $A \subseteq X$. First, $0 \leq \mu^*(A \cap B) \leq \mu^*(A \cap E) \leq \mu^*(E) = 0$, so $\mu^*(A \cap B) = 0$. Now,

$$\begin{aligned} A \cap B^c &\subseteq A \\ \mu^*(A \cap B^c) &\leq \mu^*(A) \\ \mu^*(A \cap B) + \mu^*(A \cap B^c) &\leq \mu^*(A). \end{aligned}$$

  However, because μ^* is subadditive, we have that $\mu^*(A \cap B) + \mu^*(A \cap B^c) \geq \mu^*(A)$. It follows that $\mu^*(A \cap B) + \mu^*(A \cap B^c) = \mu^*(A)$, and B is also μ^*-measurable. □

- **Theorem 35.** *Let μ be a premeasure on an algebra $\mathcal{A}$ on X. Let μ^* be the outer measure generated by μ, and let $\mathcal{A}^*$ be the collection of μ^*-measurable sets. Then, $\mathcal{A} \subseteq \mathcal{A}^*$, and $(X, \mathcal{A}^*, \mu^*)$ is complete measure space with the property that $\mu^*(E) = \mu(E)$ for all $E \in \mathcal{A}$.*

  Proof. That μ^* is a σ-algebra on $\mathcal{A}^*$ follows from Theorem 32. That $(X, \mathcal{A}^*, \mu^*)$ is a complete measure space follows from Proposition 34. That $\mu^*(E) = \mu(E)$ for all $E \in \mathcal{A}$ follows from Lemma 29. It remains to show that $\mathcal{A} \subseteq \mathcal{A}^*$.

  Let $E \in \mathcal{A}$. Let $A \subseteq X$. Let (E_j) be a collection of sets in $\mathcal{A}$ that covers A. According to Lemma 28, we can choose (E_j) so that

$$\sum_{j=1}^{\infty} \mu^*(E_j) = \sum_{j=1}^{\infty} \mu(E_j) \leq \mu^*(A) + \varepsilon$$

  for any arbitrary $\varepsilon > 0$. Because (E_j) covers A, it follows that $(E_j \cap E)$ covers $A \cap E$, and $(E_j \cap E^c)$ covers $A \cap E^c$. As a result,

$$\begin{aligned} \mu^*(A \cap E) &\leq \sum_{j=1}^{\infty} \mu(E_j \cap E) = \sum_{j=1}^{\infty} \mu^*(E_j \cap E), \\ \mu^*(A \cap E^c) &\leq \sum_{j=1}^{\infty} \mu(E_j \cap E^c) = \sum_{j=1}^{\infty} \mu^*(E_j \cap E^c). \end{aligned}$$

  Adding the above two inequalities, we have that

$$\begin{aligned} \mu^*(A \cap E) + \mu^*(A \cap E^c) &\leq \sum_{j=1}^{\infty} \mu^*(E_j \cap E) + \sum_{j=1}^{\infty} \mu^*(E_j \cap E^c) \\ &= \sum_{j=1}^{\infty} \big[\mu^*(E_j \cap E) + \mu^*(E_j \cap E^c)\big]. \end{aligned}$$

  Because $E_j \cap E$ and $E_j \cap E^c$ are disjoint, we have that

$$\mu^*(A \cap E) + \mu^*(A \cap E^c) \leq \sum_{j=1}^{\infty} \big[\mu^*(E_j \cap E) + \mu^*(E_j \cap E^c)\big] = \sum_{j=1}^{\infty} \mu(E_j) \leq \mu^*(A) + \varepsilon.$$

  As ε is arbitrary, it follows that $\mu^*(A \cap E) + \mu^*(A \cap E^c) \leq \mu(A)$. Moreover, because μ^* is subadditive, we have that $\mu^*(A \cap E) + \mu^*(A \cap E^c) \geq \mu^*(A)$. It follows that E is measurable, and $\mathcal{A} \subseteq \mathcal{A}^*$. □

- **Theorem 36 (Hahn's extension theorem).** *If μ is a σ-finite premeasure on an algebra $\mathcal{A}$, then μ^* is the unique extension of μ that is a measure on $\mathcal{A}^*$.*

Proof. Let ν be an extension of μ that is a measure on $\mathcal{A}^*$. We have that $\nu(E) = \mu(E) = \mu^*(E)$ for all $E \in \mathcal{A}$. We need to show that $\nu(E) = \mu^*(E)$ for all $E \in \mathcal{A}^*$.

Let $E \subseteq \mathcal{A}^*$. Let (E_j) be a sequence of sets that covers E such that $E_j \in \mathcal{A}$ for all j. We have that

$$\nu(E) \le \nu\bigg(\bigcup_{j=1}^{\infty} E_j\bigg) \le \sum_{j=1}^{\infty} \nu(E_j) = \sum_{j=1}^{\infty} \mu(E_j).$$

It follows that $\nu(E)$ is a lower bound of the set $\{\sum \mu(E_j) : E_j \in \mathfrak{C}(E)\}$. As a result, $\nu(E) \le \mu^*(E)$ for any $E \in \mathcal{A}^*$.

Since μ is σ-finite, there exists a sequence (E_j) with each $E_j \in \mathcal{A}$ such that $\bigcup_j E_j = X$ and $\mu(E_j) = \mu^*(E_j) = \nu(E_j)$ is finite for all j. Applying Lemma 33 to (E_j), we have a disjoint sequence (H_j) that covers X. Morever, $\mu(H_j) \le \mu(E_j) < \infty$.

For each H_j, we have that

$$\begin{aligned}
\mu^*(H_j) &= \nu(H_j) && (H_j \in \mathcal{A}) \\
\mu^*((H_j \cap E) \cup (H_j \cap E^c)) &= \nu((H_j \cap E) \cup (H_j \cap E^c)) && \\
\mu^*(H_j \cap E) + \mu^*(H_j \cap E^c) &= \nu(H_j \cap E) + \nu(H_j \cap E^c) && (\text{both } \mu^* \text{ and } \nu \text{ are additive})
\end{aligned}$$

Because $\mu^*(H_j \cap E) \ge \nu(H_j \cap E)$ and $\mu^*(H_j \cap E^c) \ge \nu(H_j \cap E^c)$, the equality can hold only if $\mu^*(H_j \cap E) = \nu(H_j \cap E)$ and $\mu^*(H_j \cap E^c) = \nu(H_j \cap E^c)$.

Lastly,

$$\mu^*(E) = \mu^*(E \cap X) = \mu^*\bigg(E \cap \bigcup_{j=1}^{\infty} H_j\bigg) = \mu^*\bigg(\bigcup_{j=1}^{\infty} (E \cap H_j)\bigg) = \sum_{j=1}^{\infty} \mu^*(E \cap H_j).$$

Similarly,

$$\nu(E) = \sum_{j=1}^{\infty} \nu(E \cap H_j).$$

Because we just show that $\mu^*(H_j \cap E) = \nu(H_j \cap E)$ for all j, it follows that $\mu^*(E) = \nu(E)$. $\square$

3.4 Lebesgue Measure on the Real Line

- Recall that the length function ℓ is defined on $\mathcal{F}$, the collection of all finite unions of intervals of the form

  $$(a, b], (-\infty, b], (a, \infty), (-\infty, \infty),$$

  We have that $\mathcal{F}$ is an algebra (Lemma 24) and ℓ is a premeasure on it (Lemma 26).

- We can apply Theorem 32 to generate a measure space $(\mathbb{R}, \mathcal{F}^*, \ell^*)$, which is complete by construction (Proposition 34). Moreover, ℓ is σ-finite because the real line can be covered by intervals of length 1. Thus, by Theorem 36, ℓ^* is the only extension of ℓ on $\mathcal{F}^*$.

- The elements of the σ-algebra $\mathcal{F}^*$ are called the **Lebesgue measurable sets**. The measure ℓ^* is called the **Lebesgue measure** on $\mathbb{R}$.

- The smallest σ-algebra containing $\mathcal{F}$ is the Borel algebra $\mathcal{B}$.The restriction of Lebesgue measure to the Borel sets is called the **Borel measure**.

- Because $\mathcal{B} \subseteq \mathcal{F}^*$, we have that $\mathcal{F}^*$ is more extensive than $\mathcal{B}$ because it contains more sets. However, they can be practically ignored.

Proposition 37. *Let A be a Lebesgue measurable subset of $\mathbb{R}$. There exists a Borel measurable subset B of $\mathbb{R}$ such that $A \subseteq B$, and $\ell^*(B - A) = 0$.*

Proof. Let A be a Lebesgue measurable subset of $\mathbb{R}$. Let us assume for now that $\ell^*(A)$ is finite. For arbitrary $\varepsilon > 0$ and for each $n \in \mathbb{N}$, let $(E_{n,j})$ be a sequence of sets in $\mathcal{F}$ such that $(E_{n,j})$ covers A and $\sum \ell(E_{n,j}) \geq \ell^*(A) + \varepsilon/n$. Let $B_n = \bigcup E_{n,j}$. We have that B_n is a Borel set because it is a countable union of intervals. Moreover, $\ell^*(B_n) \leq \sum \ell(B_{n,j}) \leq \ell^*(A) + \varepsilon/n$. Take $B = \bigcap_{n=1}^{\infty} B_n$. We have that B is a Borel set because it is a countable intersection of Borel sets. We also have that $A \subseteq B$ because each B_n covers A. Moreover, let $B'_n = \bigcup_{i=1}^{n} B_i$. We have that $\ell^*(B'_1)$ is finite, and the sequence B'_n is decreasing, and $\ell^*(B'_n) \leq \ell^*(B_n) \leq \mu^*(A) + \varepsilon/n$ By Lemma 7, we have that

$$\ell^*(B) = \lim_{n\to\infty} \ell^*(B'_n) \leq \lim_{n\to\infty} \ell^*(B_n) \leq \lim_{n\to\infty} \left(\ell^*(A) + \frac{\varepsilon}{n} \right) = \ell^*(A).$$

However, because $A \subseteq B$, it follows that $\ell^*(B) \geq \ell^*(A)$, and thus $\ell^*(B) = \ell^*(A)$. Because $\ell^*(B)$ is finite, it follows that $\ell^*(B - A) = \ell^*(B) - \ell^*(A) = 0$.

Let us now remove the assumption that $\ell^*(A)$ is finite. Let $A_k = A \cap [k, k+1)$ for any $n \in \mathcal{Z}$. We have that the A_k's are disjoint. Moreover $\ell^*(A_k) \leq \ell^*([k, k+1)) = 1$. Because A_k is finite, there exists a Borel set B_k such that $A_k \subseteq B_k$ and $\ell^*(B_k) = \ell^*(A_k)$ and $\ell^*(B_k - A_k) = 0$. Take $B = \bigcup_{k=-\infty}^{\infty} B_k$. It follows that B is a Borel set because it is a countable union of Borel sets. Moreover, $A \subseteq B$ obviously. Lastly, $B - A = \bigcap_{k=-\infty}^{\infty} (B_k - A_k)$. Let $C_j = \bigcup_{k=-j}^{j} (B_j - A_j)$. We have that C_j is an increasing sequence of sets. By Lemma 7,

$$\ell^*(B - A) = \ell^* \left(\bigcup_{j=0}^{\infty} C_j \right) = \lim_{j\to\infty} \ell^*(C_j) = 0.$$

We are done. $\square$

3.5 Lebesgue Measures in Higher Dimensions

- Note that the construction in this section can be extended to $\mathbb{R}^k$.
- Here, the length of intervals becomes the volume of "rectangles" of the form

$$I_1 \times I_2 \times \cdots \times I_k$$

where I_j is an interval of the forms in (1). The volume is then defined to be

$$v(I_1 \times I_2 \times \cdots \times I_k) = \ell(I_1)\,\ell(I_2)\,\cdots\,\ell(I_n).$$

- The measure generated in this way is called the **Lebesgue measure on** $\mathbb{R}^k$. There are also **Borel measure on** $\mathbb{R}^k$ and **Borel algebra on** $\mathbb{R}^k$. All the results proved thus far do hold on them.

4 Measurable Functions

An important component of measure theory is the definition of integrals of functions and the study of their properties. Measurable functions are nice functions upon which integrals can be defined.

4.1 Measurable Real-Valued Functions

- Consider a fixed measurable space $(X, \mathcal{X})$.

- **Definition 38.** *A function $f : X \to \mathbb{R}$ is said to be* $\mathcal{X}$**-measurable** *(or simply* **measurable***) if, for every real number α, the set $\{x \in X : f(x) > \alpha\}$ is measurable (in other words, belongs to $\mathcal{X}$).*

- **Proposition 39.** *For a function $f : X \to \mathbb{R}$, the following statements are equivalent.*

  (a) For every $\alpha \in \mathbb{R}$, the set $A_\alpha = \{x \in X : f(x) > \alpha\}$ belongs to $\mathcal{X}$.

  (b) For every $\alpha \in \mathbb{R}$, the set $B_\alpha = \{x \in X : f(x) \leq \alpha\}$ belongs to $\mathcal{X}$.

  (c) For every $\alpha \in \mathbb{R}$, the set $C_\alpha = \{x \in X : f(x) \geq \alpha\}$ belongs to $\mathcal{X}$.

  (d) For every $\alpha \in \mathbb{R}$, the set $D_\alpha = \{x \in X : f(x) < \alpha\}$ belongs to $\mathcal{X}$.

  Proof. Because $A_\alpha = B_\alpha^c$, we have that (a) and (b) are equivalent. The same can be said for (c) and (d).

  We will now show that (a) implies (c). Let $\alpha \in \mathbb{R}$. Consider the sequence $(A_{a-1/n})$ for $n \in \mathbb{N}$, all of which belongs to $\mathcal{X}$ because of (a). So, there countable intersection $\bigcap_{n=1}^{\infty} A_{\alpha-1/n} = C$ also belongs to $\mathcal{X}$.

  Next, we will show that (c) implies (a). This is a result of the fact that $A_\alpha = \bigcup_{n=1}^{\infty} C_{\alpha+1/n}$. ☐

- A constant function $f(x) = c$ is measurable. This is because, if $c \leq \alpha$, then $x \in X : f(x) > \alpha = \emptyset$. On the other hand, if $c > \alpha$, then $x \in X : f(x) > \alpha = X$.

- If $E \in \mathcal{X}$, then the **characteristic function** χ_E, defined by

$$\chi_E(X) = \begin{cases} 1, & x \in E \\ 0, & x \notin E \end{cases},$$

  is measureable. This is because $\{x \in X : \chi_E(x) > \alpha\}$ is either X, E, or $\emptyset$ depending on whether α is $[1, \infty)$, $[0, 1)$ or $(-\infty, 0)$, respectively.

- If $X = \mathbb{R}$, and $\mathcal{X}$ is the Borel algebra $\mathcal{B}$, then any continuous function $f : \mathbb{R} \to \mathbb{R}$ is Borel measurable. This is because (α, ∞) is an open set, so $f^((\alpha, \infty)) = \{x \in \mathbb{R} : f(x) > \alpha\}$ is open as well. (This is a fact from real analysis.)

- If $X = \mathbb{R}$ and $\mathcal{X} = \mathcal{B}$, then any monotone function is Borel measurable. This is because the set $\{x \in \mathbb{R} : f(x) > \alpha\}$ are of the form $\{x \in \mathbb{R} : x > \beta\}$ or $\{x \in \mathbb{R} : x \geq \beta\}$ or $\mathbb{R}$ or $\emptyset$, all of which are Borel measurable.

- **Lemma 40.** *Let f and g be measurable real-valued functions and let c be a real number. Then the functions*

$$cf,\ f^2,\ f + g,\ fg,\ |f|,\ 1/f, \min(f, g), \max(f, g)$$

  are also measurable. For the case of $1/f$, we assume that $f(x) \neq 0$ for all x.

  Proof. **(a)** If $c = 0$, then cf is a constant function and so is measurable. If $c > 0$, then

$$\{x \in X : cf(x) > \alpha\} = \{x \in X : f(x) > \alpha/c\}.$$

  The RHS is measurable, so cf is measurable. The case where $c < 0$ can be handled similarly.

  (b) If $\alpha < 0$, then $\{x \in X : (f(x))^2 > \alpha\} = X \in \mathcal{X}$. If $\alpha \geq 0$, then

$$\{x \in X : (f(x))^2 > \alpha\} = \{x \in X : f(x) > \sqrt{a}\} \cup \{x \in X : f(x) < -\sqrt{a}\}.$$

  Both sets on the RHS are measurable, so f^2 is measurable.

(c) We have that

$$\{x \in X : f(x) + g(x) < \alpha\} = \bigcup_{q+r<b; q,r \in \mathbb{Q}} \{x \in X : f(x) < q\} \cap \{x \in X : g(x) < r\}.$$

The RHS is a countable union of measurable sets. Hence, $f + g$ is measurable.
(d) fg is measurable because $fg = \frac{1}{2}[(f+g)^2 - f^2 - g^2]$.
(e) If $\alpha < 0$, then $\{x \in X : f(x) > 0\} = X$, which is measurable. If $\alpha \geq 0$, then

$$\{x \in X : |f(x)| > \alpha\} = \{x \in X : f(x) < -\alpha\} \cup \{x \in X : f(x) < \alpha\}.$$

Both sets on the RHS are measurable. As a result, $|f|$ is measurable.
(f) If $f(x) \neq 0$ for all x, we have that

$$\{x \in X : 1/f(x) < \alpha\} = \begin{cases} \{x \in X : 1/b < f(x) < 0\} & \text{if } b < 0, \\ \{x \in X : f(x) < 0\} & \text{if } b = 0, \\ \{x \in X : f(x) < 0\} \cup \{x \in X : f(x) > 1/\alpha\} & \text{if } b > 0. \end{cases}$$

In all cases, the RHS is a measurable set. It follows that $1/f$ is also measurable if $f(x) \neq 0$ for all x.
(g) and (h) We have that

$$\{x \in X : \min(f,g)(x) < \alpha\} = \{x \in X : f(x) < \alpha\} \cup \{x \in X : g(x) < \alpha\},$$
$$\{x \in X : \max(f,g)(x) < \alpha\} = \{x \in X : f(x) < \alpha\} \cap \{x \in X : g(x) < \alpha\}.$$

So $\min(f,g)$ and $\max(f,g)$ are measurable. □

- **Definition 41.** *For any function $f : X \to \mathbb{R}$, let f^+ and f^- be non-negative functions defined by:*

$$f^+(x) = \max(f(x), 0) \qquad\qquad f^-(x) = \max(-f(x), 0).$$

We call f^+ the **positive part** *of f, and f^- the* **negative part***.*

- It should be clear that, if f is measurable, then f^+ and f^- are also measurable.

4.2 Measurable Extended Real-Valued Functions

- Working with functions with extended real values makes it more convenience to work with limits of sequences of functions.

- **Definition 42.** *An extended real-valued function on X is $\mathcal{X}$-measurable if $x \in X : f(x) > \alpha \in \mathcal{X}$ for all $\alpha \in \mathbb{R}$. The collection of extended real-valued $\mathcal{X}$-measurable functions is denoted by $M(X, \mathcal{X})$.*

- If $f \in M(X, \mathcal{X})$, then

$$\{x \in X : f(x) = \infty\} = \bigcap_{n=1}^{\infty} \{x \in X : f(x) > n\}$$
$$\{x \in X : f(x) = -\infty\} = \bigcap_{n=1}^{\infty} \{x \in X : f(x) < -n\},$$

so these sets are also in $\mathcal{X}$ automatically.

- **Lemma 43.** *An extended real-valued function f is measurable if and only if (1) the sets*

$$A = \{x \in X : f(x) = \infty\},$$
$$B = \{x \in X : f(x) = -\infty\}$$

are measurable, and (2) the real-valued function f_1 defined by

$$f_1(x) = \begin{cases} f(x), & x \in A \cup B, \\ 0, & x \notin A \cup B \end{cases}$$

is measurable.

Proof. TODO □

- The consequence to the last lemma and the lemmas in the last section is that, if f is $M(X, \mathcal{X})$, then the functions

$$cf, f^2, |f|, f^+, f^-$$

are also in $M(X, \mathcal{X})$.

- For cf, we use the convention that $0(\pm\infty) = 0$.
- For $f + g$, note that the sum is not well-defined when $f(x)$ and $g(x)$ are infinities with different signs.
- **Lemma 44.** *Let (f_n) be a sequence of functions in $M(X, \mathcal{X})$. Define*

$$f(x) = \inf_{n \geq 1} \{f_n(x)\},$$
$$F(x) = \sup_{n \geq 1} \{f_n(x)\},$$
$$f^*(x) = \liminf_{n \to \infty} \{f_n(x)\},$$
$$F^*(x) = \limsup_{n \to \infty} \{f_n(x)\}.$$

Then f, F, f^*, F^* all belong to $M(X, \mathcal{X})$.

Proof. TODO □

- **Corollary 45.** *If (f_n) is a sequence in $M(X, \mathcal{X})$ which converges to f on X, then f is also in $M(X, \mathcal{X})$.*

  Proof. We have that $f(x) = \lim f_n(x) = \liminf f_n(x)$. □

- **Lemma 46.** *If $f, g \in M(X, \mathcal{X})$, then $fg \in M(X, \mathcal{X})$ too.*

  Proof. TODO □

- The following lemma establishes the fact that a measurable non-negative function can be approximated by an increasing sequence of functions, all of which takes on a finite number of real values.

  Lemma 47. *If f is non-negative function in $M(X, \mathcal{X})$, then there exists a sequence (φ_n) in $M(X, \mathcal{X})$ such that:*

  (a) $0 \leq \varphi_n(x) \leq \varphi_{n+1}(x)$ *for all* $x \in X$ *and* $n \in \mathbb{N}$.
  (b) $f(x) = \lim \varphi_n(x)$ *for each* $x \in X$.
  (c) Each φ_n has only a fininte number of real values.

  Proof. TODO □

5 Integration

- In this section, we consider a fixed measure space $(X, \mathcal{X}, \mu)$.
- Denote the set of all $\mathcal{X}$-measurable functions by $M = M(X, \mathcal{X})$. Denote the set of all non-negative $\mathcal{X}$-measurable functions by $M^+ = M^+(X, \mathcal{X})$.
- **Definition 48.** *A real-valued function is* **simple** *if it has only a finite number of values.*
- A simple measurable function φ can be represented by

$$\varphi = \sum_{j=1}^{n} a_j \chi_{E_j}$$

  where $a_j \in \mathbb{R}$, and χ_{E_j} is the characteristic function fo a set E_j in $\mathcal{X}$.
- The **standard representation** is the representation where the a_j's are distinct, and the E_j's are disjoint, non-empty subsets of X. Moreover, we require that $X = \bigcup_{j=1}^{n} E_j$.
- **Definition 49.** *If φ is a simple function in $M^+(X, \mathcal{X})$ with the standard representation $\varphi = \sum a_j \chi_{E_j}$, we define the* **integral** *of φ with respect to μ to be the extended real number*

$$\int \varphi \, \mathrm{d}\mu = \sum_{j=1}^{n} \alpha_j \, \mu(E_j).$$

- Note that the value of the integral is always well defined because all the a_j's are positive, so we never encounter an expression of the form $\infty + (-\infty)$.
- **Lemma 50.** *If φ and ψ are simple functions in $M^+(X, \mathcal{X})$ and $c \geq 0$, then*

$$\int c\varphi \, \mathrm{d}\mu = c \int \varphi \, \mathrm{d}\mu,$$
$$\int (\varphi + \psi) \, \mathrm{d}\mu = \int \varphi \, \mathrm{d}\mu + \int \psi \, \mathrm{d}\mu.$$

  Proof. TODO □
- **Lemma 51.** *Let φ be a simple function in $M^+(X, \mathcal{X})$. For each $E \in \mathcal{X}$, define $\lambda(E)$ to be*

$$\lambda(E) = \int \varphi \, \chi_E \, \mathrm{d}\mu.$$

  Then, λ is a measure on $\mathcal{X}$.

  Proof. TODO □
- **Definition 52.** *If f belongs to $M^+(X, \mathcal{X})$, define the* **(Lebesgue) integral of f with respect to μ** *to be the extended real number*

$$\int f \, \mathrm{d}\mu = \sup \left\{ \int \varphi \, \mathrm{d}\mu \;\middle|\; \varphi \in M^+(X, \mathcal{X}) \text{ is simple, and } \varphi(x) \leq f(x) \text{ for all } x \in S \right\}$$

  For any $E \in \mathcal{X}$, define the **(Lebesgue) integral of f over E with respect to μ** *to be the extended real number*

$$\int_E f \, \mathrm{d}\mu = \int f \chi_E \, \mathrm{d}\mu.$$

- **Lemma 53.** *If $f, g \in M^+(X, \mathcal{X})$, then $\int f \,\mathrm{d}\mu \leq \int g \,\mathrm{d}\mu$.*

  Proof. TODO □

- **Lemma 54.** *If $f \in M^+(X, \mathcal{X})$, $E, F \in \mathcal{X}$, and $E \subseteq F$, then $\int_E f \,\mathrm{d}\mu \leq \int_F f \,\mathrm{d}\mu$.*

  Proof. TODO □

- **Theorem 55 (Monotone Convergence Theorem).** *If (f_n) is a monotonically increasing sequence of functions in $M^+(X, \mathcal{X})$ which converges to f, then*

$$\int f \,\mathrm{d}\mu = \lim_{n\to\infty} \int f_n \,\mathrm{d}\mu.$$

  Proof. TODO □

- **Corollary 56.** *If f belongs to $M^+(X, \mathcal{X})$ and $c \geq 0$, then $cf \in M^+(X, \mathcal{X})$, and*

$$\int cf \,\mathrm{d}\mu = c \int f \,\mathrm{d}\mu.$$

  Proof. TODO □

- **Corollary 57.** *If f and g belong to $M^+(X, \mathcal{X})$, then $f + g \in M^+(X, \mathcal{X})$, and*

$$\int (F + g) \,\mathrm{d}\mu = \int f \,\mathrm{d}\mu + \int g \,\mathrm{d}\mu.$$

  Proof. TODO □

- **Theorem 58 (Fatou's Lemma).** *If (f_n) is a sequence of functions in $M^+(X, \mathcal{X})$, then*

$$\int \left(\liminf_{n\to\infty} f_n\right) \mathrm{d}\mu \leq \liminf_{n\to\infty} \int f_n \,\mathrm{d}\mu.$$

  Proof. TODO □

- **Corollary 59.** *If f belongs to $M^+(X, \mathcal{X})$, and λ is defined on $\mathcal{X}$ by*

$$\lambda(E) = \int_E f \,\mathrm{d}\mu,$$

  then λ is a measure.

  Proof. TODO □

- **Corollary 60.** *Let f belong to $M^+(X, \mathcal{X})$. Then, $f(x) = 0$ μ-almost everyone on X if and only if $\int f \,\mathrm{d}\mu = 0$.*

  Proof. TODO □

- **Corollary 61.** *Let f belong to $M^+(X, \mathcal{X})$, and let λ be defined as in Corollary 59. Then, the measure λ is absolutely continuous with respect to μ in the sense that if $E \in \mathcal{X}$, then $\mu(E) = 0$, then $\lambda(E) = 0$.*

  Proof. TODO □

- **Corollary 62.** *If (f_n) is a monotonically increasing sequence of functions in $M^+(X, \mathcal{X})$ which converges μ-almost everywhere on X to a function f in $M^+(X, \mathcal{X})$, then*

$$\int f \, \mathrm{d}\mu = \lim_{n\to\infty} \int f_n \, \mathrm{d}\mu.$$

  Proof. TODO $\square$

- **Corollary 63.** *If (g_n) be a sequence of functions in $M^+(X, \mathcal{X})$, then*

$$\int \left(\sum_{n=1}^{\infty} g_n \right) \mathrm{d}\mu = \sum_{n=1}^{\infty} \left(\int g_n \, \mathrm{d}\mu \right).$$

  Proof. TODO $\square$

6 Integrable Functions

- **Definition 64.** *The collection $L = L(X, \mathcal{X}, \mu)$ of* **integrable functions** *consists of all real-valued $\mathcal{X}$-measurable functions f defined on X, such that both the positive and negative parts (f^+ and f^-) of f have finite integrals with respective to μ.*

- **Definition 65.** *If $f \in L(X, \mathcal{X}, \mu)$, the* **integral of f with respect to μ** *is defined to be*

$$\int f \, \mathrm{d}\mu = \int f^+ \, \mathrm{d}\mu + \int f^- \, \mathrm{d}\mu.$$

  If $E \in \mathcal{X}$, define

$$\int_E f \, \mathrm{d}\mu = \int_E f^+ \, \mathrm{d}\mu + \int_E f^- \, \mathrm{d}\mu.$$

- **Definition 66.** *Let $\mathcal{X}$ be a σ-algebra on a set X. A function $\mu : \mathcal{X} \to \mathbb{R}$ is said to be a* **charge** *on $\mathcal{X}$ if the following properties are satisfied.*

  1. *$\mu(\emptyset) = 0$.*
  2. *μ is countably additive. This is, for a sequence (E_n) of disjoint sets, it holds that*

$$\mu\left(\bigcup_{n=1}^{\infty} E_n \right) = \sum_{n=1}^{\infty} \mu(E_n).$$

- The difference between a charge and a measure is that a measure is always non-negative, but a charge can be negative. Moreover, a charge cannot take infinite values.

- **Lemma 67.** *If f belongs to L, and $\lambda(E) = \int_E f \, \mathrm{d}\mu$, then λ is a charge.*

  Proof. TODO $\square$

- The function λ defined as in the above lemma is often called the **indefinite integral of f with respect to μ**.

- Since λ is a charge, if (E_n) is a disjoint sequence of sets in $\mathcal{X}$ whose union is E, then

$$\int_E f \, \mathrm{d}\mu = \sum_{n=1}^{\infty} \int_{E_n} f \, \mathrm{d}\mu.$$

  In other words, **the indefinite integral of a function in L is countably additive**.

- **Theorem 68.** *A measurable function f belongs fo L if and only if $|f|$ belons to L. Moreover,*

$$\left|\int f \,\mathrm{d}\mu\right| \leq \int |f| \,\mathrm{d}\mu.$$

Proof. f belongs to L if and only if f^+ and f^- belong to M^+. Now, we have that $|f|^+ = |f| = f^+ + f^-$, and $|f|^- = 0$. Hence, $|f|$ also belongs to L. The converse can be proven in a similar fasion. Moreover,

$$\left|\int f \,\mathrm{d}\mu\right| = \left|\int f^+ \,\mathrm{d}\mu - \int f^- \,\mathrm{d}\mu\right| \leq \int f^+ \,\mathrm{d}\mu + \int f^- \,\mathrm{d}\mu = \int |f| \,\mathrm{d}\mu$$

as required. □

- **Corollary 69.** *If f is measurable, g is integrable, and $|f| \leq |g|$, then f is integrable, and*

$$\int |f| \,\mathrm{d}\mu \leq \int |g| \,\mathrm{d}\mu.$$

Proof. First, we have that $|f| \in M^+$, so $\int |f| \,\mathrm{d}\mu$ is well defined. Moreover, we know that it has a finite value because we can apply Lemma 46 to $|f|$ and $|g|$ to conclude that $\int |f| \,\mathrm{d}\mu \leq \int |g| \,\mathrm{d}\mu$. Now, we can apply Theorem 68 to conclude that f is also integrable. □

- **Theorem 70.** *A constant multiple cf and a sum $f + g$ of integrable functions are integrable, and the integrals are given by:*

$$\int cf \,\mathrm{d}\mu = c \int f \,\mathrm{d}\mu,$$
$$\int (f+g) \,\mathrm{d}\mu = \int f \,\mathrm{d}\mu + \int g \,\mathrm{d}\mu.$$

Proof. TODO □

- **Theorem 71 (Lebesgue dominated convergence theorem).** *Let (f_n) be a sequence of integrable functions which converges almost everywhere to a real-valued measurable function f. If there exists an integrable function g such that $|f_n| \geq g$ for all n, then f in integrable and*

$$\int f \,\mathrm{d}\mu = \lim_{n\to\infty} f_n \mathrm{d}\mu.$$

Proof. TODO □

7 Integrals That Involve Limits

- In this section, let f denote a function with signature $X \times [a, b] \to \mathbb{R}$. Also, assume that $f(x, t)$ is $\mathcal{X}$-measurable for each $t \in [a, b]$.

- **Corollary 72.** *Suppose that, for some t_0 in $[a, b]$, we have that*

$$f(x, t_0) = \lim_{t\to t_0} f(x, t)$$

for each $x \in X$, and that there exists an integrable function g such that $|f(x,t)| \leq g(x)$. Then,

$$\int f(x, t_0) \,\mathrm{d}\mu(x) = \lim_{t\to t_0} \int f(x, t) \,\mathrm{d}\mu(x).$$

Proof. Just apply the dominated convergence theorem. □

- **Corollary 73.** *Let the function $t \mapsto f(x,t)$ be continuous on $[a,b]$ for all $x \in X$. If there is an integrable function g on X such that $|f(x,t)| \leq g(x)$, then the function F defined by*

$$F(t) = \int f(x,t)\,\mathrm{d}\mu(x)$$

is continuous for $t \in [a,b]$.

Proof. Apply the last corollary. □

- **Corollary 74.** *Suppose that for some $t_0 \in [a,b]$, the function $x \mapsto f(x,t_0)$ is integrable on X. Suppose that $\partial f/\partial t$ exists on $X \times [a,b]$. Moreover, let there be an integrable function g on X such that*

$$\left|\frac{\partial f}{\partial t}(x,t)\right| \leq g(x).$$

Then, the function

$$F(t) = \int f(x,t)\,\mathrm{d}\mu(x)$$

is differentiable on $[a,b]$, and

$$\frac{\mathrm{d}F}{\mathrm{d}t}(t) = \frac{\mathrm{d}}{\mathrm{d}t}\int f(x,t)\,\mathrm{d}\mu(x) = \int \frac{\partial f}{\partial t}(x,t)\,\mathrm{d}\mu(x).$$

Proof. TODO □

- **Corollary 75.** *Let the function $t \mapsto f(x,t)$ be continuous on $[a,b]$ for all $x \in X$. Let there be an integrable function g on X such that $|f(x,t)| \leq g(x)$. Let $F(t) = \int f(x,t)\,\mathrm{d}\mu(x)$. We have that*

$$\int_a^b F(t)\,\mathrm{d}t = \int_a^b \left(\int f(x,t)\,\mathrm{d}\mu(x) \right) \mathrm{d}t = \int \left(\int_a^b f(x,t)\,\mathrm{d}t \right) \mathrm{d}\mu(x)$$

where the integrals with respect to t are Reimann integrals.

Proof. TODO □

8 Product Measures and Double Integrals

- In this section, we show that the Cartesian product of two measurable spaces can be made into a measurable space in a natural fashion.

- Note that this gives us another way to construct a measure on $\mathbb{R}^k$.

  - If we force the measure to agree on rectangles, though, the measure would agree with the Lebesgue measure on the Borel sets in $\mathbb{R}^k$ by the Hahn's extension theorem.

- **Definition 76.** *If $(X,\mathcal{X})$ and $(Y,\mathcal{Y})$ are measurable spaces, then the set of the form $A \times B$ with $A \in \mathcal{X}$ and $B \in \mathcal{B}$ are called a* **measurable rectangle** *or simply* **rectangle** *in $Z = X \times Y$. Let $\mathcal{Z}_0$ denote the set of all finite unions of rectangles in Z.*

- It can be shown that each element of $\mathcal{Z}_0$ can be written as a finite union of disjoin rectangles in Z. (Just use the inclusion–exclusion principle.)

- **Lemma 77.** *The collection $\mathcal{Z}_0$ is an algebra on Z.*

  Proof. ($\emptyset \in \mathcal{Z}_0$) We have that $\emptyset \in \mathcal{X}$ and $\emptyset \in \mathcal{Y}$. So, $\emptyset = \emptyset \times \emptyset \in \mathcal{Z}_0$.

  ($Z \in \mathcal{Z}_0$) Since $X \in \mathcal{X}$ and $Y \in \mathcal{Y}$, we have that $Z = X \times Y \in \mathcal{Z}_0$.

  (Closure under complementation) If $C \in \mathcal{Z}_0$, then $C = A \times B$ for some $A \in \mathcal{X}$ and $B \in \mathcal{Y}$. It follows that $C^c = (A^c \times B) \cup (A \times B^c) \cup (A^c \times B^c)$, which is a finite unions of rectangles in Z, so $C^c \in \mathcal{Z}_0$.

  Closure under finite union follows from definition of $\mathbb{Z}_0$. □

- **Definition 78.** *Let $(X, \mathcal{X})$ and $(Y, \mathcal{Y})$ be measurable spaces. Let $\mathcal{Z} = \mathcal{X} \times \mathcal{Y}$ denote the σ-algebra of subsets of $Z = X \times Y$ generated by rectangles $A \times B$ with $A \in \mathcal{X}$ and $B \in \mathcal{Y}$. A set in $\mathcal{Z}$ is called a* **$\mathcal{Z}$-measurable set** *or as a* **measurable subset** *of Z.*

- **Theorem 79 (Product measure theorem).** *Let $(X, \mathcal{X}, \mu)$ and $(Y, \mathcal{Y}, \nu)$ be measure spaces. There exists a measure π defined on $\mathcal{Z} = \mathcal{X} \times \mathcal{Y}$ with*

  $$\pi(A \times B) = \mu(A)\,\nu(B)$$

  for all $A \in \mathcal{X}$ and $B \in \mathcal{Y}$. If the measure is σ-finite, then then it is unique. The measure π is called the **product** *of μ and ν.*

  Proof. TODO □

- **Definition 80.** *If E is a subset of $Z = X \times Y$, and $x \in X$, then the* **x-section** *of E is the set*

  $$E_x = \{y \in Y : (x, y) \in E\}.$$

  Similary, the **y-section** *is the set*

  $$E^y = \{x \in X : (x, y) \in E\}.$$

- **Lemma 81.** *If E is a measurable subset of E, then every section of E is measurable.*

  Proof. TODO □

- **Definition 82.** *For any $f : Z \to \overline{\mathbb{R}}$, and $x \in X$, the* **x-section** *of f is the function f_x defined on Y as*

  $$f_x(y) = f(x, y)$$

  for all $y \in Y$. The **y-section** *is defined as*

  $$f^y(x) = f(x, y)$$

  for all $x \in X$.

- **Lemma 83.** *If $f : Z \to \overline{\mathbb{R}}$ is a measurable function, then every section of f is measurable.*

  Proof. TODO □

- **Lemma 84.** *Let $(X, \mathcal{X}, \mu)$ and $(Y, \mathcal{Y}, \nu)$ be σ-finite measure spaces. If $E \in \mathcal{Z} = \mathcal{X} \times \mathcal{Y}$, then the; functions defined by*

  $$f(x) = \nu(E_x),$$
  $$g(y) = \mu(E^y)$$

  are measurable and

  $$\int_X f \,\mathrm{d}\mu = \pi(E) = \int_Y g \,\mathrm{d}\nu.$$

Proof. TODO. This requires the monotone class theorem (Theorem 23). ☐

- **Theorem 85 (Tonelli's theorem).** *Let $(X, \mathcal{X}, \mu)$ and $(Y, \mathcal{Y}, \nu)$ be σ-finite measure spaces. Let π be the product of μ and ν, which is a measure on $\mathcal{Z} = \mathcal{X} \times \mathcal{Y}$. Let $F : Z \to \overline{\mathbb{R}}$ be a non-negative measurable function. Then, the functions defined on X and Y by*

$$f(x) = \int_Y F_x \, \mathrm{d}\nu,$$
$$g(y) = \int_X F^y \, \mathrm{d}\mu$$

are measurable. Moreover,

$$\int_X f \, \mathrm{d}\mu = \int_Z F \, \mathrm{d}\pi = \int_Y g \, \mathrm{d}\nu.$$

Equivalently,

$$\int_X \left(\int_Y F_x \, \mathrm{d}\nu \right) \mathrm{d}\mu = \int_Z F \, \mathrm{d}\pi = \int_Y \left(\int_X F^y \, \mathrm{d}\mu \right) \mathrm{d}\nu.$$

Proof. TODO. This requires the monotone class theorem (Theorem 23). ☐

- **Theorem 86 (Fubini's theorem).** *Let $(X, \mathcal{X}, \mu)$ and $(Y, \mathcal{Y}, \nu)$ be σ-finite measure spaces. Let π be the product of μ and ν, which is a measure on $\mathcal{Z} = \mathcal{X} \times \mathcal{Y}$. Let $F : Z \to \mathbb{R}$ be an integrable function with respect to π. Then, the extended real value functions defined almost everywhere by*

$$f(x) = \int_Y F_x \, \mathrm{d}\nu,$$
$$g(y) = \int_X F^y \, \mathrm{d}\mu$$

have finite integrals. Moreover,

$$\int_X f \, \mathrm{d}\mu = \int_Z F \, \mathrm{d}\pi = \int_Y g \, \mathrm{d}\nu.$$

Equivalently,

$$\int_X \left(\int_Y F_x \, \mathrm{d}\nu \right) \mathrm{d}\mu = \int_Z F \, \mathrm{d}\pi = \int_Y \left(\int_X F^y \, \mathrm{d}\mu \right) \mathrm{d}\nu.$$

Proof. TODO ☐

9 Signed Measures and Their Decompositions

- **Definition 87.** *Let $(X, \mathcal{X})$ be a measurable space. A function $\lambda : \mathcal{X} \to \overline{\mathbb{R}}$ is said to be a* **signed measures** *if the following conditions are satisfied.*

  (a) $\lambda(\emptyset) = 0$.

  (b) λ *attains at most one of the values* ∞ *and* $-\infty$.

  (c) λ *is countably additive. In other words, if (A_n) is a sequence of disjoint sets in $\mathcal{X}$, then*

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

- The condition (b) is introduced to avoid undefined quantities such as $\infty - \infty$.
- A signed measure is very similar to a charge, but it can take an infinite value while a charge cannot. Note that a charge is a signed measure, but not the other way around.
- It can be shown that, if (E_n) is an increasing sequence of sets in $\mathcal{X}$, then

$$\lambda\left(\bigcup_{n=1}^{\infty} E_n\right) = \lim_{n\to\infty} \lambda(E_n).$$

  Similarly, if (F_n) is a decreasing sequence of sets in $\mathcal{X}$, then

$$\lambda\left(\bigcap_{n=1}^{\infty} F_n\right) = \lim_{n\to\infty} \lambda(F_n).$$

- **Definition 88.** *Let λ be a signed measure on $(X, \mathcal{X})$. A set P in $\mathcal{X}$ is said to be* **positive** *with respect to λ if $\lambda(E \cap P) \geq 0$ for any $E \in \mathcal{X}$. A set $N \in \mathcal{X}$ is said to be* **negative** *with respect to λ if $\lambda(E \cap N) \leq 0$ for all $E \in \mathcal{X}$. A set $M \in \mathcal{X}$ is said to be a* **null set** *for λ if $\lambda(E \cap M) = 0$ for all $E \in \mathcal{X}$.*
- It can be shown that a measurable subset of a positive set is positive, and the union of two positive sets is also positive.
- **Theorem 89 (Hahn decomposition theorem).** *If λ is a signed measure on $\mathcal{X}$, then there exist sets P and N in $\mathcal{X}$ with $X = P \cup N$, $P \cap N = \emptyset$, and such that P is positive and N is negative with respect to λ.*

  Proof. TODO □

- A pair (P, N) of measurable sets in the above theorem is said to form a **Hahn decomposition** of X with respect to λ.
- In general, Hahn decompositions are not unique. In fact, if M is a null set for λ, then $(P \cup M, N - M)$ and $(P - M, N \cup M)$ are also Hahn decompositions of X w.r.t λ.
- **Lemma 90.** *If (P_1, N_1) and (P_2, N_2) are Hahn decompositions of X w.r.t λ, and $E \in \mathcal{X}$, then*

$$\begin{aligned}\lambda(E \cap P_1) &= \lambda(N \cap P_2),\\ \lambda(E \cap N_1) &= \lambda(N \cap N_2).\end{aligned}$$

  Proof. TODO □

- **Definition 91.** *Let λ be a signed measure on $\mathcal{X}$, and let (P, N) be a Hahn decomponent w.r.t λ. The* **positive** *and* **negative variations** *of λ are the non-negative measures λ^+, λ^- defined by:*

$$\begin{aligned}\lambda^+(E) &= \lambda(E \cap P),\\ \lambda^-(E) &= -\lambda(E \cap N).\end{aligned}$$

  The **total variation** *of λ, denoted by $|\lambda|$, is defined as:*

$$|\lambda|(E) = \lambda^+(E) + \lambda^-(E).$$

- By Lemma 90, the positive and negative variations are well defined and do not depend on the particular choice of Hahn decomposition.

- We also have that

$$\lambda(E) = \lambda^+(E) - \lambda^-(E),$$

and this result is encapsulated in the following theorem.

Theorem 92 (Jordan decomposition theorem). *If λ is a signed measure on $\mathcal{X}$, then there exists measures λ^+ and λ^- on $\mathcal{X}$, at least one of which is finite, such that the following conditions are satisifed.*

(a) $\lambda = \lambda^+ - \lambda^-$.

(b) There exists sets M and N such that $M \cup N = X$, $M \cap N = \emptyset$, $\lambda^+(N) = 0$, and $\lambda^-(M) = 0$.

(c) If $\lambda = \mu - \nu$ where μ and ν are measures, then $\mu(E) \geq \lambda^+(E)$ and $\nu(E) \geq \lambda^-(E)$.

Proof. TODO □

- Let f be an integrable function with respect to a measure μ on $\mathcal{X}$. We can define

$$\lambda(E) = \int_E f \, \mathrm{d}\mu.$$

Lemma 67 tells us that λ is a charge.

- **Lemma 93.** *If $f \in L(X, \mathcal{X}, \mu)$ and $\lambda(E) = \int_E f \, \mathrm{d}\mu$, then λ^+, λ^- and $|\lambda|$ are given by:*

$$\lambda^+(E) = \int_E f^+ \, \mathrm{d}\mu, \qquad \lambda^-(E) = \int_E f^- \, \mathrm{d}\mu, \qquad |\lambda|(E) = \int_E |\lambda| \, \mathrm{d}\mu.$$

Proof. TODO □

10 Radon–Nikodym Theorem

- From Corollary 59, we can create a measure λ from a non-negative extended real-valued measurable function f by integrating with respect to an existing measure.
- The Radon–Nikodym theorem is the converse of the above corollary. It indicates when a measure λ can be expressed as an integration of a function f with respect to μ. Hence, it is useful in defining probability density functions.
- A necessary and sufficient condition for the theorem to hold is given below.

Definition 94. *A measure λ on $\mathcal{X}$ is said to be* **absolutely continuous** *with respect to a measure μ on $\mathcal{X}$ if $\lambda(E) = 0$ for all set $E \in \mathcal{X}$ such that $\mu(E) = 0$. In this case, we write $\lambda \ll \mu$. A signed measure λ is* **absolsution continuous** *with respect to a signed measure μ if the total variation $|\lambda|$ is absolutely continuous with respect to $|\mu|$.*

- **Lemma 95.** *Let λ and μ be finite measures on $\mathcal{X}$. Then, $\lambda \ll \mu$ if and only if, for every $\varepsilon > 0$, there exists a $\delta(\varepsilon) > 0$ such that $\lambda(E) < \varepsilon$ for all E such that $\mu(E) < \delta(\varepsilon)$.*

Proof. TODO □

- **Theorem 96 (Random–Nikodym theorem).** *Let λ and μ be σ-finite measures defined on $\mathcal{X}$, and suppose that λ is absolutely continous with respect to μ. Then, there exists a function $f \in M^+(X, \mathcal{X})$ such that*

$$\lambda(E) = \int_E f \, \mathrm{d}\mu.$$

Moreover, the function f is uniquely determined μ-almost everywhere.

Proof. TODO $\square$

- The function f above is often called the **Random–Nikodym derivative** of λ with respective to μ, and it is denoted by $\mathrm{d}\lambda/\mathrm{d}\mu$.

References

[Bartle, 1995] Bartle, R. G. (1995). *The Elements of Integration and Lebesgue Measure*. John Wiley & Sons, New York.

[Billingsley, 1995] Billingsley, P. (1995). *Probability and Measure*. John Wiley and Sons, third edition.

[Hunter, 2011] Hunter, J. K. (2011). Measure theory. `https://www.math.ucdavis.edu/~hunter/measure_theory/measure_notes.pdf`. Accessed: 2021-12-28.