McGill University - Fall 2023

Math 454: Honours Analysis III

Ritchie Yu

December, 2024

Lectures delivered by Prof. Linan Chen.

Contents

0 Preliminary concepts

0.1 Topology

Definition 0.1 (Topological space)**.** Let X be a set. Then a topology $\mathcal{T}$ is a collection of subsets of X such that

1. $\emptyset, X \in \mathcal{T}$
2. Any arbitrary union of elements in $\mathcal{T}$ is also in $\mathcal{T}$
3. Any finite intersection of elements in $\mathcal{T}$ is also in $\mathcal{T}$

The elements in $\mathcal{T}$ are called the *open sets* of X, and $(X, \mathcal{T})$ is called a *topological space*.

Definition 0.2 (Basis for a topology)**.** Let X be a set. Then a basis $\mathcal{B}$ for a topology on X is a collection of subsets of X such that

1. If $x \in X$, then there exists $B \in \mathcal{B}$ such that $x \in B$
2. If $x \in X$ and there exists $B_1, B_2 \in \mathcal{B}$ with $x \in B_1 \cap B_2$, then there exists B_3 such that $x \in B_3 \subseteq B_1 \cap B_2$.

The *topology generated by* $\mathcal{B}$ can then be defined. We say $U \subseteq X$ is *open in* X if for all $x \in U$, there exists $B \in \mathcal{B}$ such that $x \in B \subseteq U$.

Definition 0.3 (Standard topology on $\mathbb{R}$)**.** Let $\mathcal{B}$ be the collection of all open intervals on the real line.

$$\mathcal{B} := \{(a, b) : a, b \in \mathbb{R}, a < b\}$$

Then $\mathcal{B}$ generates the *standard topology* on $\mathbb{R}$.

1 Motivation

1.1 Limitations of the Riemann integral

In this course we will develop the theory of Lebesgue integration. At its core, Lebesgue integration allows for a more general notion of integration than Riemann integration. Recall the classic slogan of Riemann integration:

$$\int_a^b f(x)\,dx \text{ computes the "area under the graph of } f\text{"}$$

Formally, $\int_a^b f(x)\,dx$ is defined using the Riemann upper ($\overline{\int_a^b} f(x)\,dx$) and lower ($\underline{\int_a^b} f(x)\,dx$) integrals. We set

$$\overline{\int_a^b} f(x)\,dx := \inf\Big(\sum_{i=1}^{n} \sup_{x\in[x_{i-1},x_i]} f(x)\cdot(x_i - x_{i-1})\Big)$$
$$\underline{\int_a^b} f(x)\,dx := \sup\Big(\sum_{i=1}^{n} \inf_{x\in[x_{i-1},x_i]} f(x)\cdot(x_i - x_{i-1})\Big)$$

The outer infimum and supremum are each taken over all partitions of the interval $[a,b]$. Any partition $(x_0, x_1, \cdots, x_n)$ satisfies $a = x_0 < x_1 < \cdots < x_n = b$.

Definition 1.1. Let $f : [a,b] \to \mathbb{R}$ be a function. Then f is Riemann integrable over $[a,b]$ if

$$\overline{\int_a^b} f(x)\,dx = \underline{\int_a^b} f(x)\,dx =: \int_a^b f(x)\,dx$$

However, given a function f, we can conclude its Riemann integrability only if it "behaves nicely". For example:

1. If f is continuous on $[a,b]$, then f is Riemann integrable on $[a,b]$.
2. If f is monotonic on $[a,b]$, then f is Riemann integrable on $[a,b]$.
3. If f is bounded on $[a,b]$, and continuous except at possibly finitely many points, then f is Riemann integrable on $[a,b]$.

If a function is not nice, then it need not be Riemann integrable. Consider the Dirichlet function, for example, defined as

$$f(x) = \begin{cases} 0 & x \in \mathbb{Q} \cap [a,b] \\ 1 & x \in \mathbb{Q}^c \cap [a,b] \end{cases}$$

Due to the density of $\mathbb{Q}$ in $\mathbb{R}$, then $\sup_{x\in[x_{i-1},x_i]} f(x) = 1$ for any subinterval $[x_{i-1}, x_i]$. Thus the Riemann upper integral evaluates as

$$\overline{\int_a^b} f(x)\,dx = \inf\Big(\sum_{i=1}^{n} \sup_{x\in[x_{i-1},x_i]} f(x)\cdot(x_i - x_{i-1})\Big) = \inf\Big(\sum_{i=1}^{n}(x_i - x_{i-1})\Big) = b - a$$

Similarly, due to the density of $\mathbb{Q}^c$ in $\mathbb{R}$, then $\inf_{x\in[x_{i-1},x_i]} f(x) = 0$. Thus the Riemann lower integral evaluates as

$$\underline{\int_a^b} f(x)\,dx = \sup\Big(\sum_{i=1}^{n} \inf_{x\in[x_{i-1},x_i]} f(x)\cdot(x_i - x_{i-1})\Big) = \sup\Big(\sum_{i=1}^{n} 0\cdot(x_i - x_{i-1})\Big) = 0$$

The notion of Riemann integral is therefore restrictive – it cannot handle the complexity of certain functions, such as the Dirichlet function.

1.2 Basic idea behind the Lebesgue integral

The idea is to slice the graph of a function f horizontally, rather than vertically. This produces sets A_i where

$$A_i := \{x \in [a,b] : y_i \leq f(x) < y_{i+1}\}$$

Each A_i is the pre-image of f such that f stays within the unique slice indexed by i. The contribution of slice i to the area under the graph of f is then approximately

$$y_i \cdot (\text{size of } A_i)$$

We can then think of $\sum_i y_i \cdot (\text{size of } A_i)$ as an approximate notion of integral. **However, what do we mean by the size of $\mathbf{A_i}$?** The answer is not so trivial – we cannot easily calculate the "length" of A_i since each A_i may not even be the union of intervals. The problem of measuring sizes of sets led to the development of measure theory, which we explore in the first part of this course. **But first, what sets do we even want to measure?** The answer to this second question represents the convergence of mainstream opinion over the last century, and constitutes the foundation of probability theory.

2 Measure theory

2.1 Sigma algebra

Definition 2.1 (Sigma algebra)**.** Let X be a space (non-empty set) and let $\mathcal{F}$ be a collection of subsets of X. Then $\mathcal{F}$ is called a σ-algebra of subsets of X if

1. $X \in \mathcal{F}$
2. If $A \in \mathcal{F}$, then $A^c \in F$ (closed under complement)
3. If $\{A_n : n \geq\} \subseteq \mathcal{F}$, then $\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$ (closed under countable union)

Remark. For any theory, we first need some sort of minimal structure, and σ-algebras provide this structure. We will use σ-algebras $\mathcal{F}$ to specify the collection of subsets of X for which we define measure. **In other words, $\mathcal{F}$ describes the sets we want to measure**.

Proposition 2.1 (σ-algebra properties)**.** The properties below follow by definition.

4. $\emptyset \in \mathcal{F}$
5. If $\{A_n : n \geq 1\} \subseteq \mathcal{F}$, then $\bigcap_{n=1}^{\infty} A_n \in \mathcal{F}$ (closed under countable intersection)
6. If $A_1, \ldots, A_N \in \mathcal{F}$, then $\bigcup_{n=1}^{N} A_n \in \mathcal{F}$ and $\bigcap_{n=1}^{N} A_n \in \mathcal{F}$
7. If $A, B \in \mathcal{F}$ then $A \setminus B$, $B \setminus A$, $A \triangle B \in \mathcal{F}$

Proof. Using our definition of σ-algebra:

4. Since $X \in \mathcal{F}$, then $X^c = \emptyset \in \mathcal{F}$.

5. Since $A_n \in \mathcal{F}$ for all $n \geq 1$, then $A_n^c \in \mathcal{F}$ for all $n \geq 1$ by closure under complement. By closure under countable union, we have $\bigcup_{n=1}^{\infty} A_n^c \in \mathcal{F}$. Thus $\left(\bigcup_{n=1}^{\infty} A_n^c\right)^c \in \mathcal{F}$ by closure under complement. From De Morgan's laws we conclude $\bigcap_{n=1}^{\infty} A_n \in \mathcal{F}$.

6. This follows from closure under countable union/intersection, by respectively setting $A_n = \emptyset$ or $A_n = X$ for all $n > N$.

7. Note that $A \setminus B = A \cap B^c$, so the result follows by closure under finite intersection. The cases of $B \setminus A$ and $A \triangle B = (A \setminus B) \cup (B \setminus A)$ follow similarly.

■

Remark (1). Intuitively speaking, we can always do *countably* many set operations and stay inside the σ-algebra. If we do *uncountably* many set operations, then the result is unclear at best.

Remark (2). We say a σ-algebra $\mathcal{F}_1$ is bigger than $\mathcal{F}_2$ if $\mathcal{F}_2 \subseteq \mathcal{F}_1$. Thus $\{X, \emptyset\}$ is the smallest σ-algebra while 2^X is the biggest σ-algebra. As another example, $\mathcal{F}_A$ is smaller than $\mathcal{F}_{A,B}$.

Definition 2.2. Let X be a space and let C be a collection of subsets of X. Then the σ-algebra generated by C, denoted by $\sigma(C)$, satisfies:

1. $C \subseteq \sigma(C)$
2. If $\mathcal{F}'$ is a σ-algebra such that $C \subseteq \mathcal{F}'$, then $\sigma(C) \subseteq \mathcal{F}'$

Remark. In other words, $\sigma(C)$ is the smallest σ-algebra of X that is a superset of C.

Proposition 2.2. The properties below follow by definition.

1. $\sigma(C) = \bigcap\{\mathcal{F} : \mathcal{F} \text{ is a } \sigma\text{-algebra such that } C \subseteq \mathcal{F}\}$
2. If C is a σ-algebra, then $\sigma(C) = C$.
3. If C_1, C_2 are two collections of subsets of X, and $C_1 \subseteq C_2$, then $\sigma(C_1) \subseteq \sigma(C_2)$.

Proof.

1. Let $M \coloneqq \{\mathcal{F} : \mathcal{F} \text{ is a } \sigma\text{-algebra with } C \subseteq \mathcal{F}\}$. We claim that $F \coloneqq \bigcap_{\mathcal{F} \in M} \mathcal{F}$ is a σ-algebra of subsets of X. To see this:

   (i) Since $\emptyset, X \in \mathcal{F}$ for all $\mathcal{F} \in F$, then $\emptyset, X \in F$.
   (ii) Let $A \in F$. Then $A \in \mathcal{F}$ for all $\mathcal{F} \in M$, so $A^c \in \mathcal{F}$ for all $\mathcal{F} \in M$. Therefore $A^c \in F$.
   (iii) Let $\{A_n : n \geq 1\} \subseteq F$. Then $\{A_n : n \geq 1\} \subseteq \mathcal{F}$ for all $\mathcal{F} \in M$. Thus $\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$ for all $\mathcal{F} \in M$. Therefore $\bigcup_{n=1}^{\infty} A_n \in F$.

   Next, we show that F satisfies the definition of σ-algebra generated by a collection C of subsets of X.

   (i) $C \subseteq F$ since $C \subseteq \mathcal{F}$ for all $\mathcal{F} \in M$.

(ii) Let F' be a σ-algebra with $C \subseteq F'$. Then $F \subseteq F'$ since $F' \in M$ and the intersection is taken over M.

We conclude that $\sigma(C) = F$.

2. The proposal is that $C = \sigma(C)$, where C is already a σ-algebra of subsets of X. We need to show the σ-algebra C meets the definition of σ-algebra generated by a collection C of subsets of X.

   (i) Obviously $C \subseteq C$.

   (ii) Let $\mathcal{F}'$ be a σ-algebra such that $C \subseteq \mathcal{F}'$. But then $C \subseteq \mathcal{F}'$.

3. Let C_1, C_2 be two collections of subsets of X, and suppose $C_1 \subseteq C_2$. Then $C_2 \subseteq \sigma(C_2)$ by part one of Definition 2.2. But then $C_1 \subseteq \sigma(C_2)$ implies $\sigma(C_1) \subseteq \sigma(C_2)$ by part two of Definition 2.2.

■

Remark (1). The first property says that in a sense, $\sigma(C)$ is the "most efficient" σ-algebra.

Remark (2). It is a fact that the intersection of any arbitrary collection of σ-algebras is again a σ-algebra. However, this is **not** true for unions of σ-algebras. For example, let $\mathcal{F}_1 = \{X, \emptyset, A, A^c\}$ and $\mathcal{F}_2 = \{X, \emptyset, B, B^c\}$ where $A \cap B = \emptyset$. Then $\mathcal{F}_1 \cup \mathcal{F}_2$ is not a σ-algebra, since $A \cup B \notin \mathcal{F}_1 \cup \mathcal{F}_2$.

Definition 2.3 (Borel σ-algebra)**.** The Borel σ-algebra $\mathcal{B}_{\mathbb{R}}$ is generated by all open intervals. We set

$$\mathcal{B}_{\mathbb{R}} = \sigma(\{\text{open subsets of } \mathbb{R}\})$$

Remark (1). It is a fact that if $G \subseteq \mathbb{R}$ is open, then we can write $G = \bigcup_{n=1}^{\infty} I_n$ where I_n are finite open intervals.

Remark (2). It is a fact that we can also write $\mathcal{B}_{\mathbb{R}} = \sigma(\{(a,b) : a, b \in \mathbb{R}, a < b\})$. The generator of the Borel σ-algebra is *not unique*.

Proposition 2.3. We can generate the Borel σ-algebra flexibly with different generators.

1. $\mathcal{B}_{\mathbb{R}} = \sigma\big(\{(a,b] : a, b \in \mathbb{R}, a < b\}\big)$
2. $\mathcal{B}_{\mathbb{R}} = \sigma\big(\{[a,b] : a, b \in \mathbb{R}, a < b\}\big)$
3. $\mathcal{B}_{\mathbb{R}} = \sigma\big(\{[a,b) : a, b \in \mathbb{R}, a < b\}\big)$
4. $\mathcal{B}_{\mathbb{R}} = \sigma\big(\{(-\infty, c) : c \in \mathbb{R}\}\big)$
5. $\mathcal{B}_{\mathbb{R}} = \sigma\big(\{(-\infty, c] : c \in \mathbb{R}\}\big)$
6. $\mathcal{B}_{\mathbb{R}} = \sigma\big(\{(c, \infty) : c \in \mathbb{R}\}\big)$
7. $\mathcal{B}_{\mathbb{R}} = \sigma\big(\{[c, \infty) : c \in \mathbb{R}\}\big)$

Proof. We will show that $\mathcal{B}_{\mathbb{R}} := \sigma(\{(a,b) : a, b \in \mathbb{R}, a < b\}) = \sigma(\{[a,b) : a, b \in \mathbb{R}, a < b\})$. Let $C_1 = \{(a,b) : a, b \in \mathbb{R}, a < b\}$ and let $C_2 = \{[a,b) : a, b \in \mathbb{R}, a < b\}$. The strategy is to first show $C_1 \subseteq \sigma(C_2)$. From here it follows that $\sigma(C_1) \subseteq \sigma(C_2)$, since $\sigma(C_1)$ is the *smallest* σ-algebra that supersets C_1. Second, we show $C_2 \subseteq \sigma(C_1)$, from which it follows $\sigma(C_2) \subseteq \sigma(C_1)$. This gives the desired conclusion that $\sigma(C_1) = \sigma(C_2)$.

Now let us show $C_1 \subseteq \sigma(C_2)$. Let $a, b \in \mathbb{R}$ with $a < b$ and **take a sequence of intervals** $[a + \frac{1}{n}, b)$ for $n \in \mathbb{N}$. We claim that $(a,b) = \bigcup_{n=1}^{\infty}[a + \frac{1}{n}, b)$. To see this, let $x \in (a,b)$. Then $a < x < b$. By the **Archimedean property**, there exists $N \in \mathbb{N}$ such that $0 < \frac{1}{N} < x - a$. Thus $a + \frac{1}{N} < x < b \implies x \in [a + \frac{1}{N}, b) \implies x \in \bigcup_{n=1}^{\infty}[a + \frac{1}{n}, b)$. So $(a,b) \subseteq \bigcup_{n=1}^{\infty}[a + \frac{1}{n}, b)$. Now let $x \in \bigcup_{n=1}^{\infty}[a + \frac{1}{n}, b)$. Then there exists $N \in \mathbb{N}$ such that $x \in [a + \frac{1}{N}, b)$. Thus $a + \frac{1}{N} \leq x < b \implies a < x < b \implies x \in (a,b)$. So $(a,b) \supseteq \bigcup_{n=1}^{\infty}[a + \frac{1}{n}, b)$. We have shown $(a,b) = \bigcup_{n=1}^{\infty}[a+\frac{1}{n}, b)$, and thus by **closure under countable union** we have $(a,b) \in \sigma(C_2)$. Since $a, b \in \mathbb{R}$ were arbitrary chosen to satisfy $a < b$, then $C_1 \subseteq \sigma(C_2)$. This implies $\sigma(C_1) \subseteq \sigma(C_2)$.

It remains to be shown that $C_2 \subseteq \sigma(C_1)$. Let $a, b \in \mathbb{R}$ with $a < b$ and **take a sequence of intervals** $(a - \frac{1}{n}, b)$ for $n \in \mathbb{N}$. We claim that $[a,b) = \bigcap_{n=1}^{\infty}(a - \frac{1}{n}, b)$. To see this, let $x \in [a,b)$. Then $a \leq x < b \implies a - \frac{1}{n} < a \leq x < b$ for all $n \in \mathbb{N}$. Thus $x \in \bigcap_{n=1}^{\infty}(a - \frac{1}{n}, b)$, which gives $[a,b) \subseteq \bigcap_{n=1}^{\infty}(a - \frac{1}{n}, b)$. Now let $x \in \bigcap_{n=1}^{\infty}(a - \frac{1}{n}, b)$ so that $a - \frac{1}{n} < x < b$ for all $n \in \mathbb{N}$. Suppose for contradiction that $a > x$. By the **Archimedean property**, there exists $N \in \mathbb{N}$ such that $0 < \frac{1}{N} < a - x$. Thus $a - \frac{1}{N} > x$ and contradiction arises. We conclude $a \leq x$ and so $x \in [a,b)$. Therefore $[a,b) = \bigcap_{n=1}^{\infty}(a - \frac{1}{n}, b)$ and by **closure under countable intersection** we have $[a,b) \in \sigma(C_1)$. Since $a, b \in \mathbb{R}$ were arbitrary chosen to satisfy $a < b$, then $C_2 \subseteq \sigma(C_1)$. This implies $\sigma(C_2) \subseteq \sigma(C_1)$.

It follows readily that $\sigma(C_1) = \sigma(C_2)$.

■

Definition 2.4 (Borel set)**.** If $A \in \mathcal{B}_{\mathbb{R}}$, then A is called a Borel set.

Remark (1). **All types of intervals are Borel sets**, due to the previous proposition. Any set produced by countable operations on intervals is thus also a Borel set.

Remark (2). All singletons $\{x\}$ with $x \in \mathbb{R}$ are Borel sets, since we can write $\{x\} = \bigcap_{n=1}^{\infty}(x - \frac{1}{n}, x + \frac{1}{n})$. It is easy to see $\{x\} \subseteq \bigcap_{n=1}^{\infty}(x - \frac{1}{n}, x + \frac{1}{n})$. For the other direction, let $y \in \bigcap_{n=1}^{\infty}(x - \frac{1}{n}, x + \frac{1}{n})$ and use contradiction (with Archimedean property) to show $x \leq y$ and $y \leq x$.

Remark (3). Since all singletons are Borel sets, then all finite and countable sets are Borel sets, for example $\mathbb{Q} \in \mathcal{B}_{\mathbb{R}}$.

2.2 Measures

Definition 2.5 (Measurable space)**.** Given a space X and a σ-algebra $\mathcal{F}$ (of subsets of X), we call $(X, \mathcal{F})$ a measurable space.

Remark. For example, $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ is a measurable space. A measurable space is analogous in syntax to a topological space.

Definition 2.6 (Measure)**.** Given a measurable space $(X, \mathcal{F})$, let $\mu : \mathcal{F} \to [0, \infty]$ be a non-negative set function (assigns non-negative values to sets). Then μ is a measure if

1. $\mu(\emptyset) = 0$
2. If $\{A_n : n > 1\} \subseteq \mathcal{F}$ such that A_n are *pairwise disjoint*, then

$$\mu\Big(\bigcup_{n=1}^{\infty} A_n\Big) = \sum_{n=1}^{\infty} \mu(A_n) \qquad \text{(countable additivity)}$$

In this case, $(X, \mathcal{F}, \mu)$ is called a **measure space**.

Remark. We say μ is a:

1. *Finite* measure if $\mu(X) < \infty$.
2. *σ-finite* measure if there exists $\{A_n : n \geq 1\} \subseteq \mathcal{F}$ such that $X = \bigcup_{n=1}^{\infty} A_n$ with $\mu(A_n) < \infty$ for all $n \geq 1$.
3. *Probability* measure if $\mu(X) = 1$.

Example 2.1. Some examples of measure functions with $\mu : \mathcal{B}_{\mathbb{R}} \to [0, \infty]$.

1. $\mu(A) = \begin{cases} |A| & \text{if } A \text{ finite} \\ \infty & \text{otherwise} \end{cases}$ for all $A \in \mathcal{B}_{\mathbb{R}}$. We call μ the **counting measure**.
2. $\mu(A) = \begin{cases} 1 & \text{if } x_0 \in A \\ 0 & \text{otherwise} \end{cases}$ for some $x_0 \in \mathbb{R}$, for all $A \in \mathcal{B}_{\mathbb{R}}$. We call μ a **probability measure**, with a point mass at x_0.

Theorem 2.1 (Properties of measure)**.** Let $(X, \mathcal{F}, \mu)$ be a measure space.

1. (Finite additivity) If $A_1, \ldots, A_N$ are disjoint, then $\mu(A_1 \cup A_2 \cup \cdots \cup A_N) = \sum_{n=1}^{N} \mu(A_n)$.
2. (Monotonicity) Given $A, B \in \mathcal{F}$, if $A \subseteq B$ then $\mu(A) \leq \mu(B)$.
3. (Countable/finite subadditivity) If $\{A_n : n \geq 1\} \subseteq \mathcal{F}$, then $\mu(\bigcup_{n=1}^{\infty} A_n) \leq \sum_{n=1}^{\infty} \mu(A_n)$. Similarly, $\mu(\sum_{n=1}^{N} A_n) \leq \sum_{n=1}^{N} \mu(A_n)$.

Proof.

1. Let $A_n := \emptyset$ for all $n > N$. Then the result follows by countable additivity.
2. Notice that if $A \subseteq B$, then $B = (B \setminus A) \cup A$, where $B \setminus A$ is disjoint from A. Thus $\mu(A) \leq \mu(A) + \mu(B \setminus A) = \mu(B)$, by non-negativity of μ and finite additivity.

3. Let $B_1 = A_1$. Set $\boldsymbol{B_n := A_n \setminus \bigcup_{i=1}^{n-1} A_i}$ for $n \geq 2$. Then the following are true:

   (a) All B_n belong to $\mathcal{F}$, since they are formed under finitely many set operations.

   (b) The B_n are pairwise disjoint. To see this, let $m, n \in \mathbb{N}$ and WLOG suppose $m < n$. Then $A_m \subseteq \bigcup_{i=1}^{n-1} A_n$ implies $A_m \cap B_n = \emptyset$. But since $B_m \subseteq A_m$, we have $B_m \cap B_n = \emptyset$.

   (c) Importantly, $\bigcup_{n=1}^{\infty} A_n = \bigcup_{n=1}^{\infty} B_n$. To see this, let $x \in \bigcup_{n=1}^{\infty} B_n$. Then $x \in B_n$ for some $n \in \mathbb{N}$. So $x \in A_n \setminus \bigcup_{i=1}^{n-1} A_i$. Thus $x \in A_n$, which implies $x \in \bigcup_{n=1}^{\infty} A_n$. Now, let $x \in \bigcup_{n=1}^{\infty} A_n$ so that $x \in A_n$ for some $n \in \mathbb{N}$. We must show that $x \in B_m$ for some $m \in \mathbb{N}$. If $x \in A_n \setminus \bigcup_{i=1}^{n-1} A_i$, then set $m := n$ and we are done. Otherwise, we have $x \in \bigcup_{i=1}^{n-1} A_i$, so $x \in A_k$ for some $k \in \mathbb{N}$ with $k < n$. If $x \in A_k \setminus \bigcup_{i=1}^{k-1} A_i$, then set $m := k$. Otherwise, we repeat the argument and decrement k. Since $n \in \mathbb{N}$ is finite, this process must terminate. Thus $x \in B_m$ for some $m \in \mathbb{N}$, hence $x \in \bigcup_{n=1}^{\infty} B_m$. By double inclusion, the conclusion follows.

We have therefore constructed an auxiliary sequence of sets $\{B_n : n \geq 1\} \subseteq \mathcal{F}$ which are pairwise disjoint and satisfies $\bigcup_{n=1}^{\infty} A_n = \bigcup_{n=1}^{\infty} B_n$. Then

$$\begin{aligned}
\mu\Big(\bigcup_{n=1}^{\infty} A_n\Big) &= \mu\Big(\bigcup_{n=1}^{\infty} B_n\Big) \\
&= \sum_{n=1}^{\infty} \mu(B_n) && \text{(by countable additivity)} \\
&\leq \sum_{n=1}^{\infty} \mu(A_n) && \text{(by monotonicity of } B_n \subseteq A_n\text{)}
\end{aligned}$$

We can be more rigorous about why the final inequality holds. First, monotonicity gives an inequality for the partial sums: $\sum_{n=1}^{N} \mu(B_n) \leq \sum_{n=1}^{N} \mu(A_n)$. Then, the order limit theorem from Analysis 1 tells us that $\lim_{N\to\infty} \sum_{n=1}^{N} \mu(B_n) \leq \lim_{N\to\infty} \sum_{n=1}^{N} \mu(A_n)$. The limit of partial sums is how the symbol "$\sum_{n=1}^{\infty}$" is defined.

■

Remark. If $A \in \mathcal{F}$ with $\mu(A) = 0$, then A is called a **null set**. Note that if $\{A_n : n \geq 1\} \subseteq \mathcal{F}$ is such that $\mu(A_n) = 0$ for all $n \geq 1$, then $\bigcup_{n=1}^{\infty} \mu(A_n)$ is a null set. This follows by non-negativity and countable subadditivity of measure.

Theorem 2.2 (Continuity from below)**.** Let $(X, \mathcal{F}, \mu)$ be a measure space. Let $\{A_n : n \geq 1\} \subseteq \mathcal{F}$ be an increasing sequence of sets such that $A_n \subseteq A_{n+1}$ for all $n \in \mathbb{N}$. Then

$$\mu\Big(\bigcup_{n=1}^{\infty} A_n\Big) = \lim_{n\to\infty} \mu(A_n)$$

Note that the limit always exists, due to monotone convergence theorem when $\{\mu(A_n) : n \geq 1\}$ is bounded above. Otherwise, the limit is ∞.

Proof. Let $B_1 := A_1$ and let $\boldsymbol{B_n := A_n \setminus A_{n-1}}$ for all $n \geq 2$. Notice that $B_n = A_n \setminus \bigcup_{i=1}^{n-1} A_n$ since the sets A_n are nested increasing. Thus, part 3(c) of the the proof in Theorem 2.1 applies. Indeed, in that proof, we made no additional assumptions on the B_n besides requiring $B_n \in \mathcal{F}$ for all $n \geq 1$ (which is satisfied here).

Therefore the auxiliary sequence of sets $\{B_n : n \geq 1\} \subseteq \mathcal{F}$ is pairwise disjoint, and $\bigcup_{n=1}^{\infty} A_n = \bigcup_{n=1}^{\infty} B_n$. Furthermore, observe that $\bigcup_{n=1}^{N} B_n = A_N$. We thus have

$$\begin{aligned}
\mu(\bigcup_{n=1}^{\infty} A_n) &= \mu(\bigcup_{n=1}^{\infty} B_n) \\
&= \sum_{n=1}^{\infty} \mu(B_n) && \text{(countable additivity)} \\
&= \lim_{N\to\infty} \sum_{n=1}^{N} \mu(B_n) \\
&= \lim_{N\to\infty} \mu\Big(\bigcup_{n=1}^{N} B_n\Big) && \text{(finite additivity)} \\
&= \lim_{N\to\infty} \mu(A_N)
\end{aligned}$$

■

Theorem 2.3 (Continuity from above)**.** Let $(X, \mathcal{F}, \mu)$ be a measure space. Let $\{A_n : n \geq 1\} \subseteq \mathcal{F}$ be a decreasing sequence of sets such that $A_n \supseteq A_{n+1}$ for all $n \in \mathbb{N}$. Further assuming $\mu(A_1) < \infty$, then

$$\mu\Big(\bigcap_{n=1}^{\infty} A_n\Big) = \lim_{n\to\infty} \mu(A_n)$$

Proof. Let $B_n = A_1 \setminus A_n$ for all $n \geq 1$. Then $\{B_n : n \geq 1\} \subseteq \mathcal{F}$, the B_n are good sets we want to measure. **Notice that the B_n are increasing, since we are removing less and less from A_1.** Furthermore, notice that $\bigcup_{n=1}^{\infty} B_n = A_1 \setminus \bigcap_{n=1}^{\infty} A_n$. This follows since

$$\bigcup_{n=1}^{\infty} B_n = \bigcup_{n=1}^{\infty} A_1 \setminus A_n = A_1 \cap \bigcup_{n=1}^{\infty} A_n^c = A_1 \setminus \bigcap_{n=1}^{\infty} A_n$$

Then we have

$$\begin{aligned}
\mu(\bigcup_{n=1}^{\infty} B_n) &= \lim_{n\to\infty} \mu(B_n) && \text{(continuity from below)} \\
&= \lim_{n\to\infty} \mu(A_1 \setminus A_n) \\
&= \lim_{n\to\infty} \mu(A_1) - \mu(A_n) && \text{(finite additivity)} \\
&= \mu(A_1) - \lim_{n\to\infty} \mu(A_n) && \text{(limit exists by MCT)}
\end{aligned}$$

Thus

$$\begin{aligned}
\mu(\bigcap_{n=1}^{\infty} A_n) &= \mu(A_1) - \mu(A_1 \setminus \bigcap_{n=1}^{\infty} A_n) && \text{(finite additivity)} \\
&= \mu(A_1) - \mu(\bigcup_{n=1}^{\infty} B_n) \\
&= \lim_{n\to\infty} \mu(A_n)
\end{aligned}$$

Crucially, we have relied on $\mu(A_1) < \infty$ to ensure everything below is finite, which thus makes the implications valid.

$$\mu(A_1) = \mu(A_1 \setminus A_n) + \mu(A_n) \implies \mu(A_1 \setminus A_n) = \mu(A_1) - \mu(A_n)$$
$$\mu(A_1) = \mu(A_1 \setminus \bigcap_{n=1}^{\infty} A_n) + \mu(\bigcap_{n=1}^{\infty} A_n) \implies \mu(\bigcap_{n=1}^{\infty} A_n) = \mu(A_1) - \mu(A_1 \setminus \bigcap_{n=1}^{\infty} A_n)$$

■

Remark. The hypothesis $\mu(A_1) < \infty$ is necessary. For example, take the counting measure μ for $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ defined as

$$\mu(A) = \begin{cases} |A| \text{ if } A \text{ finite} \\ \infty \text{ otherwise} \end{cases}$$

Let $E_n = \{n, n+1, \dots\}$ for all $n \in \mathbb{N}$ so that $\mu(E_n) = \infty$ for all $n \geq 2$. Thus $\lim_{n\to\infty} \mu(E_n) = \infty$. But then $\mu(\bigcap_{n=1}^{\infty} E_n) = \mu(\emptyset) = 0$, so the equality fails to hold.

2.3 Construction of Lebesgue measure

2.3.1 Part I of the construction

Definition 2.7 (Outer measure)**.** For all $A \subseteq \mathbb{R}$, we let

$$m^*(A) = \inf\left\{ \sum_{n=1}^{\infty} l(I_n) : A \subseteq \bigcup_{n=1}^{\infty} I_n, \text{ where } I_n \text{ are open intervals} \right\}$$

We call m^* the Lebesgue outer measure.

Proposition 2.4. For all $A \subseteq \mathbb{R}$, we have $m^*(A) \geq 0$ and $m^*(\emptyset) = 0$.

Proof. Let $A \subseteq \mathbb{R}$. Then $m^*(A) \geq 0$ follows by non-negativity of interval length. On the other hand, $m^*(\emptyset) = 0$ since *any* open interval is a cover for $\emptyset$. Let $\varepsilon > 0$. Then define the sequence of intervals $I_n = (-\frac{\varepsilon}{2^{n+1}}, \frac{\varepsilon}{2^{n+1}})$. This forms a cover of $\emptyset$ with $\sum_{n=1}^{\infty} l(I_n) = \varepsilon$.

■

Proposition 2.5 (Monotonicity)**.** If $A, B \subseteq \mathbb{R}$ with $A \subseteq B$, then $m^*(A) \leq m^*(B)$.

Proof. Let $A, B \subseteq \mathbb{R}$ with $A \subseteq B$. Let $M = \{\sum_{n=1}^{\infty} l(I_n) : \{I_n : n \geq 1\}$ is an open-interval cover of $A\}$. Similarly, let $N = \{\sum_{n=1}^{\infty} l(I_n) : \{I_n : n \geq 1\}$ is an open-interval cover of $B\}$ Then $N \subseteq M$, since any open-interval cover of B is also one for A. Assume towards contradiction that $\inf M > \inf N$. Then $\inf M$ is not a lower bound for N, so $\inf M > x$ for some $x \in N$. Contradiction arises since $x \in M$, and $\inf M$ is a lower bound of M. Therefore $\inf M \leq \inf N$.

■

Proposition 2.6 (Countable subadditivity)**.** If $\{A_n : n \geq 1\}$ is a sequence of subsets of $\mathbb{R}$

Proof. If $m^*(A_n) = \infty$ for some $n \in \mathbb{N}$, then $m^*(\bigcup_{n=1}^{\infty} A_n) = \infty$ by monotonicity. Furthermore, $\sum_{n=1}^{\infty} m^*(A_n) = \infty$. Thus the inequality holds.

We may assume WLOG that $m^*(A_n)$ is finite for all $n \in \mathbb{N}$. Let $\varepsilon > 0$. For each A_n, take a sequence $\{I_n : n \geq 1\}$ such that $A_n \subseteq \bigcup_{i=1}^{\infty} I_{n,i}$ and

$$m^*(A_n) \leq \sum_{n=1}^{\infty} l(I_{n,i}) \leq m^*(A_n) + \frac{\varepsilon}{2^n}$$

The second inequality comes from the definition of infimum: recall that if $b \in \mathbb{R}$ is a lower bound for $B \subseteq \mathbb{R}$, then $b = \inf B \iff \forall \varepsilon > 0\ \exists x \in B$ s.t. $x - \varepsilon < b$. Therefore

$$\bigcup_{n=1}^{\infty} A_n \subseteq \bigcup_{n=1}^{\infty} \bigcup_{n=1}^{\infty} I_{n,i}$$

This implies

$$\begin{aligned}
m^*\Big(\bigcup_{n=1}^{\infty} A_n\Big) &\leq \sum_{n,i} l(I_{n,i}) && \text{(monotonicity)} \\
&\overset{(*)}{=} \sum_{n=1}^{\infty} \sum_{i=1}^{\infty} l(I_{n,i}) && \text{(unconditional convergence)} \\
&\leq \sum_{n=1}^{\infty} (m^*(A_n) + \frac{\varepsilon}{2^n}) \\
&\overset{(**)}{=} \sum_{n=1}^{\infty} m^*(A_n) + \varepsilon
\end{aligned}$$

Since $\varepsilon > 0$ is arbitrary, then countable subadditivity follows.

■

Remark. Unconditional convergence means a series will converge to the same limit for any rearrangement of the summands. It is a fact that a convergent series is unconditionally convergent if and only if it is absolutely convergent. Since our series consists of non-negative terms, any convergent series is absolutely convergent and hence unconditionally convergent. Furthermore, if the series diverges to ∞, then so must any other rearrangement. Otherwise, we would obtain the contradiction that one rearrangement converges to a finite value, while another one does not. Thus the equality of $(*)$ holds. The last equality $(**)$ holds because if $\sum_{n=1}^{\infty} m^*(A_n) = \infty$, then the infinite series $\sum_{n=1}^{\infty}(m^*(A_n) + \frac{\varepsilon}{2^n})$ also converges to ∞ by comparison. If $\sum_{n=1}^{\infty} m^*(A_n) < \infty$, then existence of the separated limits ensures existence of $\sum_{n=1}^{\infty}(m^*(A_n) + \frac{\varepsilon}{2^n})$.

Proposition 2.7 (Outer measure of interval)**.** If $I \subseteq \mathbb{R}$ is an interval, then $m^*(I) = l(I)$.

Proof. **First, we prove the result for a closed and bounded interval** $I = [a, b]$ for $a, b \in \mathbb{R}$. For all $\varepsilon > 0$, let $I_1 := (a - \frac{\varepsilon}{2}, b + \frac{\varepsilon}{2})$ be an open-interval cover for I. Since $I \subseteq I_1$, then $m^*(I) \leq l(I_1) = b - a + \varepsilon = l(I) + \varepsilon$. Therefore the inequality $m^*(I) \leq l(I)$ holds since $\varepsilon > 0$ is arbitrary.

For the reverse inequality, we must show $l(I)$ is a lower bound on the length-sums of all open-interval covers. In other words, for all open-interval coverings $\{I_n : n \geq 1\}$ of I we have

$$\sum_{n=1}^{\infty} l(I_n) \geq l(I)$$

From there it would follow that $m^*(A) \geq l(I)$. Let $\{I_n : n \geq 1\}$ be an arbitary open-interval cover of I. WLOG assume $l(I_n) < \infty$ for all $n \geq 1$, since otherwise the inequality holds trivially. Since $I = [a, b]$ is **compact** by the Heine-Borel theorem, then I is covered by finitely many I_n. Among this finite collection, we may extract $I_1, \ldots, I_N$ such that

$$a_1 < a, b < b_N, \text{ and } a_n < b_{n-1} \quad \forall 2 \leq n \leq N$$

Thus

$$\begin{aligned}
\sum_{n=1}^{\infty} l(I_n) &\geq \sum_{n=1}^{N} l(I_n) \\
&= b_1 - a_1 + \sum_{n=2}^{N}(b_n - a_n) \\
&\geq b_1 - a_1 + \sum_{n=2}^{N}(b_n - b_{n-1}) \\
&= b_N - a_1 && \text{(telescoping sum)} \\
&\geq b - a \\
&= l(I)
\end{aligned}$$

So the result holds for a closed and bounded interval I. **Next, we prove the result for a finite interval** I. Let I be any finite interval with endpoints $a, b \in \mathbb{R}$, $a < b$. Then set $I' := [a + \frac{\varepsilon}{2}, b - \frac{\varepsilon}{2}]$, and set $I'' := [a - \frac{\varepsilon}{2}, b + \frac{\varepsilon}{2}]$, which are closed and bounded. We have by monotonicity

$$b - a - \varepsilon = m^*(I') \leq m^*(I) \leq m^*(I'') = b - a + \varepsilon$$

Since ε is arbitrary, then $m^*(I) = b - a = l(I)$. **Finally, we prove the result for an infinite interval** I. Let I be an infinite interval taking $l(I) = \infty$. Let $M > 0$. Then there exists $I_M \subseteq I$ with I_M closed and bounded with $l(I) = M$. Thus

$$m^*(I) \geq m^*(I_M) = M$$

Since M is arbitrary then $m^*(I) = \infty$.

■

Corollary 2.3.1. For all $x \in \mathbb{R}$, the singleton $\{x\}$ has outer measure zero. Furthermore, any countable $C \subseteq \mathbb{R}$ also has outer measure zero.

Proof. Let $x \in \mathbb{R}$. Let $\varepsilon > 0$. Write $I := [x - \frac{\varepsilon}{2}, x + \frac{\varepsilon}{2}]$ such that $\{x\} \subseteq I$. By monotonicity $m^*(\{x\}) \leq m^*(I) = l(I) = \varepsilon$. Thus $0 \leq m^*(\{x\}) \leq 0$.

Now let $C \subseteq \mathbb{R}$ be at most countable. If C has finitely many elements, we can write $C = \bigcup_{n=1}^{N} \{x_n\}$ where $x_n \in C$. If C has infinitely many elements, then write $C = \bigcup_{n=1}^{\infty} \{x_n\}$ where $x_n \in C$. Either way, finite/countable subadditivity gives $0 \leq m^*(C) \leq 0$, since the singletons have outer measure zero.

■

Proposition 2.8 (Translation invariance)**.** For all $A \subseteq \mathbb{R}$, we have $m^*(A + x) = m^*(A)$ for any choice of $x \in \mathbb{R}$.

Proof. Observe that $\{I_n : n \geq 1\}$ is an open-interval cover of A if and only if $\{I_n + x : n \geq 1\}$ is an open-interval cover of $A + x$. Furthermore, $l(I_n) = l(I_n + x)$, so the infimum over each set of open-interval coverings is identical.

■

Proposition 2.9 (Outer measure is outer regular)**.** For all $A \subseteq \mathbb{R}$ we have

$$m^*(A) = \inf\{m^*(B) : B \subseteq \mathbb{R} \text{ is open and } A \subseteq B\}$$

Proof. For notation sake, we will simply write $\inf\{m^*(B)\}$ to denote the right hand side. For all open sets $B \subseteq \mathbb{R}$ with $A \subseteq B$ we have by monotonicity that $m^*(A) \leq m^*(B)$. Thus $m^*(A) \leq \inf\{m^*(B)\}$.

To show the reverse inequality, WLOG assume $m^*(A) < \infty$ since otherwise the inequality holds trivially. Let $\varepsilon > 0$. Then there exists $\{I_n : n \geq 1\}$ an open-interval cover of A such that

$$m^*(A) + \varepsilon > \sum_{n=1}^{\infty} l(I_n)$$

Now let $B = \bigcup_{n=1}^{\infty} I_n$ which is an open subset of $\mathbb{R}$. Then

$$\begin{aligned} m^*(B) &= m^*(\bigcup_{n=1}^{\infty} I_n) \\ &\leq \sum_{n=1}^{\infty} m^*(I_n) \\ &= \sum_{n=1}^{\infty} l(I_n) \\ &< m^*(A) + \varepsilon \end{aligned}$$

Since ε was arbitrary then $m^*(B) \leq m^*(A)$. But then

$$m^*(A) \geq m^*(B) \geq \inf\{m^*(B)\}$$

∎

Remark. This describes the notion that measure of a set A can be approximated with a sequence of sets B_n which wrap tighter and tighter around A.

Proposition 2.10. If $A = A_1 \cup A_2 \subseteq \mathbb{R}$ with $d(A_1, A_2) > 0$, then $m^*(A) = m^*(A_1) + m^*(A_2)$.

Proof. Clearly $m^*(A) \leq m^*(A_1) + m^*(A_2)$ by finite subadditivity. For the reverse inequality, assume WLOG that $m^*(A)$ is finite, since otherwise the inequality trivially holds.

Let $\varepsilon > 0$. Let $\delta := d(A_1, A_2) > 0$. Then there exists $\{I_n : n \geq 1\}$ an open-interval cover of A such that $m^*(A) + \varepsilon > \sum_{n=1}^{\infty} l(I_n)$. Since I_n must be finite intervals (because we assumed $m^*(A) < \infty$)

∎

Proposition 2.11. If $A = \bigcup_{n=1}^{\infty} J_k$ where J_k are almost disjoint intervals (at most sharing endpoints) then $m^*(A) = \sum_{k=1}^{\infty} l(J_k)$.

2.3.2 Part II of the construction

Definition 2.8 (m^*-measurable)**.** Let $A \subseteq \mathbb{R}$. Then A is m^*-measurable if for all $B \subseteq \mathbb{R}$

$$m^*(B) = m^*(B \cap A) + m^*(B \cap A^c)$$

Remark. We can think of A as a "cookie cutter". If "mass is conserved" no matter what kinds of dough we cut, then A is "good" in the sense of m^*-measurable.

Theorem 2.4 (Carathéodory's Theorem)**.** There are two aspects to this theorem.

1. Let $\mathcal{M} = \{A \subseteq \mathbb{R} : A \text{ is } m^*\text{-measurable}\}$. Then $\mathcal{M}$ is a σ-algebra of subsets of $\mathbb{R}$.
2. Define $m : \mathcal{M} \to [0, \infty]$ by setting $m(A) = m^*(A)$ for all $A \in \mathcal{M}$. Then m is a measure on $\mathbb{R}$, called the *Lebesgue measure*.

If $A \in \mathcal{M}$, then we say A is Lebesgue measurable.

2.4 Properties of the Lebesgue Measure

2.4.1 Translation invariance of Lebesgue measure

Proposition 2.12 (Translation invariance)**.**

1. $\mathcal{M}$ is translation invariant: $\forall A \in \mathcal{M}\ \forall x \in \mathbb{R}$ we have $A + x \in \mathcal{M}$.
2. m is translation invariant: $\forall A \in \mathcal{M}\ \forall x \in \mathbb{R}$ we have $m(A) = m(A + x)$.

Remark. Note that we need translation invariance of $\mathcal{M}$ in order for translation invariance of m to make sense.

2.4.2 Finite open intervals are Lebesgue measurable

Theorem 2.5 (Measure of finite open interval)**.** For all $a, b \in \mathbb{R}$ with $a < b$, we have $(a, b) \in \mathcal{M}$ and $m((a, b)) = b - a$.

Remark. In other words, every open interval is Lebesgue measurable.

Corollary 2.5.1. Recall $\mathcal{B}_{\mathbb{R}} = \sigma\{(a, b) : a, b \in \mathbb{R} \text{ with } a < b\}$. It thus follows that $\mathcal{B}_{\mathbb{R}} \subseteq \mathcal{M}$, by definition of σ-algebra generated from a collection of sets.

Remark (1). Every Borel set is Lebesgue measurable. Does there exist a Lesbesgue measurable set which is not Borel? The answer is yes and we will eventually show this.

Remark (2). The fact that $\mathcal{B}_{\mathbb{R}} \subseteq \mathcal{M}$ is **very useful for proofs**. It means we have a wide array of sets at our disposal for which we know are Lebesgue measurable (e.g., open, closed, compact).

2.4.3 Regularity of Lebesgue measure

Theorem 2.6 (Regularity properties)**.** m has the following properties:

1. $\forall A \in \mathcal{M}, \forall \varepsilon > 0, \exists G$-open such that $A \subseteq G$ and $m(G \setminus A) \leq \varepsilon$
2. $\forall A \in \mathcal{M}, \forall \varepsilon > 0, \exists F$-closed such that $F \subseteq A$ and $m(A \setminus F) \leq \varepsilon$

3. $\forall A \in \mathcal{M}$, we have $m(A) = \inf\{m(G) : G\text{-open and } A \subseteq G\}$ (outer regularity)

4. $\forall A \in \mathcal{M}$, we have $m(A) = \sup\{m(K) : K\text{-compact and } K \subseteq A\}$ (inner regularity)

5. $\forall A \in \mathcal{M}$ where $m(A) < \infty, \forall \varepsilon > 0, \exists K$-compact such that $m(A \setminus K) \leq \varepsilon$

6. $\forall A \in \mathcal{M}$ where $m(A) < \infty, \forall \varepsilon > 0, \exists \{I_1, \ldots, I_N\}$ such that $m(A \triangle \bigcup_{n=1}^{N} I_n) \leq \varepsilon$.

Note that for property (6), the finite collection of intervals $I_1, \ldots, I_N$ is not necessarily a cover for A.

2.4.4 Completeness of Lebesgue measure

Theorem 2.7. $(\mathbb{R}, \mathcal{M}, m)$ is a *complete* measure space in the sense that for all $A \subseteq \mathbb{R}$, if there exists $B \in \mathcal{M}$ such that $A \subseteq B$ and $m(B) = 0$, then $A \in \mathcal{M}$ and $m(A) = 0$.

Remark (1). In other words, any subset of a null set is a null set. This firstly requires that any subset of a null set is *measurable*, and second that the measure is zero.

Remark (2). An equivalent statement for completeness is that for all $E \subseteq \mathbb{R}$, if there exists $F, G \in \mathcal{M}$ such that

$$F \subseteq E \subseteq G \text{ and } m(G \setminus F) = 0$$

Then $E \in \mathcal{M}$ and $m(E) = m(G) = m(F)$.

2.4.5 Uniqueness of Lebesgue measure

Proposition 2.13. Up to rescaling, m is the unique (non-trivial) measure on $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ that is **finite on compact sets and translation invariant**. That is, if μ is another such measure on $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$, then $\mu = c \cdot m$ for some $c > 0$. In particular, we can compute c as $c = \mu((0, 1))$.

Remark (1). We need to check that if $cm(A) := c \cdot (m(A))$ for all $A \in \mathcal{B}_{\mathbb{R}}$, then cm is indeed a measure on $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$.

Remark (2). Here, we are restricting m to the Borel σ-algebra (smaller than the σ-algebra of m^*-measurable sets).

Remark (3). The measure $m \equiv 0$ trivially satisfies translation invariance so it is uninteresting.

Theorem 2.8 (Dynkin's π-d Theorem)**.** Given a space X, let C be a collection of subsets of X. Then C is called a π-system of subsets of X if

$$A, B \in C \implies A \cap B \in C \qquad \text{(closed under finite intersection)}$$

Let $\mathcal{F} := \sigma(C)$. Assume μ_1, μ_2 are **finite measures** on $(X, \mathcal{F})$ such that

1. $\mu_1(X) = \mu_2(X)$

2. $\mu_1 = \mu_2$ on the π-system C

Then $\mu_1 = \mu_2$ on all of $\mathcal{F}$:

$$\mu_1(A) = \mu_2(A) \quad \forall A \in \mathcal{F}$$

Remark (1). That is, we only need to capture equality of measure on the generating π-system to achieve equality of measure on the entire σ-algebra. It is often much easier to verify equality of measure on π-systems, which is why this theorem is so useful.

Remark (2). Dynkin's π-d theorem is general to all measures.

Example 2.2. Take the Borel measurable space $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$. Then $\mathcal{C} := \{\emptyset\} \cup \{(a, b) : a, b \in \mathbb{R}, a < b\}$ is a generating π-system of $\mathcal{B}_{\mathbb{R}}$ (**not unique**).

For all $n \geq 1$, $\mathcal{B}_{[-n,n]} = \sigma(\mathcal{C})_{\cap[-n,n]}$ is the Borel σ-algebra of subsets of $[-n, n]$. We can say $\sigma(\mathcal{C})_{\cap[-n,n]}$ is a σ-algebra (of subsets of $[-n, n]$) due to a result from homework 1. Due to another result from homework 1, we can say

$$\sigma(\mathcal{C}_{\cap[-n,n]}) = \sigma(\mathcal{C})_{\cap[-n,n]}$$

Thus $\mathcal{B}_{[-n,n]} = \sigma(\mathcal{C}_{\cap[-n,n]})$. So $\mathcal{C}_{\cap[-n,n]} = \{A \cap [-n, n] : A \in \mathcal{C}\}$ is a generating π-system of $\mathcal{B}_{[-n,n]}$.

Proposition 2.14. Let μ be a measure on $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ such that $\mu(I) = l(I)$ for any interval $I \subseteq \mathbb{R}$. Then $\mu = m$, that is to say μ must be the Lebesgue measure.

2.4.6 Scaling property of Lebesgue Measure

Proposition 2.15 (Scaling property)**.** m has the scaling property. Namely, for all $A \in \mathcal{M}$, for all $c \in \mathbb{R}$, we have

$$cA = \{cx : x \in A\} \in \mathcal{M} \text{ and } m(cA) = |c| m(A)$$

In particular, m is reflection symmetric: For all $A \in \mathcal{M}$ we have $-A \in \mathcal{M}$ and $m(-A) = m(A)$.

2.5 Relation between Borel and Lebesgue σ-algebra

Recall that $(\mathbb{R}, \mathcal{M}, m)$ is a complete measure space. That is, $\forall B \subseteq \mathbb{R}$, if $\exists A \in \mathcal{M}$ with $m(A) = 0$ such that $B \subseteq A$, then $B \in \mathcal{M}$ and $m(B) = 0$.

Definition 2.9. Given a measure space $(X, \mathcal{F}, \mu)$, consider

$$\mathcal{N} = \{B \subseteq X : \exists A \in \mathcal{F} \text{ with } \mu(A) = 0 \text{ s.t. } B \subseteq A\}$$

Then, $\overline{\mathcal{F}} := \sigma(\mathcal{F} \cup \mathcal{N})$ is called the **completion** of $\mathcal{F}$ with respect to μ. If $\mathcal{N} \subseteq \mathcal{F}$, then $\overline{\mathcal{F}} = \mathcal{F}$, in which case $\mathcal{F}$ is already complete.

Proposition 2.16 (Equivalent formulation of completion)**.** Given a measure space $(X, \mathcal{F}, \mu)$, then

$$\overline{\mathcal{F}} = \{F \subseteq X : \exists E, G \in \mathcal{F} \text{ s.t. } E \subseteq F \subseteq G \text{ and } \mu(G \setminus E) = 0\}$$

Definition 2.10 (Extension of μ to $\overline{\mathcal{F}}$)**.** For all $F \in \overline{\mathcal{F}}$, suppose there exists $E, G \in \mathcal{F}$ such that

$$E \subseteq F \subseteq G \text{ with } \mu(G \setminus E) = 0$$

Then we define $\mu(F) := \mu(E) = \mu(G)$. The set function $\mu : \overline{\mathcal{F}} \to [0, \infty]$ is again a measure. We call $(X, \overline{\mathcal{F}}, \mu)$ the **completion** of $(X, \mathcal{F}, \mu)$. Then $(X, \overline{\mathcal{F}}, \mu)$ is a **complete measure space** in the sense that

$$\forall A \subseteq X, \text{ if } \exists B \in \overline{F} \text{ with } \mu(B) = 0 \text{ s.t. } A \subseteq B, \text{ then } A \in \overline{F} \text{ and } \mu(A) = 0$$

Theorem 2.9. $(\mathbb{R}, \mathcal{M}, m)$ is the completion of $(\mathbb{R}, \mathcal{B}_{\mathbb{R}}, m)$.

2.6 Cantor Set

We have seen that countable sets are Lebesgue measurable with measure zero. This raises a question: Does there exist a Lebesgue measurable set of measure zero, that is *uncountable*? The answer is yes, and the Cantor set is one such example.

Definition 2.11 (Cantor set)**.** Let $C_0 := [0, 1]$. Then define a sequence of sets where we iteratively remove the middle-third of each interval. For example, we set

$$C_1 := \left[0, \frac{1}{3}\right] \cup \left[\frac{2}{3}, 1\right]$$

Then

$$C_2 := \left[0, \frac{1}{9}\right] \cup \left[\frac{2}{9}, \frac{1}{3}\right] \cup \left[\frac{2}{3}, \frac{7}{9}\right] \cup \left[\frac{8}{9}, 1\right]$$

And so on. Then $C := \bigcap_{n=1}^{\infty} C_n$ is the **Cantor set**.

Proposition 2.17. The Cantor set is closed.

Proof. For every $n \geq 1$, C_n is the union of 2^n many disjoint closed intervals. Thus C_n is closed as we are taking a *finite union* of closed intervals. Therefore C is closed, since we are taking a countable intersection of closed sets.

■

Proposition 2.18. The Cantor set has measure zero.

Proof. For each C_n, there are 2^n disjoint closed intervals of length $\frac{1}{3^n}$. Thus $m(C_n) = \left(\frac{2}{3}\right)^n$. Since C_n forms a decreasing sequence of sets, and $m(C_1) < \infty$, then by continuity from above we have $m(C) = \lim_{n\to\infty} m(C_n) = 0$.

■

Proposition 2.19. The Cantor set is uncountable.

Proof. Let $x \in [0,1]$. Then there exists $(a_n)_{n\geq 1}$ such that $a_n \in \{0,1,2\}$ for all $n \geq 1$ with

$$x = \sum_{n=1}^{\infty} a_n \cdot \frac{1}{3^n}$$

But note that we may also write

$$C = \{x \in [0,1] : x \text{ admits a ternary expansion } (a_n)_{n\geq 1} \text{ with } a_n \in \{0,2\}\ \forall n \geq 1\}$$

This holds essentially because the "middle" of the interval gets removed each time. Now let $f : C \to [0,1]$ be defined for all $x \in C$ with corresponding ternary expansion $(a_n)_{n\geq 1}$. We set

$$f(x) = \sum_{n=1}^{\infty} \frac{a_n}{2} \cdot \frac{1}{2^n}$$

Now let $y \in [0,1]$. Then there exists binary expansion $(b_n)_{n\geq 1}$ such that $y = \sum_{n=1}^{\infty} b_n \cdot \frac{1}{2^n}$, where $b_n \in \{0,1\}$ for all $n \geq 1$. But then there is some $x \in [0,1]$ with ternary expansion $(2b_n)_{n\geq 1}$. Since $2b_n \in \{0,2\}$ for all $n \geq 1$, then $x \in C$ due to the above equivalent definition of C. Thus $y = f(x)$ for this particular x, where $x = \sum_{n=1}^{\infty}(2b_n) \cdot \frac{1}{3^n}$. Therefore C has cardinality no smaller than $[0,1]$, hence C is uncountable. Furthermore, since $C \subseteq [0,1]$, then C has cardinality no bigger than $[0,1]$ (as an injection exists from C to $[0,1]$, namely the identity function).

We conclude C has the same cardinality as $[0,1]$.

■

2.7 Devil's Staircase

Definition 2.12 (Cantor-Lebesgue function)**.** Recall the Cantor set $C = \bigcap_{n=1}^{\infty} C_n$, which is uncountable, closed, with measure zero. The Cantor-Lebesgue function, or Devil's staircase, is defined for all $x \in [0,1]$ as

$$f(x) = \begin{cases} \sum_{n=1}^{\infty} \frac{a_n}{2} \frac{1}{2^n} & \text{if } x \in C \text{ and } x = \sum_{n=1}^{\infty} a_n \frac{1}{3^n} \text{ for } a_n \in \{0,2\} \\ \sup\{f(y) : y \in C \text{ and } y < x\} & \text{if } x \notin C \end{cases}$$

Remark (1). In other words, any time an interval is removed in the construction of the Cantor set, assign constant value over that interval.

Remark (2). Very interestingly, the fact that C is measurable with $m(C) = 0$ implies that

$$1 = m([0,1]) = m([0,1] \cap C) + m([0,1] \setminus C) = m([0,1] \setminus C)$$

So the function is constant over a set with measure same as $[0,1]$. It only changes over C, a set of measure zero, yet still manages to increase from 0 to 1. We look away, and it's magically changed when we look back.

Proposition 2.20 (Properties of Cantor-Lebesgue function)**.** The following hold for the Devil's staircase

1. $f(0) = 0, f(1) = 1, f \equiv \frac{1}{2}$ on $(\frac{1}{3}, \frac{2}{3})$, $f \equiv \frac{1}{4}$ on $(\frac{1}{9}, \frac{2}{9})$, and so on...
2. $f : [0,1] \to [0,1]$ is a surjection
3. f is non-decreasing
4. f is continuous

2.8 Construction of a non-measurable set

Question: Does there exist $A \subseteq \mathbb{R}$ that is non-measurable? Yes, under the Axiom of Choice. The non-measurable set N we construct below is called the Vitali set.

Axiom of Choice
If Σ is a collection of non-empty sets, then there is a function $\mathcal{S} : \Sigma \to \bigcup_{A \in \Sigma} A$ such that for all $A \in \Sigma$ we have $S(A) \in A$. We call $\mathcal{S}$ a *selection function* and refer to $S(A)$ as a *representative* of A.

Now, define an equivalence relation $\sim$ on $[0,1]$. For all $a, b \in [0,1]$, we say

$$a \sim b \iff a - b \in \mathbb{Q}$$

Let E_a be the equivalence class containing a. Note that each E_a contains at least two elements due to density of $\mathbb{Q}$ in $\mathbb{R}$. This is because we can add an arbitrarily small rational to a and thus construct another element of E_a.

Set Σ to be the collection of all equivalence classes.

By the Axiom of Choice, we can select exactly one s_a from each E_a for all $E_a \in \Sigma$.

Proposition 2.21. Let $N := \{s_a : s_a \text{ is } \textbf{the} \text{ representative of } E_a \in \Sigma, \text{ for all } E_a \in \Sigma\}$. Then N is a non-measurable set.

Proof. Assume for the sake of contradiction that $N \in \mathcal{M}$. Let $\{q_k : k \geq 1\}$ be an enumeration of $[-1,1] \cap \mathbb{Q}$. For all $k \geq 1$ set $N_k := N + q_k$. By **translation invariance**, we have $\boldsymbol{N_k \in \mathcal{M}}$ and $\boldsymbol{m(N_k) = m(N)}$.

First, we claim the N_k are disjoint, that is, $N_k \cap N_l$ for $k \neq l$. Assume not. Then there exists $q_k \neq q_l$ and $s_a, s_b \in N$ such that $s_a + q_k = s_b + q_l$. So

$$s_a - s_b = q_l - q_k \in \mathbb{Q}$$

Thus s_a and s_b are in the same equivalence class. **Then $s_a = s_b$ since using the Axiom of Choice we only took one representative.** So $q_k = q_l$, which is a contradiction.

Next, we claim $[0,1] \subseteq \bigcup_{n=1}^{\infty} N_k$. Take $x \in [0,1]$. Then $x \sim s_a$ for some $s_a \in N$ due to density of $\mathbb{Q}$ in $\mathbb{R}$. So $x - s_a \in \mathbb{Q}$. Since $x - s_a \in [-1,1]$, then $x - s_a = q_k$ for some $k \geq 1$. **Note this is why we enumerated rationals over $[-1,1]$, to make such a statement.** Therefore $x \in N_k$ as desired.

Next, by default $\bigcup_{k=1}^{\infty} N_k \in [-1,2]$. We have shown the N_k to be disjoint. By **countable additivity** we thus have

$$m\Big(\bigcup_{k=1}^{\infty} N_k\Big) = \sum_{n=1}^{\infty} m(N_k)$$

So then $m\big(\bigcup_{k=1}^{\infty} N_k\big) \in [1,3]$. But this is a contradiction whether $m(N) = 0$ or $m(N) > 0$ (the sum would be zero or infinite).

■

Proposition 2.22. For all $A \in \mathcal{M}$, if $m(A) > 0$, then there exists $B \subseteq A$ such that $B \notin \mathcal{M}$.

Given that $\mathcal{B}_{\mathbb{R}} \subseteq \mathcal{M}$, this naturally raises a question: Does there exists $A \in \mathcal{M}$ with $A \notin \mathcal{B}_{\mathbb{R}}$? The answer is yes.

3 Integration theory

3.1 Measurable function

First some remarks on notation and syntax.

1. We consider functions f defined on $\mathbb{R}$ that are *extended* real-valued ($\overline{\mathbb{R}}$-valued). Otherwise, if $-\infty < f(x) < \infty$ for all $x \in \mathbb{R}$, then we say f is finite valued ($\mathbb{R}$-valued).

2. Further, for all $a \in \mathbb{R}$, we write the inverse image $f^{-1}([-\infty, a)) = \{x \in \mathbb{R} : -\infty \leq f(x) < a\}$. For clarity, we summarize the notation as

   $$\{f < a\} := f^{-1}\big([-\infty, a)\big) = \{x \in \mathbb{R} : -\infty \leq f(x) < a\}$$

   Similarly, we write

   $$\{f > a\} := f^{-1}\big((a, \infty]\big) = \{x \in \mathbb{R} : a < f(x) \leq \infty\}$$

   In general, for all $B \subseteq \mathbb{R}$ we write

   $$\{f \in B\} := f^{-1}(B) = \{x \in \mathbb{R} : x \in B\}$$

Recall that inverse image is very compatible with set operations. In fact

1. $f^{-1}(B^c) = f^{-1}(B)^c \qquad \forall B \subseteq \mathbb{R}$
2. $f^{-1}(A \cap B) = f^{-1}(A) \cap f^{-1}(B) \qquad \forall A, B \subseteq \mathbb{R}$
3. $f^{-1}(A \cup B) = f^{-1}(A) \cup f^{-1}(B) \qquad \forall A, B \subseteq \mathbb{R}$

Properties 2 and 3 extend to countable unions/intersections.

Definition 3.1 (Measurable function)**.** Let $f : \mathbb{R} \to \overline{\mathbb{R}}$. Then f is measurable if for all $a \in \mathbb{R}$

$$\{f < a\} \in \mathcal{M}$$

In other words, f is measurable if the inverse image of f over $[-\infty, a)$ is measurable.

Remark. Since this definition by itself is not quite powerful enough, the following propositions help build up the notion of measurable function.

Proposition 3.1. The following are equivalent definitions of measurable function for $\overline{\mathbb{R}}$-valued functions.

1. $\forall a \in \mathbb{R} \quad f^{-1}([a, \infty]) \in \mathcal{M}$
2. $\forall a \in \mathbb{R} \quad f^{-1}((a, \infty]) \in \mathcal{M}$
3. $\forall a \in \mathbb{R} \quad f^{-1}([-\infty, a]) \in \mathcal{M}$

Proof. To prove 3, it suffices to see that

$(\implies)$ $\forall a \in \mathbb{R} \quad f^{-1}([-\infty, a)) = \bigcup_{n=1}^{\infty} f^{-1}([-\infty, a - \frac{1}{n}])$

$(\impliedby)$ $\forall a \in \mathbb{R} \quad f^{-1}([-\infty, a]) = \bigcap_{n=1}^{\infty} f^{-1}([-\infty, a + \frac{1}{n}))$

(Prove the rest as exercise)

■

Proposition 3.2. Let f be $\mathbb{R}$-valued. Then f is measurable

$$\iff \forall a, b \in \mathbb{R}, a < b, f^{-1}((a, b)) \in \mathcal{M}$$

$$\iff \forall a, b \in \mathbb{R}, a < b, f^{-1}((a, b]) \in \mathcal{M}$$

$$\iff \forall a, b \in \mathbb{R}, a < b, f^{-1}([a, b)) \in \mathcal{M}$$

$$\iff \forall a, b \in \mathbb{R}, a < b, f^{-1}([a, b]) \in \mathcal{M}$$

Proof. We show equivalence with the third statement from the top. Assume f is measurable and finite-valued. Let $a, b \in \mathbb{R}$ with $a < b$. Then $f^{-1}((-\infty, a)), f^{-1}((-\infty, b)) \in \mathcal{M}$. So

$$\begin{aligned} f^{-1}([a,b)) &= f^{-1}((-\infty,a)^c \cap (-\infty,b)) \\ &= f^{-1}((-\infty,a))^c \cap f^{-1}((-\infty,b)) \in \mathcal{M} \end{aligned}$$

For the reverse direction, let $a \in \mathbb{R}$. Observe that $(-\infty, a) = \bigcup_{n=1}^{\infty}[-n, a)$. Thus

$$f^{-1}((-\infty,a)) = f^{-1}\left(\bigcup_{n=1}^{\infty}[-n,a)\right) = \bigcup_{n=1}^{\infty} f^{-1}([-n,a)) \in \mathcal{M}$$

■

Remark (1). Why can't we claim these equivalence statements for $f : \mathbb{R} \to \overline{\mathbb{R}}$? The issue is with the "reverse direction" above. Fix $a \in \mathbb{R}$, where the objective is to show $f^{-1}([-\infty, a)) \in \mathcal{M}$. But then no countable union of sets of the form $[x, a)$, with $x \in \mathbb{R}$, will allow us to achieve $[-\infty, a)$. Thus we cannot show $f^{-1}([-\infty, a)) \in \mathcal{M}$.

However, the "forward direction" does follow. To demonstrate, let $a, b \in \mathbb{R}$ with $a < b$. Then

$$f^{-1}([a,b)) = f^{-1}([-\infty,a)^c \cap [-\infty,b)) = f^{-1}([-\infty,a))^c \cap f^{-1}([-\infty,b)) \in \mathcal{M}$$

Remark (2). This result is shown later to hold for functions that are finite *almost everywhere*.

Proposition 3.3. Consider the Borel σ-algebra of subsets of $\overline{\mathbb{R}}$, defined as

$$\mathcal{B}_{\overline{\mathbb{R}}} := \sigma(\mathcal{B}_{\mathbb{R}} \cup \{\{\infty\}, \{-\infty\}\})$$

Then $\mathcal{B}_{\overline{\mathbb{R}}} = \sigma(\{[-\infty, a) : a \in \mathbb{R}\})$.

Proof. First, for any $a \in \mathbb{R}$ notice that $[-\infty, a) = (-\infty, a) \cup \{-\infty\}$. Thus $\{[-\infty, a) : a \in \mathbb{R}\} \subseteq \mathcal{B}_{\overline{\mathbb{R}}}$. So $\sigma(\{[-\infty, a) : a \in \mathbb{R}\}) \subseteq \mathcal{B}_{\overline{\mathbb{R}}}$.

Next, observe that

$$\bigcap_{n=1}^{\infty}[-\infty,n) = \{-\infty\}$$
$$\overline{\mathbb{R}} \setminus \bigcup_{n=1}^{\infty}[-\infty,n) = \{\infty\}$$

So $\{-\infty\}, \{\infty\} \in \sigma(\{[-\infty, a) : a \in \mathbb{R}\})$. For any $a \in \mathbb{R}$, we have $(-\infty, a) = [-\infty, a) \setminus \{-\infty\}$. So $\{(-\infty, a) : a \in \mathbb{R}\} \subseteq \sigma(\{[-\infty, a) : a \in \mathbb{R}\})$. Then

$$\mathcal{B}_{\mathbb{R}} = \sigma(\{(-\infty,a) : a \in \mathbb{R}\}) \subseteq \sigma(\{[-\infty,a) : a \in \mathbb{R}\})$$

We conclude that $\mathcal{B}_{\overline{\mathbb{R}}} \subseteq \sigma(\{[-\infty, a) : a \in \mathbb{R}\})$.

■

Remark. This gives an alternative characterization of the extended Borel σ-algebra. This characterization is especially useful to prove the next proposition. *Here, we are drawing a relation between* $\mathcal{B}_{\overline{\mathbb{R}}}$ *and the sets* $[-\infty, a)$ *that define measurable function.*

Proposition 3.4. $f : \mathbb{R} \to \overline{\mathbb{R}}$ is measurable $\iff$ for all $B \in \mathcal{B}_{\overline{\mathbb{R}}}$ we have $f^{-1}(B) \in \mathcal{M}$. If f is furthermore finite-valued, then f is measurable $\iff$ for all $B \in \mathcal{B}_{\mathbb{R}}$ $f^{-1}(B) \in \mathcal{M}$.

Proof. First we show " $\Longleftarrow$ ". This is immediate since $[-\infty, a) \in \mathcal{B}_{\overline{\mathbb{R}}}$ for all $a \in \mathbb{R}$. Thus applying the hypothesis we have $f^{-1}([-\infty, a)) \in \mathcal{M}$ for all $a \in \mathbb{R}$. So f is measurable.

Next, we show " $\Longrightarrow$ ". Take $C := \{[-\infty, a) : a \in \mathbb{R}\}$, where we have $\sigma(C) = \mathcal{B}_{\overline{\mathbb{R}}}$. In Assignment 2, we showed inverse image has the really nice property that given such a collection C, then

$$f^{-1}(\sigma(C)) = \sigma(f^{-1}(C))$$

This was proved via the "good set" principle. Thus

$$\begin{aligned} f^{-1}(\mathcal{B}_{\overline{\mathbb{R}}}) &= f^{-1}(\sigma(C)) \\ &= \sigma(f^{-1}(C)) \\ &= \sigma(f^{-1}(\{[-\infty, a) : a \in \mathbb{R}\})) \\ &= \sigma(\{f^{-1}([-\infty, a)) : a \in \mathbb{R}\}) \\ &\subseteq \mathcal{M} && \text{(as generator is a subset of } \mathcal{M}) \end{aligned}$$

Note the generator of the second last line is a subset of $\mathcal{M}$ by the hypothesis that $f : \mathbb{R} \to \overline{\mathbb{R}}$ is measurable. So for all $B \in \mathcal{B}_{\overline{\mathbb{R}}}$ we have $f^{-1}(B) \in \mathcal{M}$. ∎

Remark (1). This makes measurability of functions more general.

Remark (2). To show the case where f is finite-valued, simply replace $\mathcal{B}_{\overline{\mathbb{R}}}$ with $\mathcal{B}_{\mathbb{R}} = \sigma(\{(-\infty, a) : a \in \mathbb{R}\})$ in the above argument.

Proposition 3.5. Given $f : \mathbb{R} \to \overline{\mathbb{R}}$, define

$$f_{\mathbb{R}}(x) := \begin{cases} f(x) & -\infty < f(x) < \infty \\ 0 & f(x) \in \{-\infty, \infty\} \end{cases}$$

We call $f_{\mathbb{R}}$ the finite-valued component of f. Then f is measurable *if and only if*

$$\forall B \in \mathcal{B}_{\mathbb{R}} \;\; f_{\mathbb{R}}^{-1}(B) \in \mathcal{M} \text{ and } \{f = -\infty\} \in \mathcal{M} \text{ and } \{f = \infty\} \in \mathcal{M}$$

Proof. To show " $\Longleftarrow$ ", let $a \in \mathbb{R}$. Then

$$f^{-1}([-\infty, a)) = f^{-1}(\{-\infty\}) \cup f^{-1}((-\infty, a)) = f^{-1}(\{-\infty\}) \cup f_{\mathbb{R}}^{-1}((-\infty, a)) \in \mathcal{M}$$

To show " $\Longrightarrow$ ", let $B \in \mathcal{B}_{\mathbb{R}}$. Then

$$\begin{aligned} f_{\mathbb{R}}^{-1}(B) &= \{x \in \mathbb{R} : f_{\mathbb{R}}(x) \in B\} \\ &= \{x \in \mathbb{R} : f(x) \in (-\infty, \infty) \textbf{ and } f(x) \in B\} \cup \{x \in \mathbb{R} : f(x) \in \{-\infty, \infty\} \textbf{ and } 0 \in B\} \\ &= \Big(\{f \in (-\infty, \infty)\} \cap \{f \in B\}\Big) \cup \{x \in \mathbb{R} : f(x) \in \{-\infty, \infty\} \textbf{ and } 0 \in B\} \\ &\in \mathcal{M} \end{aligned}$$

The sets $\{f \in (-\infty, \infty)\}$, $\{f \in B\}$ are measurable due to Proposition 3.4. As for $\{x \in \mathbb{R} : f(x) \in \{-\infty, \infty\}$ **and** $0 \in B\}$, if indeed $0 \in B$, then the set is measurable (again by Proposition 3.4). Otherwise, if $0 \notin B$, the empty set is still measurable.

■

Remark. This means we can take a two-step approach to show measurability, where we first consider the finite-valued component of f, then the infinite-valued component.

Definition 3.2. Suppose a statement is true for every $x \in A$ where $A \in \mathcal{M}$ such that $m(A^c) = 0$. Then we say the statement "is true a.e." (almost everywhere) or "is true for a.e. x" (almost every x).

Proposition 3.6. If $f : \mathbb{R} \to \overline{\mathbb{R}}$ is measurable and $f = g$ a.e., then g is measurable.

Remark. This was proved in Assignment 2.

Corollary 3.0.1. If f is finite-valued a.e., then

$$\begin{aligned} f \text{ is measurable } &\iff f_{\mathbb{R}} \text{ is measurable} && \text{(since } f = f_{\mathbb{R}} \text{ a.e.)} \\ &\iff \forall a, b \in \mathbb{R}, a < b, f^{-1}((a,b)) \in \mathcal{M} && \text{(since } f_{\mathbb{R}} \text{ is finite-valued)} \end{aligned}$$

Remark. We pass f to finite-valued $f_{\mathbb{R}}$ then apply Proposition 3.2.

Proposition 3.7. If $f \equiv c$ (i.e., f is a constant function), then f is measurable.

Proof. Let $a \in \mathbb{R}$. Then

$$f^{-1}([-\infty, a)) = \begin{cases} \mathbb{R} & \text{if } a > c \\ \emptyset & \text{if } a \leq c \end{cases}$$

Which is measurable in either case.

■

Proposition 3.8. If $f = \mathbb{1}_A$ for some $A \subseteq \mathbb{R}$, then f is measurable *if and only if* $A \in \mathcal{M}$. Here, f is the *characteristic function* defined as

$$f(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases}$$

Proof. Let $a \in \mathbb{R}$. Then

$$f^{-1}([-\infty, a)) = \begin{cases} \mathbb{R} & \text{if } a > 1 \\ A^c & \text{if } 0 < a \leq 1 \\ \emptyset & \text{if } a \leq 0 \end{cases}$$

Which is measurable if and only if $A^c \in M$, if and only if $A \in \mathcal{M}$.

■

Proposition 3.9. If f is a finite-valued and continuous function on $\mathbb{R}$, then f is measurable.

Proof. Recall that $f : \mathbb{R} \to \mathbb{R}$ is continuous if and only if for all open sets $G \subseteq \mathbb{R}$ we have $f^{-1}(G)$ is open.

Then $f^{-1}((a, b))$ is open for all $a, b \in \mathbb{R}$ with $a < b$. Therefore f is measurable.

We can also prove this by using the fact that $\mathcal{B}_{\mathbb{R}} = \sigma(\{G : G \text{ is open in } \mathbb{R}\})$. Then

$$\begin{aligned} f^{-1}(\mathcal{B}_{\mathbb{R}}) &= f^{-1}\Big(\sigma(\{G : G \text{ is open in } \mathbb{R}\})\Big) \\ &= \sigma(f^{-1}\{G : G \text{ is open in } \mathbb{R}\}) \\ &= \sigma(\{f^{-1}(G) : G \text{ is open in } \mathbb{R}\}) \\ &\subseteq \mathcal{B}_{\mathbb{R}} \\ &\subseteq \mathcal{M} \end{aligned}$$

■

Remark. This is a useful fact. The fact that $f^{-1}(\mathcal{B}_{\mathbb{R}}) \subseteq \mathcal{B}_{\mathbb{R}}$ is a stronger statement than simply $f^{-1}(\mathcal{B}_{\mathbb{R}}) \subseteq \mathcal{M}$. Furthermore, if f^{-1} (the inverse function of f) exists and is continuous, then for all $B \in \mathcal{B}_{\mathbb{R}}$ we have $f(B) \in \mathcal{B}_{\mathbb{R}}$. This follows since $f(B) = (f^{-1})^{-1}(B)$, where the RHS takes the inverse image of the inverse function of f.

Proposition 3.10. If $f : \mathbb{R} \to \mathbb{R}$ is measurable and $g : \mathbb{R} \to \mathbb{R}$ is continuous, then $g \circ f : \mathbb{R} \to \mathbb{R}$ is measurable.

Proof. Let $a \in \mathbb{R}$. Then

$$\begin{aligned} (g \circ f)^{-1}((-\infty, a)) &= \{x \in \mathbb{R} : g(f(x)) \in (-\infty, a)\} \\ &= \{x \in \mathbb{R} : f(x) \in g^{-1}((-\infty, a))\} \\ &= f^{-1}(g^{-1}((-\infty, a))) \\ &\in \mathcal{M} \end{aligned}$$

We used the fact that $(-\infty, a)$ is open, hence continuity ensures $g^{-1}((-\infty, a))$ is open, and thus also Borel. So then taking the inverse image with f returns a measurable set.

■

Remark. The order of composition matters (counter example in Assignment 2). Taking $f \circ g$ would lead to $g^{-1}(f^{-1}((-\infty, a)))$ which is not necessarily measurable.

Proposition 3.11. If f is measurable, then

1. cf is measurable for all $c \in \mathbb{R}$, where $cf(x) = c \cdot f(x)$ for $x \in \mathbb{R}$
2. $|f|$ is measurable, where $|f|(x) = |f(x)|$ for $x \in \mathbb{R}$

3. f^k is measurable for all $k \in \mathbb{N}$, where $f^k(x) = (f(x))^k$ for $x \in \mathbb{R}$

Proof. We prove 3. Let $k \in \mathbb{R}$ and let $a \in \mathbb{R}$. Observe that

$$\left(f^k\right)^{-1}\left([-\infty, a)\right) = \begin{cases} f^{-1}\left([-\infty, a^{1/k})\right) & \text{if } k \text{ odd} \\ \emptyset & \text{if } k \text{ even and } a \leq 0 \\ f^{-1}\left((-a^{1/k}, a^{1/k})\right) & \text{if } k \text{ even and } a > 0 \end{cases}$$

The case where k is even and $a > 0$ follows since

$$\left(f^k\right)^{-1}\left((-\infty, a)\right) = \left(f^k\right)^{-1}\left([0, a)\right) = \{x \in \mathbb{R} : f^k(x) \in [0, a)\} = \{x \in \mathbb{R} : f(x) \in (-a^{1/k}, a^{1/k})\}$$

■

Proposition 3.12. Let f, g be finite measurable functions. Then

1. $f + g$
2. $f \vee g := \max\{f, g\}$
3. $f \wedge g := \min\{f, g\}$

Are all measurable functions. To clarify, $(f \vee g)(x) = \max\{f(x), g(x)\}$ for all $x \in \mathbb{R}$.

Proof. We prove 1. Let $a \in \mathbb{R}$. Then

$$\begin{aligned} (f+g)^{-1}([-\infty, a)) &= (f+g)^{-1}((-\infty, a)) \\ &= \{x \in \mathbb{R} : f(x) + g(x) < a\} \\ &= \{x \in \mathbb{R} : f(x) < a - g(x)\} \\ &= \bigcup_{q \in \mathbb{Q}} \{x \in \mathbb{R} : f(x) < q < a - g(x)\} && \text{(standard trick)} \\ &= \bigcup_{q \in \mathbb{Q}} \Big(\{x \in \mathbb{R} : f(x) < q\} \cap \{x \in \mathbb{R} : g(x) < a - q\}\Big) && \in \mathcal{M} \end{aligned}$$

Next, we prove 2. It suffices to observe that $f \vee g = \frac{1}{2}\big(|f - g| + (f + g)\big)$, which is measurable. To prove 3, notice that $f \wedge g = \min\{f, g\} = -\max\{-f, -g\} = (-f) \vee (-g)$.

■

Remark (1). We need f, g to be finite for $f(x) + g(x)$ to be well-defined. Similarly, we need this assumption for $f \vee g$ and $f \wedge g$.

Remark (2). Here, we "wedge" some number q between $f(x)$ and $a - g(x)$, where q is independent of f and g. This number comes from a countable set, so taking union keeps us within the σ-algebra.

Remark (3). In general measure spaces, this trick may no longer apply.

Proposition 3.13. Let f, g be finite measurable functions. Then

1. $f - g$ is measurable
2. $(f+g)^2, (f-g)^2$ are measurable
3. fg is measurable

With the next corollary, we define functions f^+ and f^-.

Corollary 3.0.2. If f is measurable, then $f^+ := f \vee 0 = \max\{f, 0\}$ is measurable. Furthermore, $f^- := -(f \wedge 0) = \max\{-f, 0\}$ is measurable.

Remark (1). Note that both f^- and f^+ are non-negative.

Remark (2). It is appropriate to write $f = f^+ - f^-$, since $\infty - \infty$ will never occur. As well, we have $|f| = f^+ + f^-$.

Remark (3). More strongly, $f \wedge k$ is measurable for all $k \in \mathbb{R}$.

Proposition 3.14. Let $\{f_n : n \geq 1\}$ be a sequence of measurable functions. Then

1. $\sup_{n\in\mathbb{N}} f_n$
2. $\inf_{n\in\mathbb{N}} f_n$
3. $\limsup_{n\to\infty} f_n$
4. $\liminf_{n\to\infty} f_n$

Are all measurable functions.

1. *Proof.* Let $a \in \mathbb{R}$. We need to show $\{\sup_{n\in\mathbb{N}} f_n \leq a\} \in \mathcal{M}$. To see this,

$$\begin{aligned}
& x \in \{\sup_{n\in\mathbb{N}} f_n \leq a\} \\
\iff & \sup_{n\in\mathbb{N}} f_n(x) \leq a \\
\iff & f_n(x) \leq a && (\forall n \in \mathbb{N}) \\
\iff & x \in \{f_n \leq a\} && (\forall n \in \mathbb{N}) \\
\iff & x \in \bigcap_{n=1}^{\infty} \{f_n \leq a\}
\end{aligned}$$

And we know $\bigcap_{n=1}^{\infty} \{f_n \leq a\} \in \mathcal{M}$.

■

2. *Proof.* For an easy proof, observe that $\inf_{n\in\mathbb{N}} f_n = -\sup_{n\in\mathbb{N}}(-f_n)$.

Alternatively, we have

$$\begin{aligned}
&x \in \{\inf_{n\in\mathbb{N}} f_n < a\} \\
\iff &\inf_{n\in\mathbb{N}} f_n(x) < a \\
\iff &\exists n \in \mathbb{N} \text{ s.t. } f_n(x) < a \\
\iff &x \in \bigcup_{n=1}^{\infty} \{f_n < a\}
\end{aligned}$$

So $\{\inf_{n\in\mathbb{N}} f_n < a\} = \bigcup_{n=1}^{\infty}\{f_n < a\} \in \mathcal{M}$.

■

3. *Proof.* Observe that $\sup\{f_k(x) : k \geq n\}$ is a decreasing sequence in n. Therefore

$$\limsup_{n\to\infty} f_k = \lim_{n\to\infty} \sup_{k\geq n} f_k = \inf_{n\geq 1} \sup_{k\geq n} f_k$$

But $g_n := \sup_{k\geq n} f_k$ is measurable for all $n \geq 1$ (we proved this for the case $n = 1$, but it easily generalizes to arbitrary $n \in \mathbb{N}$ by adjusting where we start the intersection). Therefore $\inf_{n\geq 1} g_n$ is measurable. So $\limsup_{n\to\infty} f_n$ is measurable.

■

4. *Proof.* We use the same idea by observing that $\inf\{f_k(x) : k \geq n\}$ is an increasing sequence in n. Therefore

$$\liminf_{n\to\infty} f_k = \lim_{n\to\infty} \inf_{k\geq n} f_k = \sup_{n\geq 1} \inf_{k\geq n} f_k$$

Then $\liminf_{n\to\infty} f_n$ is measurable.

■

Remark. The first proof fails if we study $\{\sup_{n\in\mathbb{N}} f_n < a\}$ instead, since we then cannot claim "$\sup_{n\in\mathbb{N}} f_n(x) < a \iff f_n(x) < a$ for all $n \in \mathbb{N}$". For example take $f_n \equiv 1 - \frac{1}{n}$, where $f_n(x) < 1$ for all $n \in \mathbb{N}$, but $\sup_{n\in\mathbb{N}} f_n(x) = 1$.

Similarly, the second proof fails if we study $\{\sup_{n\in\mathbb{N}} f_n \leq a\}$ instead. This happens because we cannot claim "$\inf_{n\in\mathbb{N}} f_n(x) \leq a \iff \exists n \in \mathbb{N}$ s.t. $f_n(x) \leq a$". Take $f_n \equiv 1 + \frac{1}{n}$, then for all $x \in \mathbb{R}$ we have $\inf_{n\in\mathbb{N}} f_n(x) = 1$, but $f_n(x) > 1$ for all $n \in \mathbb{N}$.

Proposition 3.15. Let $\{f_n : n \geq 1\}$ be a sequence of measurable functions. Then the following sets are in $\mathcal{M}$.

1. $\{\lim_{n\to\infty} f_n$ exists in $\mathbb{R}\}$
2. $\{\lim_{n\to\infty} f_n = c\}$ for all $c \in \mathbb{R}$
3. $\{\lim_{n\to\infty} f_n = \infty\}$

4. $\{\lim_{n\to\infty} f_n = -\infty\}$

Moreover, suppose $\lim_{n\to\infty} f_n$ exists a.e. (in $\mathbb{R}$ or as $\pm\infty$). Define $f = \lim_{n\to\infty} f_n$ a.e. Then f is measurable.

Proof.

1. Using the Cauchy characterization of limit, then $\lim_{n\to\infty} f_n(x)$ exists in $\mathbb{R}$ if and only if

   $$\forall m \in \mathbb{N}, \exists N \in \mathbb{N}, \forall n, l \geq N \quad |f_n(x) - f_l(x)| < \frac{1}{m}$$

   Therefore

   $$\{\lim_{n\to\infty} f_n \text{ exists in } \mathbb{R}\} = \bigcap_{m=1}^{\infty} \bigcup_{N=1}^{\infty} \bigcap_{n=N}^{\infty} \bigcap_{l=N}^{\infty} \{|f_n - f_l| < \frac{1}{m}\}$$

   Taking difference and absolute value of measurable functions preserves measurability. Furthermore, we are using countably many set operations. Hence the set of interest is measurable.

2. Notice this is NOT the same statement as the first one. Fixing $c \in \mathbb{R}$, we have $x \in \{\lim_{n\to\infty} f_n = c\}$ if and only if $\lim_{n\to\infty} f_n(x) = c$. But this is true if and only if

   $$\forall m \geq 1, \exists N \in \mathbb{N}, \forall n \geq N \quad |f_n(x) - c| < \frac{1}{m}$$

   So then

   $$\{\lim_{n\to\infty} f_n = c\} = \bigcap_{m=1}^{\infty} \bigcup_{N=1}^{\infty} \bigcap_{n=N}^{\infty} \{|f_n - c| < \frac{1}{m}\}$$

   Hence the set of interest is measurable.

3. We have $x \in \{\lim_{n\to\infty} f_n = \infty\}$ if and only if $\lim_{n\to\infty} f_n = \infty$. But this is true if and only if

   $$\forall M > 0, \exists N \in \mathbb{N}, \forall n \geq N \quad f_n(x) > M$$

   So then

   $$\{\lim_{n\to\infty} = \infty\} = \bigcap_{M=1}^{\infty} \bigcup_{N=1}^{\infty} \bigcap_{n=N}^{\infty} \{f_n > M\}$$

   Hence the set of interest is measurable.

4. Similarly, we have

   $$\{\lim_{n\to\infty} = -\infty\} = \bigcap_{M=1}^{\infty} \bigcup_{N=1}^{\infty} \bigcap_{n=N}^{\infty} \{f_n < -M\}$$

   Hence the set of interest is measurable.

■

Remark. We can also prove the first statement by noting that $\lim_{n\to\infty} f_n(x)$ exists in $\mathbb{R}$ if and only if

$$\limsup_{n\to\infty} f_n(x) = \liminf_{n\to\infty} f_n(x) \textbf{ and } \limsup_{n\to\infty} f_n(x), \liminf_{n\to\infty} f_n(x) \in (-\infty, \infty)$$

We require the limit superior and inferior to be finite for the reverse implication to hold. Then

$$\begin{aligned}\{\lim_{n\to\infty} f_n \text{ exists in } \mathbb{R}\} =& \{\limsup_{n\to\infty} f_n - \liminf_{n\to\infty} f_n = 0\}\\ &\cap \{-\infty < \limsup_{n\to\infty} f_n < \infty\}\\ &\cap \{-\infty < \liminf_{n\to\infty} f_n < \infty\}\end{aligned}$$

Which is measurable with a little more work (use the flexibility of inverse image under intersection and union).

3.2 Approximation by simple functions

In this section, the main result is that given some measurable function f, we can always find a sequence of "simple functions" ψ_n such that $\lim_{n\to\infty} \psi_n(x) = f(x)$ for all $x \in \mathbb{R}$.

3.2.1 STEP 1

Let $f : \mathbb{R} \to \overline{\mathbb{R}}$ be a measurable function. Then $f = f^+ - f^-$. Recall that f^+, f^- are non-negative and measurable.

3.2.2 STEP 2

For all $n \geq 1$, we set

1. $f_n^+ := (f^+ \wedge n)\mathbb{1}_{[-n,n]}$
2. $f_n^- := (f^- \wedge n)\mathbb{1}_{[-n,n]}$

In particular, f_n^+ is

1. **Non-negative**
2. **Measurable**: Since $f^+ \wedge n$ is measurable, $\mathbb{1}_{[-n,n]}$ is measurable as $[-n, n] \in \mathcal{M}$, and any product of measurable functions is again measurable.
3. **Bounded**: Bounded below by 0, and above by n.
4. **Zero outside of a finite measure set**

Further, we have that $f_n^+ \leq f_{n+1}^+$ for all $n \in \mathbb{N}$. And $f_n^+ \leq f$ for all $n \in \mathbb{N}$. Therefore $\lim_{n\to\infty} f_n^+(x) = f(x)$ for all $x \in \mathbb{R}$. This is seen as "lifting the cap" and "expanding the window".

All of these properties hold for f_n^- as well.

3.2.3 STEP 3

Now for $k = 0, 1 \dots, n \cdot 2^n$, define

$$A_{n,k} := \left\{ x \in [-n, n] : \frac{k}{2^n} \leq f_n^+(x) < \frac{k+1}{2^n} \right\} \in \mathcal{M}$$

Note that $\boldsymbol{A_{n,k}}$ **is measurable**. To see this, use the fact that f_n^+ is measurable and then write

$$A_{n,k} = [-n, n] \cap \left\{ f_n^+(x) \in \left[\frac{k}{2^n}, \frac{k+1}{2^n} \right] \right\}$$

Furthermore, $\boldsymbol{A_{n,k} \cap A_{n,l} = \emptyset}$ **if** $\boldsymbol{k \neq l}$. This is because the function graph is sliced into disjoint horizontal segments, and each $x \in [-n, n]$ is mapped to a unique $f(x)$.

Now define a function ϕ_n as

$$\phi_n := \sum_{k=0}^{n \cdot 2^n} \mathbb{1}_{A_{n,k}} \frac{k}{2^n}$$

So if $x \in A_{n,k}$, then $\phi_n(x) = \frac{k}{2^n}$. We call ϕ a *simple function*, as defined below. Importantly, $\{\phi_n : n \geq 1\}$ is an increasing sequence of measurable functions. It then happens that $\lim_{n\to\infty} \phi_n(x) = f^+(x)$ for all $x \in \mathbb{R}$. This is intuitive since the "steps" get finer and finer, as we lift the cap and expand the window. We prove this after the next definition.

Definition 3.3 (Simple function)**.** Let $1 \leq k \leq L \in \mathbb{N}$. Let $E_k \in \mathcal{M}$ with $m(E_k) < \infty$. Let $a_k \in \mathbb{R}$. Suppose a function ϕ has the form

$$\phi = \sum_{k=1}^{L} \mathbb{1}_{E_k} a_k$$

Then ϕ is a simple function.

Remark. **Every simple function is measurable.** This follows since $E_k \in \mathcal{M}$ gives measurability of $\mathbb{1}_{E_k}$. Thus $\mathbb{1}_{E_k} a_k$ is measurable since a_k is a constant. Then taking the summation over $k = 1, \dots, L$ also preserves measurability, since we are taking a finite sum of finite-valued measurable functions.

Proposition 3.16. Let $\{\phi_n : n \geq 1\}$ be defined as above, i.e.,

$$\phi_n := \sum_{k=0}^{n \cdot 2^n} \mathbb{1}_{A_{n,k}} \frac{k}{2^n}$$

Then $\lim_{n\to\infty} \phi_n(x) = f^+(x)$ for all $x \in \mathbb{R}$.

Proof. Let $x \in \mathbb{R}$. To start, choose M such that $|x| \leq M$. Then we are "within the window" in the sense that $f^+(x) = f^+(x)\mathbb{1}_{[-M,M]}(x)$.

Next, assume that $f^+(x) < \infty$. Then there exists $N \in \mathbb{N}$ such that $f^+(x) < N$. So we are "under the ceiling". Let $n = \max\{M, N\}$. Then

$$f^+(x) = (f^+(x) \wedge n)\mathbb{1}_{[-n,n]} = f_n^+(x)$$

Meaning we are both "within the window" and "under the ceiling" at the same time. Therefore

$$|f^+(x) - \phi_n(x)| = |f_n^+(x) - \phi_n(x)| \leq \frac{1}{2^n}$$

For every $m \geq n$, this inequality continues to hold since $m \geq \max\{M, N\}$. Therefore $\lim_{n\to\infty} \phi_n(x) = f^+(x)$.

Next, assume $f^+(x) = \infty$. Then by construction $\phi_n(x) = n$, since all values above n get truncated to n. Therefore $\lim_{n\to\infty} \phi_n(x) = \infty = f^+(x)$.

∎

Remark (1). Similarly, we can also define simple functions $\{\tilde{\phi}_n : n \geq 1\}$ which are increasing with $\tilde{\phi}_n \leq f^-$, and $\lim_{n\to\infty} \tilde{\phi}_n(x) = f^-(x)$ for all $x \in \mathbb{R}$.

Remark (2). Note that $\psi_n \leq \psi_{n+1} \leq f_n^+ \leq f^+$.

Theorem 3.1. Let f be measurable. Then there exists a sequence of simple functions $\{\psi_n : n \geq 1\}$ such that

1. $|\psi_n|$ is increasing with $|\psi_n| \leq |\psi_{n+1}|$ for all $n \in \mathbb{N}$
2. $|\psi_n| \leq |f|$ for all $n \in \mathbb{N}$
3. The sequence $\{\psi_n : n \geq 1\}$ attains f as its limit (pointwise)

$$\lim_{n\to\infty} \psi_n(x) = f(x) \quad \forall x \in \mathbb{R}$$

Proof. Set $A_{n,k}$ as before and define $B_{n,k}$ for f^- as

$$B_{n,k} := \left\{x \in [-n, n] : \frac{k}{2^n} \leq f_n^-(x) < \frac{k+1}{2^n}\right\}$$

Then define ψ_n as

$$\boldsymbol{\psi_n := \phi_n - \tilde{\phi}_n} = \sum_{k=0}^{n2^n} \mathbb{1}_{A_{n,k}} \frac{k}{2^n} - \sum_{k=0}^{n2^n} \mathbb{1}_{B_{n,k}} \frac{k}{2^n} = \sum_{k=0}^{n2^n} \frac{k}{2^n}(\mathbb{1}_{A_{n,k}} - \mathbb{1}_{B_{n,k}})$$

Which is again a simple function. The $|\psi_n|$ are increasing since

$$|\psi_n| = |\phi_n - \tilde{\phi}_n| = \phi_n + \tilde{\phi}_n \leq \phi_{n+1} + \tilde{\phi}_{n+1} = |\psi_{n+1}|$$

Furthermore

$$|\psi_n| = |\phi_n - \tilde{\phi}_n| = \phi_n + \tilde{\phi}_n \leq f^+ + f^- = |f|$$

Then

$$\lim_{n\to\infty} \psi_n(x) = \lim_{n\to\infty} \phi_n(x) - \lim_{n\to\infty} \tilde{\phi}_n(x) = f^+ - f^- = f$$

■

Remark. We could have used the triangle inequality to get $|\phi_n - \tilde{\phi}_n| \leq \phi_n + \tilde{\phi}_n$. But at least one of ϕ_n is zero for every $x \in \mathbb{R}$, so we actually have equality.

3.3 Approximation by step functions

Now, we would like to approximate a measurable function f using step functions, which are a type of simple function. **However in doing so, we lose pointwise convergence, and can only achieve almost everywhere convergence.**

Definition 3.4 (Step function). θ is called a step function if

$$\theta(x) = \sum_{k=1}^{L} a_k \mathbb{1}_{I_k}(x)$$

Where $L \in \mathbb{N}$, and for all $k \in \mathbb{N}$ we have a_k is constant, and I_k is a finite open inteval.

Theorem 3.2. Let f be measurable. Then there exists a sequence of step functions $\{\theta_n : n \geq 1\}$ such that

$$\lim_{n\to\infty} \theta_n(x) = f(x)$$

For **almost every** $x \in \mathbb{R}$.

Proof. W.L.O.G. assume f is non-negative. Otherwise, we can write $f = f^+ - f^-$ and treat each non-negative case separately.

We first establish a fact. Let $\varepsilon > 0$. Let $A \in \mathcal{M}$ with $m(A) < \infty$. Due to regularity of Lebesgue measure, there exists finite and open intervals $I_1, \ldots, I_N$ such that

$$m\Big(A \triangle \bigcup_{i=1}^{N} I_i\Big) < \varepsilon$$

We can assume the I_i are pairwise disjoint. Otherwise, we can make them so by taking the union of overlapping I_i and then renaming them. This does not affect the regularity property.

Let $\theta_A := \sum_{i=1}^{N} \mathbb{1}_{I_i}$, which is a step function. Due to disjointness of I_i then θ_A is only ever 0 or 1, and we have

$$\theta_A = \mathbb{1}_{\bigcup_{i=1}^{N} I_i}$$

Furthermore notice that

$$A \triangle \bigcup_{i=1}^{N} I_i = \{x \in \mathbb{R} : \mathbb{1}_A(x) \neq \mathbb{1}_{\bigcup_{i=1}^{N} I_i}(x)\} = \{x \in \mathbb{R} : \mathbb{1}_A(x) \neq \theta_A(x)\}$$

So therefore

$$m\Big(\{x \in \mathbb{R} : \mathbb{1}_A(x) \neq \theta_A(x)\}\Big) < \varepsilon$$

Now we proceed with the proof itself. Since f is non-negative and measurable, there exists $\{\phi_n : n \geq 1\}$ such that $\lim_{n\to\infty} \phi_n = f$. Where ϕ_n are again defined as

$$\phi_n := \sum_{k=0}^{n2^n} \frac{k}{2^n} \mathbb{1}_{A_{n,k}}$$

Recall that $A_{n,k}$ is measurable. By our fact, for every $A_{n,k}$ there exists a step function $\theta_{n,k}$ such that

$$m\Big(\{x \in \mathbb{R} : \mathbb{1}_{A_{n,k}}(x) \neq \theta_{n,k}(x)\}\Big) \leq \frac{1}{2^n(n2^n+1)}$$

Now set

$$\theta_n := \sum_{k=0}^{n2^n} \frac{k}{2^n} \theta_{n,k}$$

which is a step function since each $\theta_{n,k}$ is decomposed into non-overlapping characteristic functions, each over a finite open interval $I_{n,k,i}$.

Then set

$$E_n := \{x \in \mathbb{R} : \theta_n(x) \neq \phi_n(x)\}$$

Observe that for all $x \in \mathbb{R}$

$$\forall 0 \leq k \leq n2^n \quad \frac{k}{2^n}\mathbb{1}_{A_{n,k}}(x) = \frac{k}{2^n}\theta_{n,k}(x) \implies \phi_n(x) = \theta_n(x)$$

Taking the contrapositive, if there exists $x \in \mathbb{R}$ such that $\phi_n(x) \neq \theta_n(x)$, **then** there exists $0 \leq k \leq n2^n$ such that $\mathbb{1}_{A_{n,k}}(x) \neq \theta_{n,k}(x)$.

The point then is that $E_n \subseteq \bigcup_{k=0}^{n2^n} \{\theta_{n,k} \neq \mathbb{1}_{A_{n,k}}\}$. So therefore

$$\begin{aligned}
m(E_n) &\leq m\Big(\bigcup_{k=0}^{n2^n} \{\theta_{n,k} \neq \mathbb{1}_{A_{n,k}}\}\Big) \\
&\leq \sum_{k=0}^{n2^n} m(\{\theta_{n,k} \neq \mathbb{1}_{A_{n,k}}\}) \\
&\leq (n2^n+1)\frac{1}{2^n(n2^n+1)} \\
&= \frac{1}{2^n}
\end{aligned}$$

Now recall that in our discussion of simple functions, by construction the ϕ_n satify

$$|\phi_n(x) - f_n(x)| \leq \frac{1}{2^n} \text{ where } f_n := (f \wedge n)\mathbb{1}_{[-n,n]}$$

Define F_n as

$$F_n := \{x \in \mathbb{R} : |\theta_n(x) - f_n(x)| > \frac{1}{2^n}\}$$

If $x \in F_n$, then automatically $\theta_n(x) \neq \phi_n(x)$. Otherwise, if $\theta_n(x) = \phi_n(x)$, then we derive the contradiction $|\phi_n(x) - f_n(x)| \leq \frac{1}{2^n}$ AND $|\phi_n(x) - f_n(x)| > \frac{1}{2^n}$. Therefore $F_n \subseteq E_n$ so by monotonicity $m(F_n) \leq m(E_n) \leq \frac{1}{2^n}$.

Now, we claim that

$$\text{For a.e. } x \in \mathbb{R}, \exists m \geq 1 \text{ s.t. } \forall n \geq m \quad |\theta_n(x) - f_n(x)| \leq \frac{1}{2^n}$$

It suffices to show that the set of $x \in \mathbb{R}$ which satisfies the opposite statement has measure zero. That is, we wish to show

$$m\Big(\bigcap_{m=1}^{\infty} \bigcup_{n=m}^{\infty} \{x \in \mathbb{R} : |\theta_n(x) - f_n(x)| > \frac{1}{2^n}\}\Big) = 0$$

We have $F_n \in \mathcal{M}$ for all $n \in \mathbb{N}$ due to measurability of $|\theta_n - f_n|$. Furthermore $\sum_{n=1}^{\infty} m(F_n) \leq \sum_{n=1}^{\infty} \frac{1}{2^n} = 1 < \infty$. Therefore the Borel-Cantelli Lemma gives the desired result. Thus for a.e. $x \in \mathbb{R}$, there exists $m \geq 1$ such that for all $n \geq m$ we have

$$\theta_n = (\theta_n - f_n) + f_n \leq \frac{1}{2^n} + f_n$$

So for a.e. $x \in \mathbb{R}$ we have $\theta_n \to f_n$ as $n \to \infty$.

■

Remark. To summarize:

1. We found a sequence of *simple functions* that converges to f pointwise.
2. We dissected each ϕ_n, and showed that for each $\mathbb{1}_{A_{n,k}}$ it is composed of, we can find a *step function* $\theta_{n,k}$ that looks very similar. The set of $x \in \mathbb{R}$ where $\theta_{n,k}(x) \neq \mathbb{1}_{A_{n,k}}(x)$ has measure bounded by $\frac{1}{2^n(n2^n+1)}$.
3. We summed the $\theta_{n,k}$ (with appropriate coefficients) to produce another step function θ_n which looks very similar to ϕ_n. Indeed, the set E_n of $x \in \mathbb{R}$ where $\theta_n(x) \neq \phi_n(x)$ has measure bounded by $\frac{1}{2^n}$.
4. We then noticed that the set F_n of $x \in \mathbb{R}$ where $|\theta_n(x) - f_n(x)| > \frac{1}{2^n}$ is a subset of E_n, due to how similar ϕ_n is to f_n. This provided a bound on $m(F_n)$.
5. The Borel-Cantelli Lemma showed that the set of x which hits F_n infinitely often, has measure zero. In almost every case, x eventually forever satisfies $|\theta_n(x) - f_n(x)| \leq \frac{1}{2^n}$. This directly led to the conclusion.

Lemma 3.1 (Borel-Cantelli)**.** Let $\{F_n : n \geq 1\} \subseteq \mathcal{M}$. If $\sum_{n=1}^{\infty} m(F_n) < \infty$, then

$$m\Big(\bigcap_{m=1}^{\infty} \bigcup_{n=m}^{\infty} F_n\Big) = 0$$

Proof. Define an auxiliary sequence

$$B_m := \bigcup_{n=m}^{\infty} F_n$$

The sequence $\{B_m : m \geq 1\}$ is decreasing. Furthermore, we have

$$m(B_1) = m\Big(\bigcup_{n=1}^{\infty} F_n\Big) \leq \sum_{n=1}^{\infty} m(F_n) < \infty$$

We can then apply continuity from above to find

$$\begin{aligned}
m\Big(\bigcap_{m=1}^{\infty} \bigcup_{n=m}^{\infty} F_n\Big) &= m\Big(\bigcap_{m=1}^{\infty} B_m\Big) \\
&= \lim_{m\to\infty} m(B_m) \\
&\leq \lim_{m\to\infty} \sum_{n=m}^{\infty} m(F_n) \\
&= \lim_{m\to\infty} \Big(\sum_{n=1}^{\infty} m(F_n) - \sum_{n=1}^{m-1} m(F_n)\Big) \\
&= 0
\end{aligned}$$

■

Remark. Intuitively, this is like saying that the chance of an event "happening" infinitely often is zero. If $x \in \bigcap_{m=1}^{\infty} \bigcup_{n=m}^{\infty} F_n$, then x will appear in infinitely many F_n (x will happen infinitely often)

3.4 Convergence a.e. versus convergence in measure

Definition 3.5. Let $\{f_n\}$, f be measurable functions. Then f_n converges to f a.e., that is "$f_n \to f$ a.e.", if for a.e. $x \in \mathbb{R}$

$$\lim_{n\to\infty} f_n(x) = f(x)$$

Similarly, we have "$f_n \to f$ a.e. on A" if there exists $B \subseteq A$ with $m(B) = 0$ such that $\forall x \in A \setminus B$ we have $\lim_{n\to\infty} f_n(x) = f(x)$.

Definition 3.6. Let $\{f_n\}$, f be measurable *finite-valued* functions. Then f_n converges to f in measure, that is "$f_n \to f$ in measure", if for all $\delta > 0$

$$\lim_{n\to\infty} m\big(\{x \in \mathbb{R} : |f_n(x) - f(x)| \geq \delta\}\big) = 0$$

Similarly, we have "$f_n \to f$ in measure on A" if for all $\delta > 0$ we have $\lim_{n\to\infty} m(\{x \in A : |f_n(x) - f(x)| \geq \delta\}) = 0$.

Proposition 3.17. Let $\{f_n\}, f$ be measurable *finite-valued* functions. Let $A \in \mathcal{M}$ with $m(A) < \infty$.

If $f_n \to f$ a.e. on A, then $f_n \to f$ in measure on A.

Hence convergence a.e. is stronger than convergence in measure.

Proof. Let $\delta > 0$. Suppose we have some $x \in A$ such that for all $m \geq 1$ there exists $n \geq m$ such that $|f_n(x) - f(x)| > \delta$. Then $f_n(x)$ cannot converge to $f(x)$. So

$$\bigcap_{m=1}^{\infty} \bigcup_{n=m}^{\infty} \{x \in A : |f_n(x) - f(x)| > \delta\} \subseteq \{x \in A : \lim_{n\to\infty} f_n(x) \neq f(x)\}$$

We can only claim "$\subseteq$" because this δ represents just one possible failure mode. We may need to make δ smaller to discover other $x \in A$ for which x hits the event $|f_n(x) - f(x)| > \delta$ infinitely often.

Define B_m as

$$B_m := \bigcup_{n=m}^{\infty} \{x \in A : |f_n(x) - f(x)| > \delta\}$$

By assumption $m(\{x \in A : \lim_{n\to\infty} f_n(x) \neq f(x)\}) = 0$. Therefore $\bigcap_{n=1}^{\infty} B_m$ is a null set. We have $B_1 \subseteq A$ which gives $m(B_1) \leq m(A) < \infty$ by monotonicity. Furthermore $\{B_m : m \geq 1\}$ is a decreasing sequence. Continuity from above then ensures

$$\lim_{m\to\infty} m(B_m) = m\Big(\bigcap_{m=1}^{\infty} B_m\Big) = 0$$

Now observe that

$$\{x \in A : |f_m(x) - f(x)| > \delta\} \subseteq B_m$$

because $\{x \in A : |f_m(x) - f(x)| > \delta\}$ appears first in the union that defines B_m. Therefore

$$m(\{x \in A : |f_n(x) - f(x)| > \delta\}) \leq m(B_m) \to 0 \text{ as } n \to \infty$$

Which is our definition of convergence in measure on A.

∎

Remark (1). We need $m(A) < \infty$ to make a statement about the B_m using continuity from above. Indeed, suppose we left that assumption out. Then take $f_n := \mathbb{1}_{[n,\infty)}$ and $f = 0$. Then

$$\lim_{n\to\infty} f_n(x) = f(x) = 0 \quad \forall x \in \mathbb{R}$$

But $m(\{x \in \mathbb{R} : |f_n(x) - f(x)| = 1\}) = m([n,\infty)) = \infty$.

Remark (2). The statement

$$m\Big(\bigcap_{m=1}^{\infty} \bigcup_{n=m}^{\infty} \{x \in A : |f_n(x) - f(x)| > \delta\}\Big) = 0$$

has similar form to the conclusion of Borel-Cantelli Lemma.

Example 3.1. In general, even on a finite measure set A, convergence in measure does not imply convergence almost everywhere. Fill in this exmaple.

Proposition 3.18. Let $\{f_n\}, f$ be measurable and finite-valued. If $f_n \to f$ in measure, then there exists a subsequence $\{n_k : k \geq 1\} \subseteq \mathbb{N}$ such that $f_{n_k} \to f$ a.e. as $k \to \infty$.

Proof. Since $f_n \to f$ in measure, then for every $\delta > 0$ we have $m(\{|f_n - f| > \delta\}) \to 0$ as $n \to \infty$. For every $k \geq 1$, set $\delta := \frac{1}{k}$ and $\varepsilon := \frac{1}{k^2}$. So for each $k \geq 1$ there exists $n_k \in \mathbb{N}$ such that

$$m(\{|f_{n_k} - f| > \frac{1}{k}\}) \leq \frac{1}{k^2}$$

Now set

$$A_k := \{|f_{n_k} - f| > \frac{1}{k}\}$$

for which $\sum_{n=1}^{\infty} m(A_k) < \infty$, since the series is bounded above by a convergent p-series. Using the Borel-Cantelli Lemma, then

$$m\Big(\bigcap_{m=1}^{\infty} \bigcup_{n=m}^{\infty} A_n\Big) = 0$$

(In detail, $m(\bigcap_{m=1}^{\infty} \bigcup_{n=m}^{\infty} A_n) = \lim_{m\to\infty} m(\bigcup_{n=m}^{\infty} A_m) \leq \lim_{m\to\infty} \sum_{n=m}^{\infty} A_n = 0$.)

So for $x \in \mathbb{R}$ to hit A_k infinitely often will only happen on a set of measure zero. The complement statement is

$$\text{For a.e. } x \in \mathbb{R}, \exists m \geq 1 \text{ s.t. } \forall k \geq m \quad |f_{n_k}(x) - f(x)| \leq \frac{1}{k}$$

In other words for a.e. $x \in \mathbb{R}$ we have

$$\lim_{k\to\infty} f_{n_k}(x) = f(x)$$

■

Proposition 3.19 (Subsequence test). Let $\{f_n\}, f$ be measurable and finite-valued. Then $f_n \to f$ in measure *if and only if* for all subsequences $\{n_k\}$ there exists a subsubsequence $\{n_{k_l}\} \subseteq \{n_k\} \subseteq \mathbb{N}$ such that $f_{n_{k_l}} \to f$ in measure as $l \to \infty$.

Proof. Assume $f_n \to f$ in measure. Then all subsequences f_{n_k} must also converge to f in measure. So take $\{n_{k_l}\} := \{n_k\}$ as the subsubsquence which converges to f in measure.

Now assume for all subsequences $\{n_k\}$ there exists a subsubsequence $\{n_{k_l}\} \subseteq \{n_k\}$ such that $f_{n_{k_l}} \to f$ in measure. Suppose f_n does not converge to f in measure, for contradiction.

Before proceeding, we claim a sequence (x_n) does not converge to x if and only if there exists $\varepsilon > 0$ such that $|x_n - x| \geq \varepsilon$ for infinitely many n. Suppose (x_n) does not converge to x, then by negating the limit definition

$$\exists \varepsilon > 0, \forall N \in \mathbb{N}, \exists n \in \mathbb{N} \quad |x_n - x| \geq \varepsilon$$

Thus we can construct a subsequence $\{x_{n_k}\}$ that satisfies $|x_{n_k} - x| \geq \varepsilon$ for all $k \geq 1$.
Now suppose there exists $\varepsilon > 0$ such $|x_n - x| \geq \varepsilon$ for infinitely many n. Assume $\lim_{n\to\infty} x_n = x$, so there exists $N \in \mathbb{N}$ such that for all $n \geq N$ we have $|x_n - x| < \varepsilon$. But then $|x_n - x| \geq \varepsilon$ for *finitely* many n, and contradiction arises.

Now back to the proof. Since we assumed f_n does not converge to f in measure, then there exists $\delta > 0$ such that $m(\{|f_n - f| > \delta\})$ does not converge to 0. Applying our small fact, then

$$\exists \delta > 0, \exists \varepsilon > 0, \exists \{n_k\}, \forall k \geq 1 \quad m(\{|f_{n_k} - f|\} > \delta) \geq \varepsilon$$

If $\delta > \varepsilon$ then $m(\{|f_{n_k} - f|\} > \varepsilon) \geq m(\{|f_{n_k} - f|\} > \delta) \geq \varepsilon$ by monotonicity. If $\delta \leq \varepsilon$ then $m(\{|f_{n_k} - f|\} > \delta) \geq \varepsilon \geq \delta$. So W.L.O.G we can simplify the above statement as

$$\exists \delta > 0, \exists \{n_k\}, \forall k \geq 1 \quad m(\{|f_{n_k} - f|\} > \delta) \geq \delta$$

But by hypothesis, there is a subsubsequence $\{n_{k_l}\} \subseteq \{n_k\}$ such that $f_{n_{k_l}} \to f$ in measure. Contradiction arises.

■

3.5 Egorov's Theorem and Lusin's Theorem

Theorem 3.3 (Egorov's Theorem)**.** Let $\{f_n\}, f$ be measurable. Let $A \in \mathcal{M}$ with $m(A) < \infty$.

If $f_n \to f$ a.e. on A, then for all $\varepsilon > 0$ there exists a closed $A_\varepsilon \subseteq A$ with $m(A \setminus A_\varepsilon) \leq \varepsilon$ such that $f_n \to f$ **uniformly** on A_ε.

Proof. First, we assume f is finite-valued on A. Otherwise, we may replace A with $A \cap \{-\infty < f < \infty\}$, which is still measurable.

Fix $\varepsilon > 0$. We want to show there exists $A_\varepsilon \subseteq A$ with A_ε closed such that $m(A \setminus A_\varepsilon) \leq \varepsilon$ and $f_n \to f$ uniformly on A_ε. Recall that

$$f_n \to f \text{ uniformly on } A_\varepsilon \iff \sup_{x \in A_\varepsilon} |f_n(x) - f(x)| \to 0 \text{ as } n \to \infty$$

Fix $k \geq 1$ and for all $n \geq 1$ set

$$E_n^{(k)} := \{x \in A : \forall j \geq n, |f_j(x) - f(x)| \leq \frac{1}{k}\}$$

$E_n^{(k)}$ contains the elements x such that $f_j(x)$ "stays within the band" for all $j \geq n$. Note that $E_n^{(k)}$ is increasing in n with $E_n^{(k)} \subseteq E_{n+1}^{(k)}$. Taking their countable union

$$\bigcup_{n=1}^{\infty} E_n^{(k)} = \{x \in A : \exists n \geq 1, \forall j \geq n, |f_j(x) - f(x)| \leq \frac{1}{k}\}$$

we thus further have

$$\bigcup_{n=1}^{\infty} E_n^{(k)} \supseteq \{x \in A : \lim_{n\to\infty} f_n(x) = f(x)\} =: A'$$

Since $f_n \longrightarrow f$ a.e. by hypothesis, then $m(A') = m(A)$. We derive

$$\begin{aligned} m(A) &= m(A') \\ &\leq m\Big(\bigcup_{n=1}^{\infty} E_n^{(k)}\Big) \\ &= \lim_{n\to\infty} m(E_n^{(k)}) && \text{(by continuity from below)} \\ &\leq m(A) && (\text{since } \forall n \geq 1, E_n^{(k)} \subseteq A) \end{aligned}$$

Therefore $\lim_{n\to\infty} m(E_n^{(k)}) = m(A)$, which is true for *arbitrary* $k \geq 1$. In other words, the elements which stay within the band for sufficiently large n approaches the measure of A. This means that given any $\varepsilon > 0$, for each $k \geq 1$ there exists $n_k \geq 1$ large enough such that

$$m(A \setminus E_{n_k}^{(k)}) \overset{(*)}{=} m(A) - m(E_{n_k}^{(k)}) < \frac{1}{2^k} \cdot \frac{\varepsilon}{2}$$

The equality at $(*)$ uses our assumption $m(A) < \infty$, as it ensures $E_{n_k}^{(k)} \subseteq A$ has finite measure. Now define

$$B := A \setminus \Big(\bigcap_{k=1}^{\infty} E_{n_k}^{(k)}\Big)$$

for which, using the upper bound above, we have

$$\begin{aligned} m(B) &= m\Big(\bigcup_{k=1}^{\infty} A \setminus E_{n_k}^{(k)}\Big) \\ &\leq \sum_{k=1}^{\infty} m(A \setminus E_{n_k}^{(k)}) \\ &\leq \frac{\varepsilon}{2} \end{aligned}$$

Set $\tilde{A} = A \setminus B = \cap_{k=1}^{\infty} E_{n_k}^{(k)}$ so that

$$\tilde{A} = \{x \in A : \forall k \geq 1, \forall j \geq n_k, |f_j(x) - f(x)| \leq \frac{1}{k}\}$$

We have distilled A into a set $\tilde{A}$, such that the "bad elements" which do not eventually stay within any *arbitrary* band, have been isolated to B. By definition, we thus have $f_n \longrightarrow f$ uniformly on $\tilde{A}$. By regularity of measure, there exists closed $A_\varepsilon \subseteq \tilde{A}$ such that $m(\tilde{A} \setminus A_\varepsilon) \leq \frac{\varepsilon}{2}$. So then

$$m(A \setminus A_\varepsilon) = m(A \setminus \tilde{A}) + m(\tilde{A} \setminus A_\varepsilon) = m(B) + m(\tilde{A} \setminus A_\varepsilon) \leq \varepsilon$$

Now for a general function f, realize that we can write $A = A^\infty \cup A^{-\infty} \cup A^{\mathbb{R}}$, where $A^\infty = \{f = \infty\}$, $A^{-\infty} = \{f = -\infty\}$, $A^{\mathbb{R}} = \{-\infty < f < \infty\}$. We have that $f_n \longrightarrow f$ a.e. on each of $A^\infty, A^{-\infty}, A^{\mathbb{R}}$.

It suffices to treat the case A^{∞}, and the case for $A^{-\infty}$ is similar. For all $k \geq 1$, for all $n \geq 1$ set $E_n^{(k)} := \{x \in A^{\infty} : \forall j \geq n, f_j(x) > k\}$. Again, we are collecting the points wth eventual good behaviour. Then again, for fixed $k \geq 1$ we have $m(A^{\infty}) = \lim_{n\to\infty} m(E_n^{(k)})$. The rest of the proof is same as above.

■

Remark (1). The requirement $m(A) < \infty$ is necessary. For example, take $f_n = \mathbb{1}_{[n,\infty)}$ where $f_n \to 0$ pointwise. But for all $a \in \mathbb{R}$ we do not have uniform convergence of $f_n \to 0$ on (a, ∞).

Remark (2). In general, although we can make ε arbitraily small, Egorov's Theorem does NOT imply $f_n \to f$ uniformly a.e. on A. For example, take the interval $[0,1]$ with $f_n(x) = x^n$ and $f \equiv 0$. Then for all $x \in [0,1)$ we have $f_n(x) \to f(x)$ as $n \to \infty$. Thus $f_n \to f$ a.e. on $[0,1]$.

Theorem 3.4 (Lusin's Theorem)**.** Let f be measurable. Let $A \in \mathcal{M}$ with $m(A) < \infty$.

Then for all $\varepsilon > 0$ there exists closed $A_\varepsilon \subseteq A$ with $m(A \setminus A_\varepsilon) \leq \varepsilon$ such that $f|_{A_\varepsilon}$ is **continuous**.

Proof. Since f is measurable, there exists a sequence of step functions $\{\theta_n\}$ such that $\theta_n \to f$ a.e. on A. As θ_n is piecewise constant, then it is piecewise continuous. Given any $\varepsilon > 0$, we can build an open set E_n such that

$$\theta_n\big|_{E_n^c} \text{ is continuous, and } m(E_n) \leq \frac{\varepsilon}{2} \cdot \frac{1}{2^n}$$

The idea is to surround the finite number of discontinuities with small open sets. Then their union produces E_n. Since $\theta_n \to f$ a.e. then by Egorov's theorem, there exists a closed set $B \subseteq A$ with $m(A \setminus B) \leq \frac{\varepsilon}{2}$ such that $\theta_n \to f$ uniformly on B. Now put

$$A_\varepsilon := B \setminus \left(\bigcup_{n=1}^{\infty} E_n\right) = \bigcap_{n=1}^{\infty} (B \setminus E_n)$$

which "removes the jumps". Furthermore notice that $A_\varepsilon \subseteq A$ is closed, as $B \setminus E_n$ is closed and any intersection of closed sets is again closed. Now compute

$$\begin{aligned} m(A \setminus A_\varepsilon) &\leq m(A \setminus B) + m(B \setminus A_\varepsilon) \\ &\leq \frac{\varepsilon}{2} + m\left(\bigcup_{n=1}^{\infty} E_n\right) \\ &\leq \frac{\varepsilon}{2} + \sum_{n=1}^{\infty} m(E_n) \\ &\leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \sum_{n=1}^{\infty} \frac{1}{2^n} \\ &= \varepsilon \end{aligned}$$

Since A_ε is a subset of B, we have $\theta_n \to f$ uniformly on A_ε. Furthermore the restriction $\theta_n\big|_{A_\varepsilon}$ is continuous. Therefore $\lim_{n\to\infty} \theta_n\big|_{A_\varepsilon} = f\big|_{A_\varepsilon}$ is again continuous, as the uniform convergence of continuous functions produces a continuous function.

$\blacksquare$

Remark (1). Lusin's Theorem says that $f|_{A_\varepsilon}$, as a function on A_ε, is continuous. This is NOT the same as saying f, as a function on A, is continuous on points in A_ε.

Remark (2). For every $k \geq 1$ there exists closed $A_k \subseteq A$ such that $m(A \backslash A_k) \leq \frac{1}{k}$ with f_{A_k} continuous. And in fact, we can assume A_k is increasing, otherwise replace A_k by $\tilde{A}_k := A_1 \cup \cdots \cup A_k$. Even though we have continuity when restricted on each increasing A_k, Lusin's Theorem does NOT imply f is continuous a.e..

Remark (3). In the proof, θ_n is not continuous on $\mathbb{R}$. However, we can do some surgery on each θ_n, by joining the finitely many pieces with steep lines, to form a sequence $\{\tilde{\theta}_n\}$ such that $\tilde{\theta}_n \to f$ a.e. on $\mathbb{R}$. **So actually every measurable function is the a.e. limit of a sequence of continuous functions.**

Example 3.2. Take $f = \mathbb{1}_{Q \cap [0,1]}$. Then f is not continuous anywhere on $[0, 1]$. However, $f\big|_{Q \cap [0,1]}$ is constant and hence continuous on $Q \cap [0, 1]$.

Example 3.3. There exists a measurable function f on $[0, 1]$ such that for all $B \in \mathcal{M}$, where $B \subseteq A$ with $m(B) = 1$, we have $f\big|_B$ is not continuous. That is, f is not continuous a.e. on $[0, 1]$.

4 Construction of Lebesgue Integral

4.1 Integral of simple functions

Standard form of simple functions

Recall that simple functions have the form $\phi = \sum_{k=1}^{L} a_k \mathbb{1}_{E_k}$ where a_k are constant, $E_k \in \mathcal{M}$, and $m(E_k) < \infty$. We may always assume a_k are *distinct*, and E_k are *disjoint*, by renaming them and changing L as needed. Such a simple function is said to be in **standard form**.

Definition 4.1. Let $\phi = \sum_{k=1}^{L} a_k \mathbb{1}_{E_k}$. Then the Lebesgue integral of ϕ over $\mathbb{R}$ is

$$\int_{\mathbb{R}} \phi(x)dx := \sum_{k=1}^{L} a_k m(E_k)$$

If $A \in \mathcal{M}$, then $\mathbb{1}_A \cdot \phi = \sum_{k=1}^{L} a_k \mathbb{1}_{E_k \cap A}$ is again simple. Then the Lebesgue integral of ϕ over A is

$$\int_A \phi(x)dx := \int_{\mathbb{R}} (\mathbb{1}_A \cdot \phi)(x)dx = \sum_{k=1}^{L} a_k m(E_k \cap A)$$

Proposition 4.1. The following hold for the Lebesgue integral of simple functions

(1) If $\phi = \sum_{k=1}^{L} a_k \mathbb{1}_{E_k} = \sum_{l=1}^{M} b_l \mathbb{1}_{F_l}$, then

$$\sum_{k=1}^{L} a_k m(E_k) = \sum_{l=1}^{M} b_l m(F_l) \tag{Well-defined}$$

(2) If ϕ, ψ are simple, then for all $a, b \in \mathbb{R}$ we have $a\phi + b\psi$ is simple and

$$\int_{\mathbb{R}} a\phi + b\psi = a \int_{\mathbb{R}} \phi + b \int_{\mathbb{R}} \psi \tag{Linearity}$$

(3) If ϕ is simple, and $A, B \in \mathcal{M}$ with $A \cap B = \emptyset$, then

$$\int_{A \cup B} \phi = \int_A \phi + \int_B \psi \tag{Finite additivity}$$

(4) If ϕ, ψ are simple and $\phi \leq \psi$, then

$$\int_{\mathbb{R}} \phi \leq \int_{\mathbb{R}} \psi \tag{Monotonicity}$$

Monotonicity holds even if we only have $\phi \leq \psi$ a.e.

(5) If ϕ is simple, then

$$\left| \int_{\mathbb{R}} \phi \right| \leq \int_{\mathbb{R}} |\phi| \tag{Triangle inequality}$$

Proof.

(1) W.L.O.G. we may assume the E_k and F_l are each disjoint. Let $a_0 := 0 =: b_0$. Then set

$$E_0 := \left(\bigcup_{k=1}^{L} E_k \right)^c \text{ and } F_0 := \left(\bigcup_{l=1}^{M} F_l \right)^c$$

Then $\{E_0, \ldots, E_L\}$ and $\{F_0, \ldots, F_M\}$ each form a partition of $\mathbb{R}$. Observe that then

$$\mathbb{1}_{E_k} = \sum_{l=0}^{M} \mathbb{1}_{E_k \cap F_l} \text{ and } \mathbb{1}_{F_l} = \sum_{k=0}^{L} \mathbb{1}_{E_k \cap F_l}$$

So therefore

$$\phi = \sum_{k=1}^{L} a_k \mathbb{1}_{E_k} = \sum_{k=0}^{L} a_k \mathbb{1}_{E_k} = \sum_{k=0}^{L} \sum_{l=0}^{M} a_k \mathbb{1}_{E_k \cap F_l}$$

But also

$$\phi = \sum_{l=1}^{M} b_l \mathbb{1}_{F_l} = \sum_{l=0}^{M} b_l \mathbb{1}_{F_l} = \sum_{l=0}^{M} \sum_{k=0}^{L} b_l \mathbb{1}_{E_k \cap F_l}$$

By definition of integral on simple function, then

$$\int_{\mathbb{R}} \phi = \sum_{k=0}^{L} \sum_{l=0}^{M} a_k m(E_k \cap F_l) \text{ and } \int_{\mathbb{R}} \phi = \sum_{l=0}^{M} \sum_{k=0}^{L} b_l m(E_k \cap F_l)$$

Suppose $m(E_k \cap F_l) > 0$ for some k, l. Then we have $E_k \cap F_l \neq \emptyset$, and since $E_k \cap F_l$ are disjoint, then $a_k = b_l$. That is, the coefficients must agree over $E_k \cap F_l$ when it is non-empty. Note that when $k, l \geq 1$, we have $m(E_k \cap F_l) < \infty$ by definition of simple function. But we may have $m(E_0 \cap F_0) = \infty$, and in this case we use the convention $0 \cdot \infty = 0$. We conclude that

$$\sum_{k=1}^{L} a_k m(E_k) = \sum_{l=1}^{M} b_l m(F_l)$$

(4) Suppose $\phi = \sum_{k=1}^{L} a_k \mathbb{1}_{E_k}$ and $\psi = \sum_{l=1}^{M} b_l \mathbb{1}_{F_l}$. Again, assume W.L.O.G. that $\{E_k\}$ and $\{F_l\}$ are each disjoint, and form partitions $\{E_0, \ldots, E_L\}$, $\{F_0, \ldots, F_M\}$ of $\mathbb{R}$. Similarly, write

$$\phi = \sum_{k=1}^{L} a_k \mathbb{1}_{E_k} = \sum_{k=0}^{L} a_k \mathbb{1}_{E_k} = \sum_{k=0}^{L} \sum_{l=0}^{M} a_k \mathbb{1}_{E_k \cap F_l}$$

and

$$\psi = \sum_{l=1}^{M} b_l \mathbb{1}_{F_l} = \sum_{l=0}^{M} b_l \mathbb{1}_{F_l} = \sum_{l=0}^{M} \sum_{k=0}^{L} b_l \mathbb{1}_{E_k \cap F_l}$$

If $E_k \cap F_l \neq \emptyset$ then $a_k \leq b_k$ by the assumption that $\phi \leq \psi$. Using this fact and the definition of integral on simple function, then

$$\int_{\mathbb{R}} \phi = \sum_{k=0}^{L} \sum_{l=0}^{M} a_k m(E_k \cap F_l) \leq \sum_{l=0}^{M} \sum_{k=0}^{L} b_l m(E_k \cap F_l) = \int_{\mathbb{R}} \psi$$

■

4.2 Integral of non-negative functions

Definition 4.2. Let f be non-negative and measurable. Then

$$\int_{\mathbb{R}} f(x)dx := \sup \left\{ \int_{\mathbb{R}} \phi : \phi \text{ is simple and } \phi \leq f \right\}$$

Within the supremum, we are using the definition of integral on simple functions – call it (Def1). For now, refer to the above definition as (Def2).

Remark. Given a non-negative measurable function f, then (Def2) implies the existence of a sequence of simple functions $\{\phi_n\}$ such that $\phi_n \leq f$ and $\lim_{n\to\infty} \int_{\mathbb{R}} \phi_n = \int_{\mathbb{R}} f$. This is simply due to definition of supremum.

Proposition 4.2 (Consistency theorem)**.** For a non-negative simple function, the above definition (Def2) is equal to the previous definition of integral for simple functions (Def1).

Proof. Let ϕ be a simple function. Then $\int_{\mathbb{R}}^{(\text{Def1})} \phi \le \int_{\mathbb{R}}^{(\text{Def2})} \phi$ automatically, since $\int_{\mathbb{R}}^{(\text{Def1})} \phi$ is an element of the collection we are taking supremum over. To show $\int_{\mathbb{R}}^{(\text{Def1})} \phi \ge \int_{\mathbb{R}}^{(\text{Def2})} \phi$, it suffices to show that for all simple functions $\psi \le \phi$ we have $\int_{\mathbb{R}}^{(\text{Def1})} \psi \le \int_{\mathbb{R}}^{(\text{Def1})} \phi$. But this is exactly given by monotonicity.

∎

Theorem 4.1. Suppose f is non-negative and measurable. Let $\{\phi_n\}$ be a sequence of **increasing** simple functions such that $\lim_{n\to\infty} \phi_n = f$. Then

$$\lim_{n\to\infty} \int_{\mathbb{R}} \phi_n = \int_{\mathbb{R}} f$$

Proof. Let $\{\phi_n\}$ be a sequence of increasing simple functions that attains f in the limit. Since ϕ_n is increasing, monotonicity gives that $\int_{\mathbb{R}} \phi_n$ is increasing. Therefore $\lim_{n\to\infty} \int_{\mathbb{R}} \phi_n$ **exists**. Observing that $\phi_n \le f$ for all $n \in \mathbb{N}$, then by definition

$$\int_{\mathbb{R}}^{(\text{Def1})} \phi_n \le \int_{\mathbb{R}}^{(\text{Def2})} f \text{ for all } n \in \mathbb{N}$$

Therefore $\lim_{n\to\infty} \int_{\mathbb{R}}^{(\text{Def1})} \phi_n \le \int_{\mathbb{R}}^{(\text{Def2})} f$.

Now we must show $\lim_{n\to\infty} \int_{\mathbb{R}}^{(\text{Def1})} \phi_n \ge \int_{\mathbb{R}}^{(\text{Def2})} f$. It suffices to show that for all simple functions ψ with $\psi \le f$, we have

$$\int_{\mathbb{R}}^{(\text{Def1})} \psi \le \lim_{n\to\infty} \int_{\mathbb{R}}^{(\text{Def1})} \phi_n$$

Take such ψ, for which we can express as

$$\psi = \sum_{k=0}^{L} a_k \mathbb{1}_{E_k} \text{ where } \int_{\mathbb{R}}^{(\text{Def1})} \psi = \sum_{k=0}^{L} a_k m(E_k)$$

where $\{E_0, \ldots, E_L\}$ forms a partition of $\mathbb{R}$. Recall we are taking $a_0 = 0$ and using the convention $0 \cdot \infty = 0$. And while not needed for the proof, remember that $m(E_k) < \infty$ for all $k \ge 1$. By finite additivity then

$$\int_{\mathbb{R}} \phi_n = \int_{E_0 \cup \cdots \cup E_L} \phi_n = \sum_{k=0}^{L} \int_{E_k} \phi_n$$

Since $\lim_{n\to\infty} \int_{E_k} \phi_n$ exists for each $k \ge 0$ (as ϕ_n is increasing) then

$$\lim_{n\to\infty} \int_{\mathbb{R}} \phi_n = \sum_{k=0}^{L} \lim_{n\to\infty} \int_{E_k} \phi_n$$

It therefore suffices to show that for each $k \geq 0$ we have

$$\lim_{n\to\infty} \int_{E_k} \phi_n \geq a_k m(E_k)$$

W.L.O.G. fix some $k \geq 0$ and assume $a_k > 0$ and $m(E_k) > 0$, because otherwise the inequality trivially holds. Since $\lim_{n\to\infty} \phi_n = f \geq \psi$, then each $x \in E_k$ satisfies

$$\lim_{n\to\infty} \phi_n(x) \geq a_k$$

Let $\varepsilon > 0$ and define the auxiliary sequence

$$C_n^\varepsilon := \{x \in E_k : \phi_n(x) \geq (1-\varepsilon)a_k\}$$

Since $\phi_n \leq \phi_{n+1}$ then $C_n^\varepsilon \subseteq C_{n+1}^\varepsilon$, so the sequence C_n^ε is increasing. Furthermore

$$\bigcup_{n=1}^{\infty} C_n^\varepsilon = E_k$$

where "$\subseteq$" is obvious and "$\supseteq$" follows since given $x \in E_k$, we can take n arbitrarily large to satisfy $\phi_n(x) \geq (1-\varepsilon)a_k$. Now we compute (note that below we are using the definition of integral for step functions)

$$\begin{aligned}
\lim_{n\to\infty} \int_{E_k} \phi_n &= \lim_{n\to\infty} \int_{\mathbb{R}} \mathbb{1}_{E_k}\phi_n \\
&\geq \lim_{n\to\infty} \int_{\mathbb{R}} \mathbb{1}_{C_n^\varepsilon}\phi_n && (\text{as } \mathbb{1}_{E_k}\phi_n \geq \mathbb{1}_{C_n^\varepsilon}\phi_n) \\
&\geq \lim_{n\to\infty} \int_{\mathbb{R}} \mathbb{1}_{C_n^\varepsilon}(1-\varepsilon)a_k \\
&= \lim_{n\to\infty} (1-\varepsilon)a_k m(C_n^\varepsilon) \\
&= (1-\varepsilon)a_k \lim_{n\to\infty} m(C_n^\varepsilon) \\
&= (1-\varepsilon)a_k m(E_k) && (\text{by continuity from below})
\end{aligned}$$

Since $\varepsilon > 0$ is arbitrary, then $\lim_{n\to\infty} \int_{E_k} \phi_n \geq a_k m(E_k)$.

■

Remark. Note that the supremum definition of integral for non-negative measurable functions allows us to always *find some* sequence $\{\phi_n\}$ with $\lim_{n\to\infty} \int_{\mathbb{R}} \phi = \int_{\mathbb{R}} f$. This result, on the other hand, says that provided we have a sequence $\{\phi_n\}$ which **increases to f**, then $\lim_{n\to\infty} \int_{\mathbb{R}} \phi = \int_{\mathbb{R}} f$.

Remark. We use the fact that if $y \geq (1-\varepsilon)x$ for arbitrary $\varepsilon > 0$, then $y \geq x$. To see this, assume not. Then $y < x \implies (1-\varepsilon)y < (1-\varepsilon)x \leq y$, and $y < y$ is a contradiction.

Corollary 4.1.1. Let f be non-negative and measurable. $\forall n \geq 1$, and $\forall k = 0, 1, \ldots, 2^n$, set

$$A_{n,k} := \left\{\frac{k}{2^n} \leq f < \frac{k+1}{2^n}\right\}$$

Then

$$\int_{\mathbb{R}} f = \lim_{n\to\infty} \sum_{k=0}^{n2^n} \frac{k}{2^n} m(A_{n,k})$$

Proof. Set a sequence of simple functions $\{\phi_n\}$ as

$$\phi_n := \sum_{k=0}^{n2^n} \frac{k}{2^n} \mathbb{1}_{A_{n,k}}$$

Recall that this configuration has ϕ_n increasing with f as its limit. Then by the previous theorem, we have

$$\begin{aligned}
\int_{\mathbb{R}} f &= \lim_{n\to\infty} \int_{\mathbb{R}} \phi_n \\
&= \lim_{n\to\infty} \int_{\mathbb{R}} \sum_{k=0}^{n2^n} \frac{k}{2^n} \mathbb{1}_{A_{n,k}} \\
&= \lim_{n\to\infty} \sum_{k=0}^{n2^n} \frac{k}{2^n} \int_{\mathbb{R}} \mathbb{1}_{A_{n,k}} && \text{(by linearity)} \\
&= \lim_{n\to\infty} \sum_{k=0}^{n2^n} \frac{k}{2^n} m(A_{n,k}) && (A_{n,k} \in \mathcal{M})
\end{aligned}$$

■

Remark. This brings our discussion on approximation by simple functions to a full circle.

Proposition 4.3. Let $\phi = \sum_{k=0}^{L} a_k \mathbb{1}_{E_k}$ be a simple function. Let $A \in \mathcal{M}$ with $m(A) = 0$. Then $\int_A \phi = 0$.

Proof. We compute

$$\int_A \phi = \int_{\mathbb{R}} \mathbb{1}_A \phi = \int_{\mathbb{R}} \mathbb{1}_A \sum_{k=0}^{L} a_k \mathbb{1}_{E_k} = \int_{\mathbb{R}} \sum_{k=0}^{L} a_k \mathbb{1}_{E_k \cap A} = \sum_{k=0}^{L} a_k m(\mathbb{1}_{E_k \cap A}) = 0$$

■

Remark. Let us define $\int_A f := \int_{\mathbb{R}} f \mathbb{1}_A$ for a non-negative measurable function f, and a null set A. Then we also have $\int_A f = 0$. To see this, let $\{\phi_n\}$ be increasing with f as its limit. Then $\{\phi_n \mathbb{1}_A\}$ increases to $f \mathbb{1}_A$ and

$$\int_A f = \int_{\mathbb{R}} f \mathbb{1}_A = \lim_{n\to\infty} \int_{\mathbb{R}} \phi_n \mathbb{1}_A = \lim_{n\to\infty} \int_A \phi_n = 0$$

Corollary 4.1.2. Let f be non-negative and measurable. Let $A \in \mathcal{M}$ with $m(A) = 0$.

$$\int_{\mathbb{R}} f = \int_{\mathbb{R}} f \mathbb{1}_{\mathbb{R}\setminus A}$$

So we can change the function value on a null set while preserving the integral.

Proof. Let $\{\phi_n\}$ be an increasing sequence of simple functions such that $\lim_{n\to\infty} \phi_n = f$. Then we compute

$$\int_{\mathbb{R}} f = \lim_{n\to\infty} \int_{\mathbb{R}} \phi_n = \lim_{n\to\infty} \left(\int_{\mathbb{R}\setminus A} \phi_n + \int_A \phi_n \right) = \lim_{n\to\infty} \int_{\mathbb{R}\setminus A} \phi_n = \lim_{n\to\infty} \int_{\mathbb{R}} \phi_n \mathbb{1}_{\mathbb{R}\setminus A} = \int_{\mathbb{R}} f \mathbb{1}_{\mathbb{R}\setminus A}$$

The last equality follows since $\{\phi_n \mathbb{1}_{\mathbb{R}\setminus A}\}$ is an increasing sequence of simple functions which attains $f \mathbb{1}_{\mathbb{R}\setminus A}$ as its limit.

■

Proposition 4.4. The following are properties about the integral of non-negative measurable functions.

(1) Let f, g be non-negative and mesaurable such that $f = g$ a.e. Then

$$\int_{\mathbb{R}} f = \int_{\mathbb{R}} g \qquad \text{(Well-defined)}$$

(2) Let f, g be non-negative and measurable. Let $a, b \in \mathbb{R}$. Then

$$\int_{\mathbb{R}} (af + bg) = a \int_{\mathbb{R}} f + b \int_{\mathbb{R}} g \qquad \text{(Linearity)}$$

(3) Let f, g be non-negative and measurable. Suppose $f \leq g$ a.e. Then

$$\int_{\mathbb{R}} f \leq \int_{\mathbb{R}} g \qquad \text{(Monotonicity)}$$

Proof.

(1) Let $\{\phi_n\}$, $\{\psi_n\}$ be increasing sequences of simple functions that attain f and g as their limits, respectively. Define

$$h_n = \phi_n \mathbb{1}_{\{f=g\}} + \psi_n \mathbb{1}_{\{f\neq g\}}$$

Then $\lim_{n\to\infty} h_n = g$ pointwise. This then gives

$$\begin{aligned}
\int_{\mathbb{R}} g &= \lim_{n\to\infty} \int_{\mathbb{R}} h_n && (h_n \text{ is simple and increasing}) \\
&= \lim_{n\to\infty} \left(\int_{\{f=g\}} \phi_n + \int_{\{f\neq g\}} \psi_n \right) && (\text{by finite additivity}) \\
&= \lim_{n\to\infty} \int_{\{f=g\}} \phi_n && (\text{integral over zero measure set})
\end{aligned}$$

Meanwhile, we have

$$\begin{aligned}
\int_{\mathbb{R}} f &= \lim_{n\to\infty} \int_{\mathbb{R}} \phi_n \\
&= \lim_{n\to\infty} \left(\int_{\{f=g\}} \phi_n + \int_{\{f\neq g\}} \phi_n \right) \\
&= \lim_{n\to\infty} \int_{\{f=g\}} \phi_n && (\text{integral over zero measure set})
\end{aligned}$$

(2) Let $\{\phi_n\}$, $\{\psi_n\}$ be increasing sequences of simple functions that attain f and g as their limits, respectively. Define

$$h_n := a\phi_n + b\psi_n$$

which is simple and increasing, attaining $af + bg$ as its limit. We compute

$$\begin{aligned}
\int_{\mathbb{R}} af + bg &= \lim_{n\to\infty} \int_{\mathbb{R}} h_n \\
&= \lim_{n\to\infty} \int_{\mathbb{R}} a\phi_n + b\psi_n \\
&= \lim_{n\to\infty} \left(a \int_{\mathbb{R}} \phi_n + b \int_{\mathbb{R}} \psi_n \right) && (\text{linearity of simple functions}) \\
&= a \int_{\mathbb{R}} f + b \int_{\mathbb{R}} g
\end{aligned}$$

(3) W.L.O.G. we may assume $f \leq g$ pointwise. Otherwise, we can replace f by $f\mathbb{1}_{\{f\leq g\}}$ while preserving the same integral, since we are changing the function value only on a null set. This is a standard trick when dealing with integrals.

Then observe that

$$\left\{ \int_R \phi : \phi \text{ is simple and } \phi \leq f \right\} \subseteq \left\{ \int_{\mathbb{R}} \phi : \phi \text{ is simple and } \phi \leq g \right\}$$

Taking the supremum of both sides to get the integral, we have $\int_{\mathbb{R}} f \leq \int_{\mathbb{R}} g$.

■

Proposition 4.5. The following are further properties about the integral of non-negative measurable functions.

(i) Let f be non-negative and measurable. Then

$$\int_{\mathbb{R}} f = 0 \iff f = 0 \text{ a.e.}$$

(ii) Let f be non-negative and measurable, and let $A \in \mathcal{M}$ Then

$$\int_A f = 0 \iff m(A) = 0 \text{ or } f = 0 \text{ a.e.}$$

(iii) Let f be non-negative and measurable. If $\int_{\mathbb{R}} f < \infty$, then f is finite-valued a.e. That is,

$$m(\{f = \infty\}) = 0$$

Proof.

(i) To see " $\Longleftarrow$ ", assume $f = 0$ a.e. Then immediately $\int_{\mathbb{R}} f = 0$, since W.L.O.G. we can replace f by a function that is zero everywhere, while preserving the integral.

To see " $\Longrightarrow$ ", assume $\int_{\mathbb{R}} f = 0$. We want to show that if $A := \{f > 0\}$, then $m(A) = 0$. However, $\{f > 0\}$ by itself is hard to control. Let us form an auxiliary sequence $A_n := \{f \geq \frac{1}{n}\}$ such that $\{A_n\}$ is increasing and

$$A = \bigcup_{n=1}^{\infty} A_n$$

By continuity from below we have

$$\lim_{n\to\infty} m(A_n) = m\Big(\bigcup_{n=1}^{\infty} A_n\Big) = m(A)$$

Now assume $m(A) = \delta > 0$ for contradiction. Since $m(A_n)$ is increasing to its limit δ, then there exists $N \in \mathbb{N}$ such that $m(A_N) > \frac{\delta}{2}$. Thus $f \geq f\mathbb{1}_{A_N} \geq \frac{1}{N}\mathbb{1}_{A_N}$ since $f \geq \frac{1}{N}$ on A_N. By monotonicity we have

$$\int_{\mathbb{R}} f \geq \int_{\mathbb{R}} \frac{1}{N}\mathbb{1}_{A_N} = \frac{1}{N}m(A_N) \geq \frac{1}{N}\cdot\frac{\delta}{2} > 0$$

But this is a contradiction.

(ii) Assume $\int_A f = 0$, where $\int_A f = \int_{\mathbb{R}} f\mathbb{1}_A$. Then by (i) this implies $\mathbb{1}_A f = 0$ a.e. Thus $f = 0$ a.e. or $\mathbb{1}_A = 0$ a.e. If $\mathbb{1}_A = 0$ a.e., this implies $m(A) = 0$. For the converse, suppose $m(A) = 0$ or $f = 0$ a.e. If $f = 0$ a.e., then $\mathbb{1}_A f = 0$ a.e., so $\int_A f = 0$. Meanwhile, if $m(A) = 0$, then $\mathbb{1}_A = 0$ a.e., so $\mathbb{1}_A f = 0$ a.e.

(iii) Let $A = \{f = \infty\}$ and assume $m(A) = \delta > 0$ for contradiction. Observe that for any $n \in \mathbb{N}$

$$f \geq f\mathbb{1}_A \geq n\mathbb{1}_A$$

since we set all finite values to zero. Then by monotonicity and the fact that $n\mathbb{1}_A$ is simple, we have

$$\int_{\mathbb{R}} f \geq \int_{\mathbb{R}} n\mathbb{1}_A = n \cdot m(A) = n\delta$$

Since $n \in \mathbb{N}$ can be made arbitrarily large, it follows that $\int_{\mathbb{R}} f = \infty$. But this is a contradiction. The idea is that when $m(A) > 0$, we want to show $\int_{\mathbb{R}} f$ will be too large.

■

Remark. The converse of (iii) is definitely false. Take $f \equiv n$ for some $n \in \mathbb{N}$. So f is finite-valued everywhere. Let $L \in \mathbb{N}$ and set $A := [-L, L]$, which has $m(A) = 2L < \infty$. Then $n \geq n\mathbb{1}_A$, so by monotonicity

$$\int_{\mathbb{R}} n \geq \int_{\mathbb{R}} n\mathbb{1}_A = n \cdot m(A) = 2nL$$

Since L can be made arbitrarily large, then $\int_{\mathbb{R}} n = \infty$.

Proposition 4.6 (Markov's inequality)**.** Let f be non-negative and measurable. Let $0 < a < \infty$. Then

$$m(\{f > a\}) \leq \frac{1}{a}\int_{\mathbb{R}} f$$

In particular, if $\int_{\mathbb{R}} f < \infty$, then $\lim_{a\to\infty} m(\{f > a\}) = 0$. Especially, $m(\{f > a\})$ has $O(\frac{1}{a})$.

Proof. Let f be non-negative and measurable, and take $0 < a < \infty$. Now put

$$A_a := \{f > a\}$$

Observe that $f \geq f\mathbb{1}_{A_a} \geq a\mathbb{1}_{A_a}$. If $m(A_a) < \infty$, then we have

$$\int_{\mathbb{R}} f \geq \int_{\mathbb{R}} a\mathbb{1}_{A_a} = a \cdot m(A_a) \implies m(\{f > a\}) \leq \frac{1}{a}\int_{\mathbb{R}} f$$

Meanwhile, when $m(A_a) = \infty$, we cannot write "$\int_{\mathbb{R}} a\mathbb{1}_{A_a} = a \cdot m(A_a)$" anymore since $a\mathbb{1}_{A_a}$ is no longer a simple function. Instead, observe that for any $n \in \mathbb{N}$ we have

$$\int_{\mathbb{R}} f \geq \int_{\mathbb{R}} a\mathbb{1}_{A_a} \geq \int_{\mathbb{R}} a\mathbb{1}_{A_a\cap[-n,n]} = a \cdot m(A_a \cap [-n, n])$$

We claim that $m(A_a \cap [-n, n])$ can be made arbitrarily large. To see this, first put

$$A_a^n := A_a \cap [-n, n]$$

This defines an auxiliary sequence $\{A_a^n\}_{n\geq 1}$ which is increasing with

$$\bigcup_{n=1}^{\infty} A_a^n = \bigcup_{n=1}^{\infty} A_a \cap [-n, n] = A_a \cap \bigcup_{n=1}^{\infty}[-n, n] = A_a$$

By continuity from below, then

$$\lim_{n\to\infty} m(A_a^n) = m(A_a) = \infty$$

Thus indeed $m(A_a^n)$ can be made arbitrarily large. With this claim proven, then $\int_{\mathbb{R}} f = \infty$, so the inequality $\frac{1}{a}\int_{\mathbb{R}} f \geq m(\{f > a\})$ holds with $\infty \geq \infty$.

■

Remark. There are variations of Markov's inequality, such as

$$m(\{f > a\}) \leq \frac{1}{a^2}\int_{\mathbb{R}} f^2$$

We will discuss this more when we encounter L_p spaces.

4.3 Integral of general measurable function and notion of integrable function

Definition 4.3. Let f be a measurable function. Then we define

$$\int_{\mathbb{R}} f := \int_{\mathbb{R}} f^+ - \int_{\mathbb{R}} f^-$$

provided <u>at least one</u> of $\int_{\mathbb{R}} f^+$ or $\int_{\mathbb{R}} f^-$ is finite. We are using the definition of integral for *non-negative* functions to build a definition of integral for a *general* measurable function.

Remark. This definition is quite expected given that $f = f^+ - f^-$. However, only having $\int_{\mathbb{R}} f$ defined for a general measurable function f is not sufficient for desirable properties (linearity, monotonicity) to hold. In the spirit of how $\mathcal{M}$ contains the "good sets", we will now define "good functions" – i.e., those that are <u>integrable</u>.

Definition 4.4 (Integrable function)**.** Let f be a measurable function. Then f is integrable, denoted $f \in \mathcal{L}^1(\mathbb{R})$, if

$$\int_{\mathbb{R}} f^+ < \infty \text{ and } \int_{\mathbb{R}} f^- < \infty$$

Immediately,

$$f \in \mathcal{L}^1(\mathbb{R}) \iff \int_{\mathbb{R}} |f| < \infty$$

using the definition of integral for non-negative functions. This follows since linearity gives $\int_{\mathbb{R}} |f| = \int_{\mathbb{R}} f^+ + \int_{\mathbb{R}} f^-$. Furthermore,

$$f \in \mathcal{L}^1(\mathbb{R}) \iff \int_{\mathbb{R}} f \text{ is well-defined and finite valued}$$

using our definition of integral for general measurable functions (hence why we need to specify that it is well-defined). This follows immediately since for $\int_{\mathbb{R}} f$ to be well-defined and finite valued is equivalent to having both $\int_{\mathbb{R}} f^+$ and $\int_{\mathbb{R}} f^-$ to be finite.

Proposition 4.7. The following are properties about the integral of **integrable** functions.

(1) Let $f \in \mathcal{L}^1(\mathbb{R})$. Then

$$\left|\int_{\mathbb{R}} f\right| \leq \int_{\mathbb{R}} |f| \qquad \text{(Triangle inequality)}$$

(2) Let $f \in \mathcal{L}^1(\mathbb{R})$. Then

$$f \text{ is finite valued a.e.}$$

(3) Let $f, g \in \mathcal{L}^1(\mathbb{R})$, and let $a, b \in \mathbb{R}$. Then $af + bg \in \mathcal{L}^1(\mathbb{R})$ and

$$\int_{\mathbb{R}} af + bg = a\int_{\mathbb{R}} f + b\int_{\mathbb{R}} g \qquad \text{(Linearity)}$$

(4) Let $f \in \mathcal{L}^1(\mathbb{R})$, and let $A \in \mathcal{M}$ with $m(A) = 0$. Then

$$\int_A f = 0$$

(5) Let $f, g \in \mathcal{L}^1(\mathbb{R})$. If $f \leq g$ a.e., then

$$\int_{\mathbb{R}} f \leq \int_{\mathbb{R}} g \qquad \text{(Monotonicity)}$$

Remark. These do not necessarily hold for general measurable functions whose integral is well-defined.

Theorem 4.2 (Monotone convergence theorem (MON))**.** Let $\{f_n\}, f$ be non-negative and measurable. Suppose f_n is increasing with $\lim_{n\to\infty} f_n = f$ a.e. Then

$$\int_{\mathbb{R}} f = \lim_{n\to\infty} \int_{\mathbb{R}} f_n$$

Corollary 4.2.1. Let $\{f_n\}, f$ be measurable. Suppose f_n is increasing with $\lim_{n\to\infty} f_n = f$. Furthermore, assume $\int_{\mathbb{R}} f_1^- < \infty$. Then

$$\int_{\mathbb{R}} f = \lim_{n\to\infty} \int_{\mathbb{R}} f_n$$

Theorem 4.3 (Reverse MON)**.** Let $\{f_n\}, f$ be measurable. Suppose f_n is decreasing with $\lim_{n\to\infty} f_n = f$. Furthermore, assume $\int_{\mathbb{R}} f_1^+ < \infty$. Then

$$\int_{\mathbb{R}} f = \lim_{n\to\infty} \int_{\mathbb{R}} f_n$$

Theorem 4.4 (Fatou's Lemma)**.** Let $\{f_n\}$ be non-negative and measurable. Then

$$\int_{\mathbb{R}} \liminf_{n\to\infty} f_n \leq \liminf_{n\to\infty} \int_{\mathbb{R}} f_n$$

Corollary 4.4.1. Let $\{f_n\}$ be measurable. Suppose there exists a measurable function g such that $\int_{\mathbb{R}} g^- < \infty$ and $f_n \geq g$ for all $n \geq 1$. Then

$$\int_{\mathbb{R}} \liminf_{n\to\infty} f_n = \liminf_{n\to\infty} \int_{\mathbb{R}} f_n$$

Theorem 4.5 (Reverse Fatou's Lemma)**.** Let $\{f_n\}$ be measurable. Suppose there exists a measurable function g such that $\int_{\mathbb{R}} g^+ < \infty$ and $f_n \leq g$ for all $n \geq 1$. Then

$$\int_{\mathbb{R}} \limsup_{n\to\infty} f_n \geq \limsup \int_{\mathbb{R}} f_n$$

Theorem 4.6 (Dominated convergence theorem (DOM))**.** Let $\{f_n\}, f$ be measurable and suppose $f_n \to f$ a.e. Suppose there exists a dominating function $g \in \mathcal{L}^1(\mathbb{R})$ such that $|f_n| \leq |g|$ for all $n \geq 1$. Then $f_n \to f$ in $\mathcal{L}^1(\mathbb{R})$, that is

$$\lim_{n\to\infty} \int_{\mathbb{R}} |f_n - f| = 0$$

In particular, $\int_{\mathbb{R}} f = \lim_{n\to\infty} \int_{\mathbb{R}} f_n$.

5 Product spaces

Definition 5.1. Given a measure space $(X, \mathcal{F}, \mu)$, recall that μ is σ-finite if $\exists \{X_n : n \geq 1\} \subseteq \mathcal{F}$ such that $X_n \uparrow$ and $\bigcup_{n=1}^{\infty} X_n = X$ with $\mu(X_n) < \infty$ for all $n \geq 1$.

Definition 5.2. Let $f : X \to \overline{\mathbb{R}}$. THen f is $\mathcal{F}$-measurable if $\forall A \in \mathbb{R}$ we have $f^{-1}([-\infty, a)) \in \mathcal{F}$.

5.1 Product σ-algebra

Definition 5.3. Consider the product space $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$. Then the product σ-algebra of subsetes of $\mathbb{R}^2$, denoted $\mathcal{M} \otimes \mathcal{M}$ or $\mathcal{M}^2$, is defined

$$\mathcal{M} \otimes \mathcal{M} := \sigma(\{A \times B : A, B \in \mathcal{M}\})$$

Notably, $\mathcal{B}(\mathbb{R}^2) = \sigma(\{\text{open sets in } \mathbb{R}^2\}) \subseteq \mathcal{M}^2$

Definition 5.4 (Slice of set)**.** Let $E \subseteq \mathbb{R}^2$. Then for all $x \in \mathbb{R}$

$$E_x := \{y \in \mathbb{R} : (x, y) \in E\}$$

is the slice of E at x.

Proposition 5.1. Let $E \in \mathcal{M}^2$. Then $E_x \in \mathcal{M}$ for all $x \in \mathbb{R}$ and $E_y \in \mathcal{M}$ for all $y \in \mathbb{R}$.

Remark. That is, **product measurability implies marginal measurability.**