MATH 271
REAL VARIABLES

Fall 2013
Instructor: Stefaan Delcroix
Chapter 1

Preliminary Material

In this chapter, we review some material that will be used to develop measure theory. Note that any material from Math 171 or Math 172 may be needed.

1.1 Unions, Intersections and Complements

Let $X$ be a fixed set.

We denote by $\mathcal{P}(X)$ the set of all subsets of $X$. $\mathcal{P}(X)$ is sometimes called the power set of $X$. All sets in this section will be subsets of $X$.

Let $A, B \subseteq X$. The intersection of $A$ and $B$ (notation: $A \cap B$) and the union of $A$ and $B$ (notation: $A \cup B$) are defined as

$$A \cap B = \{x \in X : x \in A \text{ and } x \in B\}$$
$$A \cup B = \{x \in X : x \in A \text{ or } x \in B\}$$

We can generalize this to an arbitrary family of subsets. Let $I$ be an index set and $A_i \subseteq X$ for all $i \in I$. Then

$$\cap_{i \in I} A_i = \{x \in X : x \in A_i \text{ for all } i \in I\}$$
$$\cup_{i \in I} A_i = \{x \in X : x \in A_i \text{ for some } i \in I\}$$

The Distributive Laws give relations between intersections and unions:

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$
$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

For arbitrary collections, we get

$$A \cap (\cup_{i \in I} B_i) = \cup_{i \in I} (A \cap B_i)$$
$$A \cup (\cap_{i \in I} B_i) = \cap_{i \in I} (A \cup B_i)$$

Let $A \subseteq X$. The complement of $A$ (in $X$) (notation: $\tilde{A}$) is defined as

$$\tilde{A} = \{x \in X : x \notin A\}$$

We easily get that

$$\tilde{\emptyset} = X, \ \tilde{X} = \emptyset \text{ and } \tilde{\tilde{A}} = A$$
Let \( A, B \subseteq X \). Then the set difference \( A \setminus B \) (also called the relative complement of \( B \) in \( A \)) is defined as
\[
A \setminus B = A \cap \overline{B} = \{ x \in X : x \in A \text{ and } x \notin B \}
\]

*De Morgan’s Laws* give relations between complements and unions or intersections:
\[
\overline{A \cap B} = \overline{A} \cup \overline{B} \quad \text{and} \quad \overline{A \cup B} = \overline{A} \cap \overline{B}
\]

For arbitrary collections, we get
\[
\bigcap_{i \in I} \overline{A_i} = \overline{\bigcup_{i \in I} A_i} \quad \text{and} \quad \bigcup_{i \in I} \overline{A_i} = \overline{\bigcap_{i \in I} A_i}
\]

A collection of sets \( \{ A_i : i \in I \} \) is called a disjoint collection if the sets are pairwise disjoint: \( A_i \cap A_j = \emptyset \) for all \( i, j \in I \) with \( i \neq j \).

### 1.2 Functions

Let \( X \) and \( Y \) be sets.

We denote a function \( f \) with domain \( X \) and function values in \( Y \) by \( f : X \to Y \).

For \( A \subseteq X \), we define \( f(A) = \{ f(a) : a \in A \} \). So we can extend \( f \) to a function \( f : \mathcal{P}(X) \to \mathcal{P}(Y) \).

For \( B \subseteq Y \), we define \( f^{-1}(B) = \{ x \in X : f(x) \in B \} \). So \( f^{-1} : \mathcal{P}(Y) \to \mathcal{P}(X) \). If \( f \) is one-to-one and \( Y \) is the range of \( f \) (so \( f \) is onto), then \( f^{-1} \) sends singletons to singletons and we have the standard definition of \( f^{-1} \).

The following properties hold (where \( A_i \subseteq X \) and \( B_i \subseteq Y \) for all \( i \in I \) and \( B \subseteq Y \)):
\[
\begin{align*}
f \left( \bigcup_{i \in I} A_i \right) &= \bigcup_{i \in I} f(A_i) \\
f \left( \bigcap_{i \in I} A_i \right) &\subseteq \bigcap_{i \in I} f(A_i) \\
f^{-1} \left( \bigcup_{i \in I} B_i \right) &= \bigcup_{i \in I} f^{-1}(B_i) \\
f^{-1} \left( \bigcap_{i \in I} B_i \right) &= \bigcap_{i \in I} f^{-1}(B_i) \\
f^{-1} \left( \overline{B} \right) &= \overline{f^{-1}(B)}
\end{align*}
\]

Note that in the last property, the complements are taken with respect to different sets (\( Y \) on the left and \( X \) on the right).

### 1.3 Countable and Uncountable sets

**Definition 1.1**

(a) A set \( S \) is countable if either \( S \) is finite or if there exists a bijection between \( \mathbb{N} \) and \( S \).

(b) A set \( S \) is uncountable if \( S \) is not countable.
An infinite set $S$ is countable if we can enumerate the set:

$$S = s_1, s_2, s_3, \ldots$$

where every element of $S$ shows up exactly once in the enumeration.

**Example:** $\mathbb{Z}$ is countable since

$$0, 1, -1, 2, -2, 3, -3, 4, -4, \ldots$$

is an enumeration of $\mathbb{Z}$. □

Note that we can relax the bijection requirement under certain circumstances (every element of $S$ must show up at least once in the enumeration but not exactly once).

**Lemma 1.2** Let $S$ be a non-empty set. Then $S$ is countable if and only if there exists an onto map $\theta : \mathbb{N} \to S$.

**Proof:** Suppose first that $S$ is countable. If $S$ is finite, say $S = \{s_1, \ldots, s_n\}$, then the map

$$\theta : \mathbb{N} \to S : k \rightarrow \begin{cases} k & \text{if } k < n \\ n & \text{if } k \geq n \end{cases}$$

is onto. If $S$ is infinite, there exists a bijection between $S$ and $\mathbb{N}$ and hence there exists an onto map $\theta : \mathbb{N} \to S$.

Suppose next that there exists an onto map $\theta : \mathbb{N} \to S$. If $S$ is finite then clearly $S$ is countable. So assume that $S$ is infinite. We define a map $\varphi : \mathbb{N} \to S$ as follows:

- $\varphi(1) = \theta(1)$
- For $n \geq 1$, we put $\varphi(n+1) = \theta(m)$ where $m$ is the smallest value with $\theta(m) \notin \{\varphi(1), \ldots, \varphi(n)\}$.

Note that $\varphi$ is well-defined since $S$ is infinite. One easily proves that $\varphi$ is a bijection. Hence $S$ is countable. □

**Remark:** In Lemma 1.2, we can replace $\mathbb{N}$ be any infinite countable set. □

**Lemma 1.3** Let $S$ be a non-empty set. Then $S$ is countable if and only if there exists a one-to-one map $\theta : S \to \mathbb{N}$.

**Proof:** Suppose first that $S$ is countable. If $S$ is finite, say $S = \{s_1, \ldots, s_n\}$, then the map

$$\theta : S \to \mathbb{N} : s_k \rightarrow k$$

is one-to-one. If $S$ is infinite, there exists a bijection between $S$ and $\mathbb{N}$ and hence there exists a one-to-one map $\theta : S \to \mathbb{N}$.

Suppose next that there exists a one-to-one map $\theta : S \to \mathbb{N}$. Fix $x \in S$. We define a map $\varphi : \mathbb{N} \to S$ as follows:

- $\varphi(n) = x$ if $\theta^{-1}(n) = \emptyset$
- $\varphi(n) = \theta^{-1}(n)$ if $\theta^{-1}(n) \neq \emptyset$
Note that $\varphi$ is well-defined since $\theta$ is one-to-one. One easily proves that $\varphi$ is onto. Hence $S$ is countable by Lemma 1.2.

**Remark:** In Lemma 1.3, we can replace $\mathbb{N}$ be any infinite countable set. ▷

**Corollary 1.4** A subset of a countable set is countable.

**Proof:** Let $\emptyset \neq A \subseteq B$ with $B$ countable. By Lemma 1.3, there exists a one-to-one map $\theta : B \to \mathbb{N}$. Then the restriction of $\theta$ to $A$ is still one-to-one. Hence $A$ is countable by Lemma 1.3. □

**Proposition 1.5** The cartesian product of a finite number of countable sets is countable.

**Proof:** First, we prove that $\mathbb{N} \times \mathbb{N}$ is countable. We can list the elements of $\mathbb{N} \times \mathbb{N}$ in an infinite table:

$$(1,1) \quad (1,2) \quad (1,3) \quad (1,4) \quad \cdots$$

$$ (2,1) \quad (2,2) \quad (2,3) \quad (2,4) \quad \cdots$$

$$ (3,1) \quad (3,2) \quad (3,3) \quad (3,4) \quad \cdots$$

$$ (4,1) \quad (4,2) \quad (4,3) \quad (4,4) \quad \cdots$$

$$ \vdots \quad \vdots \quad \vdots \quad \vdots \quad \ddots$$

It is pretty clear that we can enumerate all the elements in this table. Hence $\mathbb{N} \times \mathbb{N}$ is countable.

Next, we prove the proposition. Let $A_1, \ldots, A_n$ be non-empty countable sets. We prove by induction on $n$ that $A_1 \times \cdots \times A_n$ is countable. This is obvious for $n = 1$. So let $n \geq 2$. Then $A = A_1 \times \cdots \times A_{n-1}$ is non-empty and countable (this is obvious for $n = 2$; use induction for $n \geq 3$). By Lemma 1.2, there exist onto maps $\theta : \mathbb{N} \to A$ and $\varphi : \mathbb{N} \to A_n$. Hence the map

$$\psi : \mathbb{N} \times \mathbb{N} \to A \times A_n : (x,y) \mapsto (\theta(x), \varphi(y))$$

is onto. Since $\mathbb{N} \times \mathbb{N}$ is infinite and countable, there exists a bijection $\pi : \mathbb{N} \to \mathbb{N} \times \mathbb{N}$. Then $\psi \circ \pi : \mathbb{N} \to A \times A_n$ is onto. So $A_1 \times \cdots \times A_n = A \times A_n$ is countable by Lemma 1.2. □

**Corollary 1.6** $\mathbb{Q}$ is countable.

**Proof:** The map $\theta : \mathbb{Z} \times \mathbb{N} \to \mathbb{Q} : (m,n) \mapsto \frac{m}{n}$ is clearly onto. Since $\mathbb{Z} \times \mathbb{N}$ is countable by Proposition 1.5, we get that $\mathbb{Q}$ is countable by Lemma 1.2. □

**Corollary 1.7** The countable union of countable sets is countable.

**Proof:** Let $I$ be a non-empty countable index set and let $A_i$ be a non-empty countable set for all $i \in I$. By Lemma 1.2, there exists an onto map $\theta : \mathbb{N} \to I$ and for each $i \in I$, there exists an onto map $\varphi_i : \mathbb{N} \to A_i$. Then the map

$$\psi : \mathbb{N} \times \mathbb{N} \to \bigcup_{i \in I} A_i : (m,n) \mapsto \varphi_{\theta(m)}(n)$$

is onto. Since $\mathbb{N} \times \mathbb{N}$ is infinite and countable by Proposition 1.5, we get that $\bigcup_{i \in I} A_i$ is countable by Lemma 1.2. □

Note that there exist uncountable sets. It is fairly easy to prove that $\mathcal{P}(\mathbb{N})$ is uncountable. Later, we will show (using measure theory) that any non-degenerate interval of real numbers is uncountable. Using Cantor’s Diagonalization process, one can show that the set of all infinite sequences from $\{0,1\}$ is uncountable. This result can in turn be used to prove that the interval $[0,1]$ is uncountable.
1.4 Axiom of Choice

In this section, we mention the Axiom of Choice. This axiom is really an axiom (it seems like it is ‘always true’ but there are some equivalent forms that are not so ‘obviously true’).

**Axiom of Choice:** Let \( I \) be an index set and \( A_i \) a non-empty set for all \( i \in I \). Then there exists a function \( f : I \to \bigcup_{i \in I} A_i \) with \( f(i) \in A_i \) for all \( i \in I \).

Another formulation of the Axiom of Choice is:

Let \( I \) be an index set and \( \{A_i : i \in I\} \) a disjoint collection of non-empty sets. Then there exists a set containing exactly one element of \( A_i \) for each \( i \in I \).

In this class, we accept the Axiom of Choice. Note that we do not need the Axiom of Choice if \( I \) is countable.

1.5 The Real Numbers

1.5.1 Axioms for the Real Numbers

The real numbers (notation: \( \mathbb{R} \)) are a complete, ordered field. They satisfy the field axioms, the order axioms and the completeness axiom.

- **Field Axioms**

  A set \( F \) with two binary operations \(+\) and \( \cdot \) is a field if

  - F1. \( \forall x, y \in F : x + y \in F \)
  - F2. \( \forall x, y \in F : x + y = y + x \)
  - F3. \( \forall x, y, z \in F : (x + y) + z = x + (y + z) \)
  - F4. \( \exists 0 \in F : \forall x \in F : x + 0 = x \)
  - F5. \( \forall x \in F : \exists y \in F : x + y = 0 \)
  - F6. \( \forall x, y \in F : xy \in F \)
  - F7. \( \forall x, y \in F : xy = yx \)
  - F8. \( \forall x, y, z \in F : (xy)z = x(yz) \)
  - F9. \( \exists 1 \in F : \forall x \in F : x \cdot 1 = x \)
  - F10. \( \forall x \in F \setminus \{0\} : \exists z \in F : xz = 1 \)
  - F11. \( \forall x, y, z \in F : x(y + z) = xy + xz \)

  The elements 0 (from Axiom F4)) and 1 (from (Axiom F9) are unique. The element \( y \) (from Axiom F5) is unique and is denoted by \(-x\). The element \( z \) (from Axiom F10) is unique and is denoted by \( x^{-1} \) or \( \frac{1}{x} \).

  Note that there are plenty of fields: \( \mathbb{Q}, \mathbb{R}, \mathbb{C}, \mathbb{Z}_p \) with \( p \) prime, etc.

- **Order Axioms**

  Let \((F, +, \cdot)\) be a field. Then \( F \) is an ordered field if there exists a subset \( P \) of \( F \) satisfying:
O1. \( \forall x, y \in P : x + y \in P \)
O2. \( \forall x, y \in P : xy \in P \)
O3. \( \forall x \in P : -x \notin P \)
O4. \( \forall x \in \mathbb{F} : x = 0 \) or \( x \in P \) or \( -x \in P \)

It is easy to see that \( 0 \notin P \) and \( 1 \in P \). This implies that an ordered field has characteristic 0: \( 1 + 1 + \cdots + 1 \) is never 0. Hence an ordered field contains isomorphic copies of \( \mathbb{N}, \mathbb{Z} \) and \( \mathbb{Q} \). \( \mathbb{R} \) and \( \mathbb{Q} \) are examples of ordered fields. The set \( P \) is the set of all positive numbers (under the usual order of real numbers). Note that \( \mathbb{C} \) is not an ordered field.

In an ordered field, we can introduce an order: we define \( x < y \) if \( y - x \in P \); we define \( x \leq y \) if \( x = y \) or if \( x < y \).

- **Completeness Axiom**

  Let \( \mathbb{F} \) be an ordered field. Let \( S \) be a subset of \( \mathbb{F} \). An element \( m \in \mathbb{F} \) is an upper bound for \( S \) if \( s \leq m \) for all \( s \in S \). An element \( m \in \mathbb{F} \) is a least upper bound for \( S \) if \( m \) is an upper bound for \( S \) and if \( m \leq b \) for any upper bound \( b \) for \( S \).

  Note that a set \( S \) has at most one least upper bound.

  A least upper bound for \( S \) is also called the supremum of \( S \) (notation: \( \text{sup} \ S \)).

  Now we can state the Completeness Axiom:

  An ordered field \( \mathbb{F} \) is complete if

  C1. Every non-empty subset of \( \mathbb{F} \) which has an upper bound has a least upper bound.

  It turns out that the real numbers are a complete, ordered field. This is not at all obvious (does there exist a complete, ordered field). Standard proofs use either Dedekind cuts or equivalence classes of rational Cauchy sequences.

  Another fact that is not obvious: any complete ordered field is isomorphic to the real numbers.

### 1.5.2 Archimedean Property and Density of the Rationals

The real numbers have the **Archimedean Property** (also called the Axiom of Eudoxus).

**Proposition 1.8 (Archimedes)** Let \( x \in \mathbb{R} \). Then there exists \( n \in \mathbb{Z} \) with \( x < n \).

**Proof:** If \( x < 0 \) we can put \( n = 0 \). So we may assume that \( x \geq 0 \). Put \( S = \{ k \in \mathbb{Z} : k \leq x \} \). Note that \( S \) is a non-empty subset of \( \mathbb{R} \) (since \( 0 \in S \)) and is bounded above (by \( x \)). By the Completeness Axiom, \( S \) has a supremum, say \( p = \text{sup} \ S \). Since \( p - \frac{1}{2} < p \), it follows that \( p - \frac{1}{2} \) is not an upper bound for \( S \). Hence \( q > p - \frac{1}{2} \) for some \( q \in S \). Then \( q + 1 > p - \frac{1}{2} + 1 = p + \frac{1}{2} > p \). So \( q + 1 \notin S \) since \( p = \text{sup} \ S \). Since \( q + 1 \in \mathbb{Z} \), it must be that \( q + 1 > x \). \( \Box \)

The Archimedean Property allows us to prove that the rational numbers are dense in the real numbers (that is, between any two real numbers is a rational number).
Proposition 1.9 Between any two real numbers there exists a rational number.

Proof: Let \(x, y \in \mathbb{R}\) with \(x < y\). Since \(y - x > 0\), it follows from the Archimedean Property that there exists \(n \in \mathbb{N}\) with \(n > \frac{1}{y - x}\). Hence \(ny - nx > 1\). Consider the set \(S = \{s \in \mathbb{Z} : s \geq yn\}\). Note that \(S\) is not empty by the Archimedean Property. Let \(m\) be the smallest element in \(S\) (note that \(m\) exists since \(S\) is bounded below). Then \(m - 1 \notin S\) and so \(m - 1 < yn\). Hence
\[
na = ny - (ny - nx) < ny - 1 \leq m - 1
\]
We get that \(nx < m - 1 < yn\) and so \(x < \frac{m - 1}{n} < y\). \(\square\)

1.6 Supremum and Infimum of a Set of Real Numbers

Definition 1.10 Let \(S\) be a non-empty set of real numbers.

(a) The supremum of \(S\) (notation: \(\sup S\)) is the least upper bound for \(S\). So \(\sup S \in \mathbb{R} \cup \{+\infty\}\).

(b) The infimum of \(S\) (notation: \(\inf S\)) is the greatest lower bound for \(S\). So \(\inf S \in \{-\infty\} \cup \mathbb{R}\).

These definitions are equivalent to the following:

- \(\sup S = \alpha \in \mathbb{R} \iff \begin{cases} (1) \forall s \in S : s \leq \alpha \\ (2) \forall \varepsilon > 0 : \exists s \in S : \alpha - \varepsilon < s \end{cases}\\ 
- \(\sup S = +\infty \iff \forall M \in \mathbb{R} : \exists s \in S : M \leq s \)
- \(\inf S = \alpha \in \mathbb{R} \iff \begin{cases} (1) \forall s \in S : \alpha \leq s \\ (2) \forall \varepsilon > 0 : \exists s \in S : s < \alpha + \varepsilon \end{cases}\\ 
- \(\inf S = -\infty \iff \forall M \in \mathbb{R} : \exists s \in S : s \leq M \)

We mention one property of suprema/infima that might not be obvious.

Proposition 1.11 Let \(S\) be a non-empty set of real numbers. Put \(-S = \{-s : s \in S\}\). Then \(\inf(-S) = -\sup S\) and \(\sup(-S) = -\inf S\).

Proof: We show that \(\inf(-S) = -\sup S\). Then \(\sup(-S) = -\inf(-(-S)) = -\inf S\).

Put \(\sup S = \alpha\). Suppose first that \(\alpha = +\infty\). Let \(M \in \mathbb{R}\). Then \(-M \leq s\) for some \(s \in S\). Hence \(-s \leq M\). So \(\inf(-S) = -\infty\). Suppose next that \(\alpha \in \mathbb{R}\). Then \(s \leq \alpha\) for all \(s \in S\). Thus \(-\alpha \leq -s\) for all \(s \in S\). Let \(\varepsilon > 0\). Then \(\alpha - \varepsilon < s\) for some \(s \in S\). Hence \(-s < -\alpha + \varepsilon\). So \(\inf(-S) = -\alpha = -\sup S\). \(\square\)
1.7 Sequences of Real Numbers

1.7.1 Limit of a Sequence

Definition 1.12 Let \( \{a_n\}_{n \geq 1} \) be a sequence of real numbers.

(a) The sequence \( \{a_n\}_{n \geq 1} \) converges to \( \alpha \in \mathbb{R} \) (notation: \( \lim_{n \to \infty} a_n = \alpha \)) if
\[
\forall \varepsilon > 0 : \exists N \in \mathbb{N} : \forall n \geq N : |a_n - \alpha| < \varepsilon
\]

(b) The sequence \( \{a_n\}_{n \geq 1} \) is a Cauchy sequence if
\[
\forall \varepsilon > 0 : \exists N \in \mathbb{N} : \forall m, n \geq N : |a_n - a_m| < \varepsilon
\]

Over the real numbers, these concepts are equivalent:

Theorem 1.13 A sequence of real numbers converges to a real number if and only if it is a Cauchy sequence.

We also want to deal with sequences converging to \( +\infty \) or \( -\infty \).

Definition 1.14 Let \( \{a_n\}_{n \geq 1} \) be a sequence of real numbers.

(a) The sequence \( \{a_n\}_{n \geq 1} \) converges to \( +\infty \) (notation: \( \lim_{n \to \infty} a_n = +\infty \)) if
\[
\forall M \in \mathbb{R} : \exists N \in \mathbb{N} : \forall n \geq N : a_n > M
\]

(b) The sequence \( \{a_n\}_{n \geq 1} \) converges to \( -\infty \) (notation: \( \lim_{n \to \infty} a_n = -\infty \)) if
\[
\forall M \in \mathbb{R} : \exists N \in \mathbb{N} : \forall n \geq N : a_n < M
\]

Definition 1.15 The extended real numbers (notation: \( \mathbb{R} \)) is the set \( \mathbb{R} \cup \{-\infty, +\infty\} \).

We mention some important properties of limits:

Proposition 1.16 Let \( \{a_n\}_{n \geq 1} \) and \( \{b_n\}_{n \geq 1} \) be sequences of real numbers that converge to extended real numbers. Then the following holds:

1. If \( a_n \leq b_n \) for all \( n \) large enough then \( \lim_{n \to \infty} a_n \leq \lim_{n \to \infty} b_n \).

2. If \( \alpha, \beta \in \mathbb{R} \) then \( \lim_{n \to \infty} (\alpha a_n + \beta b_n) = \alpha \lim_{n \to \infty} a_n + \beta \lim_{n \to \infty} b_n \) provided the right-hand side is defined (note that \( 0 \cdot (\pm \infty) = 0 \) here; undefined would be \((+\infty) + (-\infty))\).
1.7.2 Limit Superior and Limit Inferior of a Sequence

Not every sequence of extended real numbers has a limit. But every sequence of extended real numbers has a limit superior and a limit inferior.

Definition 1.17 Let \( \{x_n\}_{n \geq 1} \) be a sequence of extended real numbers.

(a) The limit inferior of the sequence \( \{x_n\}_{n \geq 1} \) (notation: \( \underline{\lim} \ x_n \)) is the extended real number
\[
\inf_{n \geq 1} \sup_{k \geq n} x_k.
\]

(b) The limit superior of the sequence \( \{x_n\}_{n \geq 1} \) (notation: \( \overline{\lim} \ x_n \)) is the extended real number
\[
\sup_{n \geq 1} \inf_{k \geq n} x_k.
\]

Remark: Let \( \{x_n\}_{n \geq 1} \) be a sequence of extended real numbers. For \( n \geq 1 \), put
\[
A_n = \inf_{k \geq n} x_k = \inf \{x_n, x_{n+1}, \ldots\} \quad \text{and} \quad B_n = \sup_{k \geq n} x_k = \sup \{x_n, x_{n+1}, \ldots\}
\]
Then \( \{A_n\}_{n \geq 1} \) is an increasing sequence of extended real numbers. So \( \{A_{n \geq 1}\} \) converges to \( \sup_{n \geq 1} A_n \).
Hence
\[
\underline{\lim} \ x_n = \sup_{n \geq 1} A_n = \lim_{n \to \infty} A_n
\]
Similarly, \( \{B_n\}_{n \geq 1} \) is a decreasing sequence of extended real numbers. So \( \{B_{n \geq 1}\} \) converges to \( \inf_{n \geq 1} B_n \).
Hence
\[
\overline{\lim} \ x_n = \inf_{n \geq 1} B_n = \lim_{n \to \infty} B_n
\]

There are equivalent forms of these definitions, depending on whether the limit superior/inferior is finite or infinite.

Proposition 1.18 Let \( \{x_n\}_{n \geq 1} \) be a sequence of extended real numbers. Then the following holds:

(a) \( \underline{\lim} x_n = +\infty \iff \forall M \in \mathbb{R} : \exists n \in \mathbb{N} : \forall k \geq n : x_k \geq M \)

(b) \( \underline{\lim} x_n = -\infty \iff \forall M \in \mathbb{R} : \exists n \in \mathbb{N} : \forall k \geq n : x_k \leq M \)

(c) \( \underline{\lim} x_n := \alpha \in \mathbb{R} \iff \exists \varepsilon > 0 : \forall n \in \mathbb{N} : \forall k \geq n : \alpha - \varepsilon \leq x_k \)

(d) \( \overline{\lim} x_n = -\infty \iff \forall M \in \mathbb{R} : \exists n \in \mathbb{N} : \forall k \geq n : x_k \leq M \)

(e) \( \overline{\lim} x_n = +\infty \iff \forall M \in \mathbb{R} : \exists n \in \mathbb{N} : \forall k \geq n : x_k \geq M \)

(f) \( \overline{\lim} x_n := \alpha \in \mathbb{R} \iff \exists \varepsilon > 0 : \forall n \in \mathbb{N} : \forall k \geq n : \alpha + \varepsilon \geq x_k \)
Proof: We prove the statements about the limit inferior. The statements for the limit superior are proven similarly.

For \( n \geq 1 \), put \( A_n = \inf_{k \geq n} x_k = \inf \{ x_n, x_{n+1}, \ldots \} \). Then \( \lim_{n \to \infty} x_n = \sup_{n \geq 1} A_n \).

(a) Suppose first that \( \lim x_n = +\infty \). Let \( M \in \mathbb{R} \). Since \( \sup_{n \geq 1} A_n = +\infty \), there exists \( n \in \mathbb{N} \) with \( M \leq A_n = \inf_{k \geq n} x_k \). Hence \( M \leq A_n \leq x_k \) for all \( k \geq n \).

Suppose next that the given property hold. Let \( M \in \mathbb{R} \). Then there exists \( n \in \mathbb{N} \) with \( x_k \geq M \) for all \( k \geq n \). Hence \( A_n = \inf_{k \geq n} x_k \geq M \). We proved:

\[ \forall M \in \mathbb{R} : \exists n \in \mathbb{N} : A_n \geq M \]

So \( \lim x_n = \sup_{n \geq 1} A_n = +\infty \).

(b) Suppose first that \( \lim x_n = -\infty \). Then \( \sup_{n \geq 1} A_n = -\infty \). Hence \( A_n = -\infty \) for all \( n \geq 1 \). Let \( M \in \mathbb{R} \) and let \( n \geq 1 \). Since \( A_n = \inf_{k \geq n} x_k = -\infty \), there exists \( k \geq n \) with \( x_k \leq M \).

Suppose next that the given property hold. Let \( M \in \mathbb{R} \) and let \( n \geq 1 \). Then there exists \( k \geq n \) with \( x_k \leq M \). Hence \( A_n = \inf_{k \geq n} x_k \leq M \). Since this is true for all \( n \in \mathbb{N} \), we get that \( \sup_{n \geq 1} A_n \leq M \). Since this is true for all \( M \in \mathbb{R} \), we have that \( \lim x_n = \sup_{n \geq 1} A_n = -\infty \).

(c) Suppose first that \( \lim x_n = \alpha \in \mathbb{R} \). Let \( \varepsilon > 0 \). Since \( \sup_{n \geq 1} A_n = \alpha \), there exists \( n \geq 1 \) such that \( \alpha - \varepsilon < A_n = \inf_{k \geq n} x_k \). Hence \( \alpha - \varepsilon \leq x_k \) for all \( k \geq n \) and (1) holds. Let \( \varepsilon > 0 \) and let \( n \geq 1 \). Then \( A_n \leq \sup_{n \geq 1} A_n = \alpha \). If \( x_k > \alpha + \varepsilon \) for all \( k \geq n \) then \( A_n = \inf_{k \geq n} x_k \geq \alpha + \varepsilon > \alpha \), a contradiction. Hence there exists \( k \geq n \) with \( x_k \leq \alpha + \varepsilon \) and (2) holds.

Suppose next that properties (1) and (2) hold. Let \( n \geq 1 \) and let \( \varepsilon > 0 \). By (2), there exists \( k \geq n \) with \( x_k \leq \alpha + \varepsilon \). Hence \( A_n = \inf_{k \geq n} x_k \leq \alpha + \varepsilon \). Since this is true for all \( \varepsilon > 0 \), we have that \( A_n \leq \alpha \). Let \( \varepsilon > 0 \). By (1), there exists \( n \geq 1 \) such that \( \alpha - \varepsilon \leq x_k \) for all \( k \geq n \). Hence \( A_n = \inf_{k \geq n} x_k \geq \alpha - \varepsilon \). We proved:

\[ \begin{align*}
(i) & \quad \forall n \geq 1 : A_n \leq \alpha \\
(ii) & \quad \forall \varepsilon > 0 : \exists n \geq 1 : \alpha - \varepsilon \leq A_n
\end{align*} \]

Hence \( \lim x_n = \sup_{n \geq 1} A_n = \alpha \).

We mention several properties of the limit superior and limit inferior of a sequence.

**Proposition 1.19** Let \( \{ x_n \}_{n \geq 1} \) be a sequence of extended real numbers. Then the following holds:

(a) \( \lim x_n \leq \lim x_n \)

(b) \( \lim x_n = \lim x_n := \alpha \in \mathbb{R} \) if and only if the sequence \( \{ x_n \}_{n \geq 1} \) converges to \( \alpha \).

(c) \( \lim(-x_n) = -\lim x_n \) and \( \lim(-x_n) = -\lim x_n \).

**Proof:** (a) For \( n \geq 1 \), put \( A_n = \inf_{k \geq n} x_k \) and \( B_n = \sup_{k \geq n} x_k \). Then

\[ A_n \leq B_n \quad \text{for all } n \geq 1 \]

Taking the limit on both sides as \( n \) goes to \( \infty \), we get

\[ \lim_{n \to \infty} x_n = \lim_{n \to \infty} A_n \leq \lim_{n \to \infty} B_n = \lim x_n \]
(b) Suppose first that \( \lim x_n = \lim x_n := \alpha \). If \( \alpha = -\infty \) then
\[
\forall M \in \mathbb{R} : \exists n \in \mathbb{N} : \forall k \geq n : x_k \leq M
\]
by using the formulation for the limit superior. Hence \( \{x_n\}_{n \geq 1} \) converges to \( -\infty \). If \( \alpha = +\infty \) then
\[
\forall M \in \mathbb{R} : \exists n \in \mathbb{N} : \forall k \geq n : x_k \geq M
\]
by using the formulation for the limit inferior. Hence \( \{x_n\}_{n \geq 1} \) converges to \( +\infty \). So we may assume that \( \alpha \in \mathbb{R} \). Pick \( \varepsilon > 0 \). Using the formulation for the limit inferior, we have
\[
\exists n_1 \in \mathbb{N} : \forall k \geq n_1 : \alpha - \varepsilon \leq x_k
\]
Using the formulation for the limit superior, we have
\[
\exists n_2 \in \mathbb{N} : \forall k \geq n_2 : x_k \leq \alpha + \varepsilon
\]
Put \( n = \max\{n_1, n_2\} \). Pick \( k \geq n \). Then \( \alpha - \varepsilon \leq x_k \leq \alpha + \varepsilon \) and so \( |x_n - \alpha| \leq \varepsilon \). We proved:
\[
\forall \varepsilon > 0 : \exists n \in \mathbb{N} : \forall k \geq n : |x_k - \alpha| \leq \varepsilon
\]
Hence \( \{x_n\}_{n \geq 1} \) converges to \( \alpha \).

Suppose next that the sequence \( \{x_n\}_{n \geq 1} \) converges to \( \alpha \). If \( \alpha = -\infty \) then
\[
\forall M \in \mathbb{R} : \exists n \in \mathbb{N} : \forall k \geq n : x_k \leq M
\]
Hence \( \lim x_n = -\infty \) and so \( \lim x_n = -\infty \) since \( \lim x_n \leq \lim x_n \). If \( \alpha = +\infty \) then
\[
\forall M \in \mathbb{R} : \exists n \in \mathbb{N} : \forall k \geq n : x_k \geq M
\]
Hence \( \lim x_n = +\infty \) and so \( \lim x_n = +\infty \) since \( \lim x_n \geq \lim x_n \). So we may assume that \( \alpha \in \mathbb{R} \). Let \( \varepsilon > 0 \). Then there exists \( n \geq 1 \) such that \( |x_k - \alpha| < \varepsilon \) for all \( k \geq n \). Hence \( \alpha - \varepsilon < x_k < \alpha + \varepsilon \) for all \( k \geq n \). Moreover, if \( m \geq 1 \) is given, then \( \alpha - \varepsilon < x_k < \alpha + \varepsilon \) for all \( k \geq \max\{m, n\} \). So \( \lim x_n = \alpha = \lim x_n \) by the formulations for the limit inferior and the limit superior.

(c) We prove that \( \lim (-x_n) = -\lim x_n \). Then \( \lim (-x_n) = -\lim (-(-x_n)) = -\lim x_n \).
Using that \( \inf(-S) = -\sup S \), we easily get that
\[
\lim (-x_n) = \lim_{n \to \infty} \inf_{k \geq n} (-x_k) = \lim_{n \to \infty} (-\sup_{k \geq n} x_k) = -\lim_{n \to \infty} \sup_{k \geq n} x_k = -\lim x_n
\]
We conclude by proving some properties of the limit inferior/superior of two sequences.

**Proposition 1.20** Let \( \{x_n\}_{n \geq 1} \) and \( \{y_n\}_{n \geq 1} \) be two sequences of real numbers. Then the following holds:

(a) If \( x_n \leq y_n \) for all \( n \) big enough then \( \lim x_n \leq \lim y_n \) and \( \lim x_n \leq \lim y_n \).

(b) \( \lim x_n + \lim y_n \leq \lim (x_n + y_n) \leq \lim x_n + \lim y_n \leq \lim (x_n + y_n) \leq \lim x_n + \lim y_n \) unless a sum is undefined (like \((+\infty) + (-\infty))\).

**Proof:** (a) We prove that \( \lim x_n \leq \lim y_n \). The proof of \( \lim x_n \leq \lim y_n \) is similar. By assumption, there exists \( n_1 \in \mathbb{N} \) with \( x_n \leq y_n \) for all \( n \geq n_1 \).
We consider three cases.
1. \( \lim y_n = +\infty \)
   Then clearly \( \lim x_n \leq \lim y_n \).

2. \( \lim y_n = -\infty \)
   Let \( M \in \mathbb{R} \). Then we have
   \[ \exists n_2 \in \mathbb{N} : \forall k \geq n_2 : y_k \leq M \]
   Put \( n = \max\{n_1, n_2\} \). Let \( k \geq n \). Then \( x_k \leq y_k \leq M \). We proved:
   \[ \forall M \in \mathbb{R} : \exists n \in \mathbb{N} : \forall k \geq n : x_k \leq M \]
   Hence \( \lim x_n = -\infty = \lim y_n \).

3. \( \lim y_n = \beta \in \mathbb{R} \)
   If \( \lim x_n = -\infty \) then clearly \( \lim x_n \leq \lim y_n \). So we may assume that \( \lim x_n \neq -\infty \). Let \( \varepsilon > 0 \).
   Since \( \lim y_n = \beta \), we have:
   \[ \exists n_2 \in \mathbb{N} : \forall k \geq n_2 : y_k < \beta + \varepsilon \]
   Hence \( x_k \leq y_k < \beta + \varepsilon \) for all \( k \geq \max\{n_1, n_2\} \). So \( \lim x_n \neq +\infty \) (indeed, we proved: \( \exists M \in \mathbb{R}, \exists n \in \mathbb{N} : \forall k \geq n : x_k < M \) which is the negation of \( \forall M \in \mathbb{R}, \forall n \in \mathbb{N} : \exists k \geq n : x_k \geq M \)).
   Hence \( \lim x_n = \alpha \in \mathbb{R} \). Then \( \alpha - \varepsilon < x_k \) for some \( k \geq \max\{n_1, n_2\} \). For this \( k \), we have
   \[ \alpha - \varepsilon < x_k \leq y_k < \beta + \varepsilon \]
   So we proved
   \[ \forall \varepsilon > 0 : \alpha - \varepsilon < \beta + \varepsilon \]
   Taking the limit as \( \varepsilon \to 0^+ \), we get
   \[ \lim x_n = \alpha \leq \beta = \lim y_n \]

(b) First, we prove that \( \lim (x_n + y_n) \leq \lim x_n + \lim y_n \). The proof of \( \lim x_n + \lim y_n \leq \lim (x_n + y_n) \) is similar.

We consider three cases.

1. \( \lim x_n = +\infty \) or \( \lim y_n = +\infty \)
   Then \( \lim x_n + \lim y_n = +\infty \). Hence \( \lim (x_n + y_n) \leq \lim x_n + \lim y_n \).

2. \( \lim x_n = -\infty \) or \( \lim y_n = -\infty \)
   Then \( \lim x_n + \lim y_n = -\infty \). Assume that \( \lim x_n = -\infty \). We consider two subcases.

   2a. \( \lim y_n = -\infty \)
   Let \( M \in \mathbb{R} \). Since \( \lim x_n = -\infty \), we have
   \[ \exists n_1 \in \mathbb{N} : \forall k \geq n_1 : x_k \leq \frac{M}{2} \]
   Since \( \lim y_n = -\infty \), we have
   \[ \exists n_2 \in \mathbb{N} : \forall k \geq n_2 : y_k \leq \frac{M}{2} \]
Put $n = \max\{n_1, n_2\}$. Let $k \geq n$. Then $x_k + y_k \leq \frac{M}{2} + \frac{M}{2} = M$. We proved:

$$\forall M \in \mathbb{R} : \exists n \in \mathbb{N} : \forall k \geq n : x_k + y_k \leq M$$

Hence $\lim (x_n + y_n) = -\infty = \lim x_n + \lim y_n$.

2b. $\lim y_n = \beta \in \mathbb{R}$

Let $M \in \mathbb{R}$. Since $\lim x_n = -\infty$, we have

$$\exists n_1 \in \mathbb{N} : \forall k \geq n_1 : x_k \leq M - \beta - 1$$

Since $\lim y_n = \beta$,

$$\exists n_2 \in \mathbb{N} : \forall k \geq n_1 : y_k \leq \beta + 1$$

Put $n = \max\{n_1, n_2\}$. Let $k \geq n$. Then $x_k + y_k \leq M - \beta - 1 + \beta + 1 = M$. We proved:

$$\forall M \in \mathbb{R} : \exists n \in \mathbb{N} : \forall k \geq n : x_k + y_k \leq M$$

Hence $\lim (x_n + y_n) = -\infty = \lim x_n + \lim y_n$.

3. $\lim x_n = \alpha \in \mathbb{R}$ and $\lim y_n = \beta \in \mathbb{R}$

If $\lim (x_n + y_n) = -\infty$ then clearly $\lim (x_n + y_n) \leq \lim x_n + \lim y_n$. So we may assume that $\lim (x_n + y_n) \neq -\infty$. Let $\varepsilon > 0$. Since $\lim x_n = \alpha$, we have

$$\exists n_1 \in \mathbb{N} : \forall k \geq n_1 : x_k < \alpha + \varepsilon$$

Since $\lim y_n = \beta$, we have

$$\exists n_2 \in \mathbb{N} : \forall k \geq n_1 : y_k < \beta + \varepsilon$$

Put $n = \max\{n_1, n_2\}$. Let $k \geq n$. Then $x_k + y_k < \alpha + \beta + 2\varepsilon$. Hence $\lim (x_n + y_n) \neq +\infty$ (indeed, we proved: $\exists M \in \mathbb{R}, \exists n \in \mathbb{N} : \forall k \geq n : x_k + y_k < M$ which is the negation of $\forall M \in \mathbb{R}, \forall n \in \mathbb{N} : \exists k \geq n : x_k + y_k \geq M$). So $\lim (x_n + y_n) = \gamma \in \mathbb{R}$. Then $\gamma - \varepsilon < x_k + y_k$ for some $k \geq n$. For this $k$, we get

$$\gamma - \varepsilon < x_k + y_k < \alpha + \beta + 2\varepsilon$$

So we proved:

$$\forall \varepsilon > 0 : \gamma - \varepsilon < \alpha + \beta + 2\varepsilon$$

Taking the limit as $\varepsilon \to 0^+$, we get

$$\lim (x_n + y_n) = \gamma \leq \alpha + \beta = \lim x_n + \lim y_n$$

Next, we prove that $\lim x_n + \lim y_n \leq \lim (x_n + y_n)$. The proof of $\lim (x_n + y_n) \leq \lim x_n + \lim y_n$ is similar.

We consider three cases.

1. $\lim y_n = -\infty$

Then $\lim x_n + \lim y_n = -\infty$. Hence $\lim x_n + \lim y_n \leq \lim (x_n + y_n)$.
2. \( \lim_{n \to \infty} y_n = +\infty \)

Then \( \lim_{n \to \infty} x_n + \lim_{n \to \infty} y_n = +\infty \). We consider two subcases.

2a. \( \lim_{n \to \infty} x_n = +\infty \)

Let \( M \in \mathbb{R} \). Let \( n \in \mathbb{N} \). Since \( \lim_{n \to \infty} x_n = +\infty \), we have

\[ \exists n_1 \in \mathbb{N} : \forall k \geq n_1 : x_k > \frac{M}{2} \]

Since \( \lim_{n \to \infty} y_n = +\infty \), we get

\[ \exists k \geq \max\{n, n_1\} : y_k > \frac{M}{2} \]

For this \( k \), we find that

\[ x_k + y_k > \frac{M}{2} + \frac{M}{2} = M \]

So we proved:

\[ \forall M \in \mathbb{R}, \forall n \in \mathbb{N} : \exists k \geq n : x_k + y_k > M \]

Hence \( \lim_{n \to \infty} (x_n + y_n) = +\infty = \lim_{n \to \infty} x_n + \lim_{n \to \infty} y_n \).

2b. \( \lim_{n \to \infty} x_n = \alpha \in \mathbb{R} \)

Let \( M \in \mathbb{R} \). Let \( n \in \mathbb{N} \). Since \( \lim_{n \to \infty} x_n = \alpha \), we have

\[ \exists n_1 \in \mathbb{N} : \forall k \geq n_1 : x_k > \alpha - 1 \]

Since \( \lim_{n \to \infty} y_n = +\infty \), we get

\[ \exists k \geq \max\{n, n_1\} : y_k > M - \alpha + 1 \]

For this \( k \), we find

\[ x_k + y_k > \alpha - 1 + M - \alpha + 1 = M \]

So we proved:

\[ \forall M \in \mathbb{R}, \forall n \in \mathbb{N} : \exists k \geq n : x_k + y_k > M \]

Hence \( \lim_{n \to \infty} (x_n + y_n) = +\infty = \lim_{n \to \infty} x_n + \lim_{n \to \infty} y_n \).

3. \( \lim_{n \to \infty} y_n = \beta \in \mathbb{R} \)

We consider three subcases.

3a. \( \lim_{n \to \infty} x_n = -\infty \)

Then \( \lim_{n \to \infty} x_n + \lim_{n \to \infty} y_n = -\infty \). Hence \( \lim_{n \to \infty} x_n + \lim_{n \to \infty} y_n \leq \lim_{n \to \infty} (x_n + y_n) \).

3b. \( \lim_{n \to \infty} x_n = +\infty \)

Then \( \lim_{n \to \infty} x_n + \lim_{n \to \infty} y_n = +\infty \). Let \( M \in \mathbb{R} \). Let \( n \in \mathbb{N} \). Since \( \lim_{n \to \infty} x_n = +\infty \), we have

\[ \exists n_1 \in \mathbb{N} : \forall k \geq n_1 : x_k > M - \beta + 1 \]

Since \( \lim_{n \to \infty} y_n = \beta \), we have

\[ \exists k \geq \max\{n_1, n\} : y_k > \beta - 1 \]
For this $k$, we find
\[ x_k + y_k > M - \beta + 1 + \beta - 1 = M \]

So we proved:
\[ \forall M \in \mathbb{R}, \forall n \in \mathbb{N} : \exists k \geq n : x_k + y_k > M \]

Hence $\lim (x_n + y_n) = +\infty = \lim x_n + \lim y_n$.

3c. $\lim x_n = \alpha \in \mathbb{R}$

If $\lim (x_n + y_n) = +\infty$, then clearly $\lim x_n + \lim y_n \leq \lim (x_n + y_N)$. Hence we may assume that $\lim (x_n + y_n) \neq +\infty$.

Let $n \in \mathbb{N}$. Since $\lim x_n = \alpha$, we get that
\[ \exists n_1 \in \mathbb{N} : \forall k \geq n_1 : x_k > \alpha - 1 \]

Since $\lim y_n = \beta$, we get
\[ \exists k \geq \max\{n, n_1\} : y_k > \beta - 1 \]

So for this $k$, we find
\[ x_k + y_k > \alpha + \beta - 2 \]

So we proved:
\[ \exists M \in \mathbb{R} : \forall n \in \mathbb{N} : \exists k \geq n : x_k + y_k > M \]

which is the negation of
\[ \forall M \in \mathbb{R} : \exists n \in \mathbb{N} : \forall k \geq n : x_k + y_k \leq M \]

So $\lim (x_n + y_n) \neq -\infty$.

Hence we may assume that $\lim (x_n + y_n) = \gamma \in \mathbb{R}$. Let $\varepsilon > 0$. Since $\lim (x_n + y_n) = \gamma$, we have
\[ \exists n_1 \in \mathbb{N} : \forall k \geq n_1 : x_k + y_k < \gamma + \varepsilon \]

Since $\lim x_n = \alpha$, we get that
\[ \exists n_2 \in \mathbb{N} : \forall k \geq n_2 : x_k > \alpha - \varepsilon \]

Since $\lim y_n = \beta$, we get
\[ \exists k \geq \max\{n_1, n_2\} : y_k > \beta - \varepsilon \]

For this $k$, we find
\[ \alpha - \varepsilon + \beta - \varepsilon < x_k + y_k < \gamma + \varepsilon \]

We proved:
\[ \forall \varepsilon > 0 : \alpha + \beta - 2\varepsilon < \gamma + \varepsilon \]

Taking the limit as $\varepsilon \to 0^+$, we get
\[ \lim x_n + \lim y_n = \alpha + \beta \leq \gamma = \lim (x_n + y_n) \]

\[ \square \]
1.8 Open and Closed Sets of Real Numbers

Definition 1.21 A subset $\mathcal{O} \subseteq \mathbb{R}$ is open if
\[ \forall x \in \mathcal{O} : \exists r > 0 : (x - r, x + r) \subseteq \mathcal{O} \]

One can easily show that $\emptyset, \mathbb{R}, (a, b), (a, +\infty), (-\infty, b)$ are open sets where $a, b \in \mathbb{R}$.

The next couple of propositions deal with intersections and unions of open sets.

Proposition 1.22 The intersection of finitely many open sets is open.

Proof: Let $\{\mathcal{O}_i : 1 \leq i \leq n\}$ be a finite collection of open sets. Let $x \in \bigcap_{1 \leq i \leq n} \mathcal{O}_i$. Then $x \in \mathcal{O}_i$ for $1 \leq i \leq n$. Since $\mathcal{O}_i$ is open, we get that for $1 \leq i \leq n$, there exists $r_i > 0$ with $(x - r_i, x + r_i) \subseteq \mathcal{O}_i$. Put $r = \min\{r_1, r_2, \ldots, r_n\}$. Then $r > 0$ and $(x - r, x + r) \subseteq \bigcap_{1 \leq i \leq n} \mathcal{O}_i$. So $\bigcap_{1 \leq i \leq n} \mathcal{O}_i$ is open. □

Remark: The intersection of infinitely many open sets does not need to be open. □

Proposition 1.23 The union of arbitrarily many open sets is open.

Proof: Let $\{\mathcal{O}_i : i \in I\}$ be an arbitrary collection of open sets. Let $x \in \bigcup_{i \in I} \mathcal{O}_i$. Then $x \in \mathcal{O}_j$ for some $j \in I$. Since $\mathcal{O}_j$ is open, we get that there exists $r > 0$ with $(x - r, x + r) \subseteq \mathcal{O}_j$. So $(x - r, x + r) \subseteq \mathcal{O}_j \subseteq \bigcup_{i \in I} \mathcal{O}_i$. Hence $\bigcup_{i \in I} \mathcal{O}_i$ is open. □

It turns out that we can completely describe open sets of real numbers in terms of open intervals. Note that these intervals could be infinite.

Theorem 1.24 Every open set of real numbers is the union of a countable number of pairwise disjoint open intervals.

Proof: Let $\mathcal{O}$ be a non-empty open set of real numbers.

First, we associate with an open interval $(a_x, b_x)$ with each $x \in \mathcal{O}$. So let $x \in \mathcal{O}$. We define
\[ a_x = \inf\{y \in \mathbb{R} : y \leq x \text{ and } (y, x) \subseteq \mathcal{O}\} \quad \text{and} \quad b_x = \sup\{z \in \mathbb{R} : x \leq z \text{ and } (x, z) \subseteq \mathcal{O}\} \]

Note that these sets are not empty: since $x \in \mathcal{O}$ and $\mathcal{O}$ is open, we have that $(x - r, x + r) \subseteq \mathcal{O}$ for some $r > 0$; hence $(x - r, x) \subseteq \mathcal{O}$ and $(x, x + r) \subseteq \mathcal{O})$. Also note that $a_x \in \mathbb{R} \cup \{-\infty\}$ and $b_x \in \mathbb{R} \cup \{+\infty\}$. We prove two properties of the interval $(a_x, b_x)$.

Claim 1: $(a_x, b_x) \subseteq \mathcal{O}$

We show that $(x, b_x) \subseteq \mathcal{O}$. Let $w \in (x, b_x)$. Since $b_x$ is a supremum, there exists $z \geq x$ such that $(x, z) \subseteq \mathcal{O}$ and $w < z$. Hence $w \in (x, z) \subseteq \mathcal{O}$. So $(x, b_x) \subseteq \mathcal{O}$.

Similarly, we show that $(a_x, x) \subseteq \mathcal{O}$. Hence $(a_x, b_x) \subseteq \mathcal{O}$. 16
Claim 2: \(a_x, b_x \notin \mathcal{O}\)

Suppose that \(b_x \in \mathcal{O}\). Then \(b_x \in \mathbb{R}\). Since \(\mathcal{O}\) is open, we get that \((b_x - r, b_x + r) \subseteq \mathcal{O}\) for some \(r > 0\). By Claim 1, we have that \((x, b_x) \subseteq \mathcal{O}\). Hence \((x, b_x + r) \subseteq (x, b_x) \cup (b_x - r, b_x + r) \subseteq \mathcal{O}\). So \(b_x + r \in \{z \in \mathbb{R} : x \leq z\text{ and } (x, z) \subseteq \mathcal{O}\}\), a contradiction since \(b_x = \sup\{z \in \mathbb{R} : x \leq z\text{ and } (x, z) \subseteq \mathcal{O}\}\). Thus \(b_x \notin \mathcal{O}\).

Similarly, we show that \(a_x \notin \mathcal{O}\).

Next, we prove the proposition. Consider the collection of open intervals \(\{(a_x, b_x) : x \in \mathcal{O}\}\). Clearly \(\mathcal{O} = \bigcup_{x \in \mathcal{O}} (a_x, b_x)\). Indeed, \((a_x, b_x) \subseteq \mathcal{O}\) for all \(x \in \mathcal{O}\) and so \(\bigcup_{x \in \mathcal{O}} (a_x, b_x) \subseteq \mathcal{O}\). Let \(y \in \mathcal{O}\). Then \(y \in (a_y, b_y) \subseteq \bigcup_{x \in \mathcal{O}} (a_x, b_x)\) and so \(\mathcal{O} \subseteq \bigcup_{x \in \mathcal{O}} (a_x, b_x)\). Hence \(\mathcal{O} = \bigcup_{x \in \mathcal{O}} (a_x, b_x)\).

We show that this collection is a disjoint collection. Suppose that \((a_x, b_x) \cap (a_y, b_y) \neq \emptyset\) for some \(x, y \in \mathcal{O}\), say \(z \in (a_x, b_x) \cap (a_y, b_y)\). Then \(a_y < z < b_x\). Since \(a_y \notin \mathcal{O}\) and \((a_x, b_x) \subseteq \mathcal{O}\) and \(a_y < b_x\), we get that \(a_y \leq a_x\). Similarly, we find that \(a_x \leq a_y\) and so \(a_x = a_y\). Similarly, we get that \(b_x = b_y\). Hence \((a_x, b_x) = (a_y, b_y)\).

Finally, we show that this collection is countable. Note that \(\mathcal{O}\) is uncountable but we will end up with the same interval \((a_x, b_x)\) for different values of \(x\) (in fact, any element in \((a_x, b_x)\) will lead to the same interval). For each interval \((a_x, b_x)\), pick a rational number in \((a_x, b_x)\). Since the intervals are pairwise disjoint, different intervals will give us different rational numbers. Hence we have a one-to-one map from the collection of intervals to the rational numbers. Since the rational numbers are countable, it follows that the collection of intervals is countable. \(\square\)

**Definition 1.25** Let \(E\) be a set of real numbers and \(x \in \mathbb{R}\). Then \(x\) is a **point of closure of \(E\)** if

\[
\forall \delta > 0 : \exists y \in E : |x - y| < \delta
\]

We denote the set of of point of closure of \(E\) by \(\overline{E}\) and call it the **closure of \(E\)**.

**Proposition 1.26** The following holds:

(a) Let \(E \subseteq \mathbb{R}\) and \(x \in \mathbb{R}\). Then \(x \in \overline{E}\) if and only if \(\mathcal{O} \cap E \neq \emptyset\) for every open set \(\mathcal{O}\) containing \(x\).

(b) \(E \subseteq \overline{E}\) for all \(E \subseteq \mathbb{R}\).

(c) \(\overline{\overline{E}} = \overline{E}\) for all \(E \subseteq \mathbb{R}\).

(d) Let \(A, B \in \mathbb{R}\) with \(A \subseteq B\). Then \(\overline{A} \subseteq \overline{B}\).

**Definition 1.27** Let \(F \subseteq \mathbb{R}\). Then \(F\) is **closed** if \(F = \overline{F}\).

The next proposition gives a relation between open sets and closed sets.

**Proposition 1.28** Let \(E \subseteq \mathbb{R}\). Then \(E\) is open if and only if \(\overline{E}\) is closed.

**Proof:** Put \(F = \overline{E}\).

Suppose first that \(E\) is open. Let \(x \in \overline{E}\). If \(x \in E\), then \(\mathcal{O} := (x - r, x + r) \subseteq E\) for some \(r > 0\) and so \(\mathcal{O} \cap F = \mathcal{O} \cap \overline{E} = \emptyset\), a contradiction since \(x \in \overline{E}\) and \(\mathcal{O}\) is an open set containing \(x\). Hence \(x \notin E\). So \(x \in \overline{E} = F\). Thus \(F \subseteq E\). Hence \(\overline{F} = F\) and \(F\) is closed.
Suppose next that $F$ is closed. Let $x \in E$. Then $x \notin \widetilde{E} = F = \overline{F}$ and so

$$\exists \delta > 0 : \forall y \in F : |x - y| \geq \delta$$

Let $z \in (x - \delta, x + \delta)$. If $z \in F$, then $|x - z| \geq \delta$, a contradiction since $z \in (x - \delta, x + \delta)$. Hence $z \notin F = \widetilde{E}$. So $z \in E$. Thus $(x - \delta, x + \delta) \subseteq E$ and $E$ is open. \hfill \Box

**Remark:** Since $\widetilde{E} = E$, we get that $E$ is closed if and only if $\widetilde{E}$ is open.

The next couple of propositions deal with intersections and unions of closed sets.

**Proposition 1.29** The intersection of arbitrarily many closed sets is closed.

**Proof:** Let $\{F_i : i \in I\}$ be an arbitrary collection of closed sets. Then $\widetilde{F_i}$ is open for all $i \in I$ by Proposition 1.28. Using De Morgan’s Law for intersections, we get that $\bigcap_{i \in I} \widetilde{F_i} = \bigcup_{i \in I} F_i$ is open by Proposition 1.23. So $\bigcap_{i \in I} \widetilde{F_i}$ is closed by Proposition 1.28. \hfill \Box

**Proposition 1.30** The union of finitely many closed sets is closed.

**Proof:** Let $\{F_i : 1 \leq i \leq n\}$ be a finite collection of closed sets. Then $\widetilde{F_i}$ is open for $1 \leq i \leq n$ by Proposition 1.28. Using De Morgan’s Law for unions, we get that $\bigcup_{1 \leq i \leq n} \widetilde{F_i} = \bigcap_{1 \leq i \leq n} F_i$ is open by Proposition 1.22. So $\bigcup_{1 \leq i \leq n} \widetilde{F_i}$ is closed by Proposition 1.28. \hfill \Box

**Remark:** The union of infinitely many closed sets does not need to be closed. \hfill △

The proof of the next theorem uses the famous Heine-Borel Theorem. Recall from Math 172 that there are three equivalent definitions of a compact set $E$ ($E$ is closed and bounded; every sequence in $E$ has a subsequence that converges to some element in $E$; every open covering of $E$ has a finite subcovering).

**Theorem 1.31** [Nested Set Theorem] Let $F_1 \supseteq F_2 \supseteq F_3 \supseteq \cdots$ be a descending sequence of non-empty closed sets such that $F_1$ is bounded. Then $\bigcap_{n=1}^{\infty} F_n \neq \emptyset$.

**Proof:** Suppose that $\bigcap_{n=1}^{\infty} F_n = \emptyset$. Then using De Morgan’s Law for intersections, we get that

$$\bigcup_{n=1}^{\infty} \overline{F_n} = \bigcap_{n=1}^{\infty} F_n = \emptyset = \mathbb{R}$$

Note that $\overline{F_n}$ is open for all $n \geq 1$. Hence $\{\overline{F_n} : n \geq 1\}$ is an open covering of $F_1$. Since $F_1$ is compact, it follows from the Heine-Borel Theorem that this open covering has a finite subcovering, say $\{F_{n_1}, \ldots, F_{n_k}\}$. Put $m = \max\{n_1, \ldots, n_k\}$. Since $\{F_n\}_{n \geq 1}$ is a descending sequence of sets, it follows from De Morgan’s Law for intersections that

$$F_1 \subseteq \bigcup_{i=1}^{k} \overline{F_n} \subseteq \bigcup_{n=1}^{m} \overline{F_n} = \bigcap_{n=1}^{m} F_n = \overline{F_m}$$

So $F_1 \subseteq \overline{F_m}$. Since $F_m \subseteq F_1$, we get that $\overline{F_1} \subseteq \overline{F_m}$. Hence $\mathbb{R} = F_1 \cup \overline{F_1} \subseteq \overline{F_m}$. So $\mathbb{R} = \overline{F_m}$ and thus $\emptyset = \mathbb{R} = \overline{F_m} = F_m$, a contradiction.

Hence $\bigcap_{n=1}^{\infty} F_n \neq \emptyset$. \hfill \Box

**Remark:** This theorem is false if we drop the condition that $F_1$ is bounded or that the $F_n$’s are closed. \hfill △
1.9 Continuous Functions

**Definition 1.32** Let $E \subseteq \mathbb{R}$ and $f : E \to \mathbb{R}$ a function.

(a) Let $a \in E$. Then $f$ is continuous at $a$ if

$$\forall \varepsilon > 0 : \exists \delta > 0 : \forall x \in E : |x - a| < \delta \implies |f(x) - f(a)| < \varepsilon$$

(b) $f$ is continuous on $E$ if $f$ is continuous at $a$ for all $a \in E$.

\[\blacksquare\]

There is a relation between continuous functions and open sets.

**Theorem 1.33** Let $f : E \to \mathbb{R}$ be a function. Then $f$ is continuous on $E$ if and only if for every open set $O$, there exists an open set $O^*$ with $f^{-1}(O) = E \cap O^*$.

**Proof:** Suppose first that for every open set $O$, there exists an open set $O^*$ with $f^{-1}(O) = E \cap O^*$. Let $a \in E$. Let $\varepsilon > 0$. Put $O = (f(a) - \varepsilon, f(a) + \varepsilon)$. Since $O$ is open, we have that $f^{-1}(O) = E \cap O^*$ for some open set $O^*$. Note that $a \in f^{-1}(O)$ since $f(a) \in O$. Hence $a \in O^*$. So $(a - \delta, a + \delta) \subseteq O^*$ for some $\delta > 0$ since $O^*$ is open. Let $x \in E$ with $|x - a| < \delta$. Then $x \in E \cap (a - \delta, a + \delta) \subseteq E \cap O^* = f^{-1}(O)$. Thus $f(x) \in O = (f(a) - \varepsilon, f(a) + \varepsilon)$. So $|f(x) - f(a)| < \varepsilon$. We proved:

$$\forall \varepsilon > 0 : \exists \delta > 0 : \forall x \in E : |x - a| < \delta \implies |f(x) - f(a)| < \varepsilon$$

Hence $f$ is continuous at $a$. Since $a$ was arbitrary, we get that $f$ is continuous on $E$.

Suppose next that $f$ is continuous on $E$. Let $O$ be an open set. We associate with every element of $f^{-1}(O)$ an open set. So let $a \in f^{-1}(O)$. Then $f(a) \in O$. Hence $(f(a) - \varepsilon_a, f(a) + \varepsilon_a) \subseteq O$ for some $\varepsilon_a > 0$ since $O$ is open. Using the continuity of $f$ on $E$, we get

$$\exists \delta_a > 0 : \forall x \in E : |x - a| < \delta_a \implies |f(x) - f(a)| < \varepsilon_a$$

Put $O^* = \cup_{a \in f^{-1}(O)} (a - \delta_a, a + \delta_a)$. Then $O^*$ is open by Proposition 1.23. Let $b \in f^{-1}(O)$. Then $b \in E$ and $b \in (b - \delta_b, b + \delta_b)$. Hence

$$b \in E \cap (b - \delta_b, b + \delta_b) \subseteq E \cap (\cup_{a \in f^{-1}(O)} (a - \delta_a, a + \delta_a)) = E \cap O^*$$

So $f^{-1}(O) \subseteq E \cap O^*$. Let $x \in E \cap O^*$. Then $x \in E$ and $x \in (a - \delta_a, a + \delta_a)$ for some $a \in f^{-1}(O)$. So $|x - a| < \delta_a$ and thus $|f(x) - f(a)| < \varepsilon_a$. Hence $f(x) \in (f(a) - \varepsilon_a, f(a) + \varepsilon_a) \subseteq O$. So $x \in f^{-1}(O)$. Thus $E \cap O^* \subseteq f^{-1}(O)$. Hence $f^{-1}(O) = E \cap O^*$.
Chapter 2

Measurable Sets

In this chapter, we will generalize the concept ‘length of an interval’.

2.1 Outer Measure of a Set

We want to generalize the concept ‘length of an interval’ to more general sets: what is the ‘length’ of any kind of subset of real numbers. We use the terminology measure of a set instead of ‘length’. We want this measure to have certain properties.

Does there exist a function \( \mu : \mathcal{P}(\mathbb{R}) \rightarrow [0, +\infty] \) such that

1. If \( I \) is an interval then \( \mu(I) = l(I) \) where \( l(I) \) is the length of \( I \).

2. \( \mu \) is translation invariant: \( \mu(E + x) = \mu(E) \) for all \( E \subseteq \mathbb{R} \) and all \( x \in \mathbb{R} \).

3. \( \mu \) is countable additive: If \( \{E_n\}_{n \geq 1} \) is a sequence of disjoint sets of real numbers then \( \mu \left( \bigcup_{n=1}^{\infty} E_n \right) = \sum_{n=1}^{\infty} \mu(E_n) \).

It turns out that such a function \( \mu \) does not exist! In this course, we will relax the third condition: countable additivity will only hold for a specific class of sets (called the measurable sets).

Definition 2.1 Let \( A \subseteq \mathbb{R} \).

(a) A countable cover of open intervals for \( A \) (abbreviation: ccoi for \( A \)) is a countable collection \( \{I_n : n \in \mathbb{N}\} \) of open intervals with \( A \subseteq \bigcup_{n=1}^{\infty} I_n \).

(b) The outer measure of \( A \) (notation: \( m^*(A) \)) is defined as

\[
m^*(A) = \inf \left\{ \sum_{n=1}^{\infty} l(I_n) : \{I_n : n \in \mathbb{N}\} \text{ is a ccoi for } A \right\}
\]

We list some immediate but important properties of the outer measure.

Proposition 2.2 The outer measure has the following properties:

(a) \( 0 \leq m^*(A) \leq +\infty \) for all \( A \subseteq \mathbb{R} \).
(b) [Monotonicity] If \( A \subseteq B \subseteq \mathbb{R} \) then \( m^*(A) \leq m^*(B) \).

(c) \( m^*(\emptyset) = 0 \)

(d) If \( A \) is countable then \( m^*(A) = 0 \).

\[ \text{Proof:} \quad \text{(a) Obvious.} \]

(b) This follows from the fact that a ccoi for \( B \) is also a ccoi for \( A \) if \( A \subseteq B \).

(c) Note that \( \{(0, \varepsilon)\} \) is a ccoi for \( \emptyset \) for all \( \varepsilon > 0 \). Hence \( 0 \leq m^*(\emptyset) \leq l((0, \varepsilon)) = \varepsilon \) for all \( \varepsilon > 0 \). So \( m^*(\emptyset) = 0 \).

(d) Let \( a_1, a_2, a_3, \ldots \) be an enumeration of \( A \). Let \( \varepsilon > 0 \). Then \( \{(a_n - \frac{\varepsilon}{2n+1}, a_n + \frac{\varepsilon}{2n+1}) : n \in \mathbb{N} \} \) is a ccoi for \( A \). Hence

\[
0 \leq m^*(A) \leq \sum_{n=1}^{\infty} l \left( a_n - \frac{\varepsilon}{2n+1}, a_n + \frac{\varepsilon}{2n+1} \right) = \sum_{n=1}^{\infty} \frac{\varepsilon}{2^n} = \varepsilon
\]

So \( m^*(A) = 0 \). \[ \square \]

Recall that we would like the outer measure to have certain properties. We show that the outer measure of an interval is indeed the length of the interval (this is by no means obvious), that the outer measure is translation invariant and that the outer measure is countable subadditive (this is the closest we can get to countable additive).

Proposition 2.3 Let \( I \) be an interval. Then \( m^*(I) = l(I) \).

\[ \text{Proof:} \quad \text{Suppose first that } I \text{ is compact, say } I = [a, b]. \]

Note that \( \{(a-\varepsilon, b+\varepsilon)\} \) is a ccoi for \( [a, b] \) for all \( \varepsilon > 0 \). Hence \( m^*([a, b]) \leq l((a-\varepsilon, b+\varepsilon)) = b - a + 2\varepsilon \) for all \( \varepsilon > 0 \). So \( m^*([a, b]) \leq b - a \).

Let \( \{I_n : n \in \mathbb{N}\} \) be a ccoi of \( [a, b] \). We want to show that \( \sum_{n=1}^{\infty} l(I_n) \geq b - a \). Hence we may assume that \( I_n \) is bounded for all \( n \in \mathbb{N} \). Since \( [a, b] \) is compact, it follows from the Heine-Borel Theorem that this ccoi has a finite subcovering for \( [a, b] \), say \( \{I_1, \ldots, I_m\} \).

We will construct a finite list of bounded intervals \( \{(a_1, b_1), \ldots, (a_k, b_k)\} \) such that

- \( a_1 < a < b_1 \)
- \( a_{i+1} < b_i < b_{i+1} \) for \( i = 1, 2, \ldots, k - 1 \)
- \( b \leq b_k \)

Put \( I_j = (a_j, b_j) \) for \( 1 \leq j \leq m \). Since \( a \in [a, b] \subseteq \bigcup_{j=1}^{m} I_j \), we have that \( a \in I_j \) for some \( 1 \leq j \leq m \), say \( a \in I_1 = (a_1, b_1) \). If \( b \leq b_1 \), we stop. So assume that \( b_1 < b \). Then \( a < b_1 < b \) and so \( b_1 \in I_j = (a_j, b_j) \) for some \( 1 \leq j \leq m \). Clearly \( j \geq 2 \) (since \( b_1 < b_j \)), say \( j = 2 \); \( b_1 \in I_2 = (a_2, b_2) \). If \( b \leq b_2 \), we stop. So we may assume that \( b_2 < b \). Then \( a < b_1 < b_2 < b \) and so \( b_2 \in I_j = (a_j, b_j) \) for some \( 1 \leq j \leq m \). Clearly \( j \geq 3 \) (since \( b_1 < b_2 < b_j \)), say \( j = 3 \); \( b_2 \in I_3 = (a_3, b_3) \). Continuing this way, this process must stop, say after \( k \) steps, since we have a finite subcovering. Then \( b \leq b_k \) and we have constructed our desired list.

Then \( b_i - a_{i+1} > 0 \) for \( i = 1, \ldots, k - 1 \). Hence we get that
\[ \sum_{j=1}^{k} l(I_k) = (b_1 - a_1) + (b_2 - a_2) + \cdots + (b_k - a_k) \]
\[ = b_k + (b_{k-1} - a_k) + (b_{k-2} - a_{k-1}) + \cdots + (b_1 - a_2) - a_1 \]
\[ > b_k - a_1 \]
\[ > b - a \]

So
\[ \sum_{n=1}^{\infty} l(I_n) \geq \sum_{n=1}^{k} l(I_k) > b - a \]

Since this is true for any ccoi \( \{I_n\}_{n \geq 1} \) of \([a, b]\), we find that \( m^*([a, b]) \geq b - a \).
Hence \( m^*([a, b]) = b - a = l([a, b]) \).

Next, suppose that \( I \) is a bounded interval, say with endpoints \( a \) and \( b \). Then \( I \subseteq [a, b] \) and so
\[ m^*(I) \leq m^*([a, b]) = b - a \]
by monotonicity. For \( 0 < \varepsilon < \frac{b - a}{2} \), we have that \([a + \varepsilon, b - \varepsilon] \subseteq I \) and so
\[ m^*(I) \geq m^*([a + \varepsilon, b - \varepsilon]) = b - a - 2\varepsilon \]
by monotonicity. Since this is true for all \( 0 < \varepsilon < \frac{b - a}{2} \), we get that
\[ m^*(I) \geq b - a \]
Hence \( m^*(I) = b - a = l(I) \).

Finally, suppose that \( I \) is an unbounded interval, say \( I \) is unbounded above. Let \( a \in I \). Then for all \( n \in \mathbb{N} \), we have that \([a, a + n] \subseteq I \) and so \( m^*(I) \geq m^*([a, a + n]) = n \) by monotonicity. Since this is true for all \( n \in \mathbb{N} \), we get that \( m^*(I) = +\infty = l(I) \). \hfill \Box

**Corollary 2.4** Any non-degenerate interval of real numbers is uncountable.

**Proof:** Let \( I \) be a non-degenerate interval. Then \( l(I) \neq 0 \). By Proposition 2.3, we have that \( m^*(I) = l(I) > 0 \). It follows from Proposition 2.2(d) that \( I \) is not countable. \hfill \Box

**Proposition 2.5** The outer measure is translation invariant: \( m^*(A + y) = m^*(A) \) for all \( A \subseteq \mathbb{R} \) and for all \( y \in \mathbb{R} \).

**Proof:** Let \( A \subseteq \mathbb{R} \) and \( y \in \mathbb{R} \). Note that if \( I \) is an open interval, then \( I + y \) is an open interval and \( l(I + y) = l(y) \). It follows that if \( \{I_n\}_{n \geq 1} \) is a ccoi of \( A \) then \( \{I_n + y\}_{n \geq 1} \) is a ccoi of \( A + y \) and
\[ \sum_{n=1}^{\infty} l(I_n) = \sum_{n=1}^{\infty} l(I_n + y). \]
Hence \( m^*(A) \geq m^*(A + y) \).
Using this general inequality, we get \( m^*(A) = m^*((A + y) + (-y)) \leq m^*(A + y) \).
Thus \( m^*(A + y) = m^*(A) \). \hfill \Box

**Proposition 2.6** (Countable Subadditivity) Let \( A_n \subseteq \mathbb{R} \) for \( n \in \mathbb{N} \). Then
\[ m^*(\bigcup_{n=1}^{\infty} A_n) \leq \sum_{n=1}^{\infty} m^*(A_n) \]
**Proof:** This is obvious if \( \sum_{n=1}^{\infty} m^*(A_n) = +\infty \). Hence we may assume that \( \sum_{n=1}^{\infty} m^*(A_n) \neq +\infty \). In particular, \( m^*(A_n) \neq +\infty \) for all \( n \geq 1 \).

Let \( \epsilon > 0 \). For \( n \geq 1 \), there exists a ccoi \( \{I^n_k : k \in \mathbb{N}\} \) of \( A_n \) with

\[
\sum_{k=1}^{\infty} l(I^k_n) < m^*(A_n) + \frac{\epsilon}{2^n}
\]

since \( m^*(A_n) \) is an infimum. By Proposition 1.5, \( \{I^k_n : k, n \in \mathbb{N}\} \) is a ccoi of \( \bigcup_{n=1}^{\infty} A_n \). Hence we get

\[
m^*\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{k,n=1}^{\infty} l(I^k_n) < \sum_{k=1}^{\infty} \left( m^*(A_n) + \frac{\epsilon}{2^n} \right) = \left( \sum_{n=1}^{\infty} m^*(A_n) \right) + \epsilon
\]

Since this is true for all \( \epsilon > 0 \), we find that \( m^*\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} m^*(A_n) \). \( \square \)

**Remarks:**

(a) This result holds (of course) for a finite number of sets:

\[
m^*\left(\bigcup_{k=1}^{n} A_k\right) \leq \sum_{k=1}^{n} m^*(A_k)
\]

We can use countable subadditivity with \( A_k = \emptyset \) for all \( k > n \) since \( m^*(\emptyset) = 0 \).

(b) Note that we can not improve this proposition to have equality even if \( \{A_n\}_{n \geq 1} \) is a disjoint collection. Using the Axiom of Choice, we can construct sets \( A, B \subseteq \mathbb{R} \) with \( A \cap B = \emptyset \) and \( m^*(A \cup B) < m^*(A) + m^*(B) \).

\[
\Box
\]

### 2.2 \( \sigma \)-Algebras

In order to establish countable additivity for some special disjoint collections, we will use a structure called \( \sigma \)-algebras, which are special case of algebras.

**Definition 2.7** Let \( X \) be a set and \( \mathcal{A} \) a collection of subsets of \( X \) (so \( \mathcal{A} \subseteq \mathcal{P}(X) \)).

(a) \( \mathcal{A} \) is an algebra if

1. \( \forall A \in \mathcal{A} : \bar{A} \in \mathcal{A} \)
2. \( \forall A, B \in \mathcal{A} : A \cup B \in \mathcal{A} \)

(b) \( \mathcal{A} \) is a \( \sigma \)-algebra if

1. \( \forall A \in \mathcal{A} : \bar{A} \in \mathcal{A} \)
2. \( \forall A_1, A_2, \ldots \in \mathcal{A} : \bigcup_{n=1}^{\infty} A_n \in \mathcal{A} \)

\( \Box \)
Remarks:

(a) A \(\sigma\)-algebra is clearly an algebra but not every algebra is a \(\sigma\)-algebra.

(b) For any set \(X\), there are at least two algebras: \(\{\emptyset, X\}\) and \(\mathcal{P}(X)\). Note that \(\mathcal{P}(X)\) is actually a \(\sigma\)-algebra.

(c) Let \(\mathcal{A}\) be an algebra and \(A, B \in \mathcal{A}\). Then \(\widetilde{A}, \widetilde{B} \in \mathcal{A}\) by (1). So \(\widetilde{A} \cup \widetilde{B} \in \mathcal{A}\) by (2). Again by (1), \(\mathcal{A} \ni \widetilde{A} \cup \widetilde{B} = \widetilde{A} \cap \widetilde{B} = A \cap B\). So \(A \cap B \in \mathcal{A}\).

Similarly, if \(\mathcal{A}\) is a \(\sigma\)-algebra and \(A_n \in \mathcal{A}\) for all \(n \in \mathbb{N}\), then \(\bigcap_{n=1}^{\infty} A_n \in \mathcal{A}\).

(d) Let \(\mathcal{A}\) be an algebra. Pick \(A \in \mathcal{A}\). Then \(e_A \in \mathcal{A}\) by (1) and so \(\emptyset = A \cap e_A \in \mathcal{A}\) by (c). Hence \(X = e_\emptyset \in \mathcal{A}\) by (1).

(e) If we replace unions with intersections in conditions (2), we get equivalent definitions. \(\blacksquare\)

The next proposition shows that every collection of subsets of \(X\) generates an algebra and a \(\sigma\)-algebra.

**Proposition 2.8** Let \(X\) be a set and \(C \subseteq \mathcal{P}(X)\). Then the following holds:

(a) There exists a smallest algebra on \(X\) containing \(C\).

(b) There exists a smallest \(\sigma\)-algebra on \(X\) containing \(C\).

**Proof:** We prove (a). The proof of (b) is similar.

Let \(\mathcal{A}\) be the intersection of all the algebras containing \(C\). Note that \(\mathcal{A}\) is well-defined since \(\mathcal{P}(X)\) is an algebra containing \(C\). Moreover, \(C \subseteq \mathcal{A}\).

We show that \(\mathcal{A}\) is an algebra. Let \(A \in \mathcal{A}\). If \(B\) is an algebra containing \(C\) then \(A \in B\) and so \(\widetilde{A} \in B\). Hence \(\widetilde{A} \in \mathcal{A}\). Let \(A, B \in \mathcal{A}\). If \(B\) is an algebra containing \(C\) then \(A, B \in B\) and so \(A \cup B \in B\). Hence \(A \cup B \in \mathcal{A}\).

Finally, we show that \(\mathcal{A}\) is the smallest algebra containing \(C\). Let \(\mathcal{B}\) be an algebra containing \(C\). Since \(\mathcal{A}\) is the intersection of all the algebras containing \(C\), we have that \(\mathcal{A} \subseteq \mathcal{B}\). \(\square\)

**Remark:** This algebra is called the algebra generated by \(C\) and is sometimes denoted by \(\langle C \rangle\). The \(\sigma\)-algebra generated by \(C\) is denoted by \(\langle C \rangle_\sigma\). \(\blacksquare\)

We finish this section with a technical lemma that will be used later on in the proof of countable additivity of measurable sets.

**Lemma 2.9** Let \(\mathcal{A}\) be an algebra on \(X\) and \(\{A_n : n \in \mathbb{N}\}\) a countable collection of elements of \(\mathcal{A}\). Then there exists a countable disjoint collection \(\{B_n : n \in \mathbb{N}\}\) of elements of \(\mathcal{A}\) with \(\bigcup_{n=1}^{\infty} A_n = \bigcup_{n=1}^{\infty} B_n\).

**Proof:** Put \(A_0 = \emptyset \in \mathcal{A}\). For \(n \geq 1\), we put

\[B_n = A_n \setminus (A_0 \cup A_1 \cup A_2 \cup \cdots \cup A_{n-1}) = A_n \cap \bigcup_{k=0}^{n-1} A_k\]

Note that \(B_n \in \mathcal{A}\) for all \(n \geq 1\) since \(\mathcal{A}\) is an algebra.
Next, we show that the collection \( \{B_n : n \in \mathbb{N}\} \) is a disjoint collection. Let \( 1 \leq m < n \). Then

\[
B_n = A_n \cap \bigcup_{k=0}^{n-1} A_k = A_n \cap \tilde{A_0} \cap \tilde{A_1} \cap \cdots \cap \tilde{A_{n-1}} \subseteq \tilde{A_m} \quad \text{and} \quad B_m = A_m \cap \bigcup_{k=0}^{m-1} A_k \subseteq A_m
\]

Hence \( B_n \cap B_m = \emptyset \).

Finally, we show that \( \bigcup_{n=1}^{\infty} A_n = \bigcup_{n=1}^{\infty} B_n \). Since \( B_n \subseteq A_n \) for all \( n \geq 1 \), we have that \( \bigcup_{n=1}^{\infty} B_n \subseteq \bigcup_{n=1}^{\infty} A_n \). Let \( x \in \bigcup_{n=1}^{\infty} A_n \). Then \( x \in A_n \) for some \( n \geq 1 \). Let \( m \) be the smallest index with \( x \in A_m \). Then \( m \geq 1 \) since \( A_0 = \emptyset \) and \( x \notin A_i \) for \( 0 \leq i \leq m - 1 \). Hence \( x \notin \bigcup_{i=0}^{m-1} A_i \). So

\[
x \in A_m \cap \bigcup_{i=0}^{m-1} A_i = B_m
\]

Thus \( \bigcup_{n=1}^{\infty} A_n \subseteq \bigcup_{n=1}^{\infty} B_n \).

It follows that \( \bigcup_{n=1}^{\infty} A_n = \bigcup_{n=1}^{\infty} B_n \). \( \square \)

### 2.3 Measurable Sets

In this section, we define a class of sets for which the outer measure is countable additive.

**Definition 2.10** Let \( E \subseteq \mathbb{R} \). Then \( E \) is **measurable** if

\[
m^*(A) = m^*(A \cap E) + m^*(A \cap \bar{E}) \quad \text{for all } A \subseteq \mathbb{R}
\]

**Remarks:**

(a) Let \( A, E \subseteq \mathbb{R} \). Then \( A = A \cap \mathbb{R} = A \cap (E \cup \bar{E}) = (A \cap E) \cup (A \cap \bar{E}) \). So by countable subadditivity, we have that

\[
m^*(A) \leq m^*(A \cap E) + m^*(A \cap \bar{E})
\]

Hence we get that \( E \) is measurable if and only if

\[
m^*(A \cap E) + m^*(A \cap \bar{E}) \leq m^*(A) \quad \text{for all } A \subseteq \mathbb{R}
\]

This inequality is clearly satisfied if \( m^*(A) = +\infty \). So we finally get that \( E \) is measurable if and only if

\[
m^*(A \cap E) + m^*(A \cap \bar{E}) \leq m^*(A) \quad \text{for all } A \subseteq \mathbb{R} \text{ with } m^*(A) \neq +\infty
\]

(b) It is not at all clear that the outer measure is countable additive on measurable sets. But the definition of measurable sets does have something to do with disjoint sets (note that \( A \cap E \) and \( A \cap \bar{E} \) are disjoint). Recall that we mentioned there exist disjoint sets \( A, B \subseteq \mathbb{R} \) with \( m^*(A \cup B) < m^*(A) + m^*(B) \). That is not possible if \( A \) or \( B \) are measurable. Indeed, suppose \( B \) is measurable. Since \( A \) and \( B \) are disjoint, we get that

\[
(A \cup B) \cap B = (A \cap B) \cup (B \cap B) = \emptyset \cup B = B
\]

and

\[
(A \cup B) \cap \bar{B} = (A \cap \bar{B}) \cup (B \cap \bar{B}) = A \cup \emptyset = A
\]

So

\[
m^*(A \cup B) = m^*((A \cup B) \cap B) + m^*((A \cup B) \cap \bar{B}) = m^*(B) + m^*(A)
\]

since \( B \) is measurable.
(c) $\emptyset$ and $\mathbb{R}$ are clearly measurable.

(d) Since the definition of measurable set is symmetric in $E$ and $\bar{E}$, we get that $E$ is measurable if and only if $\bar{E}$ is measurable.

**Definition 2.11** The *Lebesgue measure of a measurable set* $E$ is the outer measure of $E$ and is denoted by $m(E)$.

**Remark:** The notation $m(E)$ implies that the set $E$ is measurable. If we do not know if a certain set is measurable, we use the $m^*$ notation (for outer measure).

The next proposition gives us a family of measurable sets.

**Proposition 2.12** Let $E \subseteq \mathbb{R}$ with $m^*(E) = 0$. Then $E$ is measurable.

**Proof:** Let $A \subseteq \mathbb{R}$. Then $A \cap E \subseteq E$ and $A \cap \bar{E} \subseteq A$. Hence

$$m^*(A \cap E) + m^*(A \cap \bar{E}) \leq m^*(E) + m^*(A) = m^*(A)$$

by monotonicity. So $E$ is measurable. $\square$

Before establishing countable additivity, we show that the collection of all measurable sets forms a $\sigma$-algebra.

**Lemma 2.13** The union of finitely many measurable sets is measurable.

**Proof:** Using induction, it is enough to prove this result for two measurable sets. So let $E_1$ and $E_2$ be measurable sets. Let $A \subseteq \mathbb{R}$. Note that

$$A \cap \bar{E}_1 \cup E_2 = A \cap \bar{E}_1 \cap \bar{E}_2$$

and

$$A \cap (E_1 \cup E_2) = (A \cap E_1) \cup (A \cap E_2) = (A \cap E_1) \cup (A \cap \bar{E}_1 \cap \bar{E}_2)$$

Hence using countable subadditivity, the measurability of $E_2$ and the measurability of $E_1$, we get

$$m^*(A \cap (E_1 \cup E_2)) + m^*(A \cap \bar{E}_1 \cup E_2) = m^*((A \cap E_1) \cup (A \cap \bar{E}_1 \cap E_2)) + m^*(A \cap \bar{E}_1 \cap \bar{E}_2) \leq m^*(A \cap E_1) + m^*(A \cap \bar{E}_1 \cap E_2) + m^*(A \cap \bar{E}_1 \cap \bar{E}_2) = m^*(A \cap E_1) + m^*(A \cap \bar{E}_1) = m^*(A)$$

So $E_1 \cup E_2$ is measurable. $\square$

**Lemma 2.14** Let $\{E_1, \ldots, E_n\}$ be a finite disjoint collection of measurable sets. Then

$$m^*(A \cap \left( \bigcup_{i=1}^{n} E_i \right)) = \sum_{i=1}^{n} m^*(A \cap E_i)$$

for all $A \subseteq \mathbb{R}$. 26
Proof: Let $A \subseteq \mathbb{R}$. We use induction on $n$. The result is obvious for $n = 1$. So assume that $n \geq 2$. Since $E_i \cap E_n = \emptyset$ for $1 \leq i \leq n - 1$, we have that

$$(A \cap (\cup_{i=1}^n E_i)) \cap E_n = A \cap ((\cup_{i=1}^n E_i) \cap E_n) = A \cap (\cup_{i=1}^n (E_i \cap E_n)) = A \cap E_n$$

and

$$(A \cap (\cup_{i=1}^n E_i)) \cap \tilde{E}_n = A \cap (\cup_{i=1}^n E_i) \cap \tilde{E}_n = A \cap (\cup_{i=1}^n (E_i \cap \tilde{E}_n)) = A \cap (\cup_{i=1}^n E_i)$$

Using the measurability of $E_n$ and induction, we find

$$m^*(A \cap (\cup_{i=1}^n E_i)) = m^*(A \cap (\cup_{i=1}^n E_i) \cap E_n) + m^*((A \cap (\cup_{i=1}^n E_i)) \cap \tilde{E}_n)$$

$$= m^*(A \cap E_n) + m^*(A \cap (\cup_{i=1}^n E_i))$$

$$= m^*(A \cap E_n) + \sum_{i=1}^{n-1} m^*(A \cap E_i)$$

$$= \sum_{i=1}^n m^*(A \cap E_i)$$

Proposition 2.15 The union of countably many measurable sets is measurable.

Proof: Let $\{E_n\}_{n \geq 1}$ be a countable collection of measurable sets. Note that we already proved that the collection of all measurable sets is an algebra (this follows from Lemma 2.13 and the fact that $E$ is measurable if and only if $\tilde{E}$ is measurable). Since we want to show that $\cup_{n=1}^\infty E_n$ is measurable, it follows from Lemma 2.9 that we may assume that $\{E_n\}_{n \geq 1}$ is a disjoint collection.

Put $E = \cup_{i=1}^\infty E_i$. For $n \geq 1$, put $F_n = \cup_{i=1}^n E_i$. Note that $F_n$ is measurable for all $n \geq 1$ by Lemma 2.13.

Let $A \subseteq \mathbb{R}$. Let $n \geq 1$. Since $F_n \subseteq E$, we get that $\tilde{E} \subseteq \tilde{F}_n$ and so

$$m^*(A \cap \tilde{E}) \leq m^*(A \cap \tilde{F}_n)$$

by monotonicity. By Lemma 2.14, we have

$$m^*(A \cap F_n) = m^*(A \cap (\cup_{i=1}^n E_i)) = \sum_{i=1}^n m^*(A \cap E_i)$$

Since $F_n$ is measurable, we get

$$m^*(A \cap \tilde{E}) + \sum_{i=1}^n m^*(A \cap E_i) \leq m^*(A \cap \tilde{F}_n) + m^*(A \cap F_n) = m^*(A)$$

So

$$m^*(A \cap \tilde{E}) + \sum_{i=1}^n m^*(A \cap E_i) \leq m^*(A) \quad \text{for all } n \geq 1$$

Taking the limit as $n \to \infty$, we find

$$m^*(A \cap \tilde{E}) + \sum_{i=1}^\infty m^*(A \cap E_i) \leq m^*(A)$$

But

$$m^*(A \cap E) = m^*(A \cap (\cup_{i=1}^\infty E_i)) = m^*(\cup_{i=1}^\infty (A \cap E_i)) \leq \sum_{i=1}^\infty m^*(A \cap E_i)$$

27
by countable subadditivity. Hence
\[ m^*(A \cap E) + m^*(A \cap \tilde{E}) \leq m^*(A \cap E) + \sum_{i=1}^{\infty} m^*(A \cap E_i) \leq m^*(A) \]

So \( E = \bigcup_{i=1}^{\infty} E_i \) is measurable. \( \square \)

**Theorem 2.16** The collection of all measurable sets forms a \( \sigma \)-algebra.

**Proof:** This follows from Proposition 2.15 and the fact that \( E \) is measurable if and only if \( \tilde{E} \) is measurable. \( \square \)

Now we can prove that countable additivity holds for disjoint measurable sets.

**Proposition 2.17 (Countable Additivity)** Let \( \{E_n\}_{n \geq 1} \) be a countable disjoint collection of measurable sets. Then
\[ m\left( \bigcup_{n=1}^{\infty} E_n \right) = \sum_{n=1}^{\infty} m(E_n) \]

**Proof:** By countable subadditivity, we have
\[ m\left( \bigcup_{i=1}^{\infty} E_i \right) \leq \sum_{i=1}^{\infty} m(E_i) \]

Let \( n \geq 1 \). By Lemma 2.14 (with ‘\( A = \mathbb{R} \)’) and monotonicity, we get
\[ m\left( \bigcup_{i=1}^{n} E_i \right) \geq m\left( \bigcup_{i=1}^{n} E_i \right) = \sum_{i=1}^{n} m(E_i) \]

So
\[ \sum_{i=1}^{n} m(E_i) \leq m\left( \bigcup_{i=1}^{\infty} E_i \right) \quad \text{for all} \ n \geq 1 \]

Taking the limit as \( n \to \infty \), we find
\[ \sum_{i=1}^{\infty} m(E_i) \leq m\left( \bigcup_{i=1}^{\infty} E_i \right) \]

Hence \( m\left( \bigcup_{i=1}^{\infty} E_i \right) = \sum_{i=1}^{\infty} m(E_i) \). \( \square \)

The following formula can be quite useful.

**Proposition 2.18 (Excision)** Let \( A, B \) be sets with \( B \subseteq A \). Suppose that \( B \) is measurable and \( m(B) \neq +\infty \). Then
\[ m^*(A \setminus B) = m^*(A) - m(B) \]
Proof: Since $B$ is measurable, we have that
\[ m^*(A) = m^*(A \cap B) + m^*(A \cap \widetilde{B}) = m(B) + m^*(A \setminus B) \]
Since $m(B) \neq +\infty$, we can subtract $m(B)$ from both sides. □

Our next goal is to show that a large family of sets (which includes the open sets and the closed sets) are measurable.

**Definition 2.19** Let $A \subseteq \mathbb{R}$. Then $A$ is a **Borel set** if $A$ is in the $\sigma$-algebra generated by the collection of all open sets.

**Remark**: Clearly, every open set is a Borel set. Since $\sigma$-algebra are closed under complements, it follows from Proposition 1.28 that every closed set is also a Borel set. □

**Lemma 2.20** The interval $(a, +\infty)$ is measurable for all $a \in \mathbb{R}$.

**Proof**: Let $a \in \mathbb{R}$. Let $A \subseteq \mathbb{R}$. Let $\{I_n\}_{n \geq 1}$ be a ccoi of $A$. For $n \geq 1$, put $I_n^1 = I_n \cap (a, +\infty)$ and
\[ I_n^2 = I_n \cap (\widetilde{(a, +\infty)}) = I_n \cap (-\infty, a]. \]
We easily get that
\[ A \cap (a, +\infty) \subseteq (\cup_{n=1}^{\infty} I_n) \cap (a, +\infty) = \cup_{n=1}^{\infty} (I_n \cap (a, +\infty)) = \cup_{n=1}^{\infty} I_n^1 \]
and
\[ A \cap (\widetilde{(a, +\infty)}) = A \cap (-\infty, a] \subseteq (\cup_{n=1}^{\infty} I_n) \cap (-\infty, a] = \cup_{n=1}^{\infty} (I_n \cap (-\infty, a]) = \cup_{n=1}^{\infty} I_n^2 \]
It follows from countable subadditivity that
\[ m^*(A \cap (a, +\infty)) + m^*(A \cap (\widetilde{(a, +\infty)})) \leq \sum_{n=1}^{\infty} m^*(I_n^1) + \sum_{n=1}^{\infty} m^*(I_n^2) = \sum_{n=1}^{\infty} (m^*(I_n^1) + m^*(I_n^2)) \]
But $I_n^1$ and $I_n^2$ are intervals and $l(I_n^1) + l(I_n^2) = l(I_n)$ for all $n \geq 1$. Indeed, let $n \geq 1$. We consider five cases for $I_n$:

<table>
<thead>
<tr>
<th>$I_n$</th>
<th>$I_n^1 = I_n \cap (a, +\infty)$</th>
<th>$I_n^2 = I_n \cap (-\infty, a]$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(-\infty, +\infty)$</td>
<td>$(a, +\infty)$</td>
<td>$(-\infty, a]$</td>
</tr>
<tr>
<td>$(-\infty, d)$</td>
<td>$\emptyset$ if $d \leq a$</td>
<td>$(-\infty, a]$ if $d \leq a$</td>
</tr>
<tr>
<td></td>
<td>$(a, d]$ if $d &gt; a$</td>
<td>$(-\infty, a]$ if $d &gt; a$</td>
</tr>
<tr>
<td>$(c, +\infty)$</td>
<td>$(a, +\infty)$ if $c &lt; a$</td>
<td>$(c, a]$ if $c &lt; a$</td>
</tr>
<tr>
<td></td>
<td>$(c, \infty)$ if $c \geq a$</td>
<td>$\emptyset$ if $c \geq a$</td>
</tr>
<tr>
<td>$(c, d)$</td>
<td>$\emptyset$ if $d \leq a$</td>
<td>$(c, d]$ if $d \leq a$</td>
</tr>
<tr>
<td></td>
<td>$(a, d]$ if $c &lt; a &lt; d$</td>
<td>$(c, a]$ if $c &lt; a &lt; d$</td>
</tr>
<tr>
<td></td>
<td>$(c, d]$ if $a \leq c$</td>
<td>$\emptyset$ if $a \leq c$</td>
</tr>
<tr>
<td>$\emptyset$</td>
<td>$\emptyset$</td>
<td>$\emptyset$</td>
</tr>
</tbody>
</table>
Hence we get
\[ m^*(A \cap (a, +\infty)) + m^*(A \cap (a, +\infty)) \leq \sum_{n=1}^{\infty} (m^*(I_n^1) + m^*(I_n^2)) = \sum_{n=1}^{\infty} (l(I_n^1) + l(I_n^2)) = \sum_{n=1}^{\infty} l(I_n) \]
So
\[ m^*(A \cap (a, +\infty)) + m^*(A \cap (a, +\infty)) \leq \sum_{n=1}^{\infty} l(I_n) \text{ for every countable set } \{I_n\}_{n \geq 1} \text{ of } A \]
Thus
\[ m^*(A \cap (a, +\infty)) + m^*(A \cap (a, +\infty)) \leq m^*(A) \]
Hence \((a, +\infty)\) is measurable. \(\square\)

**Theorem 2.21** Every Borel set is measurable.

**Proof:** Let \(\mathcal{M}\) be the collection of all measurable sets. Note that \(\mathcal{M}\) is a \(\sigma\)-algebra.
First, we prove that every open interval is measurable. Note that \(\emptyset, I = (-\infty, +\infty)\) and \(I = (a, +\infty)\) for \(a \in \mathbb{R}\) are measurable (by Lemma 2.20). So \((-\infty, a] = (a, +\infty) \in \mathcal{M}\) for all \(a \in \mathbb{R}\). Hence \((-\infty, b) = \cap_{n=1}^{\infty} (-\infty, b - \frac{1}{n}] \in \mathcal{M}\) for all \(b \in \mathbb{R}\). Thus \((a, b) = (-\infty, b) \cap (a, +\infty) \in \mathcal{M}\) for all \(a, b \in \mathbb{R}\) with \(a < b\).

Next, we show that every open set is measurable. Let \(\mathcal{O}\) be an open set. Then \(\mathcal{O}\) is a countable union of open intervals by Theorem 1.24. Since every open interval is measurable, we get that \(\mathcal{O}\) is measurable.

Finally, we show that every Borel set is measurable. Let \(\mathcal{B}\) be the collection of all Borel sets. Then \(\mathcal{B}\) is the smallest \(\sigma\)-algebra containing all the open sets. Since \(\mathcal{M}\) is a \(\sigma\)-algebra containing every open set, we have that \(\mathcal{B} \subseteq \mathcal{M}\). So every Borel set is measurable. \(\square\)

**Remark:** Not every measurable set is a Borel set.

We finish this section with a couple more properties of measurable sets.

**Proposition 2.22** Let \(E \subseteq \mathbb{R}\) be a measurable set. Then \(E + x\) is measurable for all \(x \in \mathbb{R}\).

**Proof:** We use the following properties:
- \(A \cap (B + x) = ((A - x) \cap B) + x\) for all \(A, B \subseteq \mathbb{R}\) and all \(x \in \mathbb{R}\)
- \(\widetilde{A + x} = \widetilde{A} + x\) for all \(A \subseteq \mathbb{R}\) and all \(x \in \mathbb{R}\)

Let \(x \in \mathbb{R}\). Let \(A \subseteq \mathbb{R}\). Then
\[ m^*(A \cap (E + x)) = m^*(A \cap (\widetilde{E} + x)) = m^*((A - x) \cap \widetilde{E}) + x) \]
Using Proposition 2.5 (namely the outer measure is translation invariant), the measurability of \(E\) and again Proposition 2.5, we get
\[ m^*(A \cap (E + x)) + m^*(A \cap (E + x)) = m^*((A - x) \cap E) + x) + m^*((A - x) \cap \widetilde{E}) + x) \]
\[ = m^*((A - x) \cap E) + m^*((A - x) \cap \widetilde{E}) \]
\[ = m^*(A - x) \]
\[ = m^*(A) \]
So \(E + x\) is measurable. \(\square\)
Proposition 2.23 (Continuity of Measure) Let \( \{E_n\}_{n \geq 1} \) be a sequence of measurable sets. Then the following holds:

(a) If \( E_1 \subseteq E_2 \subseteq E_3 \subseteq \cdots \) then \( m \left( \bigcup_{n=1}^{\infty} E_n \right) = \lim_{n \to \infty} m(E_n) \).

(b) If \( E_1 \supseteq E_2 \supseteq E_3 \supseteq \cdots \) and \( m(E_1) \neq +\infty \) then \( m \left( \bigcap_{n=1}^{\infty} E_n \right) = \lim_{n \to \infty} m(E_n) \).

Proof: (a) Note that \( \{m(E_n)\}_{n \geq 1} \) is an increasing sequence of extended real numbers (by monotonicity) and hence converges to an extended real number.

Put \( E = \bigcup_{n=1}^{\infty} E_n \).

Suppose first that \( m(E_k) = +\infty \) for some \( k \geq 1 \). Since \( E_k \subseteq E \), we have that \( +\infty = m(E_k) \leq m(E) \) by monotonicity. Hence \( \lim_{n \to \infty} m(E_n) = +\infty = m(E) \).

Suppose next that \( m(E_k) \neq +\infty \) for all \( k \geq 1 \).

Put \( E_0 = \emptyset \). Then \( \bigcup_{k=1}^{\infty} E_k = \bigcup_{k=1}^{\infty} (E_k \setminus E_{k-1}) \). Indeed, let \( x \in \bigcup_{k=1}^{\infty} E_k \). Then \( x \in E_k \) for some \( k \geq 1 \). Let \( n \) be minimal with \( x \in E_n \). Then \( n \geq 1 \), \( x \in E_n \) and \( x \notin E_{n-1} \). So \( x \in E_n \setminus E_{n-1} \subseteq \bigcup_{k=1}^{\infty} (E_k \setminus E_{k-1}) \). Hence \( \bigcup_{k=1}^{\infty} E_k \subseteq \bigcup_{k=1}^{\infty} (E_k \setminus E_{k-1}) \). Clearly, \( \bigcup_{k=1}^{\infty} (E_k \setminus E_{k-1}) \subseteq \bigcup_{k=1}^{\infty} E_k \) and so \( \bigcup_{k=1}^{\infty} E_k = \bigcup_{k=1}^{\infty} (E_k \setminus E_{k-1}) \). Finally, suppose \( (E_k \setminus E_{k-1}) \cap (E_l \setminus E_{l-1}) \neq \emptyset \) for some \( 1 \leq k < l \), say \( x \in (E_k \setminus E_{k-1}) \cap (E_l \setminus E_{l-1}) \). Then \( x \notin E_{l-1} \), a contradiction since \( x \in E_k \subseteq E_{k+1} \subseteq \cdots \subseteq E_{l-1} \). So \( \bigcup_{k=1}^{\infty} (E_k \setminus E_{k-1}) \) is a disjoint union.

Using countable additivity and excision (note that \( m(E_k) \neq +\infty \) for all \( k \geq 1 \)), we get

\[
m(E) = m \left( \bigcup_{k=1}^{\infty} (E_k \setminus E_{k-1}) \right) = \sum_{k=1}^{+\infty} m(E_k \setminus E_{k-1}) = \lim_{n \to \infty} \left( \sum_{k=1}^{n} (m(E_k) - m(E_{k-1})) \right) = \lim_{n \to \infty} (m(E_n) - m(E_0)) = \lim_{n \to \infty} m(E_n)
\]

(b) Note that \( \{m(E_n)\}_{n \geq 1} \) is a decreasing sequence of positive real numbers (by monotonicity and the fact that \( m(E_1) \neq +\infty \)) and hence converges to a real number.

Put \( E = \bigcap_{k=1}^{\infty} E_k \). Note that \( \bigcup_{k=1}^{\infty} (E_k \setminus E_{k+1}) = E_1 \setminus E \). Indeed, let \( x \in \bigcap_{k=1}^{\infty} (E_k \setminus E_{k+1}) \). Then \( x \in E_n \setminus E_{n+1} \) for some \( n \geq 1 \). Hence \( x \in E_n \subseteq E_1 \) and \( x \notin E_{n+1} \supseteq \bigcap_{k=1}^{\infty} E_k = E \). So \( \bigcup_{k=1}^{\infty} (E_k \setminus E_{k+1}) \subseteq E_1 \setminus E \).

Let \( x \in E_1 \setminus E \). Then \( x \in E_1 \) and \( x \notin E = \bigcap_{k=1}^{\infty} E_k \). Hence \( x \notin E_k \) for some \( k \geq 1 \). Let \( n \) be the smallest index with \( x \notin E_n \). Then \( n \geq 2 \) and so \( x \in E_{n-1} \). Thus \( x \in E_{n-1} \setminus E_n \subseteq \bigcup_{k=1}^{\infty} (E_k \setminus E_{k+1}) \).

Hence \( E_1 \setminus E \subseteq \bigcup_{k=1}^{\infty} (E_k \setminus E_{k+1}) \). It follows that \( \bigcup_{k=1}^{\infty} (E_k \setminus E_{k+1}) = E_1 \setminus E \). Similarly as in part (a), we prove that \( \{E_k \setminus E_{k+1} : k \geq 1\} \) is a disjoint collection.
Using excision, countable additivity and again excision (note that \( m(E_k) \neq +\infty \) for all \( k \geq 1 \) and \( m(E) \neq +\infty \)), we get

\[
\begin{align*}
m(E_1) - m(E) &= m(E_1 \setminus E) \\
&= m\left( \bigcup_{k=1}^{\infty} (E_k \setminus E_{k+1}) \right) \\
&= \sum_{k=1}^{\infty} m(E_k \setminus E_{k-1}) \\
&= \lim_{n \to \infty} \left( \sum_{k=1}^{n} (m(E_k) - m(E_{k+1})) \right) \\
&= \lim_{n \to \infty} (m(E_1) - m(E_{n+1})) \\
&= m(E_1) - \lim_{n \to \infty} m(E_n)
\end{align*}
\]

Since \( m(E_1) < +\infty \), we get that \( m(E) = \lim_{n \to \infty} m(E_n) \). \( \square \)

Remarks

(a) Part (a) of this theorem remains true if we remove the condition ‘measurable’

(b) Part (b) of this theorem is false if we remove the condition ‘\( m(E_1) < +\infty \)’ or if we remove the condition ‘measurable’.

2.4 Non-Measurable Sets

In this section, we construct a family of non-measurable sets, using the Axiom of Choice.

Definition 2.24 Let \( A \) be a non-empty set of real numbers.

(a) We define the rational equivalent relation (notation: \( \sim \)) on \( A \) by \( x \sim y \) if \( x - y \in \mathbb{Q} \). One easily shows that \( \sim \) is an equivalence relation on \( A \).

- A choice set for \( A \) is a set containing exactly one element of each equivalence class of \( \sim \) on \( A \).

Note that choice sets exist by the Axiom of Choice. \( \square \)

We can now construct non-measurable sets.

Theorem 2.25 Let \( A \) be a bounded set of real numbers with \( m^*(A) > 0 \). Then any choice set for \( A \) is non-measurable.

Proof: Let \( C \) be a choice set for \( A \). Since \( A \) is bounded, we have that \( A \subseteq [-N, N] \) for some \( N > 0 \). Let \( \{q_n\}_{n \geq 1} \) be an enumeration of \( [-2N, 2N] \cap \mathbb{Q} \). Put \( C_n = C + q_n \) for all \( n \geq 1 \).

We show that

\( A \subseteq \bigcup_{n=1}^{\infty} C_n \)

Let \( x \in A \). It follows from the definition of a choice set that \( x \sim c \) for some \( c \in C \). So \( x - c \in \mathbb{Q} \). Since \( x, c \in A \subseteq [-N, N] \), we get that \( x - c \in [-2N, 2N] \). So \( x - c = q_n \) for some \( n \geq 1 \). Thus
Let $x = c + q_n \in C_n$. Hence $A \subseteq \bigcup_{n=1}^{\infty} C_n$. Suppose $C_k \cap C_n \neq \emptyset$ for some $k, n \geq 1$, say $x \in C_k \cap C_n$. Then $x = a + q_n = b + q_n$ for some $a, b \in \mathbb{C}$. So $a - b = q_n - q_k \in \mathbb{Q}$. Hence $a \sim b$. Since $\mathcal{C}$ is a choice set, we have that $a = b$. So $q_k = q_n$ and $k = n$. Thus $\{C_n\}_{n \geq 1}$ is a disjoint collection.

Suppose that $\mathcal{C}$ is measurable. Then for all $n \geq 1$, we have that $C_n$ is measurable by Proposition 2.22 and $m(C_n) = m(\mathcal{C})$ by Proposition 2.5. Using monotonicity and countable additivity, we get

$$m^*(A) \leq m\left(\bigcup_{n=1}^{\infty} C_n\right) = \sum_{n=1}^{\infty} m(C_n) = \sum_{n=1}^{\infty} m(C) = \left\{ \begin{array}{ll} 0 & \text{if } m(\mathcal{C}) = 0 \\ +\infty & \text{if } m(\mathcal{C}) \neq 0 \end{array} \right.$$  

Since $m^*(A) > 0$, we must have that

$$m\left(\bigcup_{n=1}^{\infty} C_n\right) = +\infty$$

Let $x \in \bigcup_{n=1}^{\infty} C_n$. Then $x \in C_n$ for some $n \geq 1$. Hence $x = c + q_n$ for some $c \in \mathbb{C}$. So $|x| = |c + q_n| \leq |c| + |q_n| \leq N + 2N = 3N$. Thus $\bigcup_{n=1}^{\infty} C_n \subseteq [-3N, 3N]$. By monotonicity, we get that $m(\bigcup_{n=1}^{\infty} C_n) \leq m([-3N, 3N]) = 6N$, a contradiction since $m(\bigcup_{n=1}^{\infty} C_n) = +\infty$.

Hence $\mathcal{C}$ is not measurable. \hfill $\square$

The existence of non-measurable sets allows us to prove that countable additivity does not hold for all sets.

**Proposition 2.26** There exist sets $A$ and $B$ such that $A \cap B = \emptyset$ and $m^*(A \cup B) > m^*(A) + m^*(B)$.

**Proof:** Let $\mathcal{C}$ be a non-measurable set. It follows from the definition of measurable set that

$$m^*(D \cap \mathcal{C}) + m^*(D \cap \mathcal{C}) \neq m^*(D) \quad \text{for some set } D$$

Since $D = (D \cap \mathcal{C}) \cup (D \cap \mathcal{C})$, we get that $m^*(D) \leq m^*(D \cap \mathcal{C}) + m^*(D \cap \mathcal{C})$ by countable subadditivity. Hence

$$m^*(D \cap \mathcal{C}) + m^*(D \cap \mathcal{C}) > m^*(D)$$

Put $A = D \cap \mathcal{C}$ and $B = D \cap \mathcal{C}$. Then $A \cap B = \emptyset$ and $A \cup B = D$. So $m^*(A) + m^*(B) > m^*(A \cup B)$. \hfill $\square$

### 2.5 Littlewood’s First Principle

In this section, we prove some equivalent definitions of a measurable set and prove Littlewood’s First Principle: every measurable set is nearly a finite union of open intervals.

Note that the next proposition could be interpreted as an expansion of Littlewood’s First Principle: every measurable set is nearly an open set (or a closed set, a $G_\delta$-set or an $F_\sigma$-set).

**Definition 2.27** Let $A \subseteq \mathbb{R}$.

- (a) $A$ is a $G_\delta$-set if $A$ is the intersection of countably many open sets.

- (b) $A$ is an $F_\sigma$-set if $A$ is the union of countably many closed sets. \hfill $\triangleright$
Remarks:

(a) Since the family of Borel sets is a σ-algebra containing all the open sets, we get that $G_\delta$-sets and $F_\sigma$-sets are Borel sets.

(b) Every open set is a $G_\delta$-set but not every $G_\delta$-set is open.

(c) It follows from De Morgan’s Laws that the complement of a $G_\delta$-set is an $F_\sigma$-set and the complement of an $F_\sigma$-set is a $G_\delta$-set.

(d) We can define other types of Borel sets: a $G_{\delta\sigma}$-set is the union of countably many $G_\delta$-sets, a $F_{\sigma\delta}$-set is the union of countably many $F_\sigma$-sets, etc. All these sets are Borel sets. One can show that there exist Borel sets that are not of any of the types described here. ▷

**Proposition 2.28** Let $E$ be a set of real numbers. Then the following are equivalent:

(i) $E$ is measurable.

(ii) For all $\varepsilon > 0$, there exists an open set $O$ containing $E$ with $m^*(O \setminus E) < \varepsilon$.

(iii) There exists a $G_\delta$-set $G$ containing $E$ with $m^*(G \setminus E) = 0$.

(iv) For all $\varepsilon > 0$, there exists a closed set $F$ contained in $E$ with $m^*(E \setminus F) < \varepsilon$.

(v) There exists an $F_\sigma$-set $F$ contained in $E$ with $m^*(E \setminus F) = 0$.

**Proof:** (i) $\implies$ (ii): Let $\varepsilon > 0$. Suppose first that $m(E) < +\infty$. Since the outer measure is an infimum, there exists a ccoi $\{I_n\}_{n \geq 1}$ of $E$ with $\sum_{n=1}^{\infty} l(I_n) < m(E) + \varepsilon$. Put $O = \bigcup_{n=1}^{\infty} I_n$. Then $O$ is an open set containing $E$ and

$$m(O \setminus E) = m(O) - m(E) = m(\bigcup_{n=1}^{\infty} I_n) - m(E) \leq \left( \sum_{n=1}^{\infty} l(I_n) \right) - m(E) < m(E) + \varepsilon - m(E) = \varepsilon$$

by excision and countable subadditivity. Suppose next that $m(E) = +\infty$. Then $E$ is the disjoint union of countably many measurable sets of finite measure (indeed, $E = \bigcup_{k \in \mathbb{Z}} (E \cap [k, k+1))$), say $E = \bigcup_{n=1}^{\infty} E_n$ where $E_n$ is measurable and $m(E_n) < +\infty$ for all $n \in \mathbb{N}$. By the previous case, we have that for all $n \in \mathbb{N}$, there exists an open set $O_n$ containing $E_n$ with $m(O_n \setminus E_n) < \frac{\varepsilon}{2^n}$. Put $O = \bigcup_{n=1}^{\infty} O_n$. Then $O$ is an open set containing $\bigcup_{n=1}^{\infty} E_n = E$. Note that $O \setminus E \subseteq \bigcup_{n=1}^{\infty} (O_n \setminus E_n)$. Indeed, let $x \in O \setminus E$. Then $x \in O \setminus E_n$ for some $k \geq 1$. Also, $x \notin E = \bigcup_{n=1}^{\infty} E_n$. Hence $x \notin E_n$ for all $n \geq 1$. In particular, $x \notin E_k$ and so $x \in O_k \setminus E_k$.

It follows from monotonicity and countable subadditivity that

$$m(O \setminus E) \leq m(\bigcup_{n=1}^{\infty} (O_n \setminus E_n)) \leq \sum_{n=1}^{\infty} m(O_n \setminus E_n) < \sum_{n=1}^{\infty} \frac{\varepsilon}{2^n} = \varepsilon$$

(ii) $\implies$ (iii): For all $n \geq 1$, there exists an open set $O_n$ containing $E$ with $m^*(O_n \setminus E) < \frac{1}{n}$. Put $G = \bigcap_{n=1}^{\infty} O_n$. Then $G$ is $G_\delta$-set containing $E$. Since $G \subseteq O_n$, we have that $(G \setminus E) \subseteq (O_n \setminus E)$ and so

$$m^*(G \setminus E) \leq m^*(O_n \setminus E) < \frac{1}{n}$$

for all $n \geq 1$.  

34
by monotonicity. Taking the limit as \( n \to \infty \), we get that \( 0 \leq m^*(G \setminus E) \leq 0 \).

(iii) \( \implies \) (i): Let \( G \) be a \( G_\delta \)-set containing \( E \) with \( m^*(G \setminus E) = 0 \). Then \( G \setminus E \) is measurable by Proposition 2.12. Clearly \( G = (G \setminus E) \cup E \) and so \( E = G \setminus (G \setminus E) \). Since \( G \) and \( G \setminus E \) are measurable, we get that \( E \) is measurable.

(i) \( \implies \) (iv): Note that \( \tilde{E} \) is measurable. Let \( \varepsilon > 0 \). By (ii) applied to \( \tilde{E} \), we get that there exists an open set \( O \) containing \( \tilde{E} \) with \( m(O \setminus \tilde{E}) < \varepsilon \). Put \( F = \tilde{O} \). Then \( F \) is a closed set contained in \( \tilde{E} = E \). Since \( E \setminus F = E \cap F = E \cap O = O \cap \tilde{E} = O \setminus \tilde{E} \), we get that \( m(E \setminus F) = m(O \setminus \tilde{E}) < \varepsilon \).

(iv) \( \implies \) (v): Similar to (ii) \( \implies \) (iii)

(v) \( \implies \) (i): Similar to (iii) \( \implies \) (i).

\[ \square \]

**Proposition 2.29 (Littlewood 1)** Let \( E \) be a set with \( m^*(E) < +\infty \). Then \( E \) is measurable if and only if for all \( \varepsilon > 0 \), there exists a finite disjoint collection of open intervals \( \{I_k\}_{k=1}^n \) such that \( m^*(E \setminus O) + m^*(O \setminus E) < \varepsilon \) where \( O = \cup_{k=1}^n I_k \).

**Proof:** Suppose first that \( E \) is measurable. Let \( \varepsilon > 0 \). By Proposition 2.28(ii), there exists an open set \( O^* \) containing \( E \) with \( m(O^* \setminus E) < \frac{\varepsilon}{2} \). By Proposition 1.24, \( O^* \) is the union of a countable disjoint collection of open intervals, say \( O^* = \cup_{k=1}^\infty I_k \). Since \( O^* = (O^* \setminus E) \cup E \), it follows from countable additivity that \( m(O^*) = m(O^* \setminus E) + m(E) < \frac{\varepsilon}{2} + m(E) < +\infty \). So by countable additivity, we get

\[ \sum_{k=1}^\infty l(I_k) = \sum_{k=1}^\infty m(I_k) = m\left(\bigcup_{k=1}^\infty I_k\right) = m(O^*) < +\infty \]

Hence there exists \( n \in \mathbb{N} \) such that \( \sum_{k=n+1}^\infty l(I_k) < \frac{\varepsilon}{2} \). Put \( O = \cup_{k=1}^n I_k \). Since \( O \subseteq \cup_{k=1}^\infty I_k = O^* \), we have that \( O \setminus E \subseteq O^* \setminus E \) and so \( m(O \setminus E) \leq m(O^* \setminus E) < \frac{\varepsilon}{2} \) by monotonicity. Also \( E \setminus O \subseteq O^* \setminus O = (\cup_{k=1}^\infty I_k) \setminus (\cup_{k=1}^n I_k) = \cup_{k=n+1}^\infty I_k \) and so \( m(E \setminus O) \leq m(\cup_{k=n+1}^\infty I_k) = \sum_{k=n+1}^\infty m(I_k) = \sum_{k=n+1}^\infty l(I_k) < \frac{\varepsilon}{2} \) by monotonicity and countable additivity. Hence

\[ m(O \setminus E) + m(E \setminus O) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \]

Suppose next that the given property holds. Let \( \varepsilon > 0 \). Then there exists an open set \( O_1 \) (namely the union of a finite disjoint collection of open intervals) with \( m^*(E \setminus O_1) + m^*(O_1 \setminus E) < \frac{\varepsilon}{2} \). Since \( m^*(E \setminus O_1) < \frac{\varepsilon}{2} \), there exists a coo \( \{I_n\}_{n=1}^\infty \) of \( E \setminus O_1 \) with \( \sum_{n=1}^\infty l(I_n) < \frac{\varepsilon}{2} \). Put \( O_2 = \cup_{n=1}^\infty I_n \) and \( O = O_1 \cup O_2 \). Then \( O \) is an open set containing \( E \) and

\[ O \setminus E = (O_1 \cup O_2) \setminus E = (O_1 \setminus E) \cup (O_2 \setminus E) \subseteq (O_1 \setminus E) \cup O_2 \]

Moreover, \( m^*(O_2) = m^*(\cup_{n=1}^\infty I_n) \leq \sum_{n=1}^\infty l(I_n) < \frac{\varepsilon}{2} \) by countable subadditivity. We get that

\[ m^*(O \setminus E) \leq m^*(O_1 \setminus E) + m^*(O_2) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \]

by monotonicity and countable subadditivity. So \( E \) is measurable by Proposition 2.28. \( \square \)
2.6 Cantor Set

In this section, we provide an example of an uncountable set of measure zero.

Put $C_0 = [0, 1]$. We divide $C_0$ into three intervals of equal length and remove the interior of the middle interval to obtain $C_1$. So

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

We repeat this process: $C_1$ is the union of two closed disjoint intervals of length $\frac{1}{3}$. We divide each of these two intervals into three intervals of equal length and remove the interior of the middle interval. So

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

In general, for $n \geq 0$, $C_n$ is the union of $2^n$ disjoint closed intervals of length $\frac{1}{3^n}$ and $C_{n+1} \subseteq C_n$. We define the Cantor set $C$ as

$$C = \bigcap_{n=0}^{\infty} C_n$$

**Proposition 2.30** The Cantor set $C$ is an uncountable, closed set of measure zero.

**Proof:** First, we show that $C$ is closed. For all $n \geq 0$, we have that $C_n$ is closed since it is the finite union of closed intervals. Hence $C = \bigcap_{n=0}^{\infty} C_n$ is closed.

Next, we show that $m(C) = 0$. Note that $m(C_n) = \frac{2^n}{3^n}$ for all $n \geq 0$ by countable additivity. Since $C \subseteq C_n$ for all $n \geq 0$, it follows from monotonicity that

$$m(C) \leq m(C_n) = \frac{2^n}{3^n} \quad \text{for all } n \geq 0$$

Taking the limit as $n \to \infty$, we find that $m(C) = 0$.

Finally, we show that $C$ is uncountable. Note that $C \neq \emptyset$ since $0, 1 \in C$. Suppose that $C$ is countable. Let $\{c_n\}_{n=1}^{k}$ be an enumeration of $C$ (so $1 \leq k \leq +\infty$). Since $C_1$ is the union of two disjoint intervals of length $\frac{1}{3}$, it must be that one of these intervals (say $F_1$) does not contain $c_1$. Since $F_1$ is the union of two disjoint intervals of length $\frac{1}{3}$, it must be that one of these intervals (say $F_2$) does not contain $c_2$. Continuing this way, we obtain a sequence $\{F_n\}_{n=1}^{k}$ such that

- $F_n$ is non-empty and closed for all $n < k + 1$
- $F_n \subseteq C_n$ for all $n < k + 1$
- $F_{n+1} \subseteq F_n$ for all $n < k$
- $c_n \notin F_n$ for all $n < k + 1$

If $k < +\infty$, then let $F_{k+1}$ be one of the two closed intervals whose union is $F_k$, let $F_{k+2}$ be one of the two closed intervals whose union is $F_{k+1}$, etc.

It follows from the Nested Set Theorem (Theorem 1.31) that $\bigcap_{n=1}^{\infty} F_n \neq \emptyset$. So let $x \in \bigcap_{n=1}^{\infty} F_n$. Then $x \in F_n$ for all $n \geq 1$. Since $F_n \subseteq C_n$ for all $n \geq 1$, we get that $x \in \bigcap_{n=1}^{\infty} F_n \subseteq \bigcap_{n=1}^{\infty} C_n = C$. Hence $x = c_i$ for some $1 \leq i < k + 1$. But then $x = c_i \notin F_i$, a contradiction since $x \in F_n$ for all $n \geq 1$.

So $C$ is uncountable. \(\square\)
Chapter 3

Measurable Functions

3.1 Measurable Functions

In this section, we generalize the concept of a continuous function and prove some general properties of these generalized functions.

**Proposition 3.1** Let $D \subseteq \mathbb{R}$ be measurable and $f : D \rightarrow \mathbb{R}$ an extended real-valued function. Then the following are equivalent:

(i) The set $\{x \in D : f(x) > \alpha\}$ is measurable for all $\alpha \in \mathbb{R}$.

(ii) The set $\{x \in D : f(x) \geq \alpha\}$ is measurable for all $\alpha \in \mathbb{R}$.

(iii) The set $\{x \in D : f(x) < \alpha\}$ is measurable for all $\alpha \in \mathbb{R}$.

(iv) The set $\{x \in D : f(x) \leq \alpha\}$ is measurable for all $\alpha \in \mathbb{R}$.

**Proof:** Recall that the collection of measurable sets is a $\sigma$-algebra.

(i) $\implies$ (ii) Let $\alpha \in \mathbb{R}$. Then

$$\{x \in D : f(x) \geq \alpha\} = \bigcap_{n=1}^{\infty} \left\{x \in D : f(x) > \alpha - \frac{1}{n}\right\}$$

Since $\{x \in D : f(x) > \alpha - \frac{1}{n}\}$ is measurable for all $n \in \mathbb{N}$, we get that $\{x \in D : f(x) \geq \alpha\}$ is measurable.

(ii) $\implies$ (iii) Let $\alpha \in \mathbb{R}$. Then

$$\{x \in D : f(x) < \alpha\} = D \setminus \{x \in D : f(x) \geq \alpha\}$$

Since $D$ and $\{x \in D : f(x) \geq \alpha\}$ are measurable, we get that $\{x \in D : f(x) < \alpha\}$ is measurable.

(iii) $\implies$ (iv) Let $\alpha \in \mathbb{R}$. Then

$$\{x \in D : f(x) \leq \alpha\} = \bigcap_{n=1}^{\infty} \left\{x \in D : f(x) < \alpha + \frac{1}{n}\right\}$$
Since \( \{ x \in D : f(x) < \alpha + \frac{1}{n} \} \) is measurable for all \( n \in \mathbb{N} \), we get that \( \{ x \in D : f(x) \leq \alpha \} \) is measurable.

(iv) \( \implies \) (i) Let \( \alpha \in \mathbb{R} \). Then

\[
\{ x \in D : f(x) > \alpha \} = D \setminus \{ x \in D : f(x) \leq \alpha \}
\]

Since \( D \) and \( \{ x \in D : f(x) \leq \alpha \} \) are measurable, we get that \( \{ x \in D : f(x) < \alpha \} \) is measurable. \( \square \)

**Remark:** Note that \( \{ x \in D : f(x) > \alpha \} = f^{-1}((\alpha, +\infty]) \).

\[ \triangleright \]

**Definition 3.2** A function \( f : D \to \mathbb{R} \) is measurable if \( D \) is measurable and if \( f \) satisfies one (and hence all four) of the statements in Proposition 3.1.

**Proposition 3.3** Let \( f : D \to \mathbb{R} \) be a measurable function. Then the set \( \{ x \in D : f(x) = \alpha \} \) is measurable for all \( \alpha \in \mathbb{R} \).

**Proof:** Let \( \alpha \in \mathbb{R} \).

Suppose first that \( \alpha \in \mathbb{R} \). Then

\[
\{ x \in D : f(x) = \alpha \} = \{ x \in D : f(x) \geq \alpha \} \cap \{ x \in D : f(x) \leq \alpha \}
\]

Since \( \{ x \in D : f(x) \geq \alpha \} \) and \( \{ x \in D : f(x) \leq \alpha \} \) are measurable (because \( f \) is measurable), we get that \( \{ x \in D : f(x) = \alpha \} \) is measurable.

Suppose next that \( \alpha \notin \mathbb{R} \). We prove the case ‘\( \alpha = +\infty \)’ as the case ‘\( \alpha = -\infty \)’ is similar. Then

\[
\{ x \in D : f(x) = +\infty \} = \bigcap_{n=1}^{\infty} \{ x \in D : f(x) > n \}
\]

Since \( \{ x \in D : f(x) > n \} \) is measurable for all \( n \in \mathbb{N} \) (because \( f \) is measurable), we get that \( \{ x \in D : f(x) = +\infty \} \) is measurable. \( \square \)

**Remark:** Note that this property does not imply \( f \) is measurable. There exists a measurable set \( D \) and a function \( f : D \to \mathbb{R} \) such that \( \{ x \in D : f(x) = \alpha \} \) is measurable for all \( \alpha \in \mathbb{R} \) but \( f \) is not measurable. \( \triangleright \)

**Proposition 3.4** Let \( I \) be an interval and \( f : I \to \mathbb{R} \) a function that is monotone on \( I \). Then \( f \) is measurable.

**Proof:** Exercise. \( \square \)

**Proposition 3.5** Let \( D \subseteq \mathbb{R} \) be measurable and \( f : D \to \mathbb{R} \) a measurable function. Then \( |f| \) is measurable.

**Proof:** Exercise. \( \square \)

The next theorem gives a characterization of measurable functions similar to the characterization of continuous functions (Theorem 1.33).
Theorem 3.6 Let \( D \subseteq \mathbb{R} \) be measurable and \( f : D \to \mathbb{R} \) a real-valued function defined on \( D \). Then \( f \) is measurable if and only if \( f^{-1}(O) \) is measurable for every open set \( O \).

**Proof:** Suppose first that \( f^{-1}(O) \) is measurable for every open set \( O \). Let \( \alpha \in \mathbb{R} \). Then \((\alpha, +\infty)\) is an open set. Hence \( f^{-1}((\alpha, +\infty)) = \{x \in D : f(x) > \alpha\} \) is measurable. So \( f \) is measurable.

Suppose next that \( f \) is measurable. We show that \( f^{-1}(I) \) is measurable for all open intervals \( I \). Clearly \( f^{-1}(\emptyset) = \emptyset \) is measurable and \( f^{-1}((\neg \infty, +\infty)) = D \) is measurable. For \( a \in \mathbb{R} \), we have that \( f^{-1}((a, +\infty)) = \{x \in D : f(x) > a\} \) is measurable and \( f^{-1}((\neg \infty, a)) = \{x \in D : f(x) < a\} \) is measurable since \( f \) is measurable. Finally, if \( a, b \in \mathbb{R} \) with \( a < b \), then

\[
\text{is measurable since } \{x \in D : f(x) > a\} \text{ and } \cap \{x \in D : f(x) < b\} \text{ are measurable (because } f \text{ is measurable)}.
\]

Now we can show that \( f^{-1}(O) \) is measurable for every open set \( O \). Let \( O \) be an open set. By Theorem 1.24, \( O \) is the union of countably many open intervals, say \( O = \bigcup_{n=1}^{\infty} I_n \). Hence

\[
f^{-1}(O) = f^{-1}(\bigcup_{n=1}^{\infty} I_n) = \bigcup_{n=1}^{\infty} f^{-1}(I_n)
\]

is measurable since \( f^{-1}(I_n) \) is measurable for all \( n \in \mathbb{N} \). \( \square \)

This characterization allows us to prove that every continuous function on a measurable domain is a measurable function.

Corollary 3.7 Let \( D \subseteq \mathbb{R} \) be measurable and \( f : D \to \mathbb{R} \) continuous on \( D \). Then \( f \) is measurable.

**Proof:** Let \( O \) be an open set. Since \( f \) is continuous on \( D \), it follows from Theorem 1.33 that \( f^{-1}(O) = D \cap O^* \) for some open set \( O^* \). So \( f^{-1}(O) \) is measurable. Hence \( f \) is measurable by Theorem 3.6. \( \square \)

Proposition 3.8 Let \( D \subseteq \mathbb{R} \) be measurable and \( f : D \to \mathbb{R} \) a real-valued function defined on \( D \). Then \( f \) is measurable if and only if \( f^{-1}(B) \) is measurable for every Borel set \( B \).

**Proof:** Exercise. \( \square \)

Proposition 3.9 Let \( D \subseteq \mathbb{R} \) be measurable, \( g : D \to \mathbb{R} \) a measurable function, \( E \subseteq \mathbb{R} \) a Borel set containing \( g(D) \) and \( f : E \to \mathbb{R} \) a function that is continuous on \( E \). Then \( f \circ g \) is measurable.

**Proof:** Let \( O \subseteq \mathbb{R} \) be open. Then \( (f \circ g)^{-1}(O) = g^{-1}(f^{-1}(O)) \). Since \( f \) is continuous, it follows from Theorem 1.33 that \( f^{-1}(O) = E \cap O^* \) for some open set \( O^* \). So \( (f \circ g)^{-1}(O) = g^{-1}(E \cap O^*) \). Since \( E \cap O^* \) is a Borel set and \( g \) is measurable, it follows from Proposition 3.8 that \( g^{-1}(E \cap O^*) \) is measurable. So \( (f \circ g)^{-1}(O) \) is measurable. Hence \( f \circ g \) is measurable by Theorem 3.6. \( \square \)

The next proposition allows us to show that a function is measurable by looking at the restriction to measurable subsets.
**Proposition 3.10** Let $D \subseteq \mathbb{R}$ be measurable and $f : D \to \mathbb{R}$ an extended real-valued function. Suppose that $D = \bigcup_{n=1}^{\infty} D_n$ where $D_n$ is measurable for all $n \in \mathbb{N}$. Then $f$ is measurable (over $D$) if and only if $f|_{D_n}$ is measurable (over $D_n$) for all $n \in \mathbb{N}$.

**Proof:** Suppose first that $f$ is measurable over $D$. Let $n \in \mathbb{N}$. Let $\alpha \in \mathbb{R}$. Then

$$\{x \in D_n : f(x) > \alpha\} = D_n \cap \{x \in D : f(x) > \alpha\}$$

is measurable since $f$ is measurable. So $f|_{D_n}$ is measurable over $D_n$.

Suppose next that $f|_{D_n}$ is measurable for all $n \in \mathbb{N}$. Let $\alpha \in \mathbb{R}$. Since $D = \bigcup_{n=1}^{\infty} D_n$, we get that

$$\{x \in D : f(x) > \alpha\} = \bigcup_{n=1}^{\infty} \{x \in D_n : f(x) > \alpha\}$$

Note that for all $n \in \mathbb{N}$, we have that $\{x \in D_n : f(x) > \alpha\}$ is measurable since $f|_{D_n}$ is measurable. So $\{x \in D : f(x) > \alpha\}$ is measurable. Hence $f$ is measurable over $D$. \(\Box\)

**Definition 3.11** Let $D \subseteq \mathbb{R}$ be a set.

(a) Let $f_k : D \to \overline{\mathbb{R}}$ be a function for $1 \leq k \leq n$ for some $n \in \mathbb{N}$. We define the functions $\min\{f_1, \ldots, f_n\}$ and $\max\{f_1, \ldots, f_n\}$ pointwise as follows:

$$\min\{f_1, \ldots, f_n\} : D \to \overline{\mathbb{R}} : x \mapsto \min\{f_1(x), \ldots, f_n(x)\}$$

$$\max\{f_1, \ldots, f_n\} : D \to \overline{\mathbb{R}} : x \mapsto \max\{f_1(x), \ldots, f_n(x)\}$$

(b) Let $f_n : D \to \overline{\mathbb{R}}$ be a function for all $n \in \mathbb{N}$. We define the functions $\inf_{n \geq 1} f_n$, $\sup_{n \geq 1} f_n$, $\liminf_{n \geq 1} f_n$ (also denoted by $\liminf f_n$) and $\limsup_{n \geq 1} f_n$ (also denoted by $\limsup f_n$) pointwise as follows:

$$\inf_{n \geq 1} f_n : D \to \overline{\mathbb{R}} : x \mapsto \inf_{n \geq 1} f_n(x)$$

$$\sup_{n \geq 1} f_n : D \to \overline{\mathbb{R}} : x \mapsto \sup_{n \geq 1} f_n(x)$$

$$\liminf_{n \geq 1} f_n : D \to \overline{\mathbb{R}} : x \mapsto \liminf_{n \geq 1} f_n(x)$$

$$\limsup_{n \geq 1} f_n : D \to \overline{\mathbb{R}} : x \mapsto \limsup_{n \geq 1} f_n(x)$$

\(\triangleright\)

**Proposition 3.12** Let $D \subseteq \mathbb{R}$ be a measurable set and $f_n : D \to \overline{\mathbb{R}}$ a measurable function for all $n \in \mathbb{N}$. Then $\min\{f_1, \ldots, f_n\}$, $\max\{f_1, \ldots, f_n\}$, $\inf_{n \geq 1} f_n$, $\sup_{n \geq 1} f_n$, $\liminf_{n \geq 1} f_n$ and $\limsup_{n \geq 1} f_n$ are measurable.

**Proof:** Let $\alpha \in \mathbb{R}$.

Let $n \in \mathbb{N}$ and put $g = \min\{f_1, \ldots, f_n\}$. Then

$$\{x \in D : g(x) > \alpha\} = \bigcap_{k=1}^{n} \{x \in D : f_k(x) > \alpha\}$$

So $\min\{f_1, \ldots, f_n\}$ is measurable.
Let \( n \in \mathbb{N} \) and put \( g = \max\{f_1, \ldots, f_n\} \). Then
\[
\{x \in D : g(x) > \alpha\} = \bigcup_{k=1}^{n} \{x \in D : f_k(x) > \alpha\}
\]
So \( \max\{f_1, \ldots, f_n\} \) is measurable.

Put \( g = \inf_{n \geq 1} f_n \). Then
\[
\{x \in D : g(x) \geq \alpha\} = \bigcap_{n=1}^{\infty} \{x \in D : f_n(x) \geq \alpha\}
\]
So \( \inf_{n \geq 1} f_n \) is measurable.

Put \( g = \sup_{n \geq 1} f_n \). Then
\[
\{x \in D : g(x) > \alpha\} = \bigcup_{n=1}^{\infty} \{x \in D : f_n(x) > \alpha\}
\]
So \( \sup_{n \geq 1} f_n \) is measurable.

For \( n \geq 1 \), put \( g_n = \inf_{k \geq n} f_k \). Then \( g_n \) is measurable for all \( n \in \mathbb{N} \). Hence \( \lim_{n \to \infty} f_n = \sup \inf_{k \geq n} f_k = \sup_{n \geq 1} g_n \) is measurable.

Similarly, we get that \( \lim_{n \to \infty} f_n \) is measurable. \( \square \)

**Corollary 3.13** Let \( D \subseteq \mathbb{R} \) be measurable and \( f_n : D \to \mathbb{R} \) measurable for all \( n \in \mathbb{N} \). Suppose that the sequence \( \{f_n\}_{n \geq 1} \) converges pointwise to the function \( f : D \to \mathbb{R} \). Then \( f \) is measurable.

**Proof:** Note that \( f = \lim_{n \to \infty} f_n = \lim f_n \). \( \square \)

Before we tackle the problem of sums or products of measurable functions, we introduce the concept of ‘almost everywhere’. It will allow us to consider functions like \( f + g \) that are defined ‘almost everywhere’.

**Definition 3.14** A property holds **almost everywhere** on a set \( E \) if there exists a subset \( E_0 \) of \( E \) with \( m(E_0) = 0 \) such that the property holds on \( E \setminus E_0 \). We write that the property holds a.e. on \( E \).

**Proposition 3.15** Let \( f, g : D \to \mathbb{R} \) be functions such that \( f \) is measurable and \( g = f \) a.e. on \( D \). Then \( g \) is measurable.

**Proof:** Put \( D_0 = \{x \in D : f(x) \neq g(x)\} \). Then \( m^*(D_0) = 0 \) since \( g = f \) a.e. on \( D \). If \( A \subseteq D_0 \) then \( m^*(A) \leq m^*(D_0) = 0 \) by monotonicity and so \( A \) is measurable by Proposition 2.12.

Let \( \alpha \in \mathbb{R} \). Then
\[
\{x \in D : g(x) > \alpha\} = (\{x \in D : f(x) > \alpha\} \setminus D_0) \cup \{x \in D_0 : g(x) > \alpha\}
\]
Note that \( \{x \in D : f(x) > \alpha\} \) is measurable since \( f \) is measurable, and \( \{x \in D_0 : g(x) > \alpha\} \) is measurable since it is a subset of \( D_0 \). Hence \( \{x \in D : g(x) > \alpha\} \) is measurable. So \( g \) is measurable. \( \square \)
Proposition 3.16 Let $D \subseteq \mathbb{R}$ be measurable and $f : D \rightarrow \mathbb{R}$ a function such that $f$ is continuous a.e. on $D$. Then $f$ is measurable.

**Proof:** Exercise. \qed

Proposition 3.17 Let $f, g : D \rightarrow \mathbb{R}$ be real-valued measurable functions. Then the following holds:

(a) $\lambda f + \mu g$ is measurable for all $\lambda, \mu \in \mathbb{R}$.

(b) $fg$ is measurable.

(c) If $g \neq 0$ on $D$ then $\frac{f}{g}$ is measurable.

**Proof:** (a) Let $\lambda \in \mathbb{R}$. We show that $\lambda f$ is measurable. If $\lambda = 0$ then $\lambda f$ is continuous on $D$ and hence measurable. If $\lambda > 0$ then for all $\alpha \in \mathbb{R}$, we have that

$$\{x \in D : (\lambda f)(x) > \alpha\} = \left\{x \in D : f(x) > \frac{\alpha}{\lambda}\right\}$$

and so $\lambda f$ is measurable. If $\lambda < 0$ then for all $\alpha \in \mathbb{R}$, we have that

$$\{x \in D : (\lambda f)(x) > \alpha\} = \left\{x \in D : f(x) < \frac{\alpha}{|\lambda|}\right\}$$

and so $\lambda f$ is measurable.

Next, we show that $f + g$ is measurable. Let $\alpha \in \mathbb{R}$. Then

$$\{x \in D : (f + g)(x) > \alpha\} = \bigcup_{q \in \mathbb{Q}} (\{x \in D : f(x) > q\} \cap \{x \in D : g(x) > \alpha - q\}) \quad (*)$$

Indeed, if $x \in D$ such that $f(x) > q$ and $g(x) > \alpha - q$ for some $q \in \mathbb{Q}$, then $(f + g)(x) = f(x) + g(x) > q + \alpha - q = \alpha$. Conversely, let $x \in D$ with $(f + g)(x) > \alpha$. Then $f(x) + g(x) > \alpha$ and so $f(x) > \alpha - g(x)$. By the density of the rational numbers, there exists $q \in \mathbb{Q}$ with $f(x) > q > \alpha - g(x)$. So $f(x) > q$ and $g(x) > \alpha - q$.

Since $f$ and $g$ are measurable, it follows from (*) that $f + g$ is measurable.

Now we can prove (a). Let $\lambda, \mu \in \mathbb{R}$. Then $\lambda f$ and $\mu g$ are measurable by the first part of the proof. Hence $\lambda f + \mu g$ is measurable by the second part of the proof.

(b) First, we show that $f^2$ is measurable. Let $\alpha \in \mathbb{R}$. If $\alpha < 0$ then $\{x \in D : f^2(x) > \alpha\} = D$ is measurable. If $\alpha \geq 0$ then

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

is measurable since $f$ is measurable. So $f^2$ is measurable.

To prove (b), note that

$$fg = \frac{1}{2}[(f + g)^2 - f^2 - g^2]$$

By (a), $f + g$ is measurable. By what we just proved in (b), $(f + g)^2$, $f^2$ and $g^2$ are measurable. Then $fg$ is measurable by (a).
Suppose \( f \) and \( g \) are two extended-real valued functions defined on some measurable set set \( D \). Then \( f + g \) is not defined at the points where we get a sum of the form \( \infty - \infty \). However, if \( f \) and \( g \) are finite a.e. on \( D \), say \( D_0 = \{ x \in D : f(x) = \pm \infty \text{ or } g(x) = \pm \infty \} \), then \( f + g \) is defined on \( D \setminus D_0 \). We assume that \( f + g \) is somehow (and it does not matter how) defined on \( D_0 \) so that \( f + g \) is defined on \( D \). Since a function defined on a set of measure zero is always measurable, we see that \( f + g \) is measurable on \( D \) if and only if \( f + g \) is measurable on \( D \setminus D_0 \).

Similarly, we have that \( fg \) is defined on \( D \setminus D_0 \). If \( fg \) is defined somehow on \( D \) then \( fg \) is measurable on \( D \) if and only if \( fg \) is measurable on \( D \setminus D_0 \).

If \( \lambda \in \mathbb{R} \) then \( \lambda f \) is defined on \( D \) unless \( \lambda = 0 \); if \( \lambda = 0 \) then we define \( \lambda f = 0 \) on \( D \).

**Proposition 3.18** Let \( f, g : D \to \mathbb{R} \) be extended real-valued measurable functions that are finite a.e. on \( D \). Then the following holds:

(a) \( \lambda f + \mu g \) is measurable for all \( \lambda, \mu \in \mathbb{R} \).

(b) \( fg \) is measurable.

(c) If \( g \neq 0 \) on \( D \) then \( \frac{f}{g} \) is measurable.

**Proof:** Let \( D_0 = \{ x \in D : f(x) = \pm \infty \text{ or } g(x) = \pm \infty \} \). Since

\[
D_0 = \{ x \in D : f(x) = \pm \infty \} \cup \{ x \in D : g(x) = \pm \infty \}
\]

and \( f \) and \( g \) are finite a.e. on \( D \), we have that \( m(D_0) = 0 \). Note that \( f|_{D \setminus D_0} \) and \( g|_{D \setminus D_0} \) are measurable real-valued functions.

(a) Let \( \lambda, \mu \in \mathbb{R} \). By Proposition 3.17, \( \lambda f + \mu g \) is measurable on \( D \setminus D_0 \). Since \( m(D_0) = 0 \), we have that \( \lambda f + \mu g \) is measurable on \( D_0 \). Hence \( \lambda f + \mu g \) is measurable on \( (D \setminus D_0) \cup D_0 = D \).

(b) (c) Similar as above. \( \square \)

### 3.2 The Cantor-Lebesgue Function

In this section, we introduce the Cantor-Lebesgue function and disprove several ‘conjectures’ about measurable functions and measurable sets.

Recall the definition of the Cantor set \( C \).

For \( k \in \mathbb{N} \), put \( \mathcal{O}_k = [0, 1] \setminus C_k \). Then for \( k \in \mathbb{N} \), we have that \( \mathcal{O}_k \subseteq \mathcal{O}_{k+1} \) and \( \mathcal{O}_k \) is the disjoint union of \( 1 + 2 + 2^2 + \cdots + 2^{k-1} = 2^k - 1 \) open intervals. Put \( \mathcal{O} = \bigcup_{k=1}^\infty \mathcal{O}_k \). Then \( C = [0, 1] \setminus \mathcal{O} \). We proved that \( \mathcal{O} \) is dense in \( [0, 1] \).

We define the Cantor-Lebesgue function \( \varphi : [0, 1] \to \mathbb{R} \) as follows:

Let \( k \in \mathbb{N} \). Then on \( \mathcal{O}_k \), \( \varphi \) is defined as the increasing function which is constant on the \( 2^k - 1 \) open intervals of \( \mathcal{O}_k \) and takes on the values

\[
\frac{1}{2^k}, \frac{2}{2^k}, \frac{3}{2^k}, \ldots, \frac{2^k - 2}{2^k}, \frac{2^k - 1}{2^k}
\]

We extend \( \varphi \) to all of \( [0, 1] \) as follows:
\[ \varphi(0) = 0 \]
\[ \varphi(x) = \sup\{ \varphi(t) : t \in \mathcal{O} \cap [0, x] \} \text{ if } 0 < x \leq 1 \]

Note that \( \varphi \) is well-defined on \( \mathcal{O} \). Let \( k \in \mathbb{N} \). Let \( I_1, I_2, \ldots, I_{2^k-1} \) be the open intervals of \( \mathcal{O}_k \) in 'increasing' order. Let \( J_1, J_2, \ldots, J_{2^k} \) be the \( 2^k \) open intervals of \( \mathcal{O}_{k+1} \setminus \mathcal{O}_k \) in 'increasing' order. Then

\[ J_1, I_1, J_2, I_2, \ldots, J_{2^k-1}, I_{2^k-1}, J_{2^k} \]

are the \( 2^{k+1} - 1 \) open intervals of \( \mathcal{O}_{k+1} \) in 'increasing' order. Now let \( x \in I_j \) for some \( 1 \leq j \leq 2^k - 1 \).

Using the definition of \( \varphi \) on \( \mathcal{O}_k \), we get that \( \varphi(x) = \frac{1}{2^k} \). Using the definition of \( \varphi \) on \( \mathcal{O}_{k+1} \), we get that \( \varphi(x) = \frac{2^{j-1}}{2^k} \). So \( \varphi \) is well-defined on \( \mathcal{O} \).

Note that \( \varphi \) is non-decreasing on \( \mathcal{O} \): if \( x, y \in \mathcal{O} \) with \( x \leq y \) then \( x, y \in \mathcal{O}_k \) for some \( k \in \mathbb{N} \) and hence \( \varphi(x) \leq \varphi(y) \).

If \( 0 < x \leq 1 \) then \( \mathcal{O} \cap [0, x] \neq \emptyset \) since \( \mathcal{O} \) is dense in \([0, 1]\). So the definition of \( \varphi(x) \) makes sense on \([0, 1]\).

The 'old' definition of \( \varphi \) on \( \mathcal{O} \) coincides with the 'new' definition of \( \varphi \) on \([0, 1]\): if \( x \in \mathcal{O} \) then \( \varphi(x) = \max\{ \varphi(t) : t \in \mathcal{O} \cap [0, x] \} \) since \( \varphi \) is non-decreasing on \( \mathcal{O} \).

Finally, since \( 0 < \varphi(t) < 1 \) for all \( t \in \mathcal{O} \), we get that \( 0 < \varphi(x) \leq 1 \) for all \( x \in (0, 1] \).

**Proposition 3.19** The Cantor-Lebesgue function \( \varphi \) is a non-decreasing, continuous function that maps \([0, 1]\) onto \([0, 1]\). Moreover, \( \varphi \) is differentiable on \( \mathcal{O} \) and \( \varphi' = 0 \) on \( \mathcal{O} \).

**Proof:** Let \( x, y \in [0, 1] \) with \( x \leq y \). If \( x = 0 \) then \( \varphi(x) = 0 \leq \varphi(y) \). So we may assume that \( x > 0 \). Then \( \mathcal{O} \cap [0, x] \subseteq \mathcal{O} \cap [0, y] \). Hence

\[ \varphi(x) = \sup\{ \varphi(t) : t \in \mathcal{O} \cap [0, x] \} \leq \sup\{ \varphi(t) : t \in \mathcal{O} \cap [0, y] \} = \varphi(y) \]

So \( \varphi \) is non-decreasing on \([0, 1]\).

Let \( x_0 \in [0, 1] \). Suppose first that \( x_0 \in \mathcal{O} \). Then \( x_0 \in \mathcal{O}_k \) for some \( k \in \mathbb{N} \). Hence \( \varphi \) is constant on an open interval containing \( x_0 \). Clearly, \( \varphi \) is differentiable at \( x_0 \) with \( \varphi'(x_0) = 0 \) and \( \varphi \) is continuous at \( x_0 \). Suppose next that \( x_0 \notin \mathcal{O} \). Assume first that \( 0 < x_0 < 1 \). Since \( \mathcal{O} \) is dense in \([0, 1]\), there exist \( k_0 \in \mathbb{N} \) and \( a, b \in \mathcal{O}_{k_0} \) with \( 0 < a < x_0 < b < 1 \). Let \( k \geq k_0 \). Then \( x \) lies between two 'consecutive' intervals in \( \mathcal{O}_k \). Hence there exist \( a_k, b_k \in \mathcal{O}_k \) with \( a_k < x_0 < b_k \) and \( \varphi(b_k) = \varphi(a_k) + \frac{1}{2^k} \). Since \( \varphi \) is non-decreasing, this implies that \( |\varphi(x) - \varphi(x_0)| \leq \frac{1}{2^k} \) if \( a_k < x < b_k \). So \( \varphi \) is continuous at \( x_0 \). Assume finally that \( x_0 = 0 \) (the case \( x_0 = 1 \) is similar). For \( k \in \mathbb{N} \), let \( a_k \) be an element in the 'first' interval of \( \mathcal{O}_k \). Then \( 0 \leq x_0 < a_k \) and \( \varphi(a_k) = \frac{1}{2^k} \) for all \( k \in \mathbb{N} \). Since \( \varphi \) is non-decreasing, this implies that \( \varphi \) is continuous at \( x_0 = 0 \).

We show that \( \varphi(1) = 1 \). We know that \( \varphi(1) \leq 1 \). Let \( k \in \mathbb{N} \) and \( b_k \) an element in the 'last' interval of \( \mathcal{O}_k \). Since \( \varphi \) is non-decreasing, we get that \( \frac{2^{k-1}}{2^k} = \varphi(b_k) \leq \varphi(1) \leq 1 \). Considering the limit as \( k \to +\infty \), we get that \( \varphi(1) = 1 \).

Since \( \varphi(0) = 0 \) and \( \varphi(1) = 1 \) and \( \varphi \) is continuous and non-decreasing on \([0, 1]\), we get that \( \varphi([0, 1]) = [0, 1] \) by the Intermediate Value Theorem. \( \square \)

**Proposition 3.20** The function \( \psi : [0, 1] \to \mathbb{R} : x \to \varphi(x) + x \) has the following properties:

(a) \( \psi \) is strictly increasing and continuous on \([0, 1]\).

(b) \( \psi([0, 1]) = [0, 2] \).
(c) $\psi(C)$ is measurable and $m(\psi(C)) = 1$.

(d) There exists a subset $E$ of $C$ such that $E$ is measurable and $\psi(E)$ is non-measurable.

**Proof**: (a) Obvious.

(b) Since $\psi(0) = \phi(0) + 0 = 0 + 0 = 0$ and $\psi(1) = \phi(1) + 1 = 1 + 1 = 2$ and $\psi$ is strictly increasing on $[0, 1]$, we have that $\psi([0, 1]) \subseteq [0, 2]$. Since $\psi(0) = 0$ and $\psi(1) = 2$ and $\psi$ is continuous, it follows from the Intermediate Value Theorem that $[0, 2] \subseteq \psi([0, 1])$.

(c) (d) We show that $\psi(O')$ and $\psi(F')$ are measurable for any open set $O'$ and any closed set $F'$. Since $\psi$ is a strictly increasing continuous function defined on an interval, it has a continuous inverse. Put $g = \psi^{-1} : [0, 2] \to [0, 1]$. Let $O'$ be an open set. Since $g$ is continuous, we have that $\psi(O') = g^{-1}(O') = [0, 2] \cap O^*$ for some open set $O^*$ by Theorem 1.33. So $\psi(O')$ is measurable. Similarly, if $F'$ is a closed set, then $\psi(F') = [0, 2] \cap F^*$ for some closed set $F^*$ and so $\psi(F')$ is measurable.

In particular, $\psi(O)$ and $\psi(C)$ are measurable. Since $[0, 1] = O \cup C$ and $\psi$ is one-to-one, we get that

$$[0, 2] = \psi([0, 1]) = \psi(O) \cup \psi(C)$$

Since $\psi(O)$ and $\psi(C)$ are measurable, it follows from countable additivity that

$$2 = m(\psi(O)) + m(\psi(C)) \quad (*)$$

Note that

$$m(O) = m([0, 1] \setminus C) = m([0, 1]) - m(C) = 1 - 0 = 1$$

by excision. Put $O = \bigcup_{k=1}^{\infty} I_k$ where $\{I_k\}_{k \geq 1}$ are the disjoint open intervals that make up $O$. Fix $k \in \mathbb{N}$. Note that $I_k$ is one of the intervals of $O_{k_0}$ for some $k_0$. Hence $\phi$ is constant on $I_k$. Let $c_k$ be this constant value: $\phi(x) = c_k$ for all $x \in I_k$. Then $\psi(x) = \phi(x) + x = c_k + x$ for all $x \in I_k$. Hence $\psi(I_k) = I_k + c_k$. Since the outer measure is translation-invariant, we get that

$$m(\psi(I_k)) = m(I_k + c_k) = m(I_k)$$

Using that $O = \bigcup_{k=1}^{\infty} I_k$, that $\psi$ is one-to-one and countable additivity, we find that

$$m(\psi(O)) = m \left( \psi \left( \bigcup_{k=1}^{\infty} I_k \right) \right) = m \left( \bigcup_{k=1}^{\infty} \psi(I_k) \right) = \sum_{k=1}^{\infty} m(\psi(I_k)) = \sum_{k=1}^{\infty} m(I_k) = m \left( \bigcup_{k=1}^{\infty} I_k \right) = m(O) = 1$$

It follows from $(*)$ that $m(\psi(C)) = 1$, which proves (c).

Since $\psi(C)$ is bounded and $m(\psi(C)) > 0$, it follows from Theorem 2.25 that $\psi(C)$ contains a non-measurable subset $D$. Put $E = \psi^{-1}(D)$. Then $E \subseteq C$ and so $m^*(E) \leq m(C) = 0$ by monotonicity. Hence $m^*(E) = 0$ and $E$ is measurable. But $\psi(E) = D$ is non-measurable, which proves (d).

We finish this section by proving that not every measurable set is a Borel set.

**Proposition 3.21** There exists a measurable set that is not a Borel set.

**Proof**: Recall that $\psi^{-1} : [0, 2] \to [0, 1]$ is continuous and hence measurable. By Proposition 3.20(d), there exists a measurable subset $E$ of $C$ and a non-measurable subset $D$ of $[0, 2]$ with $\psi(E) = D$. If $E$ is a Borel set, then $D = \psi(E) = (\psi^{-1})^{-1}(E)$ is measurable by Proposition 3.8, a contradiction. Hence $E$ is not a Borel set. 

□
3.3 The Simple Approximation Theorem

In this section, we prove that every measurable function is the pointwise limit of a sequence of simple functions.

Definition 3.22

(a) Let $D \subseteq \mathbb{R}$. The characteristic function of $D$ (notation: $\chi_D$) is the function

$$
\chi_D : \mathbb{R} \to \mathbb{R} : x \to \begin{cases} 
1 & \text{if } x \in D \\
0 & \text{if } x \notin D
\end{cases}
$$

(b) Let $D \subseteq \mathbb{R}$ be measurable and $\varphi : D \to \mathbb{R}$ a function. Then $\varphi$ is a simple function if $\varphi$ is measurable and assumes only a finite number of function values.

Remark: Let $D_1, \ldots, D_n$ be measurable subsets of $\mathbb{R}$ and $a_1, \ldots, a_n \in \mathbb{R}$. Then it is quite easy to check that $\varphi = \sum_{i=1}^{n} a_i \chi_{D_i}$ is a simple function (on any measurable domain).

Definition 3.23 Let $D \subseteq \mathbb{R}$ and $f : E \to \mathbb{R}$ be a function. We define the positive part of $f$ (notation: $f^+$) and the negative part of $f$ (notation: $f^-$) as

$$
f^+ = \max\{f, 0\} \quad \text{and} \quad f^- = \max\{-f, 0\}
$$

Remarks:

(a) Note that $f^+$ and $f^-$ are nonnegative measurable functions if $f$ is measurable.

(b) Note that $f = f^+ - f^-$ and $|f| = f^+ + f^-$. △

Lemma 3.24 Let $f : D \to \mathbb{R}$ be a measurable function and $m, M \in \mathbb{R}$ with $m \leq f \leq M$ on $D$. Then for all $\varepsilon > 0$, there exist simple functions $\varphi_\varepsilon$ and $\psi_\varepsilon$ defined on $D$ such that

$$
m \leq \varphi_\varepsilon \leq f \leq \psi_\varepsilon \leq M \quad \text{and} \quad 0 \leq \psi_\varepsilon - \varphi_\varepsilon < \varepsilon \quad \text{on } D
$$

Proof: Let $\varepsilon > 0$. Let $m = y_0 < y_1 < \cdots < y_{n-1} < y_n = M$ be a partition of $[m, M]$ with $y_k - y_{k-1} < \varepsilon$ for $1 \leq k \leq n$. Put $I_k = [y_{k-1}, y_k)$ for $1 \leq k \leq n - 1$ and $I_n = [y_{n-1}, y_n]$. Put $D_k = f^{-1}(I_k)$ for $1 \leq k \leq n$. Note that for $1 \leq k \leq n - 1$, we have that

$$
D_k = f^{-1}(I_k) = \{x \in D : y_{k-1} \leq f(x) < y_k\} = \{x \in D : f(x) \geq y_{k-1}\} \cap \{x \in D : f(x) < y_k\}
$$

and

$$
D_n = f^{-1}(I_n) = \{x \in D : y_{n-1} \leq f(x) \leq y_n\} = \{x \in D : f(x) \geq y_{n-1}\} \cap \{x \in D : f(x) \leq y_n\}
$$

Since $f$ is measurable, we have that $D_k$ is measurable for $1 \leq k \leq n$. △
Note that $D = \cup_{k=1}^{n} D_k$ since $f(D) \subseteq [m, M] = \cup_{k=1}^{n} I_k$. We put

$$\varphi_\varepsilon = \sum_{k=1}^{n} y_{k-1} \chi_{D_k} \quad \text{and} \quad \psi_\varepsilon = \sum_{k=1}^{n} y_k \chi_{D_k} \quad \text{on } D$$

Then $\varphi_\varepsilon$ and $\psi_\varepsilon$ are simple functions. Let $x \in D$. Then $x \in D_j$ for a unique $1 \leq j \leq n$. Since $D = \cup_{k=1}^{n} D_k$, we get that $\varphi_\varepsilon(x) = y_{j-1}$ and $\psi_\varepsilon(x) = y_j$. Note that $m \leq y_{j-1}$, $y_j \leq M$ and $y_{j-1} \leq f(x) \leq y_j$ since $x \in D_j$. Also, $0 \leq y_j - y_{j-1} < \varepsilon$. Hence

$$m \leq \varphi_\varepsilon(x) \leq f(x) \leq \psi_\varepsilon(x) \leq M \quad \text{and} \quad 0 \leq \psi_\varepsilon(x) - \varphi_\varepsilon(x) < \varepsilon \quad \text{for all } x \in D \quad \Box$$

**Theorem 3.25 (Simple Approximation Theorem)** Let $f : D \to \mathbb{R}$ be a measurable function. Then there exists a sequence of simple functions $\{f_n\}_{n \geq 1}$ defined on $D$ such that $\{f_n\}_{n \geq 1}$ converges pointwise to $f$ on $D$ and $|f_n| \leq |f|$ on $D$ for all $n \in \mathbb{N}$.

If $f \geq m$ on $D$ for some $m \in \mathbb{R}$, we may choose $f_n \geq m$ on $D$ for all $n \in \mathbb{N}$ (instead of $|f_n| \leq |f|$ on $D$ for all $n \in \mathbb{N}$) and $\{f_n\}_{n \geq 1}$ non-decreasing.

**Proof:** First, we prove this theorem assuming $f \geq m$ on $D$ for some $m \in \mathbb{R}$.

Let $n \in \mathbb{N}$. Put $D_n = \{x \in D : f(x) \leq n\}$. Then $D_n$ is measurable since $f$ is measurable. Suppose first that $D_n = \emptyset$. Put $g_n = \max\{m, n\}$ on $D$. Then $g_n$ is a simple function and $m \leq g_n \leq f$ on $D$. Suppose next that $D_n \neq \emptyset$ (note this implies that $n \geq m$). Then $f|_{D_n}$ is a bounded measurable function. By Lemma 3.24, there exist simple functions $a_n$ and $b_n$ defined on $D_n$ with

$$m \leq a_n \leq f \leq b_n \leq n \quad \text{and} \quad 0 \leq b_n - a_n < \frac{1}{n} \quad \text{on } D_n$$

Hence

$$0 \leq f - a_n \leq b_n - a_n < \frac{1}{n} \quad \text{on } D_n$$

We define $g_n$ on $D$ by extending $a_n$ to $D$:

$$g_n : D \to \mathbb{R} : x \to \begin{cases} 
    a_n(x) & \text{if } x \in D_n \\
    n = \max\{m, n\} & \text{if } x \notin D_n
\end{cases}$$

Then $g_n$ is still a simple function and $m \leq g_n \leq f$ on $D$.

For $n \in \mathbb{N}$, put $f_n = \max\{g_1, \ldots, g_n\}$. Then $f_n$ is a simple function and $m \leq f_n \leq f_{n+1} \leq f$ on $D$ for all $n \in \mathbb{N}$. We show that the sequence $\{f_n\}_{n \geq 1}$ converges to $f$ on $D$. Let $x_0 \in D$.

Suppose first that $f(x_0) = +\infty$. Then $g_n(x_0) = \max\{m, n\}$ for all $n \in \mathbb{N}$. So $f_n(x_0) = \max\{m, n\}$ for all $n \in \mathbb{N}$. Hence $\{f_n(x_0)\}_{n \geq 1}$ converges to $f(x_0)$.

Suppose next that $f(x_0)$ is finite. Let $N \in \mathbb{N}$ with $f(x_0) < N$. Then $x_0 \in D_n$ for all $n \geq N$. So

$$0 \leq f(x_0) - g_n(x_0) < \frac{1}{n} \quad \text{for all } n \geq N$$

Since $g_n(x_0) \leq f_n(x_0) \leq f(x_0)$, we get that

$$0 \leq f(x_0) - f_n(x_0) < \frac{1}{n} \quad \text{for all } n \geq N$$

Again, we see that $\{f_n(x_0)\}_{n \geq 1}$ converges to $f(x_0)$. 

47
Next, we prove the general case of the theorem. Applying the first part of our proof to $f^+$ and $f^-$ (with $m = 0$), we get that there exist non-decreasing sequences of nonnegative simple functions $\{a_n\}_{n \geq 1}$ and $\{b_n\}_{n \geq 1}$ such that $\{a_n\}_{n \geq 1}$ converges to $f^+$ on $D$ and $\{b_n\}_{n \geq 1}$ converges to $f^-$ on $D$. Put $f_n = a_n - b_n$ for all $n \in \mathbb{N}$. Then $f_n$ is still a simple function for all $n \in \mathbb{N}$ and $\{f_n\}_{n \geq 1}$ converges to $f^+ - f^- = f$ on $D$. Moreover, for all $n \in \mathbb{N}$, we have that

$$|f_n| = |a_n - b_n| \leq |a_n| + |b_n| = a_n + b_n \leq f^+ + f^- = |f| \quad \text{on } D \quad \Box$$

### 3.4 Littlewood’s Second and Third Principle

In this section, we prove Littlewood’s Second Principle:

Every measurable function is nearly continuous.

and Littlewood’s Third Principle:

Every pointwise convergent sequence of measurable functions is nearly uniformly convergent.

#### 3.4.1 Littlewood’s Third Principle

In this subsection, we prove one version of Littlewood’s Third Principle: Egoroff’s Theorem.

**Theorem 3.26 (Egoroff, Littlewood III)** Let $D \subseteq \mathbb{R}$ be a measurable set with $m(D) < +\infty$, $f_n : D \to \mathbb{R}$ a measurable function for all $n \in \mathbb{N}$ and $f : D \to \mathbb{R}$ a real-valued function such that $\{f_n\}_{n \geq 1}$ converges to $f$ on $D$. Then for all $\varepsilon > 0$, there exists a measurable subset $A$ of $D$ such that $m(A) < \varepsilon$ and $\{f_n\}_{n \geq 1}$ converges uniformly to $f$ on $D \setminus A$.

**Proof:** Note that $f$ is measurable (since it is the pointwise limit of a sequence of measurable functions). Hence $|f - f_k|$ is defined and measurable on $D$ since $f$ is real-valued.

First, we prove the following claim:

For all $\varepsilon_1, \varepsilon_2 > 0$, there exist $N \in \mathbb{N}$ and a measurable subset $A$ of $D$ such that $m(A) < \varepsilon_1$ and $|f - f_n| < \varepsilon_2$ on $D \setminus A$ for all $n \geq N$.

Indeed, let $\varepsilon_1, \varepsilon_2 > 0$. Note that $\{x \in D : |f(x) - f_k(x)| < \varepsilon_2\}$ is measurable for all $k \in \mathbb{N}$. Hence

$$D_n = \{x \in D : |f(x) - f_k(x)| < \varepsilon_2 \quad \text{for all } k \geq n\} = \bigcap_{k=n}^{\infty} \{x \in D : |f(x) - f_k(x)| < \varepsilon_2\}$$

is measurable for all $n \in \mathbb{N}$. Clearly, $\{D_n\}_{n \geq 1}$ is an increasing sequence and $\bigcup_{n=1}^{\infty} D_n = D$ (indeed, let $x_0 \in D$; since $\{f_n(x_0)\} \to f(x_0) \in \mathbb{R}$, there exists $n \in \mathbb{N}$ with $|f_k(x_0) - f(x_0)| < \varepsilon_2$ for all $k \geq n$). It follows from the Continuity of the Measure that $\{m(D_n)\}_{n \geq 1} \to m(D)$. Since $m(D) < +\infty$, there exists $N \in \mathbb{N}$ with $|m(D_n) - m(D)| < \varepsilon_1$ for all $n \geq N$. In particular, $0 \leq m(D) - m(D_N) < \varepsilon_1$. Put $A = D \setminus D_N$. By excision, $m(A) = m(D) - m(D_N) < \varepsilon_1$. Moreover, since $D_N = \{x \in D : |f(x) - f_n(x)| < \varepsilon_2 \quad \text{for all } n \geq N\}$, we have that $|f - f_n| < \varepsilon_2$ on $D_N = D \setminus A$ for all $n \geq N$, which proves the claim.
Next, we prove the theorem. Let \( \varepsilon > 0 \).

Let \( n \in \mathbb{N} \). Using the claim with \( \varepsilon_1 = \frac{\varepsilon}{2^n} \) and \( \varepsilon_2 = \frac{1}{n} \), there exist an \( N_n \in \mathbb{N} \) and a measurable subset \( A_n \) of \( D \) such that \( m(A_n) < \frac{\varepsilon}{2^n} \) and \( |f - f_k| < \frac{1}{n} \) on \( D \setminus A_n \) for all \( k \geq N_n \).

Put \( A = \bigcup_{n=1}^{\infty} A_n \). Then

\[
m(A) = m \left( \bigcup_{n=1}^{\infty} A_n \right) \leq \sum_{n=1}^{\infty} m(A_n) \leq \sum_{n=1}^{\infty} \frac{\varepsilon}{2^n} = \varepsilon
\]

by countable subadditivity.

We show that \( \{f_n\}_{n \geq 1} \) converges uniformly to \( f \) on \( D \setminus A \). Let \( \varepsilon' > 0 \). Pick \( m \in \mathbb{N} \) with \( \frac{1}{m} < \varepsilon' \). Let \( k \geq N_m \). Then \( |f_k - f| < \frac{1}{m} \) on \( D \setminus A_m \). Since \( A_m \subseteq A \), we get that \( D \setminus A \subseteq D \setminus A_m \). Hence \( |f_k - f| < \varepsilon' \) on \( D \setminus A \).

\[ \square \]

**Remark:** We may assume that \( D \setminus A \) is closed in Egoroff’s Theorem. Indeed, let \( \varepsilon > 0 \). By Egoroff’s Theorem, there exists a measurable subset \( B \) of \( D \) such that \( m(B) < \frac{\varepsilon}{2} \) and \( \{f_n\}_{n \geq 1} \) converges uniformly to \( f \) on \( D \setminus B \). By Proposition 2.28(i)(iv), there exists a closed subset \( F \) of \( D \setminus B \) with \( m((D \setminus B) \setminus F) < \frac{\varepsilon}{2} \). Put \( A = D \setminus F \). Then \( D \setminus A = F \) is closed,

\[
m(A) = m(B \cup ((D \setminus B) \setminus F)) = m(B) + m((D \setminus B) \setminus F) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon
\]

and \( \{f_n\}_{n \geq 1} \) converges uniformly to \( f \) on \( D \setminus A \) since \( D \setminus A = F \subseteq D \setminus B \).

\[ \triangleright \]

### 3.4.2 Littlewood’s Second Principle

In this subsection, we prove one version of Littlewood’s Second Principle: Lusin’s Theorem.

**Lemma 3.27** Let \( F_1, F_2, \ldots, F_n \) be pairwise disjoint closed sets and \( g_i : F_i \to \mathbb{R} \) continuous on \( F_i \) for \( 1 \leq i \leq n \). Put \( F = \bigcup_{i=1}^{n} F_i \) and

\[
g : F \to \mathbb{R} : x \to g_i(x) \text{ if } x \in F_i
\]

Then \( g \) is continuous on \( F \).

**Proof:** Note that \( g \) is well-defined since the \( F_i \)'s are pairwise disjoint. Let \( x_0 \in F \). Then \( x_0 \in F_i \) for a unique \( 1 \leq i \leq n \), say \( x_0 \in F_1 \). Let \( \varepsilon > 0 \). Since \( g_1 \) is continuous on \( F_1 \), we have:

\[
\exists \delta_1 > 0 : \forall x \in F_1 : |x - x_0| < \delta_1 \implies |g_1(x) - g_1(x_0)| < \varepsilon
\]

Because the \( F_i \)'s are pairwise disjoint, we have that \( x_0 \notin \bigcup_{i=2}^{n} F_i \). Hence \( x_0 \in \bigcup_{i=2}^{n} F_i = \bigcap_{i=2}^{n} F_i \), which is open since the \( F_i \)'s are closed. So \( (x_0 - \delta_2, x_0 + \delta_2) \subseteq \bigcup_{i=2}^{n} F_i \) for some \( \delta_2 > 0 \). Hence \( \big(x_0 - \delta_2, x_0 + \delta_2\big) \cap \big(\bigcup_{i=2}^{n} F_i\big) = \emptyset \). Put \( \delta = \min\{\delta_1, \delta_2\} \). Let \( x \in F \) with \( |x - x_0| < \delta \). Since \( \delta \leq \delta_2 \), we have that \( x \in (x_0 - \delta_2, x_0 + \delta_2) \) and so \( x \notin \big(\bigcup_{i=2}^{n} F_i\big) \). Hence \( x \in F_1 \). So \( |g(x) - g(x_0)| = |g_1(x) - g_1(x_0)| < \varepsilon \). Thus \( g \) is continuous on \( F \).

\[ \square \]

**Lemma 3.28** Let \( F \subseteq \mathbb{R} \) be closed and \( f : F \to \mathbb{R} \) continuous. Then there exists a continuous function \( g : \mathbb{R} \to \mathbb{R} \) with \( f = g \) on \( F \).

**Proof:** Exercise.
Lemma 3.29 Let $D \subseteq \mathbb{R}$ be measurable and $f : D \to \mathbb{R}$ a simple function. Then for all $\varepsilon > 0$, there exist a measurable subset $A$ of $D$ and a continuous function $g : \mathbb{R} \to \mathbb{R}$ such that $m(A) < \varepsilon$, $D \setminus A$ is closed and $f = g$ on $D \setminus A$.

Proof: Let $a_1, \ldots, a_n$ be the distinct function values of $f$ on $D$. For $1 \leq i \leq n$, put $D_i = \{x \in D : f(x) = a_i\}$. Then $D_i$ is measurable for all $1 \leq i \leq n$ and $D = \bigcup_{i=1}^{n} D_i$.

Let $\varepsilon > 0$. For $1 \leq i \leq n$, it follows from Proposition 2.28(i)(iv) that there exists a closed subset $F_i \subseteq D_i$ with $m(D_i \setminus F_i) < \frac{\varepsilon}{n}$. Put $F = \bigcup_{i=1}^{n} F_i$ and $A = \bigcup_{i=1}^{n} (D_i \setminus F_i)$. Then $A \subseteq D$, $F = D \setminus A$ and

$$m(A) \leq \sum_{i=1}^{n} m(D_i \setminus F_i) < \sum_{i=1}^{n} \frac{\varepsilon}{n} = \varepsilon$$

For $1 \leq i \leq n$, $f$ is constant on $F_i$ and hence continuous on $F_i$. So $f$ is continuous on $F$ by Lemma 3.27. By Lemma 3.28, there exists a continuous function $g : \mathbb{R} \to \mathbb{R}$ with $f = g$ on $F = D \setminus A$. $\square$

Theorem 3.30 (Lusin, Littlewood II) Let $D \subseteq \mathbb{R}$ be measurable and $f : D \to \mathbb{R}$ measurable. Then for all $\varepsilon > 0$, there exist a measurable subset $A$ of $D$ and a continuous function $g : \mathbb{R} \to \mathbb{R}$ such that $m(A) < \varepsilon$ and $f = g$ on $D \setminus A$. Moreover, we can choose $A$ such that $D \setminus A$ is closed.

Proof: We prove this theorem assuming $m(D) < \infty$ and leave the general case as an exercise. Let $\varepsilon > 0$. By the Simple Approximation Theorem, there exists a sequence of simple functions $\{f_n\}_{n \geq 1}$, defined on $D$, that converges to $f$ on $D$. For $n \in \mathbb{N}$, it follows from Lemma 3.29 that there exist a measurable subset $A_n \subseteq D$ and a continuous function $g_n : \mathbb{R} \to \mathbb{R}$ such that $m(A_n) < \frac{\varepsilon}{2n+1}$, $D \setminus A_n$ is closed and $f_n = g_n$ on $D \setminus A_n$. By the remark after Egoroff’s Theorem (here we use that $m(D) < \infty$), there exists a measurable subset $A_0 \subseteq D$ such that $m(A_0) < \frac{\varepsilon}{2}$, $D \setminus A_0$ is closed and $\{f_n\}_{n \geq 1}$ converges uniformly to $f$ on $D \setminus A_0$. Put $A = \bigcup_{n=0}^{\infty} A_n$. Then $A \subseteq D$,

$$m(A) \leq \sum_{n=0}^{\infty} m(A_n) < \sum_{n=0}^{\infty} \frac{\varepsilon}{2n+1} = \varepsilon$$

and $D \setminus A = \bigcap_{n=0}^{\infty} (D \setminus A_n)$ so $D \setminus A$ is closed. Since $f_n = g_n$ on $D \setminus A_n$ and $D \setminus A \subseteq D \setminus A_n$, we get that $f_n$ is continuous on $D \setminus A$ for all $n \geq 1$. Since $\{f_n\}_{n \geq 1}$ converges uniformly to $f$ on $D \setminus A_0$ and $D \setminus A \subseteq D \setminus A_0$, we see that $\{f_n\}_{n \geq 1}$ converges uniformly to $f$ on $D \setminus A$. Since $f_n$ is continuous on $D \setminus A$ for all $n \geq 1$ and the convergence is uniform, we have that $f$ is continuous on $D \setminus A$. Since $D \setminus A$ is closed, it follows from Lemma 3.28 that there exists a continuous function $g : \mathbb{R} \to \mathbb{R}$ with $f = g$ on $D \setminus A$. $\square$
Chapter 4

The Lebesgue Integral

4.1 Lebesgue Integral of Nonnegative Simple Functions

Definition 4.1 Let \( \varphi : E \to \mathbb{R} \) be a simple function.

(a) Let \( a_1, \ldots, a_n \) be the different function values of \( \varphi \). For \( 1 \leq i \leq n \), put \( A_i = \{ x \in E \mid \varphi(x) = a_i \} \).

Note that \( A_i \) is measurable for all \( 1 \leq i \leq n \). Then

\[ \varphi = \sum_{i=1}^{n} a_i \chi_{A_i} \]

This is called the canonical representation or canonical form of \( \varphi \).

(b) Suppose \( \varphi \) is nonnegative. The Lebesgue integral of \( \varphi \) over \( E \) (notation: \( \int_E \varphi \) ) is

\[ \int_E \varphi = \sum_{i=1}^{n} a_i m(A_i) \]

where \( \varphi = \sum_{i=1}^{n} a_i \chi_{A_i} \) is the canonical representation of \( \varphi \) and \( a_i m(A_i) = 0 \) if \( a_i = 0 \). □

Remark: It follows immediately that we don’t have to write down 0 as a function value of \( \varphi \) in the canonical form of \( \varphi \) to calculate \( \int_E \varphi \) : if \( a_n = 0 \) then \( \varphi = \sum_{i=1}^{n-1} a_i \chi_{A_i} \) and \( \int_E \varphi = \sum_{i=1}^{n-1} a_i m(A_i) \). □

Lemma 4.2 Let \( E \) be a measurable set, \( C_1, C_2, \ldots, C_m \) disjoint measurable subsets of \( E \) and \( c_1, \ldots, c_m \) nonnegative real numbers. Put \( \varphi = \sum_{j=1}^{m} c_j \chi_{C_j} \). Then \( \int_E \varphi = \sum_{i=1}^{m} c_j m(C_j) \) where \( c_j m(C_j) = 0 \) if \( c_j = 0 \).

Proof: We may assume that \( C_j \neq \emptyset \) for \( j = 1, 2, \ldots, m \) and that \( E = \bigcup_{j=1}^{m} C_j \). Let \( \varphi = \sum_{i=1}^{n} a_i \chi_{A_i} \) be the canonical representation of \( \varphi \). For \( i = 1, 2, \ldots, n \), put \( \Omega_i = \{ 1 \leq j \leq m \mid c_j = a_i \} \).

Since the \( a_i \)'s are distinct, we have that the \( \Omega_i \)'s are disjoint. Pick \( 1 \leq j \leq m \). Pick \( x \in C_j \). Then there exists a unique \( 1 \leq i \leq n \) with \( x \in A_i \). Hence \( a_i = \varphi(x) = c_j \) and \( j \in \Omega_i \). So \( \{ \Omega_i \mid i = 1, 2, \ldots, n \} \) forms a partition of \( \{1, 2, \ldots, m\} \).

Pick \( 1 \leq i \leq n \). For all \( j \in \Omega_i \) and all \( x \in C_j \), we have that \( \varphi(x) = c_j = a_i \) and so \( x \in A_i \). Hence \( \cup_{j \in \Omega_i} C_j \subseteq A_i \). Pick \( x \in A_i \). Then there exists a unique \( 1 \leq j \leq m \) with \( x \in C_j \). So \( a_i = \varphi(x) = c_j \). Hence \( j \in \Omega_i \) and \( A_i \subseteq \cup_{j \in \Omega_i} C_j \).
So $A_i = \cup_{j \in \Omega} C_j$. Hence $m(A_i) = \sum_{j \in \Omega} m(C_j)$. We now get that
\[
\sum_{j=1}^{m} c_j m(C_j) = \sum_{i=1}^{n} \sum_{j \in \Omega_i} c_j m(C_j) = \sum_{i=1}^{n} \sum_{j \in \Omega_i} a_i m(C_j) = \sum_{i=1}^{n} a_i \sum_{j \in \Omega_i} m(C_j) = \sum_{i=1}^{n} a_i m(A_i) = \int_E \varphi \quad \Box
\]

**Proposition 4.3** Let $\varphi, \psi : E \to \mathbb{R}$ be nonnegative simple functions. Then the following holds:

(a) If $\varphi = \sum_{i=1}^{n} a_i \chi_{A_i}$, where $A_i \subseteq E$ is measurable and $a_i \geq 0$ for $i = 1, 2, \ldots, n$, then $\int_E \varphi = \sum_{i=1}^{n} a_i m(A_i)$ where $a_i m(A_i) = 0$ if $a_i = 0$.

(b) $\int_E (a \varphi + c \psi) = a \int_E \varphi + c \int_E \psi$ for all $a, c \geq 0$.

(c) If $E = E_1 \cup E_2$ where $E_1, E_2$ are measurable, then $\int_E \varphi = \int_{E_1} \varphi + \int_{E_2} \varphi$.

(d) If $m(E) = 0$ then $\int_E \varphi = 0$.

(e) If $\varphi = \psi$ a.e. on $E$ then $\int_E \varphi = \int_E \psi$.

(f) If $\varphi \leq \psi$ a.e. on $E$ then $\int_E \varphi \leq \int_E \psi$.

**Proof:**

(b) Let $\varphi = \sum_{i=1}^{m} a_i \chi_{A_i}$, and $\psi = \sum_{j=1}^{n} c_j \chi_{C_j}$ be the canonical forms. By definition, we get that
\[
\int_E \varphi = \sum_{i=1}^{m} a_i m(A_i) \quad \text{and} \quad \int_E \psi = \sum_{j=1}^{n} c_j m(C_j)
\]

Put $B_{ij} = A_i \cap C_j$ for $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$. Then $\{B_{ij} \mid 1 \leq i \leq m, 1 \leq j \leq n\}$ forms a partition of $E$. Moreover, $A_i = \cup_{j=1}^{m} B_{ij}$ and so $m(A_i) = \sum_{j=1}^{n} m(B_{ij})$ for $i = 1, 2, \ldots, m$. Similarly, $C_j = \cup_{i=1}^{m} B_{ij}$ and so $m(C_j) = \sum_{i=1}^{m} m(B_{ij})$ for $j = 1, 2, \ldots, n$. Note that
\[
\varphi = \sum_{i=1}^{m} \sum_{j=1}^{n} a_i \chi_{B_{ij}} \quad \text{and} \quad \psi = \sum_{i=1}^{m} \sum_{j=1}^{n} c_j \chi_{B_{ij}} \quad \text{and so} \quad a \varphi + c \psi = \sum_{i=1}^{m} \sum_{j=1}^{n} (a a_i + c c_j) \chi_{B_{ij}}
\]

By Lemma 4.2, we get that
\[
\int_E (a \varphi + c \psi) = \sum_{i=1}^{m} \sum_{j=1}^{n} (a a_i + c c_j) m(B_{ij}) = a \sum_{i=1}^{m} a_i \sum_{j=1}^{n} m(B_{ij}) + c \sum_{j=1}^{n} c_j \sum_{i=1}^{m} m(B_{ij}) = a \sum_{i=1}^{m} a_i m(A_i) + c \sum_{j=1}^{n} c_j m(C_j) = a \int_E \varphi + c \int_E \psi
\]

(a) Note that $\chi_{A_i}$ is a nonnegative simple function on $E$ and $\int_E \chi_{A_i} = m(A_i)$ for $i = 1, 2, \ldots, n$. Hence by (b), we get
\[
\int_E \left( \sum_{i=1}^{n} a_i \chi_{A_i} \right) = \sum_{i=1}^{n} a_i \int_E \chi_{A_i} = \sum_{i=1}^{n} a_i m(A_i)
\]

(c) Let $\varphi = \sum_{i=1}^{n_j} a_{ij} \chi_{A_{ij}}$ be the canonical form of $\varphi$ on $E_j$ for $j = 1, 2$. Since the $A_{ij}$’s are pairwise disjoint, we get
\[
\varphi = \sum_{j=1}^{2} \sum_{i=1}^{n_j} a_{ij} \chi_{A_{ij}}
\]
By (a), we get that
\[ \int_E \varphi = \sum_{j=1}^{2} \sum_{i=1}^{n_j} a_{ij} m(A_{ij}) = \sum_{j=1}^{2} \int_{E_j} \varphi. \]

(d) Let \( \varphi = \sum_{i=1}^{n} a_{i} \chi_{A_{i}} \) be the canonical form of \( \varphi \) on \( E \). Since \( A_{i} \subseteq E \) and \( m(E) = 0 \), we get that \( m(A_{i}) = 0 \) for \( i = 1, 2, \ldots, n \). Hence by definition, we get that
\[ \int_E \varphi = \sum_{i=1}^{n} a_{i} m(A_{i}) = 0. \]

(e) Put \( D = \{ x \in E | \varphi(x) \neq \psi(x) \} \). Then \( m(D) = 0 \). Note that \( \varphi|_{E \setminus D} = \psi|_{E \setminus D} \). Using (c) and (d), we find
\[ \int_E \varphi = \int_{E \setminus D} \varphi + \int_D \varphi = \int_{E \setminus D} \varphi = \int_{E \setminus D} \psi + \int_D \psi = \int_E \psi. \]

(f) Similarly as in (e), we may assume that \( \varphi \leq \psi \) on \( E \). Note that \( \psi - \varphi \) is a nonnegative simple function. Clearly, \( \int_E (\psi - \varphi) \geq 0 \). Hence by (b), we have
\[ \int_E \psi = \int_E [\varphi + (\psi - \varphi)] = \int_E \varphi + \int_E (\psi - \varphi) \geq \int_E \varphi. \]

4.2 Lebesgue Integral of Nonegative Measurable Functions

**Definition 4.4** : Let \( f : E \to \mathbb{R} \) be a nonnegative measurable function.

(1) We define the set \( \Phi_f \) as follows :
\[ \Phi_f = \{ \varphi | \varphi \text{ is a nonnegative simple function on } E \text{ and } \varphi \leq f \text{ on } E \} \]

(2) We define the Lebesgue integral of \( f \) over \( E \) (notation : \( \int_E f \)) as
\[ \int_E f = \sup_{\varphi \in \Phi_f} \int_E \varphi. \]

**Remark** : This definition of \( \int_E f \) is consistent with our ‘old’ definition if \( f \) is a nonnegative simple function. Indeed, suppose that \( f \) is a nonnegative simple function. Then \( f \in \Phi_f \) and so
\[ \sup_{\varphi \in \Phi_f} \int_E \varphi \geq \int_E f. \]
On the other hand, if \( \varphi \in \Phi_f \), then \( \varphi \leq f \) on \( E \) and so \( \int_E \varphi \leq \int_E f \) by Proposition 4.3(f). Since this is true for all \( \varphi \in \Phi_f \), we get
\[ \sup_{\varphi \in \Phi_f} \int_E \varphi \leq \int_E f. \]
\[ \sup_{\varphi \in \Phi_f} \int_E \varphi = \int_E f. \]

**Proposition 4.5** Let \( f, g : E \to \mathbb{R} \) be nonnegative measurable functions. Then the following holds

(a) \( \int_D f = \int_E (f \cdot \chi_D) \) for all measurable \( D \subseteq E \).

(b) \( \int_E (\lambda f) = \lambda \int_E f \) for all \( \lambda \geq 0 \).

(c) If \( E = E_1 \cup E_2 \) where \( E_1, E_2 \) are measurable, then \( \int_E f = \int_{E_1} f + \int_{E_2} f \).

(d) If \( m(E) = 0 \) then \( \int_E f = 0 \).
If $f = g$ a.e. on $E$ then $\int_E f = \int_E g$.

If $f \leq g$ a.e. on $E$ then $\int_E f \leq \int_E g$.

**Proof:** (a) Let $D \subseteq E$ be measurable.

Pick $\varphi \in \Phi_{f,D}$. Put $\psi : E \to \mathbb{R} : x \to \begin{cases} \varphi(x) & \text{if } x \in D \\ 0 & \text{if } x \in E \setminus D \end{cases}$. Then $\psi \in \Phi_{f \cdot \chi_D,E}$. If $\varphi = \sum_{i=1}^{n} a_i \chi_{A_i}$ is the canonical form of $\varphi$ over $D$, then $\psi = \sum_{i=1}^{n} a_i \chi_{A_i}$ over $E$. So

$$\int_D \varphi = \int_E \psi \leq \sup_{\psi \in \Phi_{f \cdot \chi_D,E}} \int_E \psi = \int_E (f \cdot \chi_D)$$

Since this is true for all $\varphi \in \Phi_{f,D}$, we get

$$\int_D f = \sup_{\varphi \in \Phi_{f,D}} \int_D \varphi \leq \int_E (f \cdot \chi_D) \quad (\%)$$

Pick $\varphi \in \Phi_{f \cdot \chi_D,E}$. Put $\psi = \varphi |_D$. Then $\psi \in \Phi_{f,D}$. If $\psi = \sum_{i=1}^{n} a_i \chi_{A_i}$ is the canonical form of $\psi$ over $D$, then $\varphi = \sum_{i=1}^{n} a_i \chi_{A_i}$ over $E$. So

$$\int_E \varphi = \int_D \psi \leq \sup_{\psi \in \Phi_{f,D}} \int_D \psi = \int_D f$$

Since this is true for all $\varphi \in \Phi_{f \cdot \chi_D,E}$, we get that

$$\int_E (f \cdot \chi_D) = \sup_{\varphi \in \Phi_{f \cdot \chi_D,E}} \int_E \varphi \leq \int_D f \quad (\%\%)$$

From (\%) and (\%\%), it follows that $\int_D f = \int_E (f \cdot \chi_D)$.

(b) We may assume that $\lambda > 0$. Pick $\varphi \in \Phi_f$. Then $\lambda \varphi \in \Phi_{\lambda f}$. By Proposition 4.3(b), we get that

$$\int_E (\lambda f) = \sup_{\psi \in \Phi_{\lambda f}} \int_E \psi \geq \int_E (\lambda \varphi) = \lambda \int_E \varphi$$

Since this is true for all $\varphi \in \Phi_f$, we get that

$$\int_E (\lambda f) \geq \sup_{\varphi \in \Phi_f} \left( \lambda \int_E \varphi \right) = \lambda \sup_{\varphi \in \Phi_f} \int_E \varphi = \lambda \int_E f \quad (*)$$

Since (*) is true for all $\lambda > 0$ and all nonnegative measurable functions $f$, we can replace $\lambda$ by $\frac{1}{\lambda}$ and $f$ by $\lambda f$ to get

$$\int_E f = \int_E \left( \frac{1}{\lambda} (\lambda f) \right) \geq \frac{1}{\lambda} \int_E (\lambda f) \quad \text{and so} \quad \lambda \int_E f \geq \int_E (\lambda f) \quad (***)$$

From (*) and (**), it follows that $\int_E (\lambda f) = \lambda \int_E f$.

(c) Pick $\varphi \in \Phi_f$. Then $\varphi |_{E_i} \in \Phi_{f,E_i}$ for $i = 1, 2$. Using Proposition 4.3(c), we get

$$\int_E \varphi = \int_{E_1} \varphi |_{E_1} + \int_{E_2} \varphi |_{E_2} \leq \sup_{\psi \in \Phi_{f,E_1}} \int_{E_1} \psi + \sup_{\psi \in \Phi_{f,E_2}} \int_{E_2} \psi = \int_{E_1} f + \int_{E_2} f$$
Since this is true for all \( \varphi \in \Phi_f \), we get

\[
\int_E f = \sup_{\varphi \in \Phi_f} \int_E \varphi \leq \int_{E_1} f + \int_{E_2} f \quad (\#)
\]

Pick \( \varphi_i \in \Phi_{f,E_i} \) for \( i = 1, 2 \). Define \( \varphi : E \to \mathbb{R} : x \to \begin{cases} \varphi_1(x) & \text{if } x \in E_1 \\ \varphi_2(x) & \text{if } x \in E_2 \end{cases} \). Note that this is well-defined since \( E = E_1 \cup E_2 \). Then \( \varphi \in \Phi_f \) and \( \varphi|_{E_i} = \varphi_i \) for \( i = 1, 2 \). Using Proposition 4.3(c), we get that

\[
\int_{E_1} \varphi_1 + \int_{E_2} \varphi_2 = \int_{E_1} \varphi|_{E_1} + \int_{E_2} \varphi|_{E_2} = \int_E \varphi \leq \sup_{\psi \in \Phi_f} \int_E \psi = \int_E f
\]

Since this is true for all \( \varphi_1 \in \Phi_{f,E_1} \) and all \( \varphi_2 \in \Phi_{f,E_2} \), we get that

\[
\int_{E_1} f + \int_{E_2} f = \sup_{\varphi_1 \in \Phi_{f,E_1}} \int_{E_1} \varphi_1 + \sup_{\varphi_2 \in \Phi_{f,E_2}} \int_{E_2} \varphi_2 \leq \int_E f \quad (##)
\]

From (\#) and (##), it follows that \( \int_E f = \int_{E_1} f + \int_{E_2} f \).

(d) Pick \( \varphi \in \Phi_f \). By Proposition 4.3(d), we have that \( \int_E \varphi = 0 \). Since this is true for all \( \varphi \in \Phi_f \), we get that

\[
\int_E f = \sup_{\varphi \in \Phi_f} \int_E \varphi = \sup_{\varphi \in \Phi_f} 0 = 0
\]

(e) Put \( D = \{ x \in E \mid f(x) \neq g(x) \} \). Then \( m(D) = 0 \). Note that \( f = g \) on \( E \setminus D \). Using (c) and (d), we get that

\[
\int_E f = \int_{E \setminus D} f + \int_D f = \int_{E \setminus D} g = \int_{E \setminus D} g + \int_D g = \int_E g
\]

(f) Similarly as in (e), we may assume that \( f \leq g \) on \( E \). Pick \( \varphi \in \Phi_f \). Since \( f \leq g \) on \( E \), we get that \( \varphi \in \Phi_g \). Hence

\[
\int_E g = \sup_{\psi \in \Phi_g} \int_E \psi \geq \int_E \varphi
\]

Since this is true for all \( \varphi \in \Phi_f \), we have that

\[
\int_E g \geq \sup_{\varphi \in \Phi_f} \int_E \varphi = \int_E f
\]

Remark: At this point, we cannot prove linearity for the Lebesgue integral of nonnegative measurable functions. We will need to develop quite a bit of machinery (Fatou’s Lemma and the Monotone Convergence Theorem) before we can prove linearity.

4.3 Fatou’s Lemma, Monotone Convergence Theorem

Lemma 4.6 (Fatou’s Lemma) Let \( f, f_n : E \to \overline{\mathbb{R}} \) be nonnegative measurable functions for all \( n \geq 1 \) such that \( f = \lim_{n \to \infty} f_n \) a.e. on \( E \). Then

\[
\int_E f \leq \lim_{n \to \infty} \int_E f_n
\]
Proof: We may assume that $f = \lim f_n$ on $E$.

Pick $\varphi \in \Phi_f$ and $0 < t < 1$. We will show that $t \int_E \varphi \leq \lim \int_E f_n$. So we may assume that $\varphi \neq 0$. Then we can write $\varphi = \sum_{k=1}^{m} a_k \chi_{A_k}$ where $a_1, a_2, \ldots, a_m > 0$ and $A_1, A_2, \ldots, A_m$ are disjoint measurable subsets of $E$.

For $n \geq 1$ and $1 \leq k \leq m$, we define

$$B_{kn} = \{x \in A_k | f_l(x) > ta_k \text{ for all } l \geq n\} = \bigcap_{l=n}^{+\infty} \{x \in A_k | f_l(x) > ta_k\}$$

Hence $\{B_{kn}\}_{n \geq 1}$ is an increasing sequence of measurable subsets of $A_k$ for $1 \leq k \leq m$.

CLAIM: $A_k = \bigcup_{n=1}^{+\infty} B_{kn}$ for $1 \leq k \leq m$.

Proof: Pick $1 \leq k \leq m$. Clearly, $\bigcup_{n=1}^{+\infty} B_{kn} \subseteq A_k$. Pick $x \in A_k$. Note that

$$\lim f_n(x) = f(x) \geq \varphi(x) = a_k > ta_k$$

Suppose first that $f(x) = +\infty$. Then $\lim f_n(x) = +\infty$ and so

$$\forall M > 0 : \exists n \in \mathbb{N} : \forall l \geq n : f_l(x) > M$$

Using this with $\{M = ta_k\}$, we get that there exists $n \in \mathbb{N}$ such that $f_l(x) > ta_k$ for all $l \geq n$. So $x \in B_{kn}$.

Suppose next that $f(x) < +\infty$. Then $\lim f_n(x) < +\infty$ and so

$$\forall \epsilon > 0 : \exists n \in \mathbb{N} : \forall l \geq n : f_l(x) > \lim f_n(x) - \epsilon$$

Using this with $\{\epsilon = \lim f_n(x) - ta_k\}$, we get that there exists $n \in \mathbb{N}$ such that $f_l(x) > ta_k$ for all $l \geq n$. Again $x \in B_{kn}$.

Hence $A_k \subseteq \bigcup_{n=1}^{+\infty} B_{kn}$. □

By the Continuity of the Measure, we get that $m(A_k) = \lim_{n \to \infty} m(B_{kn})$ for $1 \leq k \leq m$.

Pick $n \geq 1$. Put $\psi_n = \sum_{k=1}^{m} ta_k \chi_{B_{kn}}$. Pick $x \in E$. If $x \in B_{kn}$ for some $1 \leq k \leq m$, then $x \in A_k$ and $f_n(x) > ta_k$. So $\psi_n(x) = ta_k < f_n(x)$. If $x \notin B_{kn}$ for all $1 \leq k \leq m$, then $\psi_n(x) = 0 < f_n(x)$. Hence $\psi_n \in \Phi_{f_n}$ and so by Proposition 4.3(a), we get

$$\int_E f_n = \sup_{\psi \in \Phi_{f_n}} \int_E \psi \geq \int_E \psi_n = \int_E \left( \sum_{k=1}^{m} ta_k \chi_{B_{kn}} \right) = \sum_{k=1}^{m} ta_k m(B_{kn})$$

Since this is true for all $n \geq 1$, we can apply the limit inferior to both sides. Using the fact that $m(A_k) = \lim_{n \to \infty} m(B_{kn})$ for $1 \leq k \leq m$ and Proposition 4.3(a), we get

$$\lim \int_E f_n \geq \lim \left( \sum_{k=1}^{m} ta_k m(B_{kn}) \right) = \lim_{n \to +\infty} \left( \sum_{k=1}^{m} ta_k m(B_{kn}) \right) = t \sum_{k=1}^{m} a_k m(A_k) = t \int_E \varphi$$

Since this is true for all $0 < t < 1$, we can take the limit as $t$ goes to $1^-$ on both sides:

$$\lim \int_E f_n \geq \lim_{t \to 1^-} \left( t \int_E \varphi \right) = \int_E \varphi$$

56
Since this is true for all $\varphi \in \Phi_f$, we finally get
\[
\lim \int_E f_n \geq \sup_{\varphi \in \Phi_f} \int_E \varphi = \int f
\]

Remark: Another way of viewing Fatou’s Lemma: $\int_E \lim \phi_n \leq \lim \int_E f_n$. ▷

**Theorem 4.7** Let $f, f_n : E \to \mathbb{R}$ be nonnegative measurable functions for all $n \geq 1$ such that $\{f_n\}_{n \geq 1} \to f$ a.e. on $E$ and $f_n \leq f$ a.e. on $E$ for all $n \geq 1$. Then $\{\int_E f_n\}_{n \geq 1} \to \int_E f$.

**Proof:** Since $f_n \leq f$ a.e. on $E$, we get by Proposition 4.5(f) that
\[
\int_E f_n \leq \int_E f \quad \text{for all } n \geq 1
\]
Applying the limit superior to both sides, we find
\[
\lim \int_E f_n \leq \int_E f \quad (\ast)
\]
By Fatou’s Lemma, we get
\[
\int_E f \leq \lim \int_E f_n \quad (\ast\ast)
\]
Combining $(\ast)$ and $(\ast\ast)$, we find
\[
\lim \int_E f_n \leq \int_E f \leq \lim \int_E f_n
\]
Hence $\lim \int_E f_n = \lim \int_E f_n = \int_E f$ and $\{\int_E f_n\}_{n \geq 1} \to \int_E f$. □

**Corollary 4.8** (Monotone Convergence Theorem) Let $f, f_n : E \to \mathbb{R}$ be nonnegative measurable functions for all $n \geq 1$ such that $f_1 \leq f_2 \leq f_3 \leq \cdots$ and $\{f_n\}_{n \geq 1} \to f$ a.e. on $E$. Then $\{\int_E f_n\}_{n \geq 1} \to \int_E f$.

**Proof:** Since $\{f_n\}_{n \geq 1}$ is an increasing sequence of functions converging to $f$ a.e. on $E$, we have that $f_n \leq f$ a.e. on $E$ for all $n \geq 1$. Hence the result follows from Theorem 4.7. □

Remark: Another way of viewing the last theorem and corollary: $\int_E \lim_{n \to +\infty} f_n = \lim_{n \to +\infty} \int_E f_n$. ▷

**Proposition 4.9** Let $f, g, f_n : E \to \mathbb{R}$ be nonnegative measurable functions for all $n \geq 1$. Then the following holds:

(a) $\int_E (f + g) = \int_E f + \int_E g$

(b) $\int_E \left( \sum_{n=1}^{+\infty} f_n \right) = \sum_{n=1}^{+\infty} \int_E f_n$

(c) If $E_1, E_2, \ldots$ are disjoint measurable sets such that $E = \bigcup_{n=1}^{+\infty} E_n$, then $\int_E f = \sum_{n=1}^{+\infty} \int_{E_n} f$

57
Proof: (a) By the Simple Approximation Theorem, there exists a nondecreasing sequence of nonnegative simple functions \( \{\varphi_n\}_{n \geq 1} \) (resp. \( \{\psi_n\}_{n \geq 1} \)) that converges to \( f \) (resp. \( g \)) on \( E \). Then \( \{\varphi_n + \psi_n\}_{n \geq 1} \) is a nondecreasing sequence of nonnegative simple functions that converges to \( f + g \) on \( E \). By the Monotone Convergence Theorem, we get

\[
\left\{ \int_E \varphi_n \right\}_{n \geq 1} \to \int_E f, \quad \left\{ \int_E \psi_n \right\}_{n \geq 1} \to \int_E g \text{ and } \left\{ \int_E (\varphi_n + \psi_n) \right\}_{n \geq 1} \to \int_E (f + g)
\]

By Proposition 4.3(b), we have that \( \int_E (\varphi_n + \psi_n) = \int_E \varphi_n + \int_E \psi_n \) for all \( n \geq 1 \). Hence

\[
\left\{ \int_E (\varphi_n + \psi_n) \right\}_{n \geq 1} = \left\{ \int_E \varphi_n + \int_E \psi_n \right\}_{n \geq 1} \to \int_E f + \int_E g
\]

So \( \int_E (f + g) = \int_E f + \int_E g \).

(b) For \( k \geq 1 \), put \( S_k = \sum_{n=1}^{k} f_n \). Then \( \{S_k\}_{k \geq 1} \) is a nondecreasing sequence of nonnegative measurable functions that converges to \( \sum_{n=1}^{+\infty} f_n \) on \( E \). By the Monotone Convergence Theorem, we get that

\[
\left\{ \int_E S_k \right\}_{k \geq 1} \to \int_E \left( \sum_{n=1}^{+\infty} f_n \right)
\]

By (a), we get that \( \int_E S_k = \sum_{n=1}^{k} \int_E f_n \) for all \( k \geq 1 \). Hence

\[
\left\{ \int_E S_k \right\}_{k \geq 1} = \left\{ \sum_{n=1}^{k} \int_E f_n \right\}_{k \geq 1} \to \sum_{n=1}^{+\infty} \int_E f_n
\]

So \( \int_E \left( \sum_{n=1}^{+\infty} f_n \right) = \sum_{n=1}^{+\infty} \int_E f_n \).

(c) Note that \( f = \sum_{n=1}^{+\infty} f \chi_{E_n} \). By (b) and Proposition 4.5(a), we get

\[
\int_E f = \int_E \left( \sum_{n=1}^{+\infty} f \chi_{E_n} \right) = \sum_{n=1}^{+\infty} \int_E f \chi_{E_n} = \sum_{n=1}^{+\infty} \int_{E_n} f
\]

4.4 Lebesgue Integral of Measurable Functions

Definition 4.10:

(a) Let \( f : E \to \mathbb{R} \) be a nonnegative measurable function. Then \( f \) is **Lebesgue integrable over** \( E \) if \( \int_E f \) is finite.

(b) Let \( f : E \to \mathbb{R} \) be a measurable function. Then \( f \) is **Lebesgue integrable over** \( E \) if and only if \( f^+ \) and \( f^- \) are Lebesgue integrable over \( E \). In that case we define the **Lebesgue integral of** \( f \) **over** \( E \) (notation : \( \int_E f \)) as

\[
\int_E f = \int_E f^+ - \int_E f^-
\]

Remark : This definition of \( \int_E f \) is consistent with our ‘old’ definition if \( f \) is a nonnegative measurable function because \( f^+ = f \) and \( f^- = 0 \) if \( f \) is a nonnegative measurable function.
Lemma 4.11 Let \( f, g : E \to \mathbb{R} \) be nonnegative measurable functions such that \( g \) is integrable over \( E \) and \( f \leq g \) a.e. on \( E \). Then \( f \) is integrable over \( E \).

Proof: Since \( f \) is nonnegative, we have that \( f^{-} = 0 \) and \( f^{+} = f \). So \( \int_E f^{-} = \int_E 0 = 0 \). By Proposition 4.5(f), we have \( \int_E f^{+} = \int_E f \leq \int_E g < +\infty \). Hence \( f \) is integrable over \( E \). \( \square \)

Lemma 4.12 Let \( f : E \to \mathbb{R} \) be measurable. Then \( f \) is integrable over \( E \) if and only if \( |f| \) is integrable over \( E \).

Proof: Suppose first that \( f \) is integrable over \( E \). Then \( f^{+} \) and \( f^{-} \) are integrable over \( E \). Since \( |f| = f^{+} + f^{-} \), we get that \( \int_E |f| = \int_E (f^{+} + f^{-}) = \int_E f^{+} + \int_E f^{-} < +\infty \) by Proposition 4.9(a). So \( |f| \) is integrable over \( E \).

Suppose next that \( |f| \) is integrable over \( E \). Since \( f^{+}, f^{-} \leq |f| \) over \( E \), we get that \( f^{+} \) and \( f^{-} \) are integrable over \( E \) by Lemma 4.11. Hence \( f \) is integrable over \( E \). \( \square \)

Lemma 4.13 (Tchebychev’s Inequality) Let \( f : D \to \mathbb{R} \) be a nonnegative measurable function. For \( \alpha > 0 \), put \( D_\alpha = \{ x \in D \mid f(x) \geq \alpha \} \). Then \( m(D_\alpha) \leq \frac{1}{\alpha} \int_D f \) for all \( \alpha > 0 \).

Proof: Exercise. \( \square \)

Lemma 4.14 Let \( f : E \to \mathbb{R} \) be integrable over \( E \). Then \( f \) is finite a.e. on \( E \).

Proof: Put \( D = \{ x \in E \mid f(x) = \pm \infty \} \). By Lemma 4.12, \( |f| \) is integrable over \( E \). Pick \( n \geq 1 \). Put \( D_n = \{ x \in E \mid |f(x)| > n \} \). By Tchebychev’s Inequality, we get that \( m(D_n) \leq \frac{1}{n} \int_E |f| \). Since \( D \subseteq D_n \), we get that

\[
m(D) \leq m(D_n) \leq \frac{1}{n} \int_E |f|
\]

Since this is true for all \( n \geq 1 \), we can take the limit as \( n \) goes to \( +\infty \) on both sides. Since \( |f| \) is integrable over \( E \), we find

\[
m(D) \leq \lim_{n \to +\infty} \frac{1}{n} \int_E |f| = 0
\]

\( \square \)

Lemma 4.15 Let \( f : E \to \mathbb{R} \) be measurable and \( D \subseteq E \) with \( m(D) = 0 \). Then \( f \) is integrable over \( E \) if and only if \( f \) is integrable over \( E \setminus D \).

Proof: By Proposition 4.5(c)(d), we get that

\[
\int_E f^\pm = \int_{E \setminus D} f^\pm + \int_D f^\pm = \int_{E \setminus D} f^\pm
\]

Hence \( \int_E f^\pm \) is finite if and only if \( \int_{E \setminus D} f^\pm \) is finite. \( \square \)

Remark : Let \( f, g : E \to \mathbb{R} \) be integrable over \( E \). By Lemma 4.14, \( f + g \) is well-defined a.e. on \( E \). So by Lemma 4.15, we can investigate if \( f + g \) is integrable over \( E \), even though \( f + g \) might not be defined everywhere. \( \triangleright \)

Lemma 4.16 Let \( h_1, h_2 : E \to \mathbb{R} \) be nonnegative and integrable over \( E \). Put \( h = h_1 - h_2 \). Then \( h \) is integrable over \( E \) and \( \int_E h = \int_E h_1 - \int_E h_2 \).
Proof: We will show that $h^+ \leq h_1$ and $h^- \leq h_2$. Pick $x \in E$. If $h(x) \leq 0$, then $h^+(x) = 0 \leq h_1(x)$ and $h_2(x) = h_1(x) - h(x) \geq -h(x) = h^-(x)$. If $h(x) \geq 0$, then $h^-(x) = 0 \leq h_2(x)$ and $h_1(x) = h(x) + h_2(x) \geq h(x) = h^+(x)$. Hence $h^+ \leq h_1$ and $h^- \leq h_2$. By Lemma 4.11, $h^+$ and $h^-$ are integrable over $E$. So $h$ is integrable over $E$. Since $h^+ - h^- = h = h_1 - h_2$, we get that $h^+ + h_2 = h^- + h_1$. By Proposition 4.9(a), we find

$$\int_E h^+ + \int_E h_2 = \int_E (h^+ + h_2) = \int_E (h^- + h_1) = \int_E h^- + \int_E h_1$$

So $\int_E h = \int_E h^+ - \int_E h^- = \int_E h_1 - \int_E h_2$. □

Proposition 4.17 Let $f, g : E \to \mathbb{R}$ be integrable over $E$. Then the following holds

(a) $f$ is integrable over $D$ and $\int_D f = \int_E (f \cdot \chi_D)$ for all measurable $D \subseteq E$.

(b) $\lambda f + \mu g$ is integrable over $E$ and $\int_E (\lambda f + \mu g) = \lambda \int_E f + \mu \int_E g$ for all $\lambda, \mu \in \mathbb{R}$.

(c) If $E = \bigcup_{n=1}^{+\infty} E_n$ where $E_n$ is measurable for all $n \geq 1$, then $\int_E f = \sum_{n=1}^{+\infty} \int_{E_n} f$.

(d) If $m(E) = 0$ then $\int_E f = 0$.

(e) If $f = g$ a.e. on $E$ then $\int_E f = \int_E g$.

(f) If $f \leq g$ a.e. on $E$ then $\int_E f \leq \int_E g$.

(g) $|\int_E f| \leq \int_E |f|$.

Proof: (a) Note that $(f \cdot \chi_D)^\pm = f^\pm \cdot \chi_D$. By Proposition 4.5(a), we get that

$$\int_D f^\pm = \int_E (f^\pm \cdot \chi_D) = \int_E (f \cdot \chi_D)^\pm$$

By Proposition 4.5(f), we have that $\int_E (f^\pm \cdot \chi_D) \leq \int_E f^\pm < +\infty$. So $f^\pm$ is integrable over $D$ and $(f \cdot \chi_D)^\pm$ is integrable over $E$. Hence $f$ is integrable over $D$ and $f \cdot \chi_D$ is integrable over $E$. We easily get that

$$\int_D f = \int_D f^+ - \int_D f^- = \int_E (f \cdot \chi_D)^+ - \int_E (f \cdot \chi_D)^- = \int_E (f \cdot \chi_D)$$

(b) First, we prove that $\lambda f$ is integrable over $E$ and $\int_E (\lambda f) = \lambda \int_E f$ for all $\lambda \in \mathbb{R}$.

Assume first that $\lambda \geq 0$. Then $(\lambda f)^\pm = \lambda f^\pm$. By Proposition 4.5(b), we get that $\int_E (\lambda f)^\pm = \int_E (\lambda f^\pm) = \lambda \int_E f^\pm < +\infty$. So $(\lambda f)^\pm$ is integrable over $E$. Hence $\lambda f$ is integrable over $E$. We now get

$$\int_E (\lambda f) = \int_E (\lambda f)^+ - \int_E (\lambda f)^- = \lambda \int_E f^+ - \lambda \int_E f^- = \lambda \left( \int_E f^+ - \int_E f^- \right) = \lambda \int_E f$$

Assume next that $\lambda < 0$. Then $(\lambda f)^\pm = -\lambda f^\mp$. Similarly, we get that $\lambda f$ is integrable over $E$ and $\int_E (\lambda f) = \lambda \int_E f$.

Next, we prove that $f + g$ is integrable over $E$ and $\int_E (f + g) = \int_E f + \int_E g$. We have that

$$f + g = (f^+ + f^-) + (g^+ + g^-) = (f^+ + g^+) - (f^- + g^-)$$
By Proposition 4.9(a), \( \int_E (f^+ + g^+) = \int_E f^+ + \int_E g^+ < +\infty \). So \( f^+ + g^+ \) is integrable over \( E \). By Lemma 4.16, \( f + g \) is integrable over \( E \) and

\[
\int_E (f + g) = \int_E (f^+ + g^+) - \int_E (f^- + g^-)
\]

\[
= \left( \int_E f^+ + \int_E g^+ \right) - \left( \int_E f^- + \int_E g^- \right)
\]

\[
= \left( \int_E f^+ - \int_E f^- \right) + \left( \int_E g^+ - \int_E g^- \right)
\]

\[
= \int_E f + \int_E g
\]

(c) By (a), \( f \) is integrable over \( E_n \) for all \( n \geq 1 \). Using Proposition 4.9(c), we get

\[
\int_E f = \int_E f^+ - \int_E f^- = \left( \sum_{n=1}^{+\infty} \int_{E_n} f^+ \right) - \left( \sum_{n=1}^{+\infty} \int_{E_n} f^- \right) = \sum_{n=1}^{+\infty} \left( \int_{E_n} f^+ - \int_{E_n} f^- \right) = \sum_{n=1}^{+\infty} \int_{E_n} f
\]

(d) Using Proposition 4.5(d), we get

\[
\int_E f = \int_E f^+ - \int_E f^- = 0 - 0 = 0
\]

(e) Since \( f = g \) a.e. on \( E \), we get that \( f^\pm = g^\pm \) a.e. on \( E \). By Proposition 4.5(e), we find

\[
\int_E f = \int_E f^+ - \int_E f^- = \int_E g^+ - \int_E g^- = \int_E g
\]

(f) Note that \( g - f \geq 0 \) a.e. on \( E \). By Proposition 4.5(f), we have that \( \int_E (g - f) \geq \int_E 0 = 0 \). Using (b), we find

\[
\int_E g = \int_E ((g - f) + f) = \int_E (g - f) + \int_E f \geq \int_E f
\]

(g) Note that \( |f| = f^+ + f^- \). Hence by Proposition 4.9(a), we get

\[
\left| \int_E f \right| = \left| \int_E f^+ - \int_E f^- \right| \leq \left| \int_E f^+ \right| + \left| \int_E f^- \right| = \int_E f^+ + \int_E f^- = \int_E (f^+ + f^-) = \int_E |f| \quad \Box
\]

### 4.5 Lebesgue Convergence Theorem

**Theorem 4.18** Let \( f_n, f : E \to \mathbb{R} \) be measurable functions and \( g_n, g : E \to \mathbb{R} \) integrable over \( E \) such that \( \{f_n\}_{n \geq 1} \to f \) a.e. on \( E \), \( \{g_n\}_{n \geq 1} \to g \) a.e. on \( E \), \( \{\int_E g_n\}_{n \geq 1} \to \int_E g \) and \( |f_n| \leq g_n \) a.e. on \( E \) for all \( n \geq 1 \). Then \( f_n \) are integrable over \( E \) for all \( n \geq 1 \) and \( \{\int_E f_n\}_{n \geq 1} \to \int_E f \).

**Proof:** We may assume that the inequalities and the convergence is everywhere on \( E \) instead of a.e. on \( E \).

Pick \( n \geq 1 \). Since \( |f_n| \leq g_n \) on \( E \), we get that \( |f_n| \) is integrable over \( E \) by Lemma 4.11. So \( f_n \) is integrable over \( E \) by Lemma 4.12.
Since $|f_n| \leq g_n$ for all $n \geq 1$, we can apply the limit as $n$ goes to $+\infty$ on both sides. We get

$$|f| = \lim_{n \to +\infty} |f_n| \leq \lim_{n \to +\infty} g_n = g$$

Similarly as above, $f$ is integrable over $E$.

Since $|f_n| \leq g_n$, we get $-g_n \leq f_n \leq g_n$ for all $n \geq 1$. Hence $g_n - f_n \geq 0$ and $f_n + g_n \geq 0$ for all $n \geq 1$. Applying Fatou’s Lemma to $\{g_n - f_n\}_{n \geq 1}$ and using some properties of limit inferior and limit superior and using Proposition 4.17, we find

$$\int_E g - \int_E f = \int_E (g - f) = \int_E \lim (g_n - f_n) \leq \lim \int_E (g_n - f_n)$$

$$= \lim (\int_E g_n - \int_E f_n) \leq \overline{\lim} \int_E g_n + \lim (- \int_E f_n)$$

$$= \int_E g - \overline{\lim} \int_E f_n$$

So $\int_E g - \int_E f \leq \overline{\lim} \int_E g - \lim \int_E f_n$. Hence $\overline{\lim} \int_E f_n \leq \int_E f$.

Applying Fatou’s Lemma to $\{f_n + g_n\}_{n \geq 1}$ and using some properties of limit inferior and limit superior and using Proposition 4.17, we find

$$\int_E f + \int_E g = \int_E (f + g) = \int_E \lim (f_n + g_n)$$

$$\leq \lim \int_E (f_n + g_n) = \lim \int_E f_n + \lim \int_E g_n$$

$$= \lim \int_E f_n + \overline{\lim} \int_E g_n$$

So $\int_E f + \int_E g \leq \lim \int_E f_n + \overline{\lim} \int_E g$. Hence $\int_E f \leq \lim \int_E f_n$.

So we get that

$$\overline{\lim} \int_E f_n \leq \int_E f \leq \lim \int_E f_n$$

Hence $\overline{\lim} \int_E f_n = \overline{\lim} \int_E f_n = \int_E f$. So $\{\int_E f_n\}_{n \geq 1} \to \int_E f$. \hfill $\square$

**Corollary 4.19 (Lebesgue Convergence Theorem)** Let $f_n, f : E \to \mathbb{R}$ be measurable functions and $g : E \to \mathbb{R}$ integrable over $E$ such that $\{f_n\}_{n \geq 1} \to f$ a.e. on $E$ and $|f_n| \leq g$ a.e. on $E$ for all $n \geq 1$. Then $f_n, f$ are integrable over $E$ for all $n \geq 1$ and $\{\int_E f_n\}_{n \geq 1} \to \int_E f$.

**Proof:** This follows immediately from the previous theorem with ‘$g_n = g$’ for all $n \geq 1$. \hfill $\square$

**Remark:** Another way of viewing the last theorem and corollary : $\int_{E^{n \to +\infty}} f_n = \lim_{n \to +\infty} \int_E f_n$. \hfill $\triangleright$

### 4.6 Riemann and Lebesgue Integrals

We start by noting the following:

Let $\varphi : E \to \mathbb{R}$ be a simple function and $m(E) < +\infty$. Then $\varphi$ is integrable over $E$. Moreover, if $\varphi = \sum_{i=1}^n a_i \chi_{E_i}$ where $a_i \in \mathbb{R}$ and $E_i$ is a measurable subset of $E$, then $\int_E \varphi = \sum_{i=1}^n a_i m(E_i)$.  

62
Definition 4.20: Let $E$ be a measurable set with $m(E) < +\infty$ and $f : E \to \mathbb{R}$ a bounded function. We define the sets $L_f$ and $U_f$ as follows:

$L_f = \{ \varphi \mid \varphi$ is a simple function on $E$ and $\varphi \leq f$ on $E \}$

$U_f = \{ \psi \mid \psi$ is a simple function on $E$ and $f \leq \psi$ on $E \}$

Remark: Let $E$ be a measurable set with $m(E) < +\infty$ and $f : E \to \mathbb{R}$ a bounded function. Pick $\varphi \in L_f$ and $\psi \in U_f$. Then $\varphi \leq f \leq \psi$ on $E$. By Proposition 4.17(f), we have that $\int_E \varphi \leq \int_E \psi$. Since this is true for all $\varphi \in L_f$ and all $\psi \in U_f$, we get that $\sup_{\varphi \in L_f} \int_E \varphi \leq \inf_{\psi \in U_f} \int_E \psi$. △

Theorem 4.21: Let $E$ be a measurable set with $m(E) < +\infty$ and $f : E \to \mathbb{R}$ a bounded function. Then the following are equivalent:

(a) $f$ is measurable

(b) $f$ is Lebesgue integrable over $E$

(c) $\sup_{\varphi \in L_f} \int_E \varphi = \inf_{\psi \in U_f} \int_E \psi$

In this case, $\int_E f = \sup_{\varphi \in L_f} \int_E \varphi = \inf_{\psi \in U_f} \int_E \psi$.

Proof: (a) $\Rightarrow$ (b): Since $f$ is bounded, there exists $M > 0$ such that $|f| \leq M$ on $E$. So $-M\chi_E \leq f \leq M\chi_E$ on $E$. Since $-M\chi_E$ and $M\chi_E$ are integrable over $E$, we get that $f$ is integrable over $E$.

(b) $\Rightarrow$ (a): By definition, a function that is Lebesgue integrable over $E$ is measurable over $E$.

(a) $\Rightarrow$ (c): Since $f$ is bounded, there exists $M > 0$ such that $|f| \leq M$ on $E$.

Pick $n \geq 1$. Put

$E_{nk} = \left\{ x \in E \mid \frac{kM}{n} \leq f(x) < \frac{(k+1)M}{n} \right\}$ for $k = -n, -(n-1), \ldots, n-1, n$

Note that $E_{nk}$ is measurable for $-n \leq k \leq n$ and $E = \bigcup_{k=-n}^{n} E_{nk}$. Put

$\varphi_n = \sum_{k=-n}^{n} \frac{kM}{n} \chi_{E_{nk}}$ and $\psi_n = \sum_{k=-n}^{n} \frac{(k+1)M}{n} \chi_{E_{nk}}$

Then $\varphi_n \in L_f$ and $\psi_n \in U_f$. So

$$\sup_{\varphi \in L_f} \int_E \varphi \geq \int_E \varphi_n = \frac{M}{n} \sum_{k=-n}^{n} k m(E_{nk})$$ and $$\inf_{\psi \in U_f} \int_E \psi \leq \int_E \psi_n = \frac{M}{n} \sum_{k=-n}^{n} (k+1) m(E_{nk})$$

Hence we find

$$0 \leq \inf_{\psi \in U_f} \int_E \psi - \sup_{\varphi \in L_f} \int_E \varphi \leq \frac{M}{n} \sum_{k=-n}^{n} m(E_{kn}) = \frac{M m(E)}{n}$$

63
Since this is true for all $n \geq 1$, we can take the limit as $n$ goes to $+\infty$ on both sides. Using the fact that $m(E)$ is finite, we get
\[
0 \leq \inf_{\psi \in U_f} \int_E \psi - \sup_{\varphi \in L_f} \int_E \varphi \leq \lim_{n \to +\infty} \frac{Mm(E)}{n} = 0
\]
Hence $\sup_{\varphi \in L_f} \int_E \varphi = \inf_{\psi \in U_f} \int_E \psi$.

Pick $\varphi \in L_f$ and $\psi \in U_f$. Then $\varphi \leq f \leq \psi$ on $E$. Since $f$ is integrable over $E$, we find $\int_E \varphi \leq \int_E f \leq \int_E \psi$. Since this is true for all $\varphi \in L_f$ and all $\psi \in U_f$, we get
\[
\sup_{\varphi \in L_f} \int_E \varphi \leq \int_E f \leq \inf_{\psi \in U_f} \int_E \psi
\]
Hence $\sup_{\varphi \in L_f} \int_E \varphi = \inf_{\psi \in U_f} \int_E \psi = \int_E f$.

(c) $\Rightarrow$ (a) Note that $S := \sup_{\varphi \in L_f} \int_E \varphi = \inf_{\psi \in U_f} \int_E \psi$ is finite since $f$ is bounded and $m(E)$ is finite.

Using some properties of infimum and supremum, we get that for all $n \geq 1$, there exist $\varphi_n \in L_f$ and $\psi_n \in U_f$ such that
\[
S - \frac{1}{2n} < \int_E \varphi_n \quad \text{and} \quad S + \frac{1}{2n} > \int_E \psi_n
\]
Hence
\[
\int_E \psi_n - \int_E \varphi_n < \frac{1}{n} \quad \text{for all } n \geq 1 \quad (*)
\]

Put $\varphi = \sup_{n \geq 1} \varphi_n$ and $\psi = \sup_{n \geq 1} \psi_n$. By Theorem 3.20, we have that $\varphi$ and $\psi$ are measurable. Pick $x \in E$. Then $\varphi_n(x) \leq f(x) \leq \psi_n(x)$ for all $n \geq 1$. Hence
\[
\varphi(x) = \sup_{n \geq 1} \varphi_n(x) \leq f(x) \leq \inf_{n \geq 1} \psi_n(x) = \psi(x)
\]
So $\varphi \leq f \leq \psi$ on $E$.

Put $\Delta = \{ x \in E \mid \varphi(x) < \psi(x) \}$ and $\Delta_k = \{ x \in E \mid \varphi(x) < \psi(x) - \frac{1}{k} \}$ for all $n \geq 1$. Note that
\[
\Delta = \bigcup_{k=1}^{+\infty} \Delta_k.
\]

Pick $k \geq 1$. Since $\varphi_n \leq \varphi \leq \psi_n$ on $E$, we get that $0 \leq \psi - \varphi \leq \psi_n - \varphi_n$ for all $n \geq 1$. Using Tchebychev’s Inequality and (*), we find
\[
m(\Delta_k) \leq k \int_E (\psi - \varphi) \leq k \int_E (\psi_n - \varphi_n) < \frac{k}{n} \quad \text{for all } n \geq 1
\]
Taking the limit as $n$ goes to $+\infty$ on both sides, we get
\[
0 \leq m(\Delta_k) \leq \lim_{n \to +\infty} \frac{k}{n} = 0
\]
Since this is true for all $k \geq 1$, we have
\[
m(\Delta) = m(\bigcup_{k=1}^{+\infty} \Delta_k) \leq \sum_{k=1}^{+\infty} m(\Delta_k) = 0
\]
Hence \( \varphi = \psi \) a.e. on \( E \). Since \( \varphi \leq f \leq \psi \) on \( E \), we get that \( f = \varphi \) a.e. on \( E \). Since \( \varphi \) is measurable over \( E \), we finally get that \( f \) is measurable by Proposition 3.21.

**Definition 4.22**: A function \( \varphi : [a, b] \to \mathbb{R} \) is a step function if there exist
\[
a = x_0 < x_1 < \cdots < x_{n-1} < x_n = b
\]
and \( c_1, \ldots, c_n \in \mathbb{R} \) such that \( \varphi(x) = c_i \) if \( x_{i-1} < x < x_i \).

**Remark**: If \( \varphi \) is a step function over \([a, b] \) then \( \varphi \) is a simple function over \([a, b] \) and hence Lebesgue integrable over \([a, b] \). So \( \int_{[a, b]} \varphi \) makes sense (and is a real number).

**Definition 4.23**: Let \( f : [a, b] \to \mathbb{R} \) be a bounded function.

(a) We put
\[
RL_f = \{ \varphi \mid \varphi \text{ is a step function on } [a, b] \text{ and } \varphi \leq f \text{ on } [a, b] \} \\
RU_f = \{ \psi \mid \psi \text{ is a step function on } [a, b] \text{ and } f \leq \psi \text{ on } [a, b] \}
\]

(b) We say that \( f \) is Riemann integrable over \([a, b] \) if \( \sup_{\varphi \in RL_f} \int_{[a, b]} \varphi = \inf_{\psi \in RU_f} \int_{[a, b]} \psi \). In this case, we put
\[
\int_a^b f(x) \, dx = \sup_{\varphi \in RL_f} \int_{[a, b]} \varphi = \inf_{\psi \in RU_f} \int_{[a, b]} \psi
\]

**Remark**: Let \( \varphi : [a, b] \to \mathbb{R} \) be a step function. Then \( \varphi \) is Riemann and Lebesgue integrable over \([a, b] \) and \( \int_a^b \varphi(x) \, dx = \int_{[a, b]} \varphi \). This is just a special case of the next theorem.

**Theorem 4.24** Let \( f : [a, b] \to \mathbb{R} \) be a bounded function. If \( f \) is Riemann integrable over \([a, b] \), then \( f \) is Lebesgue integrable over \([a, b] \) and \( \int_{[a, b]} f(x) \, dx = \int_a^b f(x) \, dx \).

**Proof**: Note that \( RL_f \subseteq L_f \) and \( RU_f \subseteq U_f \). So we always have
\[
\sup_{\varphi \in RL_f} \int_{[a, b]} \varphi \leq \sup_{\varphi \in L_f} \int_{[a, b]} \varphi \leq \inf_{\psi \in U_f} \int_{[a, b]} \psi \leq \inf_{\psi \in RU_f} \int_{[a, b]} \psi
\]
Since \( f \) is Riemann integrable over \([a, b] \), we have equalities instead of inequalities:
\[
\int_a^b f(x) \, dx = \sup_{\varphi \in RL_f} \int_{[a, b]} \varphi = \sup_{\varphi \in L_f} \int_{[a, b]} \varphi = \inf_{\psi \in U_f} \int_{[a, b]} \psi = \inf_{\psi \in RU_f} \int_{[a, b]} \psi
\]
By Theorem 4.21(b)(c), we get that \( f \) is Lebesgue integrable over \([a, b] \) and
\[
\int_{[a, b]} f = \sup_{\varphi \in L_f} \int_{[a, b]} \varphi = \inf_{\psi \in U_f} \int_{[a, b]} \psi
\]
Hence \( \int_{[a, b]} f = \int_a^b f(x) \, dx \). \( \square \)
Definition 4.25: Let $f : [a, b] \rightarrow \mathbb{R}$ be a function.

(a) For all $x_0 \in [a, b]$ and all $\delta > 0$, we define

$$m_\delta(x_0) = \inf \{ f(x) \mid x \in (x_0 - \delta, x_0 + \delta) \cap [a, b] \}$$

$$M_\delta(x_0) = \sup \{ f(x) \mid x \in (x_0 - \delta, x_0 + \delta) \cap [a, b] \}$$

(b) We define the lower boundary of $f$ (notation: $m$) and the upper boundary of $f$ (notation: $M$) as follows:

$$m(x_0) = \lim_{\delta \to 0^+} m_\delta(x_0) \quad \text{and} \quad M(x_0) = \lim_{\delta \to 0^+} M_\delta(x_0) \quad \text{for all} \ x_0 \in [a, b]$$

Remark: Let $x_0 \in [a, b]$. If $0 < \delta_1 \leq \delta_2$ then

$$m_{\delta_2}(x_0) \leq m_{\delta_1}(x_0) \leq f(x_0) \leq M_{\delta_1}(x_0) \leq M_{\delta_2}(x_0)$$

So $\lim_{\delta \to 0^+} m_\delta(x_0)$ and $\lim_{\delta \to 0^+} M_\delta(x_0)$ exist and

$$m(x_0) = \lim_{\delta \to 0^+} m_\delta(x_0) = \sup_{\delta > 0} m_\delta(x_0) \leq f(x_0) \leq \inf_{\delta > 0} M_\delta(x_0) = \lim_{\delta \to 0^+} M_\delta(x_0) = M(x_0)$$

Proposition 4.26 Let $f : [a, b] \rightarrow \mathbb{R}$ be a function and $x_0 \in [a, b]$. Then $f$ is continuous at $x_0$ if and only if $m(x_0) = M(x_0)$.

Proof: Suppose first that $f$ is continuous at $x_0$. Pick $\varepsilon > 0$. Then we have

$$\exists \rho > 0 : \forall x \in [a, b] : |x - x_0| < \rho \Rightarrow |f(x) - f(x_0)| < \varepsilon$$

Pick $0 < \delta < \rho$. Pick $x \in (x_0 - \delta, x_0 + \delta) \cap [a, b]$. Then $x \in [a, b]$ and $|x - x_0| < \rho$. Hence $|f(x) - f(x_0)| < \varepsilon$ and so

$$f(x_0) - \varepsilon < f(x) < f(x_0) + \varepsilon$$

Since this is true for all $x \in (x_0 - \delta, x_0 + \delta) \cap [a, b]$, we get

$$f(x_0) \geq m_\delta(x_0) = \inf \{ f(x) \mid x \in (x_0 - \delta, x_0 + \delta) \cap [a, b] \} \geq f(x_0) - \varepsilon$$

$$f(x_0) \leq M_\delta(x_0) = \sup \{ f(x) \mid x \in (x_0 - \delta, x_0 + \delta) \cap [a, b] \} \leq f(x_0) + \varepsilon$$

Hence $|m_\delta(x_0) - f(x_0)| < \varepsilon$ and $|M_\delta(x_0) - f(x_0)| < \varepsilon$. So

$$m(x_0) = \lim_{\delta \to 0^+} m_\delta(x_0) = f(x_0) = \lim_{\delta \to 0^+} M_\delta(x_0) = M(x_0)$$

Suppose next that $m(x_0) = M(x_0)$. Then $\lim_{\delta \to 0^+} m_\delta(x_0) = m(x_0) = f(x_0) = M(x_0) = \lim_{\delta \to 0^+} M_\delta(x_0)$. Pick $\varepsilon > 0$. We have:

$$\exists \rho_1 > 0 : \forall \delta > 0 : 0 < \delta < \rho_1 \Rightarrow |m_\delta(x_0) - f(x_0)| < \varepsilon$$

$$\exists \rho_2 > 0 : \forall \delta > 0 : 0 < \delta < \rho_2 \Rightarrow |M_\delta(x_0) - f(x_0)| < \varepsilon$$
Let $0 < \delta < \min\{\rho_1, \rho_2\}$. Then

$$0 \leq f(x_0) - m_\delta(x_0) < \varepsilon \quad \text{and} \quad 0 \leq M_\delta(x_0) - f(x_0) < \varepsilon$$

Pick $x \in [a, b]$ with $|x - x_0| < \delta$. Then $x \in (x_0 - \delta, x_0 + \delta) \cap [a, b]$. So

$$f(x_0) - \varepsilon < m_\delta(x) = \inf\{f(y) \mid y \in (x_0 - \delta, x_0 + \delta) \cap [a, b]\} \leq f(x)$$

$$f(x) \leq \sup\{f(y) \mid y \in (x_0 - \delta, x_0 + \delta) \cap [a, b]\} = M_\delta(x_0) < f(x_0) + \varepsilon$$

Hence $f(x_0) - \varepsilon < f(x) < f(x_0) + \varepsilon$. So $|f(x) - f(x_0)| < \varepsilon$ and $f$ is continuous at $x = x_0$. □

**Proposition 4.27** Let $f : [a, b] \to \mathbb{R}$ be bounded, $m$ its lower boundary and $M$ its upper boundary. Then the following holds:

(a) $m$ and $M$ are measurable.

(b) Let $\varphi \in RL_f$ and $x_0 \in [a, b]$ such that $\varphi$ is continuous at $x_0$. Then $\varphi(x_0) \leq m(x_0)$.

(c) Let $\psi \in RU_f$ and $x_0 \in [a, b]$ such that $\psi$ is continuous at $x_0$. Then $M(x_0) \leq \psi(x_0)$.

(d) $\sup_{\varphi \in RL_f} \int_{[a, b]} \varphi = \int_{[a, b]} m$ and $\inf_{\psi \in RU_f} \int_{[a, b]} \psi = \int_{[a, b]} M$.

**Proof:** (a) Pick $\alpha \in \mathbb{R}$. Put $D = \{x \in [a, b] \mid m(x) > \alpha\}$. We will prove

$$\forall x_0 \in D : \exists \rho > 0 : (x_0 - \rho, x_0 + \rho) \cap [a, b] \subseteq D$$

Pick $x_0 \in D$. Let $\alpha < \beta < m(x_0)$. Then we have:

$$\exists \rho > 0 : \forall \delta > 0 : 0 < \delta \leq \rho \Rightarrow |m_\delta(x_0) - m(x_0)| < m(x_0) - \beta$$

So $m_\delta(x_0) > \beta$ for all $0 < \delta \leq \rho$. Pick $x \in [a, b]$ with $|x - x_0| < \rho$. Then $x \in (x_0 - \rho, x_0 + \rho) \cap [a, b]$ and so

$$f(x) \geq \inf\{f(y) \mid y \in (x_0 - \rho, x_0 + \rho) \cap [a, b]\} = m_\rho(x_0) > \beta$$

Pick $0 < \delta < \rho - |x - x_0|$. Pick $y \in (x - \delta, x + \delta) \cap [a, b]$. Then $|y - x_0| < \rho$ and so $f(y) > \beta$. Hence $m_\delta(x) = \inf\{f(y) \mid y \in (x - \delta, x + \delta) \cap [a, b]\} \geq \beta$. Since this is true for all $0 < \delta < \rho - |x - x_0|$, we have $m(x) = \lim_{\delta \to 0^+} m_\delta(x) \geq \beta > \alpha$. Hence $x \in D$.

So we proved the following:

$$\forall x_0 \in D : \exists \rho > 0 : (x_0 - \rho, x_0 + \rho) \cap [a, b] \subseteq D$$

Hence $D = [a, b] \cap O$ for some open set $O$. So $D$ is measurable. Hence $m$ is measurable. Similarly, $M$ is measurable.

(b) Since $\varphi \in RL_f$ is continuous at $x_0$, we have:

$$\exists \rho > 0 : \forall x \in (x_0 - \rho, x_0 + \rho) \cap [a, b] : \varphi(x_0) = \varphi(x) \leq f(x)$$

Pick $0 < \delta < \rho$. Then $\varphi(x_0) = \varphi(x) \leq f(x)$ for all $x \in (x - \delta, x + \delta) \cap [a, b]$ and so

$$m_\delta(x_0) = \inf\{f(x) \mid x \in (x_0 - \delta, x_0 + \delta) \cap [a, b]\} \geq \varphi(x_0)$$
Since this is true for all $0 < \delta < \rho$, we get

$$m(x_0) = \lim_{\delta \to 0^+} m_\delta(x_0) \geq \varphi(x_0)$$

(c) Similar as in (b).

(d) Let $\varphi \in RL_f$. Then $\varphi$ is continuous a.e. on $[a, b]$. By (b), $\varphi \leq m$ a.e. on $[a, b]$. Hence

$$\int_{[a,b]} \varphi \leq \int_{[a,b]} m$$

Since this is true for all $\varphi \in RL_f$, we get that

$$\sup_{\varphi \in RL_f} \int_{[a,b]} \varphi \leq \int_{[a,b]} m \quad (\star)$$

Pick $n \geq 1$. We partition the interval $[a, b]$ into $2^n$ intervals $I_{n1}, I_{n2}, \ldots, I_{n2^n}$ of length $\frac{b-a}{2^n}$:

$$I_{n1} = \left[ a, a + \frac{b-a}{2^n} \right] \quad \text{and} \quad I_{nk} = \left( a + (k-1)\frac{b-a}{2^n}, a + k\frac{b-a}{2^n} \right] \quad \text{for } k = 2, 3, \ldots, 2^n$$

We define $\varphi_n : [a, b] \to \mathbb{R}$ as follows:

$$\forall x \in [a, b] : \varphi_n(x) = \inf \{ f(y) \mid y \in I_{nk} \} \quad \text{if } x \in I_{nk}$$

Then $\varphi_n \in RL_f$.

Let $D$ be the set of all the partition points of all the partitions:

$$D = \left\{ a + k\frac{b-a}{2^n} \mid n \geq 1; 1 \leq k \leq 2^n - 1 \right\}$$

Note that $D$ is countable.

Pick $x_0 \in [a, b] \setminus D$. We will show that $\{\varphi_n(x_0)\}_{n \geq 1} \to m(x_0)$. Pick $\varepsilon > 0$. Since $m(x_0) = \lim_{\delta \to 0^+} m_\delta(x_0)$, there exists $\delta > 0$ such that $0 \leq m(x_0) - m_\delta(x_0) < \varepsilon$. For $n \geq 1$, let $1 \leq k_n \leq 2^n$ with $x_0 \in I_{nk_n}$. Let $N \in \mathbb{N}$ with $\frac{b-a}{2^N} < \delta$. Pick $n \geq N$. Then $I_{nk_n} \subseteq (x_0 - \delta, x_0 + \delta)$. Since $x_0 \notin D$, we have that $\varphi_n$ is continuous at $x_0$. So by (b) we get:

$$m_\delta(x_0) = \inf \{ f(x) \mid x \in (x_0 - \delta, x_0 + \delta) \cap [a, b] \} \leq \inf \{ f(x) \mid x \in I_{nk_n} \} = \varphi_n(x_0) \leq m(x_0)$$

Hence $|m(x_0) - \varphi_n(x_0)| \leq |m(x_0) - m_\delta(x_0)| < \varepsilon$. Thus $\{\varphi_n(x_0)\}_{n \geq 1} \to m(x_0)$.

Since this is true for all $x_0 \in [a, b] \setminus D$ and $D$ is countable, we have that $\{\varphi_n\}_{n \geq 1} \to m$ a.e. on $[a, b]$.

Since $f$ is bounded over $[a, b]$, there exists $\lambda > 0$ such that $-\lambda \leq f(x) \leq \lambda$ for all $x \in [a, b]$. Pick $n \geq 1$ and $x_0 \in [a, b] \setminus D$. Let $1 \leq k \leq 2^n$ with $x_0 \in I_{nk}$. Then by (b) we get

$$-\lambda \leq \inf \{ f(x) \mid x \in [a, b] \} \leq \inf \{ f(x) \mid x \in I_{nk} \} = \varphi_n(x_0) \leq m(x_0) \leq f(x) \leq \lambda$$

So $|\varphi_n| \leq \lambda \chi_{[a,b]}$ a.e. on $[a, b]$. Note that $\lambda \chi_{[a,b]}$ is integrable over $[a, b]$.

By the Lebesgue Convergence Theorem, we get

$$\left\{ \int_{[a,b]} \varphi_n \right\}_{n \geq 1} \to \int_{[a,b]} m$$
Since $\varphi_n \in RL_f$, we get

$$\sup_{\varphi \in RL_f} \int_{[a,b]} \varphi \geq \int_{[a,b]} \varphi_n \quad \text{for all } n \geq 1$$

Since this is true for all $n \geq 1$, we can take the limit as $n$ goes to $+\infty$ on both sides:

$$\sup_{\varphi \in RL_f} \int_{[a,b]} \varphi \geq \lim_{n \to +\infty} \int_{[a,b]} \varphi_n = \int_{[a,b]} m \quad (**)$$

It now follows from (*) and (**) that

$$\sup_{\varphi \in RL_f} \int_{[a,b]} \varphi = \int_{[a,b]} m$$

Similarly, we prove that

$$\inf_{\psi \in RU_f} \int_{[a,b]} \psi = \int_{[a,b]} M.$$ 

\[\square\]

**Theorem 4.28** Let $f : [a, b] \to \mathbb{R}$ be a function. Then $f$ is Riemann integrable over $[a, b]$ if and only if $f$ is bounded on $[a, b]$ and $m(\{x \in [a, b] \mid f \text{ is discontinuous at } x\}) = 0$.

**Proof:** Put $E = \{x \in [a, b] \mid f \text{ is discontinuous at } x\}$. If $f$ is bounded over $[a, b]$, we have:

\begin{align*}
f \text{ is Riemann integrable over } [a, b] & \iff \sup_{\varphi \in RL_f} \int_{[a,b]} \varphi = \inf_{\psi \in RU_f} \int_{[a,b]} \psi \quad \text{(definition)} \\
& \iff \int_{[a,b]} m = \int_{[a,b]} M \quad \text{(Proposition 4.27(d))} \\
& \iff \int_{[a,b]} (M - m) = 0 \\
& \iff M - m = 0 \text{ a.e. on } [a,b] \quad \text{(} M - m \geq 0 \text{ on } [a,b]) \\
& \iff m(E) = 0 \quad \text{(Proposition 4.26)} \quad \square
\end{align*}