## Contents

1 Measure Theory 3  
  1.1 $\sigma$-algebras ................................. 3  
  1.2 Measures ........................................ 4  
  1.3 Construction of non-trivial measures ....................... 6  
  1.4 Lebesgue measure .................................. 10  
  1.5 Measurable functions ............................... 14  

2 Integration 17  
  2.1 Integration of simple non-negative functions ................... 17  
  2.2 Integration of non-negative functions ....................... 17  
  2.3 Integration of real and complex valued functions .......... 19  
  2.4 $L^p$-spaces ....................................... 20  
  2.5 Relation to Riemann integration ........................... 22  
  2.6 Modes of convergence ................................ 23  
  2.7 Product measures ................................... 25  
  2.8 Polar coordinates ................................... 28  
  2.9 The transformation formula ............................. 31  

3 Signed and Complex Measures 35  
  3.1 Signed measures ................................... 35  
  3.2 The Radon-Nikodym theorem ............................. 37  
  3.3 Complex measures ................................... 39  

4 Functional Analysis 43  
  4.1 Review of basic functional analysis ....................... 43  
  4.2 Bounded linear functionals and dual spaces ............... 44  
  4.3 The dual of $L^p$ .................................... 45  
  4.4 The Hardy-Littlewood maximal functions ................... 47  

5 Differentiation 50  
  5.1 Lebesgue points .................................... 50  
  5.2 Complex Borel measures on $\mathbb{R}$ ..................... 53
1 MEASURE THEORY

1.1 \( \sigma \)-ALGEBRAS

Definition 1.1.1 (Algebra). Let \( X \) be a set and \( \mathcal{A} \) be a family of subsets of \( X \). We say that \( \mathcal{A} \) is an \textit{algebra on} \( X \) if

(i) \( \emptyset \in \mathcal{A} \);
(ii) if \( A \in \mathcal{A} \), then \( X \setminus A \in \mathcal{A} \);
(iii) if \( A, B \in \mathcal{A} \), then \( A \cup B \in \mathcal{A} \).

Definition 1.1.2 (\( \sigma \)-algebra). Let \( X \) be a set and \( \mathcal{A} \) be an algebra on \( X \). We say that \( \mathcal{A} \) is a \textit{\( \sigma \)-algebra on} \( X \) if for any countable collection of sets \( A_n \in \mathcal{A} \) (\( n \in \mathbb{N} \)), we have

\[
\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}.
\]

Proposition 1.1.3. Let \( \mathcal{A} \) be an algebra on \( X \). If \( A, B \in \mathcal{A} \), then

1. \( A \cap B \in \mathcal{A} \);
2. \( A \setminus B \in \mathcal{A} \).

Moreover, if \( \mathcal{A} \) is a \( \sigma \)-algebra and \( A_n \in \mathcal{A} \), then

\[
\bigcap_{n=1}^{\infty} A_n \in \mathcal{A}.
\]

Example 1.1.4. 1. For any set \( X \), \( \mathcal{A} = \mathcal{P}(X) \) is a \( \sigma \)-algebra.

2. For any set \( X \), \( \mathcal{A} = \{ A \subset X \mid A \text{ is countable or } X \setminus A \text{ is countable} \} \) is a \( \sigma \)-algebra.

3. If \( \{ \mathcal{A}_i \mid i \in I \} \) is a family of \( \sigma \)-algebras on \( X \), then so is \( \mathcal{A} = \bigcap_i \mathcal{A}_i \).

Proposition 1.1.5. Let \( \mathcal{F} \subset \mathcal{P}(X) \) be a family of subsets of \( X \). Then there is a unique \( \sigma \)-algebra \( \mathcal{A} \) such that \( \mathcal{F} \subset \mathcal{A} \subset \mathcal{A}' \), where \( \mathcal{A}' \) is any \( \sigma \)-algebra on \( X \) with \( \mathcal{F} \subset \mathcal{A}' \).

Proof. Let \( \mathfrak{F} \) be the family of \( \sigma \)-algebras \( \mathcal{A}' \) containing \( \mathcal{F} \). The appropriate \( \mathcal{A} \) is

\[
\mathcal{A} = \bigcap_{\mathcal{A}' \in \mathfrak{F}} \mathcal{A}'.
\]

\[\square\]

Definition 1.1.6 (\( \sigma \)-algebra generated by a family of sets). The \( \sigma \)-algebra \( \mathcal{A} \) in Proposition 1.1.5 is the \textit{\( \sigma \)-algebra generated by} \( \mathcal{F} \), denoted by \( \sigma(\mathcal{F}) \).

Definition 1.1.7 (Borel \( \sigma \)-algebra). Let \( (X, \mathcal{O}) \) be a topological space. The Borel \( \sigma \)-algebra on \( X \) is \( \mathcal{B}_X = \sigma(\mathcal{O}) \), the \( \sigma \)-algebra generated by the open sets in \( X \).
Example 1.1.8. 1. Every open set is Borel, and by complementation, every closed set is Borel.

2. In $\mathbb{R}$, the set $\mathbb{Q}$ is Borel, as it is a countable union of points, which are closed.

Example 1.1.9. A rectangle in $\mathbb{R}^n$ is a set of the form $[a_1, b_1] \times \cdots \times [a_n, b_n]$ with $a_i \leq b_i$ for all $i$. Let $\mathcal{R}$ be the family of all rectangles. We claim that $\sigma(\mathcal{R}) = \mathcal{B}_{\mathbb{R}^n}$.

If $R$ is a rectangle, then $R$ is closed, so $R \in \mathcal{B}_{\mathbb{R}^n}$. As $\mathcal{B}_{\mathbb{R}^n}$ is a $\sigma$-algebra, $\sigma(\mathcal{R}) \subset \mathcal{B}_{\mathbb{R}^n}$. For the other inclusion, let $U \subset \mathbb{R}^n$ be open, i.e. $U \in \mathcal{O}_{\mathbb{R}^n}$. It is a well-known exercise that $U$ can be covered by countably many rectangles, so $U \in \sigma(\mathcal{R})$. Hence $\mathcal{B}_{\mathbb{R}^n} = \sigma(\mathcal{O}) \subset \sigma(\mathcal{R})$, which completes the proof.

1.2 MEASURES

Definition 1.2.1 (Measure). A pair $(X, \mathcal{A})$ of a set $X$ and a $\sigma$-algebra $\mathcal{A}$ on $X$ is a measurable space. A (positive) measure $\mu$ on a measurable space $(X, \mathcal{A})$ is a function $\mu : \mathcal{A} \to [0, \infty]$ such that

(i) $\mu(\emptyset) = 0$;

(ii) if $A_n \in \mathcal{A}$ ($n \in \mathbb{N}$) are pairwise disjoint, then

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

The triple $(X, \mathcal{A}, \mu)$ is a measure space.

Example 1.2.2. Let $X$ be a set and $\mathcal{A} = \mathcal{P}(X)$.

1. For a given $a \in X$, the Dirac measure at $a$ is

$$\delta_a(M) = \begin{cases} 0 & a \notin M, \\ 1 & a \in M. \end{cases}$$

2. The counting measure is

$$\mu(M) = \begin{cases} |M| & M \text{ is finite}, \\ \infty & \text{otherwise}. \end{cases}$$

3. If $X$ is a topological space, then a Borel measure is a measure defined on $(X, \mathcal{B}_X)$.

Notation. 1. Write $A_n \nearrow A$ if $A_1 \subset A_2 \subset \cdots$ and

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

2. Write $A_n \searrow A$ if $A_1 \supset A_2 \supset \cdots$ and

$$A = \bigcap_{n=1}^{\infty} A_n.$$
Theorem 1.2.3. Let \((X, \mathcal{A}, \mu)\) be a measure space. Then

1. (monotonicity) if \(A, B \in \mathcal{A}\) and \(A \subset B\), then \(\mu(A) \leq \mu(B)\);

2. (countable subadditivity) if \(A_n \in \mathcal{A}\), then
   \[
   \mu\left( \bigcup_{n=1}^{\infty} A_n \right) \leq \sum_{n=1}^{\infty} \mu(A_n);
   \]

3. (continuity from below) if \(A_n \in \mathcal{A}\) and \(A_n \nearrow A\),
   \[
   \mu(A) = \lim_{n \to \infty} \mu(A_n);
   \]

4. (continuity from above) if \(A_n \in \mathcal{A}\), \(A_n \searrow A\), and \(\mu(A_1) < \infty\), then
   \[
   \mu(A) = \lim_{n \to \infty} \mu(A_n).
   \]

Proof. 1. Since \(A \subset B\), we have a decomposition \(B = A \cup (B \setminus A)\) into disjoint subsets. Then
   \[
   \mu(B) = \mu(A) + \mu(B \setminus A) \geq \mu(A).
   \]

2. Define
   \[
   B_n = A_n \setminus \bigcup_{k=1}^{n-1} A_k \subset A_n.
   \]
   By construction, the \(B_n\)'s are pairwise disjoint and \(\bigcup_n B_n = \bigcup_n A_n\). Then
   \[
   \mu\left( \bigcup_{n=1}^{\infty} A_n \right) = \mu\left( \bigcup_{n=1}^{\infty} B_n \right) = \sum_{n=1}^{\infty} \mu(B_n) = \sum_{n=1}^{\infty} \mu(A_n).
   \]

3. This time, define \(B_n = A_n \setminus A_{n-1}\) (suppose \(A_0 = \emptyset\)). We get
   \[
   \mu\left( \bigcup_{n=1}^{\infty} A_n \right) = \mu\left( \bigcup_{n=1}^{\infty} B_n \right) = \sum_{n=1}^{\infty} \mu(B_n) = \lim_{N \to \infty} \sum_{n=1}^{N} \mu(B_n) = \lim_{N \to \infty} \mu(A_N).
   \]

4. Let \(B_n = A_1 \setminus A_n\). Then \(B_n \nearrow A_1 \setminus A\) in \(\mathcal{A}\) and \(\mu(A_1) = \mu(A_n) + \mu(B_n)\), so applying continuity from below gives
   \[
   \mu(A_1) = \mu(A_1 \setminus A) + \mu(A) = \lim_{n \to \infty} \mu(B_n) + \mu(A).
   \]

   Since \(\mu(A_1)\) is finite, the result follows from writing
   \[
   \mu(A_1) = \lim_{n \to \infty} (\mu(A_1) - \mu(A_n)) + \mu(A) = \mu(A_1) + \mu(A) - \lim_{n \to \infty} \mu(A_n).
   \]

Note that continuity from above fails if we allow \(\mu(A_1) = \infty\). For example, on \(\mathbb{R}\) with the counting measure, let \(A_n = [n, \infty)\).
Definition 1.2.4 (Null set). A set \( N \in \mathcal{A} \) is a \((\mu-)null set\) if \( \mu(N) = 0 \).

Definition 1.2.5 (Complete measure space). A measure space is \textit{complete} if every subset of a null set is a null set.

Theorem 1.2.6. Let \( \mathcal{N} = \{ N \in \mathcal{A} \mid \mu(N) = 0 \} \) and
\[
\overline{\mathcal{A}} = \{ A \cup B \mid A \in \mathcal{A} \text{ and there exists } N \in \mathcal{N} \text{ such that } B \subset N \}. 
\]
Then \( \overline{\mathcal{A}} \) is a \( \sigma \)-algebra on \( X \) with \( \mathcal{A} \subset \overline{\mathcal{A}} \). Moreover, \( \mu \) can be uniquely extended to a measure \( \overline{\mu} \) on \( \overline{\mathcal{A}} \) which is complete.

Proof. First we show uniqueness, given existence. If \( \overline{\mu} \) exists and \( A \cup B \in \overline{\mathcal{A}} \) with \( B \subset N \) for some \( \mu \)-null set \( N \), we must have
\[
\mu(A) = \overline{\mu}(A) = \overline{\mu}(A \cup N) = \overline{\mu}(A) + \mu(N) = \mu(A),
\]
hence \( \overline{\mu}(A \cup B) = \mu(A) \) is fixed.

To show existence, define \( \overline{\mu} \) by \( \overline{\mu}(A \cup B) = \mu(A) \). We must show that this is a well-defined complete measure which extends \( \mu \).

To see that it is well-defined, let \( A \cup B = A' \cup B' \) with \( A, A' \in \mathcal{A} \) and \( B, B' \) subsets of null sets. The union of two null sets is a null set, so we can take \( B, B' \subset N \) for some null set \( N \). Then
\[
\overline{\mu}(A \cup B) = \mu(A) = \mu(A \cup N) = \mu(A' \cup N) = \mu(A') = \overline{\mu}(A' \cup B').
\]
The rest of the proof is omitted. \( \square \)

Definition 1.2.7 (\( \mu \)-almost everywhere). Let \((X, \mathcal{A}, \mu)\) be a measure space and \( P(x) \) be a statement about a point \( x \in X \). We say that \( P(x) \) is true \( \mu \)-almost everywhere (\( \mu \)-a.e.) if there exists a set \( N \in \mathcal{A} \) with \( \mu(N) = 0 \) and \( P(x) \) true for all \( x \in X \setminus N \).

Definition 1.2.8 (Finite / \( \sigma \)-finite). A measure \( \mu \) on \((X, \mathcal{A})\) is \textit{finite} if \( \mu(X) \) is finite. It is \textit{\( \sigma \)-finite} if \( X = \bigcup_n A_n \) with \( \mu(A_n) < \infty \) for each \( n \).

1.3 CONSTRUCTION OF NON-TRIVIAL MEASURES

Definition 1.3.1 (Outer measure). An outer measure on a set \( X \) is a function \( \mu^* : \mathcal{P}(X) \to [0, \infty] \) such that
\begin{enumerate}
\item \( \mu^*(\emptyset) = 0; \)
\item \( \mu^*(A) \leq \mu^*(B) \) whenever \( A \subset B; \)
\item \( \mu^*(\bigcup_n A_n) \leq \sum_n \mu^*(A_n). \)
\end{enumerate}

Lemma 1.3.2. Let \( \mathcal{A} \) be an algebra on \( X \).

1. Suppose that whenever \( A_n \in \mathcal{A} \) with \( A_n \nrightarrow A \), we have \( A \in \mathcal{A} \). Then \( \mathcal{A} \) is a \( \sigma \)-algebra.

2. Suppose that whenever \( A_n \in \mathcal{A} \) are pairwise disjoint, we have \( A = \bigcup_n A_n \in \mathcal{A} \). Then \( \mathcal{A} \) is a \( \sigma \)-algebra.
**Theorem 1.3.3** (Carathéodory). Let $\mu^*$ be an outer measure on $X$ and $\mathcal{A}$ be the family of sets $A \subseteq X$ such that

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

for all $T \subseteq X$. Then $\mathcal{A}$ is a $\sigma$-algebra on $X$ and $\mu = \mu^*|_\mathcal{A}$ is a complete measure on $(X, \mathcal{A})$.

**Proof.** To show that $\mathcal{A}$ is a $\sigma$-algebra, it is enough to show that $\mathcal{A}$ is an algebra which is closed under countable unions of pairwise disjoint sets. It is clear that $\emptyset \in \mathcal{A}$. Since the condition for $A \in \mathcal{A}$ is symmetric in $A$ and $A^c$, it follows that $A^c \in \mathcal{A}$. To get finite unions, let $A, B \in \mathcal{A}$. Since

$$A \cup B = (A \cap B^c) \cup (A \cap B) \cup (B \cap A^c),$$

we have

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

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

$$\geq \mu^*(T \cap (A \cup B)) + \mu^*(T \cap A^c \cap B) + \mu^*(T \cap (A \cup B)^c))
$$

$$\geq \mu^*(T),$$

hence equality holds everywhere and $A \cup B \in \mathcal{A}$.

Let $A_n \in \mathcal{A}$ be pairwise disjoint and $B_n = \bigcup_{i=1}^n A_i$. Then for any test set $T$,

$$\mu^*(T \cap B_n) = \sum_{i=1}^n \mu^*(T \cap A_i),$$

so

$$\mu^*(T) = \mu^*(T \cap B_n) + \mu^*(T \cap B_n^c) = \sum_{i=1}^n \mu^*(T \cap A_i) + \mu^*(T \cap B_n^c).$$

Letting $n \to \infty$ and writing $B = \bigcup_n B_n = \bigcup_n A_n$, we have by countable subadditivity

$$\mu^*(T) \geq \mu^*(T \cap B) + \mu^*(T \cap B^c) \geq \mu^*(T),$$

so we have the required equality for $B \in \mathcal{A}$.

Using the test condition for $T = B$ as above, we conclude

$$\mu^*(B) = \sum_{n=1}^\infty \mu^*(A_n) + \mu^*(\emptyset).$$

This means that $\mu^*$ is countably additive on $\mathcal{A}$, so $\mu$ is a measure.

To show that $\mu^*$ is complete, we must show that if $\mu^*(A) = 0$, then $A \in \mathcal{A}$. For any test set $T$,

$$\mu^*(T) \leq \mu^*(T \cap A) + \mu^*(T \cap A^c) = \mu^*(T \cap A^c) \leq \mu^*(T),$$

so equality holds everywhere and we have the required equality for $A \in \mathcal{A}$. \qed
Definition 1.3.4 ($\mu^*$-measurable). If $\mu^*$ is an outer measure on $X$, then a subset $A \subset X$ is $\mu^*$-measurable if it lies in the family $\mathcal{A}$ from the theorem.

Definition 1.3.5 (Premeasure). Let $\mathcal{A}$ be an algebra on $X$. A premeasure $\nu: \mathcal{A} \to [0, \infty]$ is a function such that

(i) $\nu(\emptyset) = 0$,

(ii) if $A_n \in \mathcal{A}$ are pairwise disjoint with $\bigcup_n A_n \in \mathcal{A}$, then

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

Lemma 1.3.6. Let $\mathcal{A}$ be an algebra on $X$ and $\nu: \mathcal{A} \to [0, \infty]$ be a premeasure. Define

$$
\mu^*(M) = \inf \left\{ \sum_{n=1}^{\infty} \nu(A_n) \mid A_n \in \mathcal{A} \text{ and } M \subset \bigcup_{n=1}^{\infty} A_n \right\}
$$

for $M \subset X$. Then $\mu^*$ is an outer measure extending $\nu$. Moreover, each $M \in \mathcal{A}$ is $\mu^*$-measurable.

Proof. First we show that $\mu^*$ is an outer measure. That $\mu^*(\emptyset) = 0$ and that $\mu^*$ is monotonic are clear, so it remains to show countable subadditivity. Let $M_n \subset X$. If some $\mu^*(M_n) = \infty$, then the result is clear, so suppose $\mu^*(M_n) < \infty$. Let $\epsilon > 0$. For each $n$, there exists a countable cover of $M_n$ by sets $A_{n,k} \in \mathcal{A}$ such that

$$
\mu^*(M_n) \leq \sum_{n=1}^{\infty} \nu(A_{n,k}) \leq \mu^*(M_n) + \frac{\epsilon}{2^n}.
$$

Then

$$
M = \bigcup_{n=1}^{\infty} M_n \subset \bigcup_{n,k \in \mathbb{N}} A_{n,k}
$$

is a countable cover of $M$ by elements of $\mathcal{A}$, so

$$
\mu^*(M) \leq \sum_{n,k=1}^{\infty} \nu(A_{n,k}) \leq \sum_{n=1}^{\infty} \mu^*(M_n) + \epsilon.
$$

Since $\epsilon > 0$ is arbitrary, we have countable subadditivity.

To show that $\mu^*$ and $\nu$ agree on $\mathcal{A}$, suppose $M \in \mathcal{A}$. Certainly we know that $\mu^*(M) \leq \nu(M)$, as $M$ itself covers $M$. In the other direction, let $A_n \in \mathcal{A}$ cover $M$ and let

$$
B_n = M \cap \left( A_n \setminus \bigcup_{i=1}^{n-1} A_i \right).
$$

Then the $B_n$ are pairwise disjoint elements of $\mathcal{A}$ with $\bigcup_n B_n = M$, so

$$
\nu(M) = \sum_{n=1}^{\infty} \nu(B_n) \leq \sum_{n=1}^{\infty} \nu(A_n).
$$
Taking the infimum over all countable covers $A_n$ of $M$, we have $\nu(M) \leq \mu^*(M)$. Finally, we show that each $M \in \mathcal{A}$ is $\mu^*$-measurable. Let $T \subset X$ and $\epsilon > 0$ be arbitrary. If $\mu^*(T) = \infty$, then

$$
\mu^*(T) = \infty \geq \mu^*(T \cap M) + \mu^*(T \cap M^c) \geq \mu^*(T),
$$

so we have equality everywhere. If $\mu^*(T) < \infty$, then pick a cover $A_n \in \mathcal{A}$ of $T$ such that

$$
\mu^*(T) \leq \sum_{n=1}^{\infty} \nu(A_n) \leq \mu^*(T) + \epsilon.
$$

Since $\nu(M \cap A_n) + \nu(M^c \cap A_n) = \nu(A_n)$, we have

$$
\mu^*(T) + \epsilon \geq \sum_{n=1}^{\infty} \nu(A_n)
= \sum_{n=1}^{\infty} \nu(M \cap A_n) + \sum_{n=1}^{\infty} \nu(M^c \cap A_n)
= \mu^*(M \cap T) + \mu^*(M^c \cap T) \geq \mu^*(T).
$$

Since $\epsilon$ is arbitrary, we have the required (Carathéodory) criterion. \qed

**Theorem 1.3.7** (Carathéodory extension theorem). Let $\mathcal{A}$ be an algebra on $X$, $\nu : A \to [0, \infty]$ be a premeasure, and $\mathcal{M} = \sigma(\mathcal{A})$. Then there exists a measure $\mu : \mathcal{M} \to [0, \infty]$ extending $\nu$. Moreover, if $\nu$ is $\sigma$-finite, then the extension $\mu$ on $\mathcal{M}$ is unique.

**Proof.** Using Theorem 1.3.3 and Lemma 1.3.6, $\nu$ defines an outer measure $\mu^*$ on $X$ extending $\nu$. If $\mathcal{B}$ is the family of $\mu^*$-measurable sets, then $\mathcal{B}$ is a $\sigma$-algebra containing $\mathcal{A}$, so $\mathcal{M} = \sigma(\mathcal{A}) \subset \mathcal{B}$. The function $\mu = \mu^*|_{\mathcal{B}}$ is then a measure extending $\nu$ through $\mu^*$.

For uniqueness, let $\tilde{\mu}$ be another measure on $\mathcal{M}$ that extends $\nu$. Let $A_n \in \mathcal{A}$ cover $M$. Then

$$
\tilde{\mu}(M) \leq \sum_{n=1}^{\infty} \tilde{\mu}(A_n) = \sum_{n=1}^{\infty} \nu(A_n).
$$

Taking the infimum over all countable covers of $M$ by sets in $\mathcal{A}$, we get $\tilde{\mu}(M) \leq \mu(M)$. To get the other direction, we first show that $\mu(M) \leq \tilde{\mu}(M)$ for all $M \in \mathcal{M}$ with $\mu(M) < \infty$. Let $\epsilon > 0$. Then we can find sets $A_n \in \mathcal{A}$ covering $M$ such that

$$
\sum_{n=1}^{\infty} \nu(A_n) \leq \mu^*(M) + \epsilon = \mu(M) + \epsilon.
$$

Define $B_n = A_n \setminus \bigcup_{i=1}^{n-1} A_i$. These sets are pairwise disjoint in $\mathcal{A}$ and $A = \bigcup_{n} A_n = \bigcup_{n} B_n$, so

$$
\mu(A) = \sum_{n=1}^{\infty} \mu(B_n) = \sum_{n=1}^{\infty} \nu(B_n) \leq \sum_{n=1}^{\infty} \nu(A_n) \leq \mu(M) + \epsilon.
$$

Since $\mu(M) < \infty$, it follows that

$$
\tilde{\mu}(A \setminus M) \leq \mu(A \setminus M) = \mu(A) - \mu(M) < \epsilon.
$$
Hence
\[ \mu(M) \leq \mu(A) = \lim_{n \to \infty} \mu \left( \bigcup_{i=1}^{n} A_i \right) = \lim_{n \to \infty} \tilde{\mu} \left( \bigcup_{i=1}^{n} A_i \right) = \tilde{\mu}(A) < \tilde{\mu}(M) + \epsilon. \]

As \(\epsilon\) is arbitrary, we have \(\mu(M) \leq \tilde{\mu}(M)\) when \(\mu(M) < \infty\).

Having shown that \(\mu(M) = \tilde{\mu}(M)\) whenever \(M\) has finite \(\mu\)-measure, we show that this is true in general. Since \(\nu\) is \(\sigma\)-finite, we can find pairwise disjoint sets \(F_n \in \mathcal{A}\) with \(\nu(F_n) = \mu(F_n) < \infty\) and \(X = \bigcup_n F_n\). For each \(M \in \mathcal{M}\), we have
\[
\mu(M) = \sum_{n=1}^{\infty} \mu(M \cap F_n) = \sum_{n=1}^{\infty} \tilde{\mu}(M \cap F_n) = \tilde{\mu}(M).
\]

1.4 LEBESGUE MEASURE

**Definition 1.4.1 (h-interval).** An h-interval (half-open interval) is a set \(I \subset \mathbb{R}\) of the form
\[
I = \begin{cases} 
(a, b] & a, b \in \mathbb{R}, a < b \\
(-\infty, b] & b \in \mathbb{R} \\
(a, \infty) & a \in \mathbb{R} \\
\mathbb{R} & \\
\emptyset.
\end{cases}
\]
The length of an h-interval \(I\) is \(l(I) = b - a\) for h-intervals of the first form, \(\infty\) for the second, third, and fourth forms, and 0 for the empty set.

**Lemma 1.4.2.** Let \(C\) be a family of subsets of \(X\) such that

(i) \(\emptyset \in C\)

(ii) \(A \cap B \in C\) whenever \(A, B \in C\)

(iii) if \(A \cap C\), then \(A^c\) is a finite union of elements of \(C\) which is pairwise disjoint.

If \(\mathcal{A}\) is the family of all subsets of \(X\) that can be represented as a finite union of pairwise disjoint sets in \(C\), then \(\mathcal{A}\) is an algebra on \(X\).

**Corollary 1.4.3.** Let \(\mathcal{A}\) be the family of all subsets of \(\mathbb{R}\) that can be written as a finite union of pairwise disjoint h-intervals. Then \(\mathcal{A}\) is an algebra on \(\mathbb{R}\) (the algebra generated by h-intervals).

**Definition 1.4.4 (h-rectangle).** An h-rectangle \(R \subset \mathbb{R}^n\) is a set of the form
\[ R = I_1 \times \cdots \times I_n \]
where each \(I_k\) is an h-interval.

**Proposition 1.4.5.** The family \(\mathcal{A}\) of all subsets of \(\mathbb{R}^n\) that can be written as a finite union of pairwise disjoint h-rectangles is an algebra on \(\mathbb{R}^n\).
Proof. We use Lemma 1.4.2 for the family $\mathcal{R}$ of h-rectangles in $\mathbb{R}^n$.

(i) We have $\emptyset = \emptyset \times \cdots \times \emptyset \in \mathcal{R}$.

(ii) If $R = I_1 \times \cdots \times I_n$ and $S = J_1 \times \cdots \times J_n$ are in $\mathcal{R}$, then

$$R \cap S = (I_1 \cap J_1) \times \cdots \times (I_n \cap J_n) \in \mathcal{R},$$

as intersections of h-intervals are h-intervals.

(iii) If $R = I_1 \times I_n \in \mathcal{R}$, then

$$I_c^i = R \setminus I_i$$

is a disjoint union of at most two h-intervals. This implies

$$R^c = \mathbb{R}^n \setminus R = \bigcup_{1 \leq i_1 < \cdots < i_k \leq n} \prod_{j=1}^n \left\{ \begin{array}{ll} I_j^i & j = i_i \\ I_j & \text{ otherwise} \end{array} \right\}$$

is a large but finite disjoint union of h-rectangles.

\[\square\]

Definition 1.4.6 (Content). If $R = I_1 \times \cdots \times I_n$ is an h-rectangle in $\mathbb{R}^n$, then the content of $R$ is

$$|R| = l(I_1) \cdot \cdots \cdot l(I_n) \in [0, \infty],$$

with the convention that $0 \cdot \infty = 0$.

Lemma 1.4.7 (Basic lemma). Let $\{R_i\}_{i \in I}$ and $\{S_j\}_{j \in J}$ be two countable families of h-rectangles in $\mathbb{R}^n$. Suppose $R_k \cap R_l = \emptyset$ for distinct $k, l \in I$ and $\bigcup_i R_i \subset \bigcup_j S_j$. Then

$$\sum_{i \in I} |R_i| \leq \sum_{j \in J} |S_j|.$$ 

Proof. See Homework 3 Problem 1. \[\square\]

Definition 1.4.8 (Premeasure on h-rectangles). Let $A$ be the algebra generated by the h-rectangles in $\mathbb{R}^n$. To define a premeasure $\nu : A \to [0, \infty]$, let $M \in A$ have the form $M = R_1 \sqcup \cdots \sqcup R_k$ for h-rectangles $R_i$. Then set $\nu(M) = \sum_{i=1}^k |R_i|$.

Proposition 1.4.9. $\nu$ is a well-defined premeasure on $A$.

Proof. To see that $\nu$ is well-defined, suppose $M = R_1 \sqcup \cdots \sqcup R_k = S_1 \sqcup \cdots \sqcup S_l$. The basic lemma then applies in both directions, so $\sum_{i=1}^k |R_i| = \sum_{j=1}^l |S_j|$.

To see that $\nu$ is a premeasure, first we have $\nu(\emptyset) = 0$. Then, given pairwise disjoint $A_n \in A$, with $A_n = R_{n,1} \sqcup \cdots \sqcup R_{n,k_n}$. If $A = \bigcup_n A_n = \bigcup_{n,k} R_{n,k}$, then applying the basic lemma in both directions again, we have

$$\nu(A) = \sum_{n,k} |R_{n,k}| = \sum_{n=1}^\infty \sum_{k=1}^{k_n} |R_{n,k}| = \sum_{n=1}^\infty \nu(A_n).$$

\[\square\]
Definition 1.4.10 (Lebesgue outer measure). The Lebesgue outer measure is the outer measure induced by \( \nu \) through Lemma 1.3.6, i.e.

\[
\lambda^*_\nu(M) = \inf \left\{ \sum_{i=1}^{\infty} \nu(A_i) \mid A_i \in \mathcal{A} \text{ and } M \subset \bigcup_i A_i \right\}.
\]

Definition 1.4.11 (Lebesgue measurable subset). A subset of \( \mathbb{R}^n \) which is \( \lambda^* \)-measurable is (Lebesgue) measurable.

Definition 1.4.12 (Lebesgue measure). The Lebesgue measure is the measure \( \lambda_n \) induced by \( \lambda^*_n \) through Theorem 1.3.3.

Proposition 1.4.13. 1. Every h-rectangle \( R \) is measurable with \( \lambda(R) = |R| \).

2. Every Borel set is measurable.

3. Lebesgue measure is complete.

Definition 1.4.14 (Rectangle). A rectangle is a compact set of the form \( R = [a_1, b_1] \times \cdots \times [a_n, b_n] \).

Proposition 1.4.15. 1. If \( R = [a_1, b_1] \times \cdots \times [a_n, b_n] \) is a rectangle, then

\[
\lambda(R) = (b_1 - a_1) \cdots (b_n - a_n).
\]

2. If \( M \subset \mathbb{R}^n \), then

\[
\lambda^*(M) = \inf \left\{ \sum_{i=1}^{\infty} \lambda(R_i) \mid R_i \text{ rectangles with } M \subset \bigcup_i R_i \right\}.
\]

Theorem 1.4.16. Let \( M \subset \mathbb{R}^n \). The following are equivalent:

1. \( M \) is a null set, i.e. \( M \) is measurable and \( \lambda(M) = 0 \);

2. for each \( \epsilon > 0 \), there is a countable cover of \( M \) by rectangles \( R_k \) such that \( \sum_k \lambda(R_k) < \epsilon \);

3. there exists a Borel set \( B \subset \mathbb{R}^n \) with \( M \subset B \) and \( \lambda(B) = 0 \).

Proof. See Homework 3 Problem 2. \( \square \)

Corollary 1.4.17. A set \( M \subset \mathbb{R}^n \) is measurable if and only if there exists a Borel set \( B \) and a null set \( N \) such that \( M = B \cup N \).

Proof. \( (\Rightarrow) \) Omitted.

\( (\Leftarrow) \) If \( \lambda(M) = \lambda^*(M) < \infty \), then for each \( k \in \mathbb{N} \), there exist rectangles \( R_{k,i} \) covering \( M \) such that

\[
\lambda(M) \leq \lambda \left( \bigcup_{i=1}^{\infty} R_{k,i} \right) \leq \sum_{i=1}^{\infty} \lambda(R_{k,i}) \leq \lambda(M) + \frac{1}{k}.
\]

Let

\[
A_k = \bigcup_i R_{k,i} \quad \text{and} \quad A = \bigcap_k A_k.
\]

12
This is Borel with $M \subset A$, so by construction, $\lambda(A) = \lambda(M)$. Thus $\lambda(A \setminus M) = 0$, so there exists a Borel set $C$ with $A \setminus M \subset C$ and $\lambda(C) = 0$. We then have

$$M = (A \setminus C) \cup (M \cap C)$$

with $B = A \setminus C$ Borel and $N = M \cap C$ null.

For the general case, let $R_k = [-k, k]^n$. If $M_k = M \cap R_k$, then each $M_k$ has finite measure, so we can find Borel sets $B_k$ and null sets $N_k$ such that $M_k = B_k \cup N_k$. Then if $B = \bigcup_k B_k$ and $N = \bigcup_k N_k$, we have that $B$ is Borel, $N$ is null, and $M = \bigcup_k M_k = B \cup N$.

\[ \square \]

**Theorem 1.4.18.** The Lebesgue measure on $\mathbb{R}^n$ is the unique measure $\lambda$ such that

(i) $\lambda$ is defined on the $\sigma$-algebra of all Lebesgue measurable subsets;

(ii) $\lambda$ is translation invariant, i.e. $\lambda(M) = \lambda(t + M)$ for all $t \in \mathbb{R}^n$ and $M \subset \mathbb{R}^n$ measurable;

(iii) $\lambda([0, 1]^n) = 1$.

**Proof.** See Homework 4 Problem 1. \[ \square \]

**Definition 1.4.19** (Regular measure). Let $X$ be a topological space, $\mathcal{A}$ be a $\sigma$-algebra containing $\mathcal{B}_X$, and $\mu : \mathcal{A} \to [0, \infty]$ be a measure. We say that

1. $\mu$ is *inner regular* if for all $A \in \mathcal{A}$,
   $$\mu(A) = \sup\{\mu(K) \mid K \subset \text{compact in } X\};$$

2. $\mu$ is *outer regular* if for all $A \in \mathcal{A}$,
   $$\mu(A) = \inf\{\mu(U) \mid U \supset \text{open in } X\};$$

3. $\mu$ is *regular* if it is inner and outer regular.

**Proposition 1.4.20.** Lebesgue measure is regular.

**Proof.** See Homework 2 Problem 3. \[ \square \]

**Proposition 1.4.21.** If $T : \mathbb{R}^n \to \mathbb{R}^n$ is a linear map and $M \subset \mathbb{R}^n$ is measurable, then $T(M)$ is measurable and $\lambda(T(M)) = |\det T| \cdot \lambda(M)$.

In particular, orthogonal transformations preserve Lebesgue measure.

**Proof.** See Homework 4 Problems 3 and 4. \[ \square \]
Example 1.4.22 (Non-measurable subset of \( \mathbb{R} \)). Define an equivalence relation on \([0, 1]\) by

\[
x \sim y \iff x - y \in \mathbb{Q}.
\]

Invoking the axiom of choice, we can pick an element from each equivalence class to get a set \( E \subset [0, 1] \). Suppose that \( E \) were measurable. Enumerate the rationals \( q_i \in [-1, 1] \) and consider the translates \( q_i + E \subset [-1, 2] \). These are disjoint by construction, so

\[
\lambda \left( \bigcup_{i=1}^{\infty} (q_i + E) \right) = \sum_{i=1}^{\infty} \lambda(q_i + E) = \infty \cdot \lambda(E).
\]

Since all of these translates lie in \([-1, 2] \), \( \lambda \left( \bigcup_{i=1}^{\infty} (q_i + E) \right) \leq 3 \). However, given any \( x \in [0, 1] \), there must be some \( y \in [0, 1] \) with \( x \sim y \), so there is a rational \( q \in [-1, 1] \) such that \( x = y + q \). This means that \( \lambda \left( \bigcup_{i=1}^{\infty} (q_i + E) \right) \geq 1 \). There is no possible value of \( \lambda(E) \) which makes \( \infty \cdot \lambda(E) \) lie in this range, so \( E \) cannot be measurable.

This can be slightly modified to prove the following result.

**Proposition 1.4.23.** Every measurable set \( M \subset \mathbb{R} \) with \( \lambda(M) > 0 \) has a non-measurable subset.

Example 1.4.24 (Lebesgue measurable set that is not Borel). Let \( c : [0, 1] \to [0, 1] \) be the Cantor function and \( f : [0, 1] \to [0, 2] \) be given by \( f(x) = c(x) + x \). Then if \( C \subset [0, 1] \) is the Cantor set, we have \( \lambda(f([0, 1] \setminus C)) = \lambda([0, 1] \setminus C) = 1 \), from which it follows that \( \mu(f(C)) = 1 \). Hence there is a non-measurable set \( N \subset f(C) \), which in particular is not Borel. Since \( f \) is strictly increasing, it maps Borel sets to Borel sets (see Homework 1 Problem 4). Thus \( f^{-1}(N) \) is not Borel, but it is a subset of \( C \), which has measure zero, so \( N \) is Lebesgue measurable (with measure zero).

### 1.5 Measurable Functions

**Definition 1.5.1** (Measurable function). Let \((X, \mathcal{A})\) and \((Y, \mathcal{B})\) be measurable spaces. A map \( f : X \to Y \) is called \((\mathcal{A}, \mathcal{B})\)-measurable if \( f^{-1}(B) \in \mathcal{A} \) for all \( B \in \mathcal{B} \).

If \( \mathcal{A} \) and \( \mathcal{B} \) are understood, we will simply say that \( f \) is measurable.

**Proposition 1.5.2.** If \( \mathcal{B} = \sigma(\mathcal{S}) \), then \( f \) is measurable if and only if \( f^{-1}(S) \in \mathcal{A} \) for all \( S \in \mathcal{S} \).

**Corollary 1.5.3.** Let \( X \) and \( Y \) be topological spaces with their Borel \( \sigma \)-algebras. If \( f \) is continuous, then \( f \) is (Borel) measurable.

**Proposition 1.5.4.** If \( f : X \to Y \) and \( g : Y \to Z \) are measurable, then \( g \circ f \) is measurable.

In the case that \( X = \mathbb{R}^n \), there are two natural \( \sigma \)-algebras to put on \( \mathbb{R}^n \), namely the Borel \( \sigma \)-algebra \( \mathcal{B} \) and the \( \sigma \)-algebra \( \mathcal{L} \) of Lebesgue measurable sets.

**Definition 1.5.5** (Measurable functions from \( \mathbb{R}^n \)). A function \( f : \mathbb{R}^n \to Y \) is Lebesgue measurable, or simply measurable, if \( f : (\mathbb{R}^n, \mathcal{L}) \to (Y, \mathcal{B}_Y) \) is \((\mathcal{L}, \mathcal{B}_Y)\)-measurable. We say \( f \) is Borel measurable if \( f : (\mathbb{R}^n, \mathcal{B}) \to (Y, \mathcal{B}_Y) \) is \((\mathcal{B}, \mathcal{B}_Y)\)-measurable.

**Remark 1.5.6.** 1. If \( f \) is Borel measurable, then \( f \) is (Lebesgue) measurable.
2. According to this convention, a function \( f : \mathbb{R} \to \mathbb{R} \) is measurable if it is \((\mathcal{L}, \mathcal{B})\)-measurable, so the domain and codomain have different \( \sigma \)-algebras. In particular, if \( f, g : \mathbb{R} \to \mathbb{R} \) are measurable, then \( g \circ f \) need not be measurable. On the other hand, if \( f \) is measurable and \( g \) is Borel measurable, then \( g \circ f \) is measurable.

**Lemma 1.5.7.** Let \((X, \mathcal{A})\) be a measurable space. Then

1. \( f : X \to \mathbb{R} \) is measurable if and only if \( f^{-1}(a, \infty) \in \mathcal{A} \) for each \( a \in \mathbb{R} \).
2. if \( u, v : X \to \mathbb{R} \) are measurable, \( Z \) is a topological space, and \( F : \mathbb{R}^2 \to Z \) is continuous, then \( F \circ (u, v) : X \to Z \) is measurable.

**Corollary 1.5.8.**

1. If \( f, g : X \to \mathbb{C} \) are measurable, then so are \( f \), \( f + g \), and \( f \cdot g \).
2. If \( f : X \to \mathbb{C} \) is measurable and \( z \) is a constant, then \( u = \text{Re} f \), \( v = \text{Im} f \), \( |f| \), \( 1/f \), and \( zf \) are measurable.

**Proposition 1.5.9.** Let \( f_n : X \to \mathbb{R} \) be measurable functions. Then \( \sup f_n \), \( \inf f_n \), \( \limsup f_n \), and \( \liminf f_n \) are measurable.

**Corollary 1.5.10.**

1. If \( f \) and \( g \) are measurable, then \( f \lor g = \max(f, g) \) and \( f \land g = \min(f, g) \) are measurable.
2. If \( f_n : X \to \mathbb{R} \) are measurable and \( f_n \to f \) pointwise, then \( f \) is measurable.

**Definition 1.5.11** (Positive and negative parts). Given \( f : X \to \mathbb{R} \), the positive part \( f^+ \) and negative part \( f^- \) are

\[
 f^+ = \max(f, 0), \quad f^- = -\min(f, 0),
\]

so that \( f = f^+ - f^- \).

**Proposition 1.5.12.** If \( f : X \to \mathbb{R} \) are measurable, then \( f^+ \) and \( f^- \) are measurable.

**Definition 1.5.13** (Characteristic function). If \( A \subset X \), then \( \chi_A : X \to \mathbb{R} \) defined by \( \chi(x \in A) = 1 \) and \( \chi(x \notin A) = 0 \) is the characteristic function (or indicator function) of \( A \).

**Proposition 1.5.14.** If \( A \subset X \) is measurable, then \( \chi_A : X \to \mathbb{R} \) is measurable.

**Definition 1.5.15** (Simple function). A simple function is a function \( f : X \to \mathbb{C} \) of the form

\[
 f = \sum_{i=1}^{n} \alpha_i \chi_{A_i},
\]

where \( \alpha_i \in \mathbb{C} \) and \( A_i \) is measurable for each \( i \).

**Notation.** Write \( f_n \nearrow f \) if \( f_1 \leq f_2 \leq \cdots \) and \( f_n \to f \) pointwise.

**Theorem 1.5.16.** Let \((X, \mathcal{A})\) be a measurable space and \( f : X \to [0, \infty] \) be a measurable function. Then there exist simple functions \( s_n : X \to [0, \infty) \) such that \( s_n \nearrow f \) and \( s_n \to f \) pointwise.
Proof. Define
\[ E_{n,k} = f^{-1}\left(\left(\frac{k}{2^n}, \frac{k+1}{2^n}\right]\right), \quad F_n = f^{-1}((2^n, +\infty]), \]
and set
\[ s_n = \sum_{k=0}^{2^{2n}-1} \frac{k}{2^n} \cdot \chi_{E_{n,k}} + 2^n \cdot \chi_{F_n}. \]
2 INTEGRATION

2.1 INTEGRATION OF SIMPL E NON-NEGATIVE FUNCTIONS

Note that every simple function has finite image set. Conversely, every measurable function which
has a finite image set is a simple function.

Definition 2.1.1 (Standard representation). If \( s : X \to \mathbb{C} \) is a simple function with image set
\( s(X) = \{ \alpha_1, \ldots, \alpha_n \} \), then the standard representation of \( s \) is

\[
s = \sum_{k=1}^{n} \alpha_k \chi_{A_k}
\]

for \( A_k = s^{-1}(\alpha_k) \).

Notation. Write \( S^+ \) be the space of all simple functions with values in \([0, \infty)\).

Definition 2.1.2 (Integral of a non-negative simple function). Given a simple function \( s \in S^+ \) in
standard representation, the Lebesgue integral of \( s \) with respect to \( \mu \) is

\[
\int s \, d\mu = \sum_{k=1}^{n} \alpha_k \mu(A_k).
\]

Proposition 2.1.3. 1. If \( s = \sum_k \beta_k \chi_{B_k} \) as a finite sum which is not necessarily the standard
representation, then \( \int s = \sum_k \beta_k \mu(B_k) \).

2. If \( s \in S^+ \) and \( c \geq 0 \), then \( \int cs = c \int s \).

3. If \( s, t \in S^+ \), then \( \int (s + t) = \int s + \int t \).

4. If \( s, t \in S^+ \) with \( s \leq t \), then \( \int s \leq \int t \).

Notation. If \( A \in \mathcal{A} \) and \( s \in S^+ \), then write

\[
\int_{A} s \, d\mu = \int s \cdot \chi_{A} \, d\mu.
\]

Proposition 2.1.4. Let \( s \in S^+ \) and define for \( A \in \mathcal{A} \)

\[
\mu(A) = \int_{A} s \, d\mu.
\]

Then \( \mu \) is a measure on \((X, \mathcal{A})\).

2.2 INTEGRATION OF NON-NEGATIVE FUNCTIONS

Notation. Let \((X, \mathcal{A}, \mu)\) be a measure space. Write \( \mathcal{L}^+(\mu) \) for the set of non-negative measurable
functions on \( X \).
2.2 Integration of non-negative functions

Definition 2.2.1 (Integral of a non-negative function). For $f \in \mathcal{L}^+$, the Lebesgue integral of $f$ with respect to $\mu$ is
\[
\int f \, d\mu = \sup \left\{ \int s \, d\mu \mid s \in \mathcal{S}^+ \text{ with } s \leq f \right\}.
\]

Proposition 2.2.2. 1. This is consistent with the previous definition for simple functions.
2. If $f, g \in \mathcal{L}^+$ and $f \leq g$, then $\int f \, d\mu \leq \int g \, d\mu$.

Theorem 2.2.3 (Lebesgue monotone convergence theorem). If $f_n, f \in \mathcal{L}^+$ and $f_n \nearrow f$, then
\[
\int f \, d\mu = \lim_{n \to \infty} \int f_n \, d\mu.
\]

Proof. Since $\{\int f_n \, d\mu\}$ is an increasing sequence of extended real numbers which is bounded above by $\int f \, d\mu$, its limit exists and is at most $\int f \, d\mu$. To show the reverse inequality, let $\alpha \in (0, 1)$ and $\phi$ be a simple function with $0 \leq \phi \leq f$. Define $A_n = \{x \in X \mid f_n(x) \geq \alpha \phi(x)\}$. By construction, $A_n \nearrow X$ and $\int f_n \, d\mu \geq \int_{A_n} f_n \, d\mu \geq \alpha \int_{A_n} \phi \, d\mu$. Taking the limit as $n \to \infty$, we have
\[
\lim_{n \to \infty} \int f_n \, d\mu \geq \lim_{n \to \infty} \alpha \int_{A_n} \phi \, d\mu = \alpha \int \phi \, d\mu.
\]
Taking the supremum over all $\phi$ and letting $\alpha \to 1^-$, we have
\[
\int f \, d\mu \leq \lim_{n \to \infty} \int f_n \, d\mu,
\]
as required. \qed

Proposition 2.2.4. 1. If $f \in \mathcal{L}^+$ and $\alpha \geq 0$, then $\int \alpha f \, d\mu = \alpha \int f \, d\mu$.
2. If $f, g \in \mathcal{L}^+$, then $\int (f + g) \, d\mu = \int f \, d\mu + \int g \, d\mu$.
3. If $f_n \in \mathcal{L}^+$, then $\int (\sum_n f_n) \, d\mu = \sum_n \int f_n \, d\mu$.
4. If $f \in \mathcal{L}^+$, then $\int f = 0$ if and only if $f = 0 \mu$-a.e.

Corollary 2.2.5. 1. If $f_n, f \in \mathcal{L}^+$ and $f_n \nearrow f$ $\mu$-a.e., then $\int f_n \, d\mu \to \int f \, d\mu$.
2. If $f, g \in \mathcal{L}^+$ and $f = g$ $\mu$-a.e., then $\int f \, d\mu = \int g \, d\mu$.

Lemma 2.2.6 (Fatou). Let $f_n : X \to [0, \infty]$ be measurable functions. Then
\[
\int \liminf_{n \to \infty} f_n \, d\mu \leq \liminf_{n \to \infty} \int f_n \, d\mu.
\]

Proof. Let
\[
g_k = \inf_{n \geq k} f_n.
\]
Then $g_k \leq f_m$ for each $m \geq k$, hence
\[
\int g_k \, d\mu \leq \inf_{m \geq k} \int f_m \, d\mu.
\]
Applying Theorem 2.2.3 to the sequence \( \{g_k\} \),

\[
\int \liminf_{n \to \infty} f_n \, d\mu = \int \lim_{k \to \infty} \inf_{m \geq k} g_k \, d\mu \leq \liminf_{n \to \infty} \int f_n \, d\mu.
\]

\[\square\]

2.3 INTEGRATION OF REAL AND COMPLEX VALUED FUNCTIONS

Definition 2.3.1 (Integrable function to \( \mathbb{R} \)). A function \( f : X \to \mathbb{R} \) is integrable (with respect to \( \mu \)) if \( f \) is measurable and \( \int |f| \, d\mu < \infty \).

Notation. The space of real-valued measurable functions is \( \mathcal{L}^1_\mathbb{R}(\mu) \), or \( \mathcal{L}^1 \) if \( \mu \) is understood.

Proposition 2.3.2. If \( f \in \mathcal{L}^1 \), then \( f^+, f^- \in \mathcal{L}^+ \) and \( \int f^+ \, d\mu, \int f^- \, d\mu < \infty \).

Definition 2.3.3 (Integral of a real-valued integrable function). For \( f \in \mathcal{L}^1_\mathbb{R} \), the integral of \( f \) is

\[
\int f \, d\mu = \int f^+ \, d\mu - \int f^- \, d\mu.
\]

Proposition 2.3.4. A function \( f : X \to \mathbb{R} \) is in \( \mathcal{L}^1_\mathbb{R} \) if and only if \( \int f^+ \, d\mu \) and \( \int f^- \, d\mu \) are finite.

Definition 2.3.5 (Integrable function to \( \mathbb{C} \)). A function \( f : X \to \mathbb{C} \) is integrable if \( f \) is measurable and \( \int |f| \, d\mu < \infty \). The space of complex-valued integrable functions is \( \mathcal{L}^1 \).

Notation. The space of complex-valued measurable functions is \( \mathcal{L}^1(\mu) \), or \( \mathcal{L}^1 \) if \( \mu \) is understood.

Proposition 2.3.6. If \( f \in \mathcal{L}^1 \), then \( \text{Re} f, \text{Im} f \in \mathcal{L}^1_\mathbb{R} \).

Definition 2.3.7 (Integral of a complex-valued integrable function / integral over a subset). For \( f \in \mathcal{L}^1 \), the integral of \( f \) is

\[
\int f \, d\mu = \int \text{Re} f \, d\mu + i \int \text{Im} f \, d\mu.
\]

If \( A \in \mathcal{A} \) and \( f \in \mathcal{L}^1 \), define

\[
\int_A f \, d\mu = \int \chi_A f \, d\mu.
\]

Theorem 2.3.8. \( \mathcal{L}^1 \) is a \( \mathbb{C} \)-vector space and the integral is a linear functional on \( \mathcal{L}^1 \).

Proposition 2.3.9. If \( f \in \mathcal{L}^1 \), then \( \int |f| \, d\mu \leq \int |f| \, d\mu \).

Proof. There exists \( \alpha \in \mathbb{C} \) with \( |\alpha| = 1 \) such that \( \int f \, d\mu = \alpha \int f \, d\mu \). Then

\[
\left| \int f \, d\mu \right| = \alpha \int f \, d\mu = \int \alpha f \, d\mu = \int \text{Re}(\alpha f) \, d\mu \leq \int \left| \text{Re}(\alpha f) \right| \, d\mu \leq \int |\alpha f| \, d\mu = \int |f| \, d\mu.
\]

\[\square\]
Theorem 2.3.10 (Lebesgue dominated convergence theorem). Let \( f_n \) be a sequence in \( \mathcal{L}^1 \) and \( f : X \to \mathbb{C} \). Suppose that \( f_n \to f \) pointwise and there exists a non-negative \( g \in \mathcal{L}^1 \) such that \( |f_n| \leq g \) for all \( n \). Then \( f \in \mathcal{L}^1 \) and \( \int |f_n - f| \, d\mu \to 0 \).

Proof. By taking real and imaginary parts, it suffices to assume that \( f_n \) and \( f \) are real-valued. Both \( g + f_n \) and \( g - f_n \) are non-negative integrable functions, so by Lemma 2.2.6,

\[
\begin{align*}
\int g \, d\mu + \int f \, d\mu &= \int \liminf_{n \to \infty} (g + f_n) \, d\mu \leq \liminf_{n \to \infty} \int (g + f_n) \, d\mu = \int g \, d\mu + \liminf_{n \to \infty} \int f_n \, d\mu \\
\int g \, d\mu - \int f \, d\mu &= \int \liminf_{n \to \infty} (g - f_n) \, d\mu \leq \liminf_{n \to \infty} \int (g - f_n) \, d\mu = \int g \, d\mu - \limsup_{n \to \infty} \int f_n \, d\mu.
\end{align*}
\]

The result follows.

Corollary 2.3.11. Let \( f_n, f : X \to \mathbb{C} \) be measurable. Suppose \( f_n(x) \to f(x) \) for \( \mu \)-a.e. \( x \in X \) and there exists a non-negative measurable function \( g : X \to [0, \infty] \) with \( \int g < \infty \) and \( |f_n| \leq g \) \( \mu \)-a.e. for each \( n \). Then the functions \( f_n, f \) are all integrable and \( \int |f_n - f| \, d\mu \to 0 \) as \( n \to \infty \).

Proposition 2.3.12. If \( \int |f_n - f| \, d\mu \to 0 \) as \( n \to \infty \), then \( \int f_n \, d\mu = \int f \, d\mu \).

2.4 \( L^p \)-SPACES

Definition 2.4.1 (\( L^p \)-spaces). Let \( 1 \leq p < \infty \). We define \( L^p(\mu) \) (or \( L^p \) when \( \mu \) is understood) as the space of all integrable functions \( f : X \to \mathbb{C} \) for which \( \int |f|^p \, d\mu < \infty \).

Notation. If \( f : X \to \mathbb{C} \) is measurable, then write

\[
\|f\|_p = \left( \int |f|^p \, d\mu \right)^{1/p}.
\]

Proposition 2.4.2. \( \|f\|_p = 0 \) if and only if \( f = 0 \) \( \mu \)-a.e.

Lemma 2.4.3. If \( a, b \geq 0 \) and \( 0 < \lambda < 1 \), then

\[
a^\lambda b^{1-\lambda} \leq \lambda a + (1-\lambda)b,
\]

with equality if and only if \( a = b \).

Proof. Since \( \log : (0, \infty) \to \mathbb{R} \) is strictly concave, Jensen’s inequality gives

\[
\lambda \log a + (1-\lambda) \log b \leq \log(\lambda a + (1-\lambda)b).
\]

Definition 2.4.4 (Conjugate exponent). If \( 1 < p < \infty \), then the unique \( q \) for which \( 1 < q < \infty \) and \( 1/p + 1/q = 1 \) is the conjugate exponent of \( p \).

Theorem 2.4.5 (Hölder’s inequality). Let \( 1 < p < \infty \) and let \( q \) be the conjugate exponent of \( p \). If \( f, g : X \to \mathbb{C} \) are measurable, then

\[
\|fg\|_1 \leq \|f\|_p \|g\|_q,
\]

with equality if and only if \( \alpha |f|^p = \beta |g|^q \) \( \mu \)-a.e. for some constants \( \alpha, \beta \) which are not both zero.
Proof. If one of \( \|f\|_p \) or \( \|g\|_q \) is 0 or \( \infty \), then the result is clear. Furthermore, scaling \( f \) and \( g \) do not change the validity of the inequality, so it suffices to consider \( \|f\|_p = \|g\|_q = 1 \). By Lemma 2.4.3 with \( a = |f(x)|^p \), \( b = |g(x)|^q \), and \( \lambda = 1/p \), we have

\[
|f(x)g(x)| \leq \frac{1}{p}|f(x)|^p + \frac{1}{q}|g(x)|^q.
\]

Integrating both sides,

\[
\|fg\|_1 \leq \frac{1}{p} \int |f|^p \, d\mu + \frac{1}{q} \int |g|^q \, d\mu = \frac{1}{p} + \frac{1}{q} = 1 = \|f\|_p \|g\|_q.
\]

Equality holds if and only if it holds \( \mu \)-a.e. in (*), which happens when \( |f|^p = |g|^q \) \( \mu \)-a.e. \( \square \)

**Theorem 2.4.6** (Minkowski’s inequality). If \( 1 \leq p < \infty \) and \( f, g \in L^p \), then

\[
\|f + g\|_p \leq \|f\|_p + \|g\|_p.
\]

Proof. If \( f + g = 0 \) \( \mu \)-a.e. or \( p = 1 \), then the result is clear. Otherwise,

\[
|f + g|^p \leq (|f| + |g|)|f + g|^{p-1},
\]

so applying Theorem 2.4.5 twice,

\[
\int |f + g|^p \, d\mu \leq \|f\|_p \|f + g|^{p-1}\|_q + \|g\|_p \|f + g|^{p-1}\|_q
\]

\[
= (\|f\|_p + \|g\|_p) \left( \int |f + g|^p \, d\mu \right)^{1/q}
\]

\[
\|f + g\|_p \leq (\|f\|_p + \|g\|_p)^{(1-q^{-1})-1/p} = \|f\|_p + \|g\|_p.
\]

\( \square \)

**Corollary 2.4.7.** Let \( 1 \leq p < \infty \). Then \( L^p \) is a \( \mathbb{C} \)-vector space on which \( \|\cdot\|_p \) is a seminorm.

**Notation.** For measurable functions \( f, g : X \to \mathbb{C} \), write \( f \sim g \) if \( f = g \) \( \mu \)-a.e. This is an equivalence relation on the set of measurable functions and on each \( L^p \).

**Definition 2.4.8** \((L^p\text{-spaces})\). Let \( 1 \leq p < \infty \). Define \( L^p(\mu) = L^p(\mu)/\sim \).

**Theorem 2.4.9.** \( L^p \) is a \( \mathbb{C} \)-vector space on which \( \|\cdot\|_p \) is a norm.

**Definition 2.4.10** (Banach space). A normed (real or complex) vector space \( V \) is a Banach space if it is complete with respect to the metric induced by the norm.

**Lemma 2.4.11.** A normed vector space is a Banach space if and only if every absolutely convergent series converges.

**Theorem 2.4.12.** \( L^p \) is a Banach space for \( 1 \leq p < \infty \).
2.5 Relation to Riemann integration

Proof. Let $f_n \in L^p$ be such that $\sum_n \|f_n\|_p = S < \infty$. Define

$$F_n = \sum_{k=1}^n f_k, \quad G_n = \sum_{k=1}^n |f_k|, \quad G = \sum_{k=1}^\infty |f_k|.$$ 

For each $n$, we have $|G_n| \leq B$, so by Theorem 2.2.3,

$$\int G^p \, d\mu = \lim_{n \to \infty} \int G_n^p \, d\mu \leq B^p.$$ 

This tells us that $G \in L^p$, so in particular $G(x) < \infty \, \mu$-a.e. If $F = \sum_k f_k$ (defined $\mu$-a.e.), we have $|F| \leq G$, so $F \in L^p$. Furthermore,

$$\left| F - \sum_{k=1}^n f_k \right| \leq (2G)^p,$$

so by Theorem 2.3.10,

$$\left\| F - \sum_{k=1}^n f_k \right\|_p^p = \int \left| F - \sum_{k=1}^n f_k \right|^p \, d\mu \to 0,$$

i.e. the series for $F$ converges in the $L^p$ norm. \qed

Proposition 2.4.13. For $1 \leq p < \infty$, the set of simple functions $s = \sum_j a_j \chi_{E_j}$ with $\mu(E_j) < \infty$ for all $j$ is dense in $L^p$.

Definition 2.4.14 (Essential supremum / $L^\infty$). Let $f : X \to \mathbb{C}$ be a measurable function. The essential supremum of $f$ is

$$\|f\|_\infty = \text{ess sup}_{x \in X} |f(x)| = \{ \lambda \geq 0 \mid \mu(\{x \in X \mid |f(x)| > \lambda \}) = 0 \}.$$ 

The space of measurable functions with finite essential supremum is $L^\infty$, and $L^\infty = L^\infty / \sim$.

Remark 2.4.15. We may regard $\infty$ as the conjugate exponent for 1 and vice versa.

Theorem 2.4.16. 1. Hölder’s inequality extends to $\{p, q\} = \{1, \infty\}$.

2. $L^\infty$ is a $C$-vector space on which $\| \cdot \|_\infty$ is a norm.

3. $L^\infty$ is a Banach space.

4. Simple functions are dense in $L^\infty$.

2.5 RELATION TO RIEMANN INTEGRATION

Notation. Given a function $f : [a, b] \to \mathbb{R}$ and a partition $P$ of $[a, b]$, write $U_P f$ and $L_P f$ for the upper and lower sums of $f$ on the partition.

Theorem 2.5.1. If $f : [a, b] \to \mathbb{R}$ is Riemann integrable, then $f$ is Lebesgue integrable and

$$\int_a^b f(x) \, dx = \int_{[a,b]} f \, d\lambda.$$
Proof. For a partition \( P \) given by \( a = t_0 < t_1 < \cdots < t_n = b \), let

\[
M_i = \sup_{x \in [t_{i-1}, t_i]} f(x) \quad \text{and} \quad m_i = \inf_{x \in [t_{i-1}, t_i]} f(x),
\]

then define simple functions

\[
\Phi_P = \sum_{i=1}^{n} M_i \chi_{(t_{j-1}, t_j)} \quad \text{and} \quad \phi_P = \sum_{i=1}^{n} m_i \chi_{(t_{j-1}, t_j)}
\]

so that \( U_P f = \int_{[a,b]} \Phi_P \, d\lambda \) and \( L_P f = \int_{[a,b]} \phi_P \, d\lambda \). By assumption, there are partitions \( P_k \), each one a refinement of the previous, such that \( U_{P_k} f \) and \( L_{P_k} f \) converge to \( \int_a^b f(x) \, dx \). Suppose \( \Phi_{P_k} \rightarrow \Phi \) and \( \phi_{P_k} \rightarrow \phi \); note that these limits are monotone. Then we have \( \phi_{P_k} \leq \phi \leq f \leq \Phi \leq \Phi_{P_k} \leq \Phi_P \) for each \( k \), so by the dominated convergence theorem,

\[
\int_{[a,b]} \phi \, d\lambda = \lim_{k \to \infty} \int_{[a,b]} \phi_{P_k} \, d\lambda = \int_{[a,b]} f(x) \, dx, \quad \int_{[a,b]} \Phi \, d\lambda = \lim_{k \to \infty} \int_{[a,b]} \Phi_{P_k} \, d\lambda = \int_{[a,b]} f(x) \, dx.
\]

Thus \( \int_{[a,b]} (\Phi - \phi) \, d\lambda = 0 \) with \( \Phi - \phi \geq 0 \), so \( \Phi = \phi \) almost everywhere. Since \( \phi \leq f \leq \Phi \), in fact \( \phi = f \) almost everywhere. As \( \phi \) is the limit of simple functions, \( \phi \) is measurable, so \( f \) is also measurable. Furthermore,

\[
\int_{[a,b]} \phi \, d\lambda \leq \int_{[a,b]} f \, d\lambda = \int_{[a,b]} \Phi \, d\lambda = \int_{[a,b]} f(x) \, dx,
\]

so \( \int_{[a,b]} f \, d\lambda = \int_{[a,b]} f(x) \, dx \).

\[ \square \]

Theorem 2.5.2. If \( f : [a, b] \to \mathbb{R} \) is bounded, then \( f \) is Riemann integrable if and only if the set of discontinuities of \( f \) has measure zero.

Proof. To be written. (Folland ex 2.23) \[ \square \]

2.6 MODES OF CONVERGENCE

Definition 2.6.1 (Convergence in measure). Let \( f_n, f : X \to \mathbb{C} \) be measurable functions. We say that \( f_n \to f \) in measure if for each \( \epsilon > 0 \),

\[
\lim_{n \to \infty} \mu ( \{ x \in X \mid |f_n(x) - f(x)| > \epsilon \} ) = 0.
\]

Example 2.6.2. Consider three modes of convergence for measurable functions \( f_n, f : X \to \mathbb{C} \):

1. \( f_n \to f \) pointwise \( \mu \)-a.e.;
2. \( f_n \to f \) in \( L^p \) for some \( 1 \leq p < \infty \);
3. \( f_n \to f \) in measure.
Given $1 \leq p < \infty$, the example on $[0, 1]$ given by
\[ f_n = n^{1/p} \chi_{[0,1/n]} \quad \text{and} \quad f = 0 \]
shows that $1 \not\Rightarrow 2$ and $3 \Rightarrow 2$. The typewriter sequence on $[0, 1]$ given by
\[ f_n = \chi_{[j/2^k,(j+1)/2^k]} \quad \text{and} \quad f = 0, \]
where $n = 2^k + j$ with $0 \leq j < 2^k$, shows that $2 \not\Rightarrow 1$ and $3 \Rightarrow 1$. The example on $\mathbb{R}$ given by
\[ f_n = \chi_{[n,n+1]} \quad \text{and} \quad f = 0 \]
shows that $1 \Rightarrow 3$. The only remaining implication, $2 \implies 3$, turns out to be correct. The implications $3 \iff 1$ and $1 \iff 3$ can also be made to work with slight modifications.

**Proposition 2.6.3.** If $f_n \rightarrow f$ in $L^p$ for some $1 \leq p < \infty$, then $f_n \rightarrow f$ in measure.

**Proof.** Let $\epsilon > 0$ and consider $E_n = \{ x \in X \mid |f_n(x) - f(x)| \geq \epsilon \}$. We have
\[ \int |f_n - f| \, d\mu \geq \int_{E_n} |f_n - f| \, d\mu \geq \epsilon \cdot \mu(E_n), \]

so $\mu(E_n) \leq \epsilon^{-1} \int |f_n - f| \, d\mu \rightarrow 0$ as $n \rightarrow \infty$. \hfill \Box

**Proposition 2.6.4.** If $f_n \rightarrow f$ in measure, then there exists a subsequence $(f_{n_j})_j$ such that $f_{n_j} \rightarrow f$ pointwise $\mu$-a.e.

**Proof.** If $f_n \rightarrow f$ in measure, then $f_n$ is also Cauchy in measure. Choose a subsequence $g_j = f_{n_j}$ such that for
\[ E_j = \{ x \in X \mid |g_j(x) - g_{j+1}(x)| \geq 2^{-j} \}, \]
we have $\mu(E_k) \leq 2^{-k}$. Setting $F_k = \bigcup_{j \geq k} E_j$, then $\mu(F_k) \leq 2^{k-1}$ and
\[ |g_j(x) - g_l(x)| \leq \sum_{i=j}^{l-1} |g_{i+1}(x) - g_i(x)| \leq 2^{1-j} \]
for $i \geq j \geq k$ and $x \not\in F_k$. This means that $\{g_j(x)\}$ is Cauchy if $x \not\in F_k$, for any $k$. Let $F = \bigcap_k F_k$. Then $\mu(F) = 0$ and $f(x) = \lim g_j(x)$ is defined $\mu$-a.e. (Define $f$ to be zero elsewhere.) We have that $f$ is measurable and $g_j \rightarrow f$ $\mu$-a.e. \hfill \Box

**Theorem 2.6.5** (Egorov). If $f_n \rightarrow f$ pointwise $\mu$-a.e. and $\mu(X) < \infty$, then for every $\epsilon > 0$, there exists $E \subset X$ such that $\mu(E) < \epsilon$ and $f_n \rightarrow f$ uniformly on $X \setminus E$.

**Proof.** Without loss of generality, suppose $f_n \rightarrow f$ everywhere. Given $n$ and $k$, let
\[ E_n(k) = \bigcup_{m \geq n} \{ x \in X \mid |f_m(x) - f(x)| \geq k^{-1} \}. \]
For fixed $k$, we have a decreasing sequence in $n$ with $\bigcap_n E_n(k) = \emptyset$, so since the measure is finite, we have $\mu(E_n(k)) \rightarrow \mu(\emptyset) = 0$. Given $\epsilon > 0$ and $k$, choose $n_k$ so that $\mu(E_{n_k}(k)) < \epsilon \cdot 2^{-k}$ and define $E = \bigcup_k E_{n_k}(k)$. Then $\mu(E) < \epsilon$ and $f_n \rightarrow f$ uniformly on $E^c$ by construction. \hfill \Box
**Corollary 2.6.6.** If \( f_n \to f \) pointwise \( \mu \)-a.e. and \( \mu(X) < \infty \), then \( f_n \to f \) in measure.

**Definition 2.6.7** (Radon measure). A Borel measure \( \mu \) on a topological space \( X \) is a Radon measure if it is finite on all compact sets, outer regular on all Borel sets, and inner regular on all open sets.

**Lemma 2.6.8** (Urysohn). Let \( X \) be a normal topological space and let \( A, B \subset X \) be disjoint non-empty closed subsets of \( X \). Then there exists a continuous function \( f : X \to [0,1] \) such that \( f(A) = \{0\} \) and \( f(B) = \{1\} \).

**Proof.** Let \( Q \subset [0,1] \) consist of all dyadic rationals. We construct a collection of open sets \( U_q \) indexed by \( Q \) with \( U_q \subset U_r \) whenever \( q \leq r \). Set \( U_1 = X \). Since \( X \) is normal, there are open sets \( U \) and \( V \) separating \( A \) and \( B \). Set \( U_0 = U \). Since \( X \) is normal, there is a set \( U_{1/2} \) with \( U_0 \subset U_{1/2} \subset U_{1/2} \subset X \setminus B \). Repeating this, we inductively define all of the sets \( U_q \) for \( q \in Q \).

The continuous function we wish to define is then

\[
f(x) = \inf\{q \in Q \mid x \in U_q\}.
\]

\( \square \)

**Theorem 2.6.9** (Lusin). Let \( \mu \) be a Radon measure on a locally compact Hausdorff space \( X \). Suppose \( f : X \to \mathbb{C} \) is measurable and \( A \subset X \) is a Borel set of finite measure. Then for each \( \epsilon > 0 \), there exists a closed \( E \subset A \) such that \( \mu(E) < \epsilon \) and \( f \) is continuous on \( A \setminus E \).

**Proof.** To be written. \( \square \)

**Corollary 2.6.10.** The space \( C_c(X) \) of continuous functions with compact support is dense in \( L^p \) for \( 1 \leq p < \infty \).

**Proof.** To be written. \( \square \)

**Remark 2.6.11.** In the case \( p = \infty \), we still have \( C_c(X) \subset L^\infty \), but \( C_c(X) \) need not be dense in \( L^\infty \). For example, in \( \mathbb{R} \) with Lebesgue measure, the ball of radius \( 1/2 \) around \( f = \chi_{[0,1]} \) contains no continuous function with compact support.

### 2.7 Product Measures

**Definition 2.7.1** (Product \( \sigma \)-algebra). Let \( (X, \mathcal{A}) \) and \( (Y, \mathcal{B}) \) be measurable spaces. The product \( \sigma \)-algebra on \( X \times Y \) is

\[
\mathcal{A} \otimes \mathcal{B} = \sigma\{A \times B \mid A \in \mathcal{A} \text{ and } B \in \mathcal{B}\}.
\]

**Definition 2.7.2** (Sections). Let \( (X, \mathcal{A}) \) and \( (Y, \mathcal{B}) \) be measurable spaces and \( E \subset X \times Y \). For \( x \in X \) and \( y \in Y \), the \( x \)-section \( E_x \) and the \( y \)-section \( E^y \) are

\[
E_x = \{y \in Y \mid (x, y) \in E\}, \quad E^y = \{x \in X \mid (x, y) \in E\}.
\]

Given \( f : X \times Y \to \mathbb{C} \), for \( x \in X \) and \( y \in Y \), the \( x \)-section \( f_x : Y \to \mathbb{C} \) and the \( y \)-section \( f^y : X \to \mathbb{C} \) are defined by

\[
f_x(y) = f^y(x) = f(x, y).
\]
Proposition 2.7.3. 1. If \( E \in \mathcal{A} \otimes \mathcal{B} \), then \( E_x \in \mathcal{B} \) and \( E^y \in \mathcal{A} \) for all \( x \in X \) and \( y \in Y \).

2. If \( f : X \times Y \to \mathbb{C} \) is measurable, then \( f_x \) and \( f^y \) are measurable.

Lemma 2.7.4 (Monotone class theorem). Let \((X, \mathcal{A})\) be a measurable space and \( \mathcal{F} \) be a family of functions \( f : X \to \mathbb{C} \) such that

(i) there exists a \( \pi \)-system \( \mathcal{P} \subset \mathcal{A} \) with \( X \in \mathcal{P} \), \( \sigma(\mathcal{P}) = \mathcal{A} \), and \( \chi_A \in \mathcal{F} \) for each \( A \in \mathcal{P} \);

(ii) \( \mathcal{F} \) is closed under linear combinations;

(iii) \( \mathcal{F} \) is closed under monotone limits.

Then \( \mathcal{F} \) contains all measurable functions \( X \to \mathbb{R} \).

Proof. See Homework 6 Problem 1. \( \square \)

Lemma 2.7.5. Let \((X, \mathcal{A}, \mu)\) and \((Y, \mathcal{B}, \nu)\) be \( \sigma \)-finite measure spaces. If \( E \in \mathcal{A} \otimes \mathcal{B} \), then the functions \( x \mapsto \nu(E_x) \) and \( y \mapsto \mu(E^y) \) are measurable and

\[
\int \nu(E_x) \, d\mu = \int \mu(E^y) \, d\nu.
\]

Proof. Suppose \( \mu \) and \( \nu \) are finite, and let \( \mathcal{F} \) be the set of all \( E \in \mathcal{A} \otimes \mathcal{B} \) for which this claim holds. For \( E = A \times B \), we have \( \nu(E_x) = \chi_A(x)\nu(B) \) and \( \mu(E^y) = \mu(A)\chi_B(y) \), so

\[
\int \nu(E_x) \, d\mu = \nu(B) \int \chi_A(x) \, d\mu = \nu(B)\mu(A),
\]

\[
\int \mu(E^y) \, d\nu = \mu(A) \int \chi_B(y) \, d\nu = \mu(A)\nu(B),
\]

showing that \( E \in \mathcal{F} \). By additivity of the integral, finite disjoint unions of rectangles are in \( \mathcal{F} \), thus it remains to show closure under monotone limits. Let \( E_n \nearrow E \) with \( E_n \in \mathcal{F} \). Then \( f_n(y) = \mu((E_n)^y) \) and \( g_n(x) = \nu((E_n)_x) \) are measurable with \( f_n \nearrow f : y \mapsto \mu(E^y) \) and \( g_n \nearrow g : x \mapsto \nu(E_x) \). Hence \( f \) and \( g \) are measurable, and by the monotone convergence theorem (2.2.3),

\[
\int \mu(E^y) \, d\nu = \lim_{n \to \infty} \int \mu((E_n)^y) \, d\nu = \lim_{n \to \infty} \int \nu((E_n)_x) \, d\mu = \int \nu(E_x) \, d\mu.
\]

Thus \( E \in \mathcal{F} \), as required to complete the proof in the finite measure case.

Applying the monotone convergence theorem again gives the result for general \( \sigma \)-finite measures. \( \square \)

Theorem 2.7.6 (Tonelli). Let \((X, \mathcal{A}, \mu)\) and \((Y, \mathcal{B}, \nu)\) be \( \sigma \)-finite measure spaces.

1. There exists a unique measure \( \omega = \mu \times \nu \) on \((X \times Y, \mathcal{A} \otimes \mathcal{B})\) such that \( \omega(A \times B) = \mu(A)\nu(B) \) whenever \( A \in \mathcal{A} \) and \( B \in \mathcal{B} \).

2. If \( f : X \times Y \to [0, \infty] \) is measurable, then the functions

\[
G(x) = \int f_x \, d\nu \quad \text{and} \quad H(y) = \int f^y \, d\mu
\]

are also non-negative measurable, and

\[
\int f \, d\omega = \int G \, d\mu = \int H \, d\nu.
\]
Proof. 1. See Homework 8 Problems 1 and 2.

2. The case where \( f \) is a characteristic function is Lemma 2.7.5, and we get non-negative simple functions by linearity of the integral. For a general measurable \( f \), take simple functions \( s_n \nearrow f \) and define \( G_n, H_n \) on the \( s_n \) as above. By the monotone convergence theorem (2.2.3), \( G_n \nearrow G \) and \( H_n \nearrow H \), so \( G \) and \( H \) are measurable. Applying it again,

\[
\int G \, d\mu = \lim_{n \to \infty} \int G_n \, d\mu = \lim_{n \to \infty} \int s_n \, d\omega = \int f \, d\omega,
\]

and similarly for \( \int H \, d\nu \).

\[\Box\]

Theorem 2.7.7 (Fubini). Let \((X,\mathcal{A},\mu)\) and \((Y,\mathcal{B},\nu)\) be \(\sigma\)-finite measure spaces and \(\omega = \mu \times \nu\).

1. If \( f \in L^1(\omega) \), then \( f_x \in L^1(\nu) \) and \( f^y \in L^1(\mu) \) for \(\mu\text{-a.e. } x \in X \) and \(\nu\text{-a.e. } y \in Y\).

2. The functions

\[
G(x) = \int f_x \, d\nu \quad \text{and} \quad H(y) = \int f^y \, d\mu
\]

are defined for \(\mu\text{-a.e. } x \in X \) and \(\nu\text{-a.e. } y \in Y\), and by making arbitrary (re-)definitions on a null set, \( G \in L^1(\mu) \) and \( H \in L^1(\nu) \).

3. We have

\[
\int f \, d\omega = \int G \, d\mu = \int H \, d\nu.
\]

Proof. If \( f \) is non-negative integrable, then by Tonelli’s theorem (2.7.6), it follows that \( G < \infty \) \(\mu\text{-a.e.} \) and \( H < \infty \) \(\nu\text{-a.e.} \), so \( f_x \in L^1(\nu) \) and \( f^y \in L^1(\mu) \) for almost every \( x \in X \) and \( y \in Y \).

Fubini’s theorem then pops out by applying Tonelli’s theorem to the positive and negative parts of the real and imaginary parts of \( f \).

\[\Box\]

Lemma 2.7.8. If \((Z,\mathcal{A},\mu)\) is the completion of \((Z,\mathcal{A},\mu)\) and \( f : Z \to \mathbb{C} \) is \(\mathfrak{p}\)-measurable, then there exists a \(\mu\)-measurable function \( g : Z \to \mathbb{C} \) such that \( f = g \mathfrak{p}\text{-a.e.} \).

Lemma 2.7.9. If \( h = 0 \) \(\omega\text{-a.e.} \) on \( X \times Y \), then \( h_x = 0 \) \(\nu\text{-a.e.} \) for \(\mu\text{-a.e. } x \in X \) and \( h^y = 0 \) \(\mu\text{-a.e.} \) for \(\nu\text{-a.e. } y \in Y \).

Theorem 2.7.10 (Fubini-Tonelli for complete measures). Let \((X,\mathcal{A},\mu)\) and \((Y,\mathcal{B},\nu)\) be \(\sigma\)-finite complete measure spaces and \((X \times Y,\mathcal{A} \otimes \mathcal{B},\mu \times \nu)\) be the completion of \((X \times Y,\mathcal{A} \otimes \mathcal{B},\mu \times \nu)\). Suppose that \( f : X \times Y \to \mathbb{C} \) is a measurable function.

1. If \( f \) is in \( L^+ \), then \( f_x \) and \( f^y \) are measurable for \(\mu\text{-a.e. } x \in X \) and \(\nu\text{-a.e. } y \in Y \). Moreover, the functions \( x \mapsto \int f_x \, d\nu \) and \( y \mapsto \int f^y \, d\mu \) are measurable.

2. If \( f \) is in \( L^1(\omega) \), then \( f_x \) and \( f^y \) are integrable for \(\mu\text{-a.e. } x \in X \) and \(\nu\text{-a.e. } y \in Y \). Moreover, the functions \( x \mapsto \int f_x \, d\nu \) and \( y \mapsto \int f^y \, d\mu \) are integrable.

In both cases,

\[
\int f \, d\omega = \int \int f(x,y) \, d\mu(x) \, d\nu = \int \int f(x,y) \, d\nu(y) \, d\mu(x).
\]
**Example 2.7.11.** Let $B_n$ be the Borel $\sigma$-algebra on $\mathbb{R}^n$ and $L_n$ be the $\sigma$-algebra of Lebesgue measurable subsets of $\mathbb{R}^n$. Define a Borel measure $\beta_n$ on $B_n$ by restricting the Lebesgue measure $\lambda_n$ to Borel sets. Since $B_{m+n} = B_m \otimes B_n$ (see Homework 5 Problem 1), uniqueness of product measure implies $\beta_{m+n} = \beta_m \times \beta_n$. In the case of Lebesgue measure, it turns out that $\lambda_{m+n} \neq \lambda_m \times \lambda_n$. However, it is the case that $\lambda_{m+n}$ is the completion of $\lambda_m \times \lambda_n$, so Fubini-Tonelli can be used for Lebesgue integrals.

**Proposition 2.7.12.** Let $f : \mathbb{R} \to [0, \infty)$ be a measurable function. Then

$$M = \{(x, y) \mid 0 \leq y \leq f(x)\} \subset \mathbb{R}^2$$

is measurable, and

$$\lambda_2(M) = \int f \, d\lambda_1.$$

**Proof.** To be written. \qed

## 2.8 Polar Coordinates

**Definition 2.8.1 (Restriction of a measure).** Let $(X, \mathcal{A}, \mu)$ be a measure space and $M \subset \mathcal{A}$ be measurable. The **restriction** of $\mu$ to $M$ is the measure space $(M, \mathcal{A}_M, \mu|_M)$, where $\mathcal{A}_M = \{A \in \mathcal{A} \mid A \subset M\}$ and $\mu|_M(A) = \mu(A)$ for $A \in \mathcal{A}_M$.

**Proposition 2.8.2.** Let $(X, \mathcal{A}, \mu)$ be a measure space, $(M, \mathcal{A}_M, \mu|_M)$ be the restriction of $\mu$ to $M \in \mathcal{A}$, and $f : M \to \mathbb{C}$ be a function. Define $\tilde{f} : X \to \mathbb{C}$ by

$$\tilde{f} = \begin{cases} 0 & x \notin M, \\ f(x) & x \in M. \end{cases}$$

Then

1. $f$ is measurable (integrable) if and only if $\tilde{f}$ is measurable (integrable);
2. if $f$ is integrable, then

$$\int_M f \, d\mu|M = \int_X \tilde{f} \, d\mu.$$

**Definition 2.8.3 (Pushforward measure).** Let $(X, \mathcal{A}, \mu)$ be a measure space, $(Y, \mathcal{B})$ be a measurable space, and $T : (X, \mathcal{A}) \to (Y, \mathcal{B})$ be measurable. The **pushforward measure** of $\mu$ by $T$ is the measure defined on $(Y, \mathcal{B})$ by

$$T_*\mu(B) = \mu(T^{-1}(B)).$$

**Proposition 2.8.4.** If $g : Y \to [0, \infty]$ is measurable, then $g \circ T : X \to [0, \infty]$ is measurable and

$$\int g \, d(T_*\mu) = \int g \circ T \, d\mu.$$
Representation in polar coordinates induces a homeomorphism \( \Phi : \mathbb{R}^n \setminus \{0\} \to (0, \infty) \times S^{n-1} \). Let \( m = \Phi_* (\beta_n) \) be the pushforward measure of \( \beta_n \) by \( \Phi \). That is, given \( A \subset (0, \infty) \times S^{n-1} \) Borel, we have
\[
m(A) = \beta_n(\Phi^{-1}(A)) = \lambda_n(\Phi^{-1}(A)).
\]

Let \( \rho \) be the Borel measure on \((0, \infty)\) defined as
\[
\rho(E) = \int_E x^{n-1} \, dx = \int_E x^{n-1} \, dx.
\]

**Theorem 2.8.6.** There exists a unique Borel measure \( \sigma \) on \( S^{n-1} \) such that \( m = \rho \times \sigma \). If \( f \) is a Borel measurable function on \( \mathbb{R}^n \) with values in \([0, \infty)\) or \( f \in L^1(\beta_n) \), then
\[
\int f \, d\lambda_n = \int_0^\infty \left( \int_{S^{n-1}} f(r\xi) r^{n-1} \, d\sigma(\xi) \right) \, dr.
\]

**Proof.** To show existence of \( \sigma \), if \( E \subset S^{n-1} \) is Borel, then let
\[
\tilde{E} = \{ r\xi \mid 0 < r < 1 \text{ and } \xi \in E \}
\]
be the open cone with base \( E \). This is Borel, and we can define a Borel measure
\[
\sigma(E) = n\lambda_n(\tilde{E}).
\]

We claim that \( m = \rho \times \sigma \). By uniqueness of the product measure, it suffices to show that \( m(A \times B) = \rho(A)\sigma(B) \) whenever \( A \subset (0, \infty) \) and \( B \subset S^{n-1} \) are Borel. We have
\[
m(A \times B) = \lambda_n(\Phi^{-1}(A \times B)) = \lambda_n(\{r\xi \mid r \in A \text{ and } \xi \in B\}).
\]

Fixing \( B \subset S^{n-1} \) Borel, it suffices to do the proof on the \( \pi \)-system of intervals \((0, \alpha)\) with \( \alpha > 0 \). We have
\[
\lambda_n(\{r\xi \mid r \in (0, \alpha) \text{ and } \xi \in B\}) = \lambda_n(\alpha \cdot \tilde{B}) = \alpha^n \lambda_n(\tilde{B}) = \frac{\alpha^n}{n} \sigma(B) = \left( \int_0^\alpha r^{n-1} \, dr \right) \sigma(B) = \rho((0, \alpha))\sigma(B).
\]

For uniqueness, suppose \( m(A \times B) = \rho(A)\sigma(B) = \rho(A)\hat{\sigma}(B) \). Picking \( A = (0, 1) \), we have \( \sigma(B) = \hat{\sigma}(B) \) for all Borel sets \( B \subset S^{n-1} \).

To show that the integral is correct, it is enough to consider \( f = \chi_M \) for \( M \subset \mathbb{R}^n \) Borel and then apply the monotone class theorem. Then \( N = M \setminus \{0\} \). By Fubini,
\[
\int_{\mathbb{R}^n} \chi_M \, d\lambda_n = \int_{\mathbb{R}^n \setminus \{0\}} \chi_N \, d\beta_n = \int_{\mathbb{R}^n \setminus \{0\}} (\chi_N \circ \Phi^{-1}) \circ \Phi \, d(\beta_n|_{\mathbb{R}^n \setminus \{0\}}) = \int_{(0, \infty) \times S^{n-1}} \chi_N(r\xi) \, dm(r, \xi) = \int_0^\infty \left( \int_{S^{n-1}} \chi_M(r\xi) r^{n-1} \, d\sigma(\xi) \right) \, dr.
\]
**Proposition 2.8.7.** 1. \( \sigma \) is invariant under rotations.

2. If \( n = 2 \) and \( E \subset S^1 \) is an arc of angle \( \alpha \), then \( \sigma(E) = \alpha \) (arc length measure).

**Lemma 2.8.8.**

\[
I_n = \int_{\mathbb{R}^n} e^{-|x|^2} \, d\lambda_n(x) = \pi^{n/2}
\]

**Proof sketch.** Use Fubini to show that \( I_{n+1} = I_n \cdot I_1 \). That \( I_1 = \sqrt{\pi} \) is a well-known argument via computing \( I_2 \) with polar coordinates. \( \square \)

**Definition 2.8.9 (\( \Gamma \) function).** The \( \Gamma \) function on positive reals is

\[
\Gamma(z > 0) = \int_0^\infty t^{z-1} e^{-t} \, dt.
\]

This admits an analytic continuation to a meromorphic function on the complex plane.

**Proposition 2.8.10.** 1. \( \Gamma(z + 1) = z\Gamma(z) \);

2. \( \Gamma(z)\Gamma(1 - z) = \pi / \sin(\pi z) \);

3. \( \Gamma(n) = (n - 1)! \) for \( n \in \mathbb{N} \);

4. \( \Gamma(1/2) = \sqrt{\pi} \).

**Proposition 2.8.11.** For \( n \in \mathbb{N} \),

\[
\sigma(S^{n-1}) = \frac{2\pi^{n/2}}{\Gamma(n/2)}
\]

If \( \mathbb{B}^n = \{ x \in \mathbb{R}^n \mid |x| < 1 \} \), then

\[
\lambda_n(\mathbb{B}^n) = \frac{\pi^{n/2}}{\Gamma(n/2 + 1)}.
\]

**Proof.** We compute

\[
I_n = \pi^{n/2} = \int_{\mathbb{R}^n} e^{-|x|^2} \, d\lambda_n(x)
\]

\[
= \int_0^\infty \left( \int_{S^{n-1}} e^{-r^2} r^{n-1} \, d\sigma(\xi) \right) \, dr
\]

\[
= \sigma(S^{n-1}) \int_0^\infty e^{-r^2} r^{n-1} \, dr
\]

\[
= \sigma(S^{n-1}) \int_0^\infty \frac{1}{2} e^{-x \cdot x^{n/2}} \, dx,
\]
which gives the first formula. For the second formula,

\[
\lambda_n(\mathbb{B}^n) = \lambda_n(\mathbb{B}^n \setminus \{0\}) = \lambda_n(\tilde{S}^{n-1}) = \frac{1}{n} \sigma(S^{n-1}) = \frac{1}{n} \cdot 2\pi^{n/2} \cdot \frac{\pi^{n/2}}{\Gamma(n/2)} = \frac{\pi^{n/2}}{\Gamma(n/2 - 1)}.
\]

\[\square\]

**Example 2.8.12.** For the first few \(n\),

\[
\lambda_1(\mathbb{B}^1) = 2, \quad \lambda_3(\mathbb{B}^3) = \frac{4\pi}{3}, \quad \lambda_4(\mathbb{B}^4) = \frac{\pi^2}{2}.
\]

For odd dimensions, it is helpful to use the functional equation to reduce the \(\Gamma\) functions to the case \(\Gamma(1/2)\).

### 2.9 THE TRANSFORMATION FORMULA

**Definition 2.9.1** (\(C^1\)-diffeomorphism). Let \(U, V \subset \mathbb{R}^n\) be open. A \(C^1\)-diffeomorphism \(T : U \to V\) is a bijection such that \(T\) and \(T^{-1}\) are both \(C^1\)-smooth.

**Lemma 2.9.2.** Let \(U \subset \mathbb{R}^n\) and \(T : U \to V \subset \mathbb{R}^m\) be \(C^1\)-smooth with \(T(x) = (y_1(x), \ldots, y_m(x))\). Then each component function \(y_i\) is \(C^1\)-smooth, and the derivative of \(T\) is the Jacobian matrix

\[
DT(x) = \begin{pmatrix}
\frac{\partial y_1}{\partial x_1}(x) & \cdots & \frac{\partial y_1}{\partial x_n}(x) \\
\vdots & \ddots & \vdots \\
\frac{\partial y_m}{\partial x_1}(x) & \cdots & \frac{\partial y_m}{\partial x_n}(x)
\end{pmatrix}.
\]

**Definition 2.9.3** (Jacobian). The Jacobian (determinant) is the determinant of the Jacobian matrix,

\[
J_T(x) = \det(DT(x)).
\]

**Lemma 2.9.4** (Chain rule). Let \(T : U \to V\) and \(S : V \to W\) be \(C^1\)-smooth. Then

\[
D(S \circ T)(x) = DS(T(x)) \circ DT(x).
\]

**Corollary 2.9.5.** If \(T : U \to V\) is a \(C^1\)-diffeomorphism, then \(DT^{-1}(T(x)) = (DT(x))^{-1}\).

**Lemma 2.9.6** (Whitney decomposition). Let \(\Omega \subseteq \mathbb{R}^n\) be non-empty and open. Then there is a countable collection of cubes \(Q_k\) such that \(\Omega = \bigcup_k Q_k\) and the \(Q_k\) have disjoint interiors. Moreover, we can ensure that

\[
diam(Q_k) \leq \text{dist}(Q_k, \partial \Omega) \leq 4 \text{diam}(Q_k).
\]
Proof. To be written.  

Theorem 2.9.7 (Transformation formula). Let $U, V \subset \mathbb{R}^n$ be open and $T : U \to V$ be a $C^1$-diffeomorphism. If $f$ is a (Lebesgue) measurable function on $V$ taking values in $[0, \infty]$, or if $f$ is an integrable function on $V$ taking values in $\mathbb{C}$, then $f \circ T$ is measurable / integrable,

$$
\int_{V} f \, d\lambda_n = \int_{U} (f \circ T) |J_T| \, d\lambda_n.
$$

Proof (outline). Since $T$ is a homeomorphism, $E \subset U$ is Borel if and only if $T(E) \subset V$ is Borel. Furthermore, since $T$ is a $C^1$-diffeomorphism, $N \subset U$ is a null set if and only if $T(N) \subset V$ is a null set. Thus $M \subset U$ is measurable if and only if $T(M)$ is measurable, and $f$ is a measurable function on $V$ if and only if $f \circ T$ is a measurable function on $U$. Since $|J_T|$ is continuous, $(f \circ T) |J_T|$ is measurable. (These statements all work with measurable replaced by integrable.)

First we show that if $Q \subset U$ is a cube, then

$$
\lambda_n(T(Q)) \leq \int_{Q} |J_T| \, d\lambda_n.
$$

For this, decompose $Q$ into small cubes $Q_1, \ldots, Q_N$ on which $T$ behaves like an affine map, i.e. for $x \in Q_i$, we have

$$
T(x) = T(x_i) + A_i(x) + o(|x - x_i|), \quad A_i(x) = T(x_i) + DT(x_i)(x - x_i).
$$

where $x_i$ is the center of the cube $Q_i$. Since the cubes overlap on sets of measure zero, which are preserved by $C^1$-diffeomorphisms, we can say

$$
\lambda_n(T(Q)) = \sum_{i=1}^{N} \lambda_n(T(Q_i)) \leq \sum_{i=1}^{N} \lambda_n(A_i(Q_i)) = \sum_{i=1}^{N} \lambda_n(Q_i) = \sum_{i=1}^{N} \int_{Q_i} |J_T(x_i)| \, d\lambda_n
$$

$$
\approx \sum_{i=1}^{N} \int_{Q_i} |J_T| \, d\lambda_n = \int_{Q} |J_T| \, d\lambda_n.
$$

Next we show that

$$
\int_{V} f \, d\lambda_n \leq \int_{U} (f \circ T) \cdot |J_T| \, d\lambda_n
$$

for all measurable $f \geq 0$ on $V$. This is true if $f$ is the characteristic function of $T(Q)$ for a cube $Q \subset U$. Since every open set in $\mathbb{R}^n$ can be decomposed into cubes with non-overlapping interiors (Whitney cube decomposition), the claim also holds for $f = \chi_W$ whenever $W \subset V$ is open. By outer regularity and a limiting argument, the claim then holds for $f = \chi_{T(M)}$ whenever $M \subset U$ is measurable. Then it holds for simple functions, and finally by monotone convergence it follows for arbitrary non-negative measurable functions. This claim implies the result with an inequality in one direction, with $g = (f \circ T) |J_T| \geq 0$. For the other direction, we apply the claim for $T^{-1}$ and $g = (f \circ T) |J_T|$. 

\[\square\]
Corollary 2.9.8. If $E \subset U$ is measurable, then $T(E) \subset V$ is measurable and

$$
\lambda_n(T(E)) = \int_E |J_T| \, d\lambda_n.
$$
3 SIGNED AND COMPLEX MEASURES

3.1 SIGNED MEASURES

Definition 3.1.1 (Signed measure). Let \((X, \mathcal{A})\) be a measurable space. A signed measure on \((X, \mathcal{A})\) is a function \(\mu : \mathcal{A} \to \mathbb{R}\) such that

(i) \(\mu(\emptyset) = 0\);

(ii) \(\mu\) takes at most one of the values \(+\infty\) or \(-\infty\);

(iii) \(\mu\) is countably additive, i.e. if \(A_n \in \mathcal{A}\) are pairwise disjoint, then

\[
\mu \left( \bigcup_{n=1}^{\infty} A_n \right) = \sum_{n=1}^{\infty} \mu(A_n);
\]

in particular, the sum always converges in \(\mathbb{R}\).

Example 3.1.2. 1. Every measure is a signed measure. For emphasis, we may refer to measures in the previous sense as positive measures.

2. Let \(\nu\) be a (positive) measure on \((X, \mathcal{A})\) and let \(f : X \to \mathbb{R}\) be measurable such that

\[
\int f^+ \, d\nu < \infty \quad \text{or} \quad \int f^- \, d\nu < \infty.
\]

Define

\[
\mu(A) = \int_A f \, d\nu = \int_A f^+ \, d\nu - \int_A f^- \, d\nu.
\]

Then \(\mu\) is a signed measure.

Definition 3.1.3 (Mutually singular measures). Let \((X, \mathcal{A})\) be a measurable space and \(\mu, \nu\) are positive measures. We say that \(\mu\) and \(\nu\) are mutually singular, written \(\mu \perp \nu\), if there exist \(E, F \in \mathcal{A}\) such that

(i) \(E\) and \(F\) partition \(X\);

(ii) \(\mu(E) = 0\) and \(\nu(F) = 0\).

Definition 3.1.4 (Totally positive set). Let \((X, \mathcal{A})\) be a measurable space and \(\mu\) be a signed measure. A set \(P \in \mathcal{A}\) is (totally) positive if \(\mu(A) \geq 0\) for all \(A \in \mathcal{A}\) with \(A \subset P\).

Lemma 3.1.5. Let \(\mu\) be a signed measure on \((X, \mathcal{A})\).

1. If \(A_n \nearrow A\) with \(A_n, A \in \mathcal{A}\), then \(\mu(A_n) \to \mu(A)\).

2. If \(A_n \searrow A\) with \(A_n, A \in \mathcal{A}\) and \(\mu(A_1) \in \mathbb{R}\), then \(\mu(A_n) \to \mu(A)\).

3. Measurable subsets of positive sets are positive and countable union of positive sets are positive.
4. If $\mu$ never takes the value $+\infty$ and $\mu(A) \neq -\infty$, then $A$ contains a positive subset $P$ with $\mu(P) \geq \mu(A)$.

Proof. 1. Omitted.

2. Omitted.

3. Omitted.

4. For each $\epsilon > 0$, we claim that there exists $A_\epsilon \in \mathcal{A}$ with $\mu(A_\epsilon) \geq \mu(A)$ such that if $B \in \mathcal{A}$ and $B \subset A_\epsilon$, then $\mu(B) \geq -\epsilon$. Suppose for the sake of contradiction that this is not the case, so then we can inductively find pairwise disjoint sets $B_k \in \mathcal{A}$ such that $\mu(B_k) < -\epsilon$. If $A_\epsilon = A$ satisfies the required properties, then we are done. Otherwise, we try $A_\epsilon = A \setminus B_1$. If that fails, we try $A_\epsilon = A \setminus (B_1 \cup B_2)$, and so on. Supposing we never find any $A_\epsilon$ with the desired properties, let $B = \bigcup B_k$. Then $\mu(B) = -\infty$ and $\mu(A) = -\infty$, since $\mu$ does not take the value $+\infty$ anywhere. This is a contradiction.

Now we use the claim with $\epsilon = 1/n$ for each $n$ to construct measurable set $A \supset A_1 \supset A_2 \supset \cdots$ such that $A_n \not\supset A_{n+1}$ and $B \subset A_n$ measurable implies $\mu(B) \geq -1/n$. If $P = \bigcap_n A_n \in \mathcal{A}$, then $\mu(P) \geq \mu(A)$ and $P$ is positive, as otherwise there exists $B \subset P$ measurable with $\mu(B) < 0$, hence $\mu(B) < -1/n$ for $n$ large, contradiction.

\[\square\]

**Theorem 3.1.6** (Hahn decomposition). Let $\mu$ be a signed measure on $(X, \mathcal{A})$. Then there exists a positive set $P$ and a negative set $N$ such that $P \cap N = \emptyset$ and $P \cup N = X$. Moreover, if $P', N'$ is another such pair, then $P \Delta P'$ and $N \Delta N'$ are total null.

Proof. Without loss of generality, suppose $\mu$ does not take $+\infty$. Let

$$s = \sup \{\mu(A) \mid A \in \mathcal{A}\} \in [0, +\infty].$$

By the lemma, there exists a sequence of positive sets $P_n$ such that $\mu(P_n) \to s$. Then $P = \bigcup_n P_n$ is positive and $\mu(P) = s < +\infty$. The set $N = X \setminus P \in \mathcal{A}$ is negative, since otherwise there exists $E \subset N$ with $\mu(E) > 0$, then $\mu(P \cup E) > s$, contradiction. This shows existence of the Hahn decomposition.

For uniqueness, $P' \setminus P \subset P' \cap N$ is positive and negative, hence total null. Similarly $P \setminus P'$ is total null, so $P \Delta P'$ is total null. The same proof works for $N \Delta N'$.

\[\square\]

**Theorem 3.1.7** (Jordan decomposition). Let $\mu$ be a signed measure on $(X, \mathcal{A})$. Then there exist unique positive measures $\mu^+$ and $\mu^-$ such that $\mu = \mu^+ - \mu^-$ and $\mu^+ \perp \mu^-$.\[\square\]

Proof. Let $X = P \cup N$ be the Hahn decomposition for $\mu$. Take $\mu^+(A) = \mu(A \cap P)$ and $\mu^-(A) = -\mu(A \cap N)$. For uniqueness, suppose $\mu = \nu^+ - \nu^-$ for positive measures $\nu^+ \perp \nu^-$. Let $E, F$ be the sets on which $\nu^+(E) \equiv 0$ and $\nu^-(E) \equiv 0$. For $A \in \mathcal{A}$ with $A \subset E$, we have

$$\mu(A) = \mu(A \cap E) + \mu(A \cap F) = \nu^+(A) - \nu^-(A) = \nu^+(A \cap E) - \nu^-(A \cap E).$$
This means \( E \) is positive for \( \mu \), and similarly \( F \) is negative for \( \mu \). Thus the splitting of \( X \) into \( E \) and \( F \) is a Hahn decomposition for \( \mu \). Uniqueness of Hahn decomposition tells us that \( E \triangle P \) and \( F \triangle N \) are totally null for \( \mu \). Hence for \( A \in \mathcal{A} \),

\[
\nu^+(A) = \nu^+(A \cap E) + \nu^+(A \cap F) = \mu(A \cap E) = \mu(A \cap P) = \mu^+(A),
\]

since \((A \cap E) \triangle (A \cap P) \subset E \triangle P\).

**Remark 3.1.10.** It is not generally true that \(|\mu|(A) = |\mu(A)|\).

### 3.2 THE RADON-NIKODYM THEOREM

**Definition 3.1.8 (Variations).** The \( \mu^+ \) and \( \mu^- \) are the positive and negative variations of \( \mu \).

**Definition 3.1.9 (Total variation).** The total variation \(|\mu|\) of \( \mu \) is the positive measure given by \(|\mu| = \mu^+ + \mu^-\).

**Example 3.2.2.** Let \((X, \mathcal{A})\) be a measurable space and \( \nu \) be a positive measure. Say \( f : X \to \mathbb{R} \) is \( \nu \)-integrable in the extended sense if \( f \) is measurable and one of \( \int f^+ \, d\nu < \infty \) or \( \int f^- \, d\nu < \infty \). Let \( \mu \) be the signed measure defined by \( \mu(A) = \int_A f \, d\nu \). For abbreviation, we write \( d\mu = f \, d\nu \).

Define \( P = \{f \geq 0\} \) and \( N = \{f < 0\} \). Then \( X = P \cup N \) and \( P \cap N = 0 \), so we have a Hahn decomposition for \( \mu \). The sets \( P' = \{f > 0\} \) and \( N' = \{f \leq 0\} \) also give a Hahn decomposition, so the set \( \{f = 0\} \) is total null. The functions \( \mu^+ \) and \( \mu^- \) are given by \( d\mu^+ = f^+ \, d\nu \) and \( d\mu^- = f^- \, d\nu \). Then \( d|\mu| = |f| \, d\nu \).

Let \((X, \mathcal{A})\) be a measurable space, \( \mu \) be a positive measure, and \( \nu \) be a signed measure. We say that \( \nu \) is absolutely continuous with respect to \( \mu \), written \( \nu \ll \mu \), if \( \mu(A) = 0 \) implies \( \nu(A) = 0 \) for each \( A \in \mathcal{A} \).

**Remark 3.2.3.**

1. If \( \nu \) is finite, then \( \nu \ll \mu \) if and only if for all \( \epsilon > 0 \), there exists \( \delta > 0 \) such that \( \mu(A) < \delta \) implies \( |\nu(A)| < \epsilon \) for all \( A \in \mathcal{A} \).

2. \( \nu \ll \mu \) if and only if \( \nu^+ \ll \mu \) and \( \nu^- \ll \mu \), or equivalently if and only if \( |\nu| \ll \mu \).

3. If \( \nu \) is a signed measure, \( \mu \) is a positive measure, \( \nu \ll \mu \), and \( \nu \perp \mu \), then \( \nu = 0 \).

**Lemma 3.2.4.** Let \( \nu, \mu \) be positive measures on \((X, \mathcal{A})\). Then either \( \nu \perp \mu \) or there exist \( \epsilon > 0 \) and \( E \in \mathcal{A} \) such that \( \mu(E) > 0 \) and \( \nu \geq \epsilon \mu \) on \( E \). (That is, \( \nu(A) \geq \epsilon \mu(A) \) for any \( A \subset E \) with \( A \in \mathcal{A} \)).

**Proof.** Let \( X = P_n \cup N_n \) be a Hahn decomposition of \( \nu - (1/n)\mu \). Define \( P = \bigcup_n P_n \) and \( N = X \setminus P = \bigcap_n P_n^c \). Then \( N \subset N_n \) and \( N_n \) is totally negative for \( \nu - (1/n)\mu \), so

\[
0 \geq \nu(N) - \frac{1}{n}\mu(N) \implies 0 \leq \nu(N) \leq \frac{1}{n}\mu(N) \to 0,
\]

so \( \nu(N) = 0 \). If \( \mu(P) = 0 \), then \( \nu \perp \mu \). Otherwise, \( \mu(P) > 0 \), so then \( \mu(P_n) > 0 \) for some \( n \). Since \( P_n \) is totally positive for \( \nu - (1/n)\mu \), we can take \( E = P_n \) and \( \epsilon = 1/n \). \( \square \)
Definition 3.2.5 (Lebesgue decomposition). Let $\mu$ be a $\sigma$-finite positive measure and $\nu$ be a $\sigma$-finite signed measure. A Lebesgue decomposition for $\nu$ is a pair of $\sigma$-finite signed measures $\lambda, \rho$ with $\nu = \lambda + \rho$ such that $\lambda \perp \mu$ and $\rho \ll \mu$.

Theorem 3.2.6 (Radon-Nikodym). Let $\mu$ be a $\sigma$-finite positive measure and $\nu$ be a $\sigma$-finite signed measure. Then $\nu$ has a unique Lebesgue decomposition $\nu = \lambda + \rho$. Moreover, there exists an extended $\mu$-integrable function $f : X \to \mathbb{R}$ such that $d\rho = fd\mu$ and $f$ is unique up to equality $\mu$-a.e.

Proof. Suppose $\mu$ and $\nu$ are finite positive measures. Let

$$\mathcal{F} = \{ f : X \to [0, \infty] \text{ measurable} \mid \int_E f \, d\mu \leq \nu(E) \text{ for all } E \in \mathcal{A} \}.$$ 

We will show that this family contains a “maximal” function, which will be the required function $f$.

First, we note that $\mathcal{F} \neq \emptyset$ since $0 \in \mathcal{F}$. If $f, g \in \mathcal{F}$, then $h = \max(f, g) \in \mathcal{F}$, as if we define

$$A = \{ x \in X \mid f(x) > g(x) \} \in \mathcal{A},$$

for $E \in \mathcal{A}$, we have

$$\int_E h \, d\mu = \int_{E \cap A} f \, d\mu + \int_{E \cap A^c} g \, d\mu \leq \nu(E \cap A) + \nu(E \cap A^c) = \nu(E).$$

Let $f_n \in \mathcal{F}$ and $f_n \not\uparrow f$. We claim that $f \in \mathcal{F}$, i.e. $\mathcal{F}$ is closed under monotone limits. This follows from the monotone convergence theorem (Theorem 2.2.3). Let

$$a = \sup \left\{ \int_X f \, d\mu \mid f \in \mathcal{F} \right\},$$

which is finite since each of these integrals is bounded by $\nu(X) < \infty$. Pick $f_n \in \mathcal{F}$ such that $\int f_n \, d\mu \to a$ as $n \to \infty$. Let $g_n = \max(f_1, \ldots, f_n) \in \mathcal{F}$. Then $\int f_n \, d\mu \leq \int g_n \, d\mu \leq a$ for all $n$ and $\int g_n \, d\mu \to a$ as $n \to \infty$. From the definition, $g_n \not\uparrow$, so we can define a pointwise limit $f(x) = \lim_{n \to \infty} g_n(x) \in [0, \infty]$ for $x \in X$. From the above, $f \in \mathcal{F}$. By the monotone convergence theorem, $a = \lim_{n \to \infty} \int g_n \, d\mu = \int f \, d\mu$, so $f$ takes the value $+\infty$ only on a set of measure zero. Redefine $f$ on this set to be zero wherever it was $+\infty$ before, so $f$ takes values in $[0, \infty)$.

Consider the finite measure $\rho$ defined as $d\rho = f \, d\mu$, meaning that

$$\rho(A) = \int_A f \, d\mu$$

for $A \in \mathcal{A}$. Since $f \in \mathcal{F}$, the difference $\lambda = \nu - \rho$ is a finite positive measure. By construction, $\rho \ll \mu$. We claim that $\lambda \perp \mu$. Otherwise, by the lemma, there exist $E \in \mathcal{A}$ and $\epsilon > 0$ such that $\mu(E) > 0$ and $\lambda \geq \epsilon \mu$ on $E$. Then

$$\nu(A) \geq \epsilon \mu(A) + \rho(A)$$

for all $A \in \mathcal{A}$ with $A \subset E$, so

$$\nu(A) \geq \epsilon \int_A \chi_E \, d\mu + \int_A f \, d\mu.$$
Hence \( f + \epsilon \chi_E \in \mathcal{F} \), but
\[
\int_X (f + \epsilon \chi_E) \, d\mu = a + \epsilon \mu(E) > a,
\]
contradicting the definition of \( a \). Thus we have shown the existence of the Lebesgue decomposition in the case that \( \mu \) and \( \nu \) are finite.

If \( \mu \) and \( \nu \) are \( \sigma \)-finite positive measures, then by suitable modification, we can decompose \( X \) into countably many pairwise disjoint measurable pieces, each of finite measure. This allows us to assemble a Lebesgue decomposition from the pieces.

Finally, in the general case where \( \nu \) need not be a positive measure, we write the Jordan decomposition \( \nu = \nu^+ - \nu^- \) and then obtain Lebesgue decompositions for the variations to be combined into a Lebesgue decomposition for \( \nu \).

For uniqueness, first suppose \( \mu \) and \( \nu \) are finite. Given two Lebesgue decompositions \( \nu = \lambda_1 + \rho_1 = \lambda_2 + \rho_2 \), one can show that these terms are all finite, so then \( \lambda = \lambda_2 - \lambda_1 = \rho_2 - \rho_1 \) is mutually singular with \( \mu \), but also \( \lambda \leq \rho_2 - \rho_1 \ll \mu \), which implies \( \lambda_1 = \lambda_2 \) and \( \rho_1 = \rho_2 \). Finally, to show essential uniqueness of \( f \), one can show that \( d(\rho_1 - \rho_2) = (f_1 - f_2) \, d\mu = 0 \) implies \( f_1 = f_2 = 0 \) almost everywhere. In the general case of \( \sigma \)-finite measures, use the partition idea again.

**Definition 3.2.7** (Radon-Nikodym derivative). Let \( \mu, \nu \) be as above and suppose \( \nu \ll \mu \). Then \( \lambda = 0 \) and \( \rho = \nu \) is a Lebesgue decomposition, so \( d\nu = d\rho = f \, d\mu \) for some \( f \) which is unique up to modification on a set of measure zero. The function \( f \) is the Radon-Nikodym derivative of \( \nu \) with respect to \( \mu \).

### 3.3 COMPLEX MEASURES

**Definition 3.3.1** (Complex measure). Let \((X, \mathcal{A})\) be a measurable space. A complex measure \( \nu \) is a function \( \nu : \mathcal{A} \to \mathbb{C} \) such that

(i) \( \nu(\emptyset) = 0 \);

(ii) if \( A_n \in \mathcal{A} \) are pairwise disjoint and \( A = \bigcup_n A_n \), then
\[
\nu(A) = \sum_{n=1}^{\infty} \nu(A_n).
\]

If \( \nu \) is a complex measure, then \( \nu = \nu_r + i \nu_i \) for signed measure \( \nu_r, \nu_i \) defined by
\[
\nu_r(A) = \text{Re} \nu(A), \quad \nu_i = \text{Im} \nu(A).
\]

Since \( \nu \) only takes finite values, both \( \nu_r \) and \( \nu_i \) take values in \( \mathbb{R} \), hence are finite signed measures.

**Theorem 3.3.2** (Radon-Nikodym for complex measures). Let \((X, \mathcal{A})\) be a measurable space, \( \mu \) be a \( \sigma \)-finite positive measure, and \( \nu \) be a complex measure. Then \( \nu = \lambda + \rho \), where \( \lambda \perp \mu \) (i.e. \( \lambda_r, \lambda_i \perp \mu \)) and \( \rho \ll \mu \) (i.e. \( \rho_r, \rho_i \ll \mu \)). Moreover, there exists \( f \in L^1(\mu) \) such that \( d\rho = f \, d\mu \). The measures \( \lambda, \rho \) are unique, and \( f \) is unique \( \mu \)-a.e.

**Proof.** Apply Radon-Nikodym to the finite signed measure \( \nu_r \) and \( \nu_i \). \( \square \)
Let $\nu$ be a complex measure on $(X, \mathcal{A})$. Then we define a set function $|\nu| : \mathcal{A} \to [0, \infty]$ given by

$$
|\nu|(A) = \sup \left\{ \sum_{n=1}^{\infty} |\nu(A_n)| \mid A_n \in \mathcal{A} \text{ pairwise disjoint and } \bigcup_{n=1}^{\infty} A_n = A \right\}.
$$

For signed measures, this agrees with our definition of the total variation (Homework 1B Problem 2). However, for complex numbers, $|\nu| \neq |\nu_r| + |\nu_i|$ in general.

**Proposition 3.3.3.** If $\nu$ is a complex measure on $(X, \mathcal{A})$, then $|\nu|$ is a finite positive measure.

**Proof.** We can find a positive measure $\mu$ on $(X, \mathcal{A})$ such that $\nu \ll \mu$ (take $\mu = |\nu_r| + |\nu_i|$ for example). By Radon-Nikodym, there exists $f \in L^1(\mu)$ such that $d\nu = fd\mu$, i.e.,

$$
\nu(A) = \int_A f \, d\mu.
$$

We claim that

$$
|\nu|(A) = \int_A |f| \, d\mu
$$

for $A \in \mathcal{A}$, i.e. $d|\nu| = |f| \, d\mu$, from which it follows that $|\nu|$ is a finite positive measure.

First, let $A \in \mathcal{A}$ and $A = \bigcup_n A_n$ be a measurable partition, i.e. a partition into measurable sets. Then

$$
\sum_{n=1}^{\infty} |\nu(A_n)| = \sum_{n=1}^{\infty} \left| \int_{A_n} f \, d\mu \right|
\leq \sum_{n=1}^{\infty} \int_{A_n} |f| \, d\mu
= \int \left( \sum_{n=1}^{\infty} |f| \chi_{A_n} \right) \, d\mu \quad \text{(monotone convergence)}
= \int_A |f| \, d\mu.
$$

To obtain the reverse inequality, let $A \in \mathcal{A}$ and $\epsilon > 0$. Since simple functions are in $L^1(\mu)$, there exists $s = \sum_k c_k \chi_{B_k}$ in standard representation such that

$$
\int |f - s| \, d\mu < \epsilon.
$$

Let $B = \bigcup_k B_k$ and $A_0 = A \cap B^c$, $A_i = A \cap B_i$ for $i = 1, \ldots, N$. We obtain a measurable partition...
of $A$ (let $A_i = \emptyset$ for $i > N$), so

\[
\sum_{i=0}^{N} |\nu(A_i)| = \sum_{i=0}^{N} \left| \int_{A_i} (f - s) \, d\mu + \int_{A_i} s \, d\mu \right|
\geq \sum_{i=0}^{N} \left| \int_{A_i} s \, d\mu \right| - \int |f - s| \, d\mu
\geq \sum_{i=0}^{N} \int_{A_i} |s| \, d\mu - \epsilon \quad \text{(s constant on each $A_i$)}
\geq \int_{A} |f| \, d\mu - 2\epsilon.
\]
4 FUNCTIONAL ANALYSIS

4.1 REVIEW OF BASIC FUNCTIONAL ANALYSIS

Throughout, let $F = \mathbb{R}$ or $F = \mathbb{C}$.

Let $X$ and $Y$ be normed vector spaces over $F$ and let $T : X \to Y$ be a linear map. Then $T$ is bounded if there exists a constant $C \geq 0$ such that $\|T(x)\| \leq C\|x\|$ for all $x \in X$.

**Theorem 4.1.1.** Let $X$ and $Y$ be normed vector spaces and $T : X \to Y$ be a linear map. Then the following are equivalent:

1. $T$ is continuous;
2. $T$ is continuous at 0;
3. $T$ is bounded.

**Proof.**

1 $\implies$ 2. Obvious.

2 $\implies$ 3. Since $T$ is linear, $T(0) = 0$, so if $T$ is continuous at 0, then there exists $\delta > 0$ such that $\|x\| \leq \delta$ implies $\|T(x)\| \leq 1$. If $u \in X$ is an arbitrary non-zero vector, then define $x = (\delta/\|u\|)u$; we have $\|x\| = \delta$, so $\|T(x)\| \leq 1 \implies \|T(u)\| \leq (1/\delta)\|u\|$.

3 $\implies$ 1. If $T$ is bounded, then there exists $C \geq 0$ such that $\|T(x)\| \leq C\|x\|$ for all $x \in X$. By linearity,

$$\|T(x) - T(y)\| = \|T(x - y)\| \leq C\|x - y\|,$$

so $T$ is Lipschitz, hence continuous.

Example 4.1.2. On the space of smooth functions on $[0,1]$ with the supremum norm, the differentiation operator is unbounded.

**Definition 4.1.3 (Operator norm).** Let $T : X \to Y$ be a bounded operator. The operator norm of $T$ is

$$\|T\| = \inf \{C \mid \|T(x)\| \leq C\|x\| \text{ for all } x \in X\}.$$

**Proposition 4.1.4.**

$$\|T\| = \sup \left\{ \frac{\|T(x)\|}{\|x\|} \mid x \in X \setminus \{0\} \right\} = \sup \{\|T(x)\| \mid x \in X \text{ and } \|x\| \leq 1\}.$$

**Proposition 4.1.5.** For vector spaces $X$ and $Y$ over $F$, the set $L(X,Y)$ of bounded linear operators $X \to Y$ is a vector space with pointwise addition and scalar multiplication. The operator norm is a norm on $L(X,Y)$.

**Proposition 4.1.6.** If $T \in L(X,Y)$ and $S \in L(Y,Z)$, then $S \circ T \in L(X,Z)$ and $\|S \circ T\| \leq \|S\| \cdot \|T\|$.
Proposition 4.1.7. Let $X, Y$ be normed vector spaces over $F$. If $Y$ is complete, then $L(X,Y)$ is complete.

Proof. Suppose $Y$ is complete and let $(T_n)$ be a Cauchy sequence in $L(X,Y)$, i.e. for all $\epsilon > 0$, there exists $N$ such that for all $n, m \geq N$, $\|T_n - T_m\| < \epsilon$. Then for each $x \in X$, it is clear that $(T_n(x))$ is a Cauchy sequence in $Y$, so it has a limit. Hence define a function $T : X \rightarrow Y$ by

$$T(x) = \lim_{n \rightarrow \infty} T_n(x).$$

4.2 BOUNDED LINEAR FUNCTIONALS AND DUAL SPACES

Let $X$ be a normed vector space over $F$. A bounded linear map $f : X \rightarrow F$ is a bounded linear functional, i.e. $f \in L(X,F)$. The space $L(X,F)$ with the operator norm is the dual space of $X$, denoted $X^*$. By construction, it is always a Banach space.

Theorem 4.2.1 (Hahn-Banach). Let $X$ be a normed vector space over $F$ and let $M \subset X$ be a linear subspace. Let $f : M \rightarrow F$ be a bounded linear functional. Then there exists a bounded linear functional $F : X \rightarrow F$ such that $F|_M = f$ and $\|F\| = \|f\|$.

Proof. First suppose $F = \mathbb{R}$. If $\|f\| = 0$, then $f = 0$ on $M$, so we can take $F = 0$ on $X$. Otherwise, without loss of generality, $\|f\| = 1$. We consider the set $Z$ of pairs $(U, g)$, where $U \subset X$ is a subspace with $M \subset U$ and $g : U \rightarrow F$ with $g|_M = f$ and $\|g\| = \|f\| = 1$. Then $Z$ can be partially ordered by saying that $(U, g) \leq (\tilde{U}, \tilde{g})$ if $U \subset \tilde{U}$ and $\tilde{g}|_U = g$. Every chain in $Z$ has an upper bound in the obvious way. Thus by Zorn’s lemma, there is a maximal element of $Z$, say $(V, F)$. We claim that $V = X$. Suppose otherwise, and let $x_0 \in X \setminus V$. Take $W$ to be the span of $V$ and $x_0$. We extend $F$ to some $G : W \rightarrow F$ by taking $G(x_0) = \alpha$ for some suitable $\alpha$ and extending linearly. This contradicts maximality of $(V, F)$, so we are done.

If $F = \mathbb{C}$, then we can also regard $X$ as a real normed space. Consider $\mathbb{R}$-linear functionals $X \rightarrow \mathbb{R}$, i.e. real-valued functionals $F$ with $F(\alpha x) = \alpha F(x)$ only required for $\alpha \in \mathbb{R}$. Given a (C-linear) functional $F : X \rightarrow \mathbb{C}$, it is immediate that $u = \text{Re } F$ is an $\mathbb{R}$-linear functional. Conversely, if we know $u : X \rightarrow \mathbb{R}$, the corresponding $F$ is

$$F(x) = u(x) - iu(ix).$$

Looking at operator norms,

$$\|u(x)\| = |\text{Re } F(x)| \leq |F(x)| \leq \|F\| \|x\|$$

for all $x$, so $\|u\| \leq \|F\|$. Conversely, for any $x \in X$, pick $\alpha \in \mathbb{C}$ with $|\alpha| = 1$ so that $\alpha F(x) \in \mathbb{R}$. Then

$$|F(x)| = |\alpha F(x)| = |F(\alpha x)| = |u(\alpha x)| \leq \|u\| |\alpha x| = \|u\| \|x\|.$$

We have thus showed that $\|u\| = \|F\|$.

In the context of Hahn-Banach, let $f : M \rightarrow \mathbb{C}$ be a $\mathbb{C}$-linear functional and $u = \text{Re } f : M \rightarrow \mathbb{R}$. Then $\|f\| = \|u\|$, and by Hahn-Banach, we can extend $u$ to an $\mathbb{R}$-linear function $u : X \rightarrow \mathbb{R}$ with $\|u\| = \|u\| = \|f\|$. Define $F : X \rightarrow \mathbb{C}$ from $u$ as above. Then $F$ is a $\mathbb{C}$-linear functional with $\|F\| = \|u\| = \|f\|$ and $F$ extends $f$. □
Corollary 4.2.2. For each $x \in X$,
\[ \|x\| = \sup\{|f(x)| \mid f \in X^* \text{ with } \|f\| \leq 1 \}, \]
and this supremum is attained as a maximum.

Proof. First let $f \in X^*$ with $\|f\| \leq 1$. Then $|f(x)| \leq \|f\||x| \leq \|x\|$, so
\[ \|x\| \geq \sup\{|f(x)| \mid f \in X^* \text{ with } \|f\| \leq 1 \}. \]

For the inequality in the other direction, without loss of generality, suppose $x \neq 0$. Let $M$ be the span of $x$, which defines a subspace of $X$, and consider the bounded linear functional $g : M \to \mathbb{F}$ given by $g(\lambda x) = \lambda \|x\|$. This has $\|g\| = 1$, and by Hahn-Banach, we can extend this to a functional $f \in X^*$ with $\|f\| = 1$ extending $g$. Then $|f(x)| = |g(x)| = \|x\|$. \qed

Remark 4.2.3. In general, $\|f\| = \sup\{|f(x)| \mid x \in X \text{ with } \|x\| \leq 1 \}$ by definition, but the supremum is not generally attained as a maximum if $X$ is infinite-dimensional.

We can define a bilinear form $X \times X^* \to \mathbb{F}$ by $\langle x, f \rangle = f(x)$. This gives us an injection $X \to X^{**}$ defined by $x \mapsto \langle x, - \rangle$. If it is in fact an isomorphism, then we say that $X$ is reflexive. Since the dual of any normed space is a Banach space, all reflexive spaces are Banach.

Example 4.2.4. Every finite-dimensional normed vector space is reflexive.

4.3 THE DUAL OF $L^p$

Let $(X, \mathcal{A}, \mu)$ be a measure space and $1 \leq p \leq \infty$. Let $q$ be the conjugate exponent of $p$. If $g \in L^q(\mu)$, then we can define
\[ \Phi_g : L^p(\mu) \to \mathbb{C}, \quad \Phi_g(f) = \int fg \, d\mu. \]

By Hölder’s inequality, it follows that $\Phi_g$ is a bounded linear functional.

Theorem 4.3.1. Let $1 \leq p < \infty$ and $q$ be the conjugate exponent of $p$, and suppose that $(X, \mathcal{A}, \mu)$ is $\sigma$-finite. Then for each bounded linear functional $\Phi : L^p(\mu) \to \mathbb{C}$, there exists $g \in L^q(\mu)$ such that $\Phi = \Phi_g$. Moreover, $g$ is unique up to a set of measure zero and $\|\Phi\| = \|g\|_q$.

Proof (outline). Let $1 \leq p < \infty$, so $1 < q \leq \infty$.

For existence, let $\Phi \in (L^p)^*$ be arbitrary. If $\mu$ is a finite measure, then $\chi_A \in L^p$ for all $A \in \mathcal{A}$, so we can define a complex measure $\nu(A) = \Phi(\chi_A)$. Indeed, $\nu(\emptyset) = 0$, and to see that $\nu$ is countable additive, let $A_n \in \mathcal{A}$ be pairwise disjoint with $A = \bigcup_n A_n$ and $B_n = A_1 \cup \cdots \cup A_n$. We use continuity of $\Phi$ to get
\[ \nu(A) = \Phi(\chi_A) = \lim_{n \to \infty} \Phi(\chi_{B_n}) \]
\[ = \lim_{n \to \infty} \Phi(\chi_{A_1} + \cdots + \chi_{A_n}) \]
\[ = \lim_{n \to \infty} \nu(A_1) + \cdots + \nu(A_n) \]
\[ = \sum_{n=1}^{\infty} \nu(A_n). \]
We have \( \nu \ll \mu \), as if \( \mu(A) = 0 \) for some \( A \in \mathcal{A} \), then \( \chi_A = 0 \) almost everywhere, i.e. \( \chi_A = 0 \) in \( L^p \), so \( \nu(A) = \Phi(\chi_A) = 0 \). By Radon-Nikodym, there exists \( g \in L^1 \) such that \( \Phi(\chi_A) = \nu(A) = \int_A g \, d\mu = \int \chi_A g \, d\mu \) for all \( A \in \mathcal{A} \). By linearity, this extends to all simple functions, i.e.

\[
\Phi(s) = \int sg \, d\mu
\]

for all simple functions \( s \in L^p \). Now let \( h \in L^\infty \subset L^p \). Since simple functions are dense in \( L^\infty \), there is a sequence \( s_n \to h \) in \( L^\infty \). Hence

\[
\|s_n - h\|_p^p = \int |s_n - h|^p \, d\mu \leq \int \|s_n - h\|_\infty^p \, d\mu = \mu(X) \|s_n - h\|_\infty^p \to 0,
\]

so \( s_n \to h \) in \( L^p \). Hence for \( h \in L^\infty \).

\[
\Phi(n) = \lim_{n \to \infty} \Phi(s_n) = \lim_{n \to \infty} \int s_n g \, d\mu = \int hg \, d\mu
\]

where to see that the last step is valid, we have

\[
\lim \sup \left| \int s_n g \, d\mu - \int hg \, d\mu \right| \leq \lim \sup \left| \int s_n - h \right| |g| \, d\mu
\]

\[
= \lim \sup \left| \int s_n - h \right| \, \|g\| \, d\mu = 0.
\]

We now claim that \( g \in L^q \). If \( 1 < p < \infty \), then \( 1 < q < \infty \). There exists a measurable function \( \alpha \) on \( X \) such that \( |\alpha| = 1 \) and \( g = \alpha|g| \) (exercise). Let \( E_n = \{|g| \leq n\} \in \mathcal{A} \) and \( f_n = \alpha|g|^{p-1} \chi_{E_n} \in L^\infty \subset L^p \). We can now apply \( \Phi \) to get

\[
\int_{E_n} |g|^q \, d\mu = \left| \int_{E_n} \alpha|g|^{q-1} g \, d\mu \right|
\]

\[
= \left| \int f_n g \, d\mu \right|
\]

\[
= |\Phi(f_n)| \leq \|\Phi\| \|f_n\|_p
\]

\[
= \|\Phi\| \left( \int_{E_n} |g|^{(q-1)p} \right)^{1/p}
\]

\[
= \|\Phi\| \left( \int_{E_n} |g|^q \right)^{1/p},
\]

so

\[
\|g\chi_{E_n}\|_q \leq \|\Phi\|.
\]

Taking the limit as \( n \to \infty \), we get \( \|g\|_q \leq \|\Phi\| < \infty \). Since \( \Phi_g = \Phi \) for simple functions, which are dense in \( L^p \), continuity gives us \( \Phi_g = \Phi \) everywhere. For the norm we have \( \|g\|_q \leq \|\Phi\| = \|\Phi_g\| \leq \|g\|_q \). In the case \( p = 1 \) and \( q = \infty \), one can show that \( \|g\|_\infty \leq \|\Phi\| \), and then the argument proceeds in the same way. If \( \mu \) is \( \sigma \)-finite, then there exist measurable sets \( E_n \to \chi \) of finite measure. Use a limiting argument and monotone convergence.

For uniqueness, suppose that \( \Phi_g = \Phi_{\tilde{g}} \) for \( g, \tilde{g} \in L^q \). Then \( \int fg \, d\mu = \int f\tilde{g} \, d\mu \) for all \( f \in L^p \), so equivalently, \( \int f \cdot (g - \tilde{g}) \, d\mu \) for all \( f \in L^p \). Then \( 0 = \int \chi_E (g - \tilde{g}) \, d\mu \) whenever \( E \in \mathcal{A} \) and \( \mu(E) < \infty \), so \( g - \tilde{g} = 0 \) almost everywhere.
Remark 4.3.2. 1. The map \( T : L^q(\mu) \to L^p(\mu)^* \) given by \( T(g) = \Phi_g \in L^p(\mu)^* \) is a well-defined surjective linear isometry which identifies \( L^p(\mu)^* \) with \( L^q(\mu) \) for \( 1 \leq p < \infty \).

2. On the other hand, \( L^\infty(\mu)^* \) need not be isomorphic with \( L^1(\mu) \). For example, \((l^\infty)^* \neq l^1\), where \( l^p \) is the \( L^p \) norm on \( \mathbb{N} \) with the counting measure.

Theorem 4.3.3 (Basic 5B-covering lemma). Let \((X,d)\) be a metric space and let \( \mathcal{B} \) be a collection of open balls in \( X \) with uniformly bounded radii. Then there exists a collection \( \tilde{\mathcal{B}} \) of pairwise disjoint balls with
\[
\bigcup_{B \in \mathcal{B}} B \subset \bigcup_{B \in \tilde{\mathcal{B}}} 5B.
\]

Proof. See Homework B4 Problem 3.

4.4 THE HARDY-LITTLEWOOD MAXIMAL FUNCTIONS

Let \( L^1_{\text{loc}} \) denote the space of locally integrable functions, i.e. \( f : \mathbb{R}^n \to \mathbb{C} \) is measurable and \( \int_B |f| < \infty \) for all balls \( B \subset \mathbb{R}^n \), or equivalently, \( \int_K |f| < \infty \) for all \( K \subset \mathbb{R}^n \) compact.

If \( f \in L^1_{\text{loc}} \), we define the centered Hardy-Littlewood maximal function by
\[
(Hf)(x) = \sup_{r > 0} \int_{B(x,r)} |f| \, d\lambda_n, \quad \int_B |f| = \frac{1}{\lambda_n(B)} \int_B |f|.
\]

There is also an uncentered Hardy-Littlewood maximal function given by
\[
(Mf)(x) = \sup_{B \ni x} \int_B |f|,
\]
the supremum ranging over open balls.

Lemma 4.4.1. 1. If \( f \in L^1_{\text{loc}} \), then \( Hf : \mathbb{R}^n \to [0, \infty] \) and \( Mf : \mathbb{R}^n \to [0, \infty] \) are measurable and \( Hf \leq Mf \).

2. The operators \( M \) and \( H \) are sublinear, i.e. \( M(f + g) \leq M(f) + M(g) \) and \( M(\alpha f) = |\alpha| M(f) \), and similarly for \( H \).

Proof. 1. It is clear that \( Hf \leq Mf \) from definition. Measurability of \( Hf \) is in Folland, so only measurability of \( Mf \) is left. It suffices to show that \( \{ x \in \mathbb{R}^n \mid (Mf)(x) > \alpha \} \) is open for each \( \alpha \in \mathbb{R} \). If \( x \in \{ Mf > \alpha \} \), then there exists an open ball \( B \) with \( x \in B \) and \( \int_B |f| > \alpha \). Then \( (Mf)(y) > \alpha \) for all \( y \in B \), and so \( B \subset \{ Mf > \alpha \} \).

Remark 4.4.2. 1. For \( f \in L^1_{\text{loc}} \), we have \(|f| \leq Hf \leq Mf \), but \( Mf \) is typically not much larger than \(|f|\). For \( f \in L^p \) with \( 1 < p \leq \infty \), we get
\[
\|Mf\|_p \leq C(n,p)\|f\|.
\]
2. For \( f \in L^p \) with \( 1 \leq p \leq \infty \), we get
\[
\lambda_n(|f| > \alpha) \leq \int_{|f| > \alpha} \frac{|f|^p}{\alpha^p} \leq \frac{C}{\alpha^p}
\]
for \( \alpha > 0 \) and \( C \) a constant independent of \( \alpha \).

3. In general, \( Mf \notin L^1 \) if \( f \in L^1 \), but \( Mf \) still satisfies a “weak-type \( L^1 \) estimate”.

**Theorem 4.4.3** (Weak-type \( L^1 \)-estimate for the Hardy-Littlewood maximal function). There exists a constant \( C = C(n) = 5^n \) such that
\[
\lambda_n\{Mf > \alpha\} \leq \frac{C}{\alpha} \|f\|
\]
for \( \alpha > 0 \) and \( f \in L^1(\mathbb{R}^n) \).

**Proof.** Let \( f \in L^1 \) and \( \alpha > 0 \). Let \( A = \{ x \in \mathbb{R}^n \mid (Mf)(x) > \alpha \} \) and pick \( x \in A \). Then we can find a ball \( B_x \subset \mathbb{R}^n \) with \( x \in B_x \) and \( (Mf)(x) \geq \int_{B_x} |f| \, d\lambda_n > \alpha \).

Let \( B = \{ B_x \mid x \in A \} \) be the collection of these balls. If \( B = B(a, r) \in \mathcal{B} \), then \( \alpha < \int_B |f| \) means that
\[
\lambda_n(B) = c_n r^n \frac{1}{\alpha} \int_B |f| \leq \frac{1}{\alpha} \|f\|_1 < \infty,
\]
so the balls of \( \mathcal{B} \) have uniformly bounded radii. Applying the 5B-covering lemma, there is a subfamily \( \tilde{\mathcal{B}} \subset \mathcal{B} \) of pairwise disjoint balls with
\[
A \subset \bigcup_{B \in \tilde{\mathcal{B}}} B \subset \bigcup_{B \in \tilde{\mathcal{B}}} 5B.
\]

Note that \( \tilde{\mathcal{B}} \) is countable, say \( \tilde{\mathcal{B}} = \{ B_i \mid i \in I \subset \mathbb{N} \} \). Then
\[
\begin{align*}
\lambda_n(A) & \leq \lambda_n\left( \bigcup_{i \in I} 5B_i \right) \\
& \leq \sum_{i \in I} \lambda_n(5B_i) \\
& = 5^n \sum_{i \in I} \lambda_n(B_i) \\
& \leq \frac{5^n}{\alpha} \sum_{i \in I} \int_{B_i} |f| \\
& = \frac{5^n}{\alpha} \int_{\bigcup B_i} |f| \\
& \leq \frac{5^n}{\alpha} \|f\|_1.
\end{align*}
\]
Lemma 4.4.4. Let \( g : \mathbb{R}^n \to \mathbb{C} \) and \( 1 \leq p \leq \infty \). Then
\[
\int |g|^p \, d\lambda_n = p \int_0^\infty \alpha^{p-1} \lambda_n \{ |g| > \alpha \} \, d\alpha.
\]

Theorem 4.4.5 (\( L^p \)-boundedness of the maximal function). Let \( n \in \mathbb{N} \) and \( 1 < p \leq \infty \). Then there exists a constant \( C = C(p, n) = \frac{2}{(p \cdot 5^n)^{1/p}} \) if \( p < \infty \) (for \( p = \infty \), we take \( C = 1 \)) such that \( \|Mf\|_p \leq C\|f\|_p \) for all \( f \in L^p(\mathbb{R}^n) \).

Proof. The \( L^\infty \) case is trivial.

The \( L^p \)-boundedness of \( M \) is derived from an interpolation technique (see Marcinkiewicz interpolation theorem). Our basic idea is to split the function \( f \in L^p \) into a “small part” \( g \) and a “large part” \( h \), then use different estimates for \( g \) and \( h \).

We have
\[
\|Mf\|_p^p = \int |Mf|^p = p \int_0^\infty \alpha^{p-1} \lambda_n \{ |g| > \alpha \} \, d\alpha.
\]

To show that \( \|Mf\|_p < \infty \), we need good estimates for \( \lambda(\alpha) = \lambda_n \{ Mf > \alpha \} \). Fix \( \alpha > 0 \). Define
\[
g = f \circ \chi_{\{|f| \leq \alpha/2\}}, \quad h = f \cdot \chi_{\{|f| > \alpha/2\}}.
\]

Then \( f = g + h \) and \( \|g\|_\infty \leq \alpha/2 \). We have
\[
\|h\|_1 = \int_{\{|f| > \alpha/2\}} |f| \leq \int_{\{|f| > \alpha/2\}} \frac{|f|^{p-1}}{(\alpha^p/2)^{p-1}} \leq \left(\frac{2}{\alpha}\right)^{p-1} \int |f|^p < \infty,
\]

so \( h \in L^1 \). Hence
\[
Mf \leq Mg + Mh \leq \frac{\alpha}{2} + Mh.
\]

This implies that
\[
\lambda(\alpha) = \lambda_n \{ Mf > \alpha \} \leq \lambda_n \{ Mh > \alpha/2 \}
\]
\[
\leq \frac{5^n}{\alpha/2} \|h\|_1
\]
\[
= \frac{2 \cdot 5^n}{\alpha} \int_{\{|f| > \alpha/2\}} |f|.
\]
Substituting,

\[
\|Mf\|_p^p \leq p \int_0^\infty \frac{\alpha^{p-1}}{\alpha} \left( \int_{\{|f| > \alpha/2\}} |f| \right) d\alpha 
\]

\[
= 2p \cdot 5^n \left( \int_0^\infty \alpha^{p-2} \chi_{\{|f(x)| > \alpha/2\}} |f| d\alpha \right) d\lambda_n(x) 
\]

\[
= 2p \cdot 5^n \left( \int_0^\infty \alpha^{p-2} \chi_{\{|f(x)| > \alpha/2\}} d\alpha \right) d\lambda_n(x) 
\]

\[
= 2p \cdot 5^n \left( \int_0^\infty \alpha^{p-2} - \alpha^p d\alpha \right) d\lambda_n(x) 
\]

\[
= \frac{2p \cdot 5^n}{p-1} \int_{\mathbb{R}^n} |f(x)| \cdot 2^{p-1} |f(x)|^{p-1} d\lambda_n(x) 
\]

\[
= \frac{2p \cdot 5^n}{p-1} \int_{\mathbb{R}^n} |f|^p = \frac{2p \cdot 5^n}{p-1} \|f\|_p^p. 
\]

5 DIFFERENTIATION

5.1 LEBESGUE POINTS

Definition 5.1.1 (Lebesgue point). Let \( f : \mathbb{R}^n \to \mathbb{C} \) be measurable. We say that \( x \in \mathbb{R}^n \) is a **Lebesgue point** of \( f \) if

\[
\lim_{r \to 0^+} \int_{B(x,r)} |f(y) - f(x)| d\lambda_n(y) = 0.
\]

Remark 5.1.2. If \( f \) is continuous, then every point is a Lebesgue point of \( f \).

Example 5.1.3. Let \( f = \chi_{[0,1]} \). Then the Lebesgue points of \( f \) are the points in \( \mathbb{R} \setminus \{0,1\} \).

Theorem 5.1.4 (Lebesgue differentiation theorem). Let \( f \in L^1_{\text{loc}}(\mathbb{R}^n) \). Then almost every point is a Lebesgue point of \( f \).

Proof. Without loss of generality, \( f \in L^1(\mathbb{R}^n) \), as the Lebesgue points of \( f \) and \( f \cdot \chi_{B(0,n)} \in L^1_{\text{loc}}(\mathbb{R}^n) \) are the same on \( B(0,n) \) for each \( n \).

For \( r > 0 \), define

\[
(T_rf)(x) = \int_{B(x,r)} |f - f(x)| d\lambda_n
\]

and

\[
Tf(x) = \limsup_{r \to 0^+} T_rf(x) \in [0,\infty].
\]

We must show that \( Tf(x) = 0 \) for almost every \( x \in \mathbb{R}^n \).
Let \( k \in \mathbb{N} \) be arbitrary. Since \( C_c(\mathbb{R}^n) \) is dense in \( L^1(\mathbb{R}^n) \), we can find \( g_k \in C_c(\mathbb{R}^n) \) such that \( \| f - g_k \|_1 < 1/k \). Since \( g_k \) is continuous, \( Tg \equiv 0 \). Let \( h_k = f - g_k \). Then \( f = g_k + h_k \) with \( \| h_k \|_1 < 1/k \). We have \( T_rf \leq T_rg_k + T_rh_k \), and so \( Tf \leq Th_k \). This gives

\[
T_rh_k(x) = \int_{B(x,r)} |h_k - h_k(x)| \, d\lambda_n
\leq \int_{B(x,r)} |h_k| \, d\lambda_n + |h_k(x)|
\leq (Mh_k)(x) + |h_k|(x).
\]

Letting \( r \to 0^+ \), we get

\[
Th_k(x) \leq (Mh_k)(x) + |h_k|(x)
\]

for all \( x \in \mathbb{R}^n \). Let \( \alpha > 0 \) be arbitrary and

\[
A(k, \alpha) = \{ Mh_k \geq \alpha/2 \} \cup \{ |h_k| > \alpha/2 \},
\]

so then \( \{ Tf < \alpha \} \subset A(k, \alpha) \) for all \( k \). By our weak-type \( L^1 \)-estimate for the maximal function and \( |h_k| \), we have

\[
\lambda_n(A(k, \alpha)) \leq \frac{5^n}{\alpha^2} \| h_k \|_1 + \frac{1}{\alpha^2} \| h_k \|_1 = \frac{1}{\alpha} \left( \frac{2(5^n + 1)}{\alpha} \right) \to 0
\]

as \( k \to \infty \). Hence

\[
\lambda_n \left( \bigcap_{k=1}^{\infty} A(k, \alpha) \right) = 0,
\]

so we conclude that \( \{ Tf > \alpha \} \) is Lebesgue measurable with measure zero. Since this is true for all \( \alpha = 1/L \) with \( L \in \mathbb{N} \), we conclude that \( \{ Tf > 0 \} \) has measure 0.

**Corollary 5.1.5.** If \( f \in L^1_{\text{loc}}(\mathbb{R}^n) \), then

\[
\lim_{r \to 0^+} \int_{B(x,r)} f = f(x)
\]

for almost every \( x \).

**Proof.** If \( x \in \mathbb{R}^n \) is a Lebesgue point, then

\[
\limsup_{r \to 0^+} \int_{B(x,r)} \left| f - f(x) \right| \leq \limsup_{r \to 0^+} \int_{B(x,r)} |f - f(x)| = 0.
\]

A family \( \{ E_r \}_{r>0} \) of measurable sets in \( \mathbb{R}^n \) *shrinks nicely* to \( x \in \mathbb{R}^n \) if

(i) \( E_r \subset B(x,r) \) for \( r > 0 \);

(ii) there exists \( \alpha > 0 \) such that \( \lambda_n(E_r) \geq \alpha \lambda_n(B(x,r)) \).
Remark 5.1.6. The sets $E_r$ need not contain $x$.

**Theorem 5.1.7.** Let $f \in L^1_{\text{loc}}(\mathbb{R}^n)$ and let $x \in \mathbb{R}^n$ be a Lebesgue point of $f$. Suppose $\{E_r\}_{r>0}$ is a family of sets shrinking nicely to $x$. Then

$$\lim_{r \to 0^+} \int_{E_r} |f(y) - f(x)| \, d\lambda_n(y).$$

**Proof.**

$$\int_{E_r} |f(y) - f(x)| \, d\lambda_n(y) = \frac{1}{\lambda_n(E_r)} \int_{E_r} |f - f(x)|$$

$$\leq \frac{1}{\lambda_n(E_r)} \int_{B(x,r)} |f - f(x)|$$

$$\leq \frac{1}{\alpha} \int_{B(x,r)} |f - f(x)| \to 0.$$ 

Let $E \subset \mathbb{R}^n$ be measurable. The (metric) density of $E$ at $x \in \mathbb{R}^n$ is

$$D_E(x) = \lim_{r \to 0^+} \frac{\lambda_n(E \cap B(x,r))}{\lambda_n(B(x,r))}$$

if the limit exists. A point with $D_E(x) = 1$ is called a Lebesgue density point of $E$.

**Theorem 5.1.8.** Let $E \subset \mathbb{R}^n$ be measurable. Then $D_E(x)$ exists for almost every $x \in \mathbb{R}^n$ and $D_E(x) = 0$ for almost every $x \in \mathbb{R}^n \setminus E$, while $D_E(x) = 1$ for almost every $x \in E$.

**Proof.** Let $f = \chi_E \in L^1_{\text{loc}}$. For each Lebesgue point $x$ of $f$,

$$\int_{B(x,r)} \chi_E = \frac{\lambda_n(E \cap B(x,r))}{\lambda_n(B(x,r))} \to \chi_E(x) \in \{0,1\},$$

as $r \to 0$. \qed

**Theorem 5.1.9.** Let $\nu$ be a complex Borel measure on $\mathbb{R}^n$ and $d\nu = d\rho + f \, d\lambda_n$ be its Lebesgue decomposition with respect to $\lambda_n$. Suppose for each $x \in \mathbb{R}^n$ we have a family $\{E_r(x)\}_{r>0}$ of Borel sets shrinking nicely to $x$. Then

$$f(x) = \lim_{r \to 0^+} \frac{\nu(E_r(x))}{\lambda_n(E_r(x))}$$

for $\lambda_n$-a.e. $x$.

**Proof.** We start by writing

$$\frac{\nu(E_r(x))}{\lambda_n(E_r(x))} = \frac{\rho(E_r(x))}{\lambda_n(E_r(x))} + \int_{E_r(x)} f \, d\lambda_n.$$
For the second term, at each Lebesgue point $x$,

$$\lim_{r \to 0^+} \int_{E_r(x)} f \, d\lambda_n = f(x).$$

For the first term,

$$\left| \frac{\rho(E_r(x))}{\lambda_n(E_r(x))} \right| \leq \frac{|\rho|(B(x, r))}{\lambda_n(B(x, r))} \leq \frac{1}{\alpha(x)} \cdot \frac{|\rho|(B(x, r))}{\lambda_n(B(x, r))}.$$

Let $\mu = |\rho|$, a finite positive Borel measure on $\mathbb{R}^n$ with $\mu \perp \lambda$. Let $A$ be a Borel set such that $\mu(A) = 0$ and $\lambda(\mathbb{R}^n \setminus A) = 0$. For each $k \in \mathbb{N}$, consider

$$F_k = \left\{ x \in A \mid \limsup_{r \to 0^+} \frac{\mu(B(x, r))}{\lambda_n(B(x, r))} > \frac{1}{k} \right\}.$$

Since $\mu$ is outer regular, for each $\epsilon > 0$, there is an open set $U_\epsilon \supset A$ such that $\mu(U_\epsilon \setminus A) < \epsilon$. Then $\mu(U_\epsilon) = \mu(U_\epsilon \setminus A) + \mu(A) < \epsilon$. For each $x \in F_k$, there exists a ball $B_x \subset U_\epsilon$ of small radius (wlog bounded by 1) such that $\mu(B_x)/\lambda_n(B_x) > 1/k$. By the 5B-covering lemma, we can find a disjoint subfamily $\{B_m \mid m \in \mathbb{N}\}$ such that

$$F_k \subset V_\epsilon = \bigcup_{x \in F_k} B_k \subset \bigcup_{m=1}^\infty 5B_m.$$

Then

$$\lambda_n(F_k) \leq \lambda_n(V_\epsilon) \leq \sum_{m=1}^\infty \lambda_n(5B_m) = 5^n \sum_{m=1}^\infty \lambda_n(B_m) \leq 5^n k \mu(U_\epsilon) \leq 5^n k \epsilon.$$

Letting $\epsilon \to 0$ and keeping $k$ fixed, we get $\lambda_n(F_k) = 0$. Hence

$$\lambda_n \left( (\mathbb{R}^n \setminus A) \cup \bigcup_{k=1}^\infty F_k \right) = 0,$$

which implies the result.

5.2 COMPLEX BOREL MEASURES ON $\mathbb{R}$

We wish to describe all complex (Borel) measures on $\mathbb{R}$. The basic idea is that given a complex measure $\mu$ on $\mathbb{R}$, we can define $F_\mu(x) = \mu((-\infty, x]) \in \mathbb{C}$. These functions will be a special type of function, namely a function of “bounded variation”.

53
5.2 Complex Borel measures on $\mathbb{R}$

**Definition 5.2.1** (Functions of bounded variation). A function $F: \mathbb{R} \to \mathbb{C}$ has *bounded variation* (is a *BV-function*) if there exists a constant $M > 0$ such that

$$\sum_{k=1}^{n} |F(x_k) - F(x_{k-1})| \leq M$$

whenever $n \in \mathbb{N}$ and $x_0 < x_1 < \cdots < x_n$.

The set of all BV-functions on $\mathbb{R}$ is denoted by $BV(\mathbb{R})$, or simply $BV$ when understood.

If $F \in BV$ and $v \in \mathbb{R}$, we define

$$T_F(x) = \sup \left\{ \sum_{k=1}^{n} |F(x_k) - F(x_{k-1})| \mid x_0 < x_1 < \cdots < x_n \leq x \right\},$$

the total variation of $F$ up to $x$.

**Theorem 5.2.2.**

1. $F \in BV$ if and only if $\text{Re} \, F \in BV$ and $\text{Im} \, F \in BV$.

2. $F: \mathbb{R} \to \mathbb{R}$ is a BV-function if and only if $F$ is the difference of two bounded increasing functions.

3. If $F \in BV$, then

$$F(x^+) = \lim_{y \to x^+} F(y) \quad \text{and} \quad F(x^-) = \lim_{y \to x^-} F(y)$$

exist for all $x \in \mathbb{R}$. Furthermore, $F(+\infty)$ and $F(-\infty)$ also exist.

4. If $F \in BV$, then $F$ has at most countably many discontinuities.