## Contents

1 Measure Theory ........................................ 3  
  1.1 \(\sigma\)-algebras .................................. 3  
  1.2 Measures ........................................... 4  
  1.3 Construction of non-trivial measures .............. 5  
  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 .................................. 40  

4 More on \(L^p\) Spaces .................................. 43  
  4.1 Bounded linear maps and dual spaces ............ 43  
  4.2 The dual of \(L^p\) ................................... 45  
  4.3 The Hardy-Littlewood maximal functions ........ 47  

5 Differentiability ...................................... 51  
  5.1 Lebesgue points .................................... 51  
  5.2 Complex Borel measures on \(\mathbb{R}\) ........... 54  
  5.3 The fundamental theorem of calculus ............ 58  

6 Functional Analysis ................................... 61  
  6.1 Locally compact Hausdorff spaces ................. 61  
  6.2 Weak topologies .................................... 62  
  6.3 Some theorems in functional analysis ............. 65  
  6.4 Hilbert spaces ..................................... 67
7 Fourier Analysis

7.1 Trigonometric series .................................................. 73
7.2 Fourier series ............................................................. 74
1 MEASURE THEORY

1.1 σ-ALGEBRAS

Definition 1.1.1 (σ-algebra). Let \(X\) be a set and \(A\) be a family of subsets of \(X\). We say that \(A\) is an algebra on \(X\) if

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

We say that \(A\) is a σ-algebra on \(X\) if closure under finite unions extends to countable unions.

Proposition 1.1.2. Let \(A\) be an algebra on \(X\). If \(A, B \in A\), then

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

Moreover, if \(A\) is a σ-algebra and \(A_n \in A\), then

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

Example 1.1.3.

1. For any set \(X\), \(A = \mathcal{P}(X)\) is a σ-algebra.
2. For any set \(X\), \(A = \{A \subset X \mid A\) is countable or \(X \setminus A\) is countable\}\) is a σ-algebra.
3. If \(\{A_i \mid i \in I\}\) is a family of σ-algebras on \(X\), then so is \(A = \bigcap_i A_i\).

Proposition 1.1.4. Let \(F \subset \mathcal{P}(X)\) be a family of subsets of \(X\). Then there is a unique σ-algebra \(A\) such that \(F \subset A \subset A'\), where \(A'\) is any σ-algebra on \(X\) with \(F \subset A'\).

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

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

Definition 1.1.5 (σ-algebra generated by a family of sets). The σ-algebra \(A\) is the σ-algebra generated by \(F\), denoted by \(\sigma(F)\).

Definition 1.1.6 (Borel σ-algebra). Let \((X, \mathcal{O})\) be a topological space. The Borel σ-algebra on \(X\) is \(B_X = \sigma(\mathcal{O})\), the σ-algebra generated by the open sets in \(X\).

Example 1.1.7.

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.
3. 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 \(i\). If \(\mathcal{R}\) is the family of all rectangles, then \(\sigma(\mathcal{R}) = B_{\mathbb{R}^n}\).
1.2 MEASURES

**Definition 1.2.1** (Measure). A pair \((X, A)\) of a set \(X\) and a σ-algebra \(A\) on \(X\) is a measurable space. A (positive) measure \(μ\) on a measurable space \((X, A)\) is a function \(μ : A \to [0, \infty]\) such that

1. \(μ(\emptyset) = 0;\)
2. if \(A_n \in A\) \((n \in \mathbb{N})\) are pairwise disjoint, then

\[
μ \left( \bigcup_{n=1}^{∞} A_n \right) = \sum_{n=1}^{∞} μ(A_n).
\]

The triple \((X, A, μ)\) is a measure space.

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

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

\[
δ_a(M) = \begin{cases} 
0 & a \notin M, \\
1 & a \in M.
\end{cases}
\]

2. The counting measure is

\[
μ(M) = \begin{cases} |M| & M \text{ is finite,} \\
∞ & \text{otherwise.}
\end{cases}
\]

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 \(A = \bigcup_n A_n\).
2. Write \(A_n \searrow A\) if \(A_1 \supset A_2 \supset \cdots\) and \(A = \bigcap_n A_n\).

**Theorem 1.2.3.** Let \((X, A, μ)\) be a measure space. Then

1. (monotonicity) if \(A, B \in A\) and \(A \subset B\), then \(μ(A) \leq μ(B)\);
2. (countable subadditivity) if \(A_n \in A\), then

\[
μ \left( \bigcup_{n=1}^{∞} A_n \right) \leq \sum_{n=1}^{∞} μ(A_n);
\]

3. (continuity from below) if \(A_n \in A\) and \(A_n \nearrow A\),

\[
μ(A) = \lim_{n \to ∞} μ(A_n);
\]

4. (continuity from above) if \(A_n \in A\), \(A_n \searrow A\), and \(μ(A_1) < ∞\), then

\[
μ(A) = \lim_{n \to ∞} μ(A_n).
\]
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

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

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 \( B_n \not\supset A_1 \setminus A \) in \( A \) 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

\[
\mu(A_1) = \lim_{n \to \infty} (\mu(A_1) - \mu(A_n)) + \mu(A) = \mu(A_1) + \mu(A) - \lim_{n \to \infty} \mu(A_n).
\]

\[\square\]

**Remark 1.2.4.** Continuity from above may fail if we allow \( \mu(A_1) = \infty \). For example, on \( \mathbb{R} \) with the counting measure, let \( A_n = [n, \infty) \). Then \( \bigcap_n A_n = \emptyset \), but \( \mu(A_n) = \infty \) for each \( n \).

**Definition 1.2.5** (Null set). Let \((X, \mathcal{A}, \mu)\) be a measure space. A **null set** is a measurable set \( N \subset X \) with \( \mu(N) = 0 \).

**Definition 1.2.6** (“Almost everywhere”). Let \( P(x) \) be a statement about points \( x \) in a measure space \((X, \mathcal{A}, \mu)\). We say that \( P(x) \) is true \( \mu \)-**almost everywhere** (or for \( \mu \)-almost every \( x \in X \)), abbreviated \( \mu \)-a.e., if there exists a null set \( N \) for which \( P(x) \) is true for all \( x \in X \setminus N \).

### 1.3 CONSTRUCTION OF NON-TRIVIAL MEASURES

Let \((X, \mathcal{A}, \mu)\) be a measure space.

**Definition 1.3.1** (Complete measure space). We say that \((X, \mathcal{A}, \mu)\) is **complete** if every subset of a null set is a null set. (In particular, every subset of a null set is measurable.)

**Theorem 1.3.2.** Let

\[
\overline{\mathcal{A}} = \{ A \cup B \mid A \in \mathcal{A} \text{ and there exists a null set } N \text{ such that } B \subset N \}.
\]

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. It is clear that $\mathcal{A} \subset \mathcal{A}$. To see that $\mathcal{A}$ is a $\sigma$-algebra, we check the axioms.

(i) Since $\mathcal{A} \subset \mathcal{A}$, we have $\emptyset \in \mathcal{A}$ and $X \in \mathcal{A}$.

(ii) Let $A \cup B \in \mathcal{A}$ with $B \subset N$ for some null set $N$. By replacing $A$ with $A \setminus N$ and $B$ with $B \cup (A \cap N)$ if necessary, we can suppose that $A \cap N = \emptyset$. Then

$$X \setminus (A \cup B) = (X \setminus A) \cap (X \setminus B) = [(X \setminus A) \cap (X \setminus N)] \cup [N \setminus B] \in \mathcal{A}.$$ 

(iii) Let $A_n \cup B_n \in \mathcal{A}$ with $B_n \subset N_n$ for null sets $N_n$, and let

$$A = \bigcup_{n=1}^{\infty} A_n, \quad B = \bigcup_{n=1}^{\infty} B_n, \quad N = \bigcup_{n=1}^{\infty} N_n.$$ 

Then $A, N \in \mathcal{A}$ and $\mu(N) \leq \sum_n \mu(N_n) = 0$, so $N$ is a null set with $B \subset N$, and

$$\bigcup_{n=1}^{\infty} (A_n \cup B_n) = A \cup B \in \mathcal{A}.$$ 

For existence of an extension $\overline{\mu}$, we attempt to define $\overline{\mu}(A \cup B) = \mu(A)$. To show existence, define $\overline{\mu}$ by $\overline{\mu}(A \cup B) = \mu(A)$. 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.

If $\overline{\mu}$ exists and $A \cup B \in \mathcal{A}$ with $B \subset N$ for some $\mu$-null set $N$, we must have

$$\mu(A) = \overline{\mu}(A) \leq \overline{\mu}(A \cup B) \leq \overline{\mu}(A \cup N) \leq \overline{\mu}(A) + \overline{\mu}(N) = \mu(A) + \mu(N) = \mu(A),$$

hence $\overline{\mu}(A \cup B) = \mu(A)$ is fixed. \hfill $\square$

**Definition 1.3.3** (Outer measure). An outer measure on a set $X$ is a function $\mu^*: \mathcal{P}(X) \to [0, \infty)$ such that

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

(ii) $\mu^*(A) \leq \mu^*(B)$ whenever $A \subset B$;

(iii) $\mu^*(\bigcup_n A_n) \leq \sum_n \mu^*(A_n)$.

**Lemma 1.3.4.** Let $\mathcal{A}$ be an algebra on $X$.

1. Suppose that whenever $A_n \in \mathcal{A}$ with $A_n \nrightarrow A$, we have $A \notin \mathcal{A}$. 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 1.3.5 (Carathéodory). Let \( \mu^* \) be an outer measure on \( X \) and \( \mathcal{A} \) be the family of sets \( A \subset X \) such that

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

for all \( T \subset X \). 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

\[
\mu^*(T) = \mu^*(T \cap A) + \mu^*(T \cap A^c)
= \mu^*(T \cap A \cap B) + \mu^*(T \cap A \cap B^c) + \mu^*(T \cap A^c \cap B) + \mu^*(T \cap A^c \cap B^c)
\geq \mu^*(T \cap (A \cup B)) + \mu^*(T \cap A^c \cap B^c)
= \mu^*(T \cap (A \cup B)) + \mu^*(T \cap (A \cup B)^c))
\geq \mu^*(T),
\]

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 \( T \),

\[
\mu^*(T) \leq \mu^*(T \cap A) + \mu^*(T \cap A^c) \leq \mu^*(T \cap A^c) \leq \mu^*(T),
\]

so equality holds everywhere and we have the required equality for \( A \in \mathcal{A} \). \( \square \)
Definition 1.3.6 ($\mu^*$-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 1.3.7 (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 1.3.8. 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. 

**Theorem 1.3.9** (Carathéodory extension theorem). Let $\mathcal{A}$ be an algebra on $X$, $\nu : \mathcal{A} \to [0, \infty]$ 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 1.3.5 and Lemma 1.3.8, $\nu$ 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 \( F_n \in A \) 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 \( \mathcal{C} \) be a family of subsets of \( X \) such that

(i) \( \emptyset \in \mathcal{C} \)

(ii) \( A \cap B \in \mathcal{C} \) whenever \( A, B \in \mathcal{C} \)

(iii) if \( A \in \mathcal{C} \), then \( A^c \) is a finite union of elements of \( \mathcal{C} \) 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 \( \mathcal{C} \), 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 \( I_i^c = R \setminus I_i \) is a disjoint union of at most two h-intervals. This implies
\[
R^c = \mathbb{R}^n \setminus R = \bigcup_{1 \leq i_1 < \cdots < i_k \leq n} \prod_{j=1}^{n} \begin{cases} I_j^i & \text{j = i for some } 1 \leq i \leq k \\ I_j & \text{otherwise} \end{cases}
\]
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
\[
|R| = l(I_1) \cdots l(I_n) \in [0, \infty],
\]
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 \( \mathcal{A} \) be the algebra generated by the h-rectangles in \( \mathbb{R}^n \). To define a premeasure \( \nu : \mathcal{A} \to [0, \infty] \), let \( M \in \mathcal{A} \) 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 \( \mathcal{A} \).

**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 \( A_n \in \mathcal{A} \), 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 1.3.8, i.e.

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

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 1.3.5.

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. \hfill \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. ($\implies$) Omitted.

($\impliedby$) 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 $A_k = \bigcup_i R_{k,i}$ and $A = \bigcap_k A_k$. 

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 \( B_X \), and \( \mu : \mathcal{A} \rightarrow [0, \infty] \) 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 \( T : \mathbb{R}^n \rightarrow \mathbb{R}^n \) 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. \( \square \)

13
**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 \( \mathcal{B} = \sigma(S) \), then \( f \) is measurable if and only if \( f^{-1}(S) \in \mathcal{A} \) for all \( S \in S \).

**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 \( \mathcal{A}^{-1}((a, \infty]) \in \mathcal{A} \) 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 \( f \), \( f + g \), and \( f \cdot g \) are measurable.

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 \( f \vee g \) and \( f \wedge g \) 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 \( f_n \uparrow f \) 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 \( s_n \uparrow f \) and \( s_n \to f \) pointwise.
Proof. Define

\[ E_{n,k} = \left( f^{-1} \left( \left[ \frac{k}{2^n}, \frac{k+1}{2^n} \right] \right) \right), \quad F_n = f^{-1}((2^n, +\infty]), \]

and set

\[ s_n = \sum_{k=0}^{2^n-1} \frac{k}{2^n} \cdot \chi_{E_{n,k}} + 2^n \cdot \chi_{F_n}. \]
2 INTEGRATION

2.1 INTEGRATION OF SIMPLE 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 \( \mathcal{S}^+ \) 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 \( s \in \mathcal{S}^+ \) 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 \( s \in \mathcal{S}^+ \) and \( c \geq 0 \), then \( \int cs = c \int s \).

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

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

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

\[
\int_A s \, d\mu = \int s \cdot \chi_A \, d\mu.
\]

Proposition 2.1.4. Let \( s \in \mathcal{S}^+ \) 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 \( X \).
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 \( f_n \uparrow f \), 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, \( A_n \uparrow X \) 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
\[
\lim_{n \to \infty} \int f_n \, d\mu \geq \lim_{n \to \infty} \alpha \int_{A_n} \phi \, d\mu.
\]
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. \( \square \)

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 \( f_n \uparrow f \) \( \mu \)-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 \( \{g_k\} \),
\[
\int \liminf_{n \to \infty} f_n \, d\mu = \int \liminf_{k \to \infty} g_k \, d\mu \leq \liminf_{k \to \infty} \inf_{m \geq k} \int f_m \, d\mu = \liminf_{n \to \infty} \int f_n \, d\mu.
\]  

### 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 \( L^1_\mathbb{R}(\mu) \), or \( L^1_\mathbb{R} \) if \( \mu \) is understood.

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

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

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

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

**Definition 2.3.7** (Integral of a complex-valued integrable function / integral over a subset). For \( f \in L^1 \), 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 \( f \in L^1 \), define
\[
\int_A f \, d\mu = \int \chi_A f \, d\mu.
\]

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

**Proposition 2.3.9.** If \( f \in L^1 \), 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
\[
|\int f \, d\mu| = |\alpha| \int f \, d\mu = \int |\alpha f| \, d\mu = \int \text{Re}(\alpha f) \, d\mu \\
\leq \int |\text{Re}(\alpha f)| \, d\mu \leq \int |\alpha f| \, d\mu = \int |f| \, d\mu.
\]

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

**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,

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

\( \square \)

**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 \( 1 \leq p < \infty \).
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. \( \square \)

Proposition 2.4.13. For \( 1 \leq p < \infty \), the set of simple functions \( s = \sum_j a_j 1_{E_j} \) 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 \textit{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 \( \mathbb{C} \)-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 $\Phi_{P_k} \to \Phi$ and $\phi_{P_k} \to \phi$; note that these limits are monotone. Then we have $\phi_{P_k} \leq \phi \leq f \leq \Phi \leq \Phi_{P_k}$ for each $k$, so by the dominated convergence theorem,
\[ \int_{[a,b]} (\Phi - \phi) \, d\lambda = 0 \quad \text{with} \quad \Phi - \phi \geq 0, \quad \text{so} \quad \Phi = \phi \text{ almost everywhere.} \]
As $\phi$ is the limit of simple functions, $\phi$ is measurable, so $f$ is also measurable. Furthermore,
\[ \int_{[a,b]} f(x) \, dx = \int_{[a,b]} \phi \, d\lambda \leq \int_{[a,b]} f \, d\lambda = \int_{[a,b]} \Phi \, d\lambda = \int_{[a,b]} f(x) \, dx, \]
so $\int_{[a,b]} f \, d\lambda = \int_{[a,b]} f(x) \, dx. \quad \square$

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

**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
\[
f_n = n^{1/p} \chi_{[0, 1/n]}
\]
shows that \( 1 \not\Rightarrow 2 \) and \( 2 \Rightarrow 3 \). 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 \( 3 \not\Rightarrow 1 \). The example on \( \mathbb{R} \) given by
\[
f_n = \chi_{[n,n+1]} \quad \text{and} \quad f = 0
\]
shows that \( 1 \not\Rightarrow 3 \). The only remaining implication, \( 2 \Rightarrow 3 \), turns out to be correct. The implications \( 3 \implies 1 \) and \( 1 \implies 3 \) can also be made to work with slight modifications.

**Proposition 2.6.3.** If \( f_n \to f \) in \( L^p \) for some \( 1 \leq p < \infty \), then \( f_n \to f \) 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 \( \mu(E_n) \leq \epsilon^{-1} \int |f_n - f| \, d\mu \to 0 \) as \( n \to \infty \).

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

**Proof.** If \( f_n \to f \) 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 \( \mu(F_k) \leq 2^{k-l} \) 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 \( x \notin F_k \). This means that \( \{ g_j(x) \} \) is Cauchy if \( x \notin F_k \), for any \( k \). Let \( F = \bigcap_k F_k \). Then \( \mu(F) = 0 \) and \( f(x) = \lim g_j(x) \) is defined \( \mu \text{-a.e.} \). (Define \( f \) to be zero elsewhere.) We have that \( f \) is measurable and \( g_j \to f \) \( \mu \text{-a.e.} \).

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

**Proof.** Without loss of generality, suppose \( f_n \to f \) 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 \( \mu(E_n(k)) \to \mu(\emptyset) = 0 \). 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 \( f_n \to f \) uniformly on \( E^c \) by construction.
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 \( A, B \subseteq X \) 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 \( Q \subseteq [0,1] \) consist of all dyadic rationals. We construct a collection of open sets \( U_q \) indexed by \( Q \) with \( U_q \subseteq U_r \) 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 \( U_0 \subseteq U_{1/2} \subseteq \overline{U_{1/2}} \subseteq X \setminus B \). Repeating this, we inductively define all of the sets \( U_q \) for \( q \in Q \).

The continuous function we wish to define is then

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

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 \( A \subseteq X \) is a Borel set of finite measure. Then for each \( \epsilon > 0 \), there exists a closed \( E \subseteq A \) such that \( \mu(E) < \epsilon \) and \( f \) is continuous on \( A \setminus E \).

Proof. To be written.

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.

Remark 2.6.11. In the case \( p = \infty \), we still have \( C_c(X) \subseteq L^\infty \), 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 \( E \subseteq X \times Y \). 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

\[
    f_x(y) = f^y(x) = f(x,y).
\]
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

1. 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 \(A \in \mathcal{P}\);
2. \(\mathcal{F}\) is closed under linear combinations;
3. \(\mathcal{F}\) 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

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

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 $\int H \, d\nu$.

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

1. If $f \in L^1(\omega)$, then $f_x \in L^1(\nu)$ and $f^y \in L^1(\mu)$ for $\mu$-a.e. $x \in X$ and $\nu$-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 $\mu$-a.e. $x \in X$ and $\nu$-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 $G < \infty \mu$-a.e. and $H < \infty \nu$-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 $f$.

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

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

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

1. If $f$ is in $L^1(\omega)$, then $f_x$ and $f^y$ are measurable for $\mu$-a.e. $x \in X$ and $\nu$-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 $\mu$-a.e. $x \in X$ and $\nu$-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 $\mathcal{B}_n$ be the Borel $\sigma$-algebra on $\mathbb{R}^n$ and $\mathcal{L}_n$ be the $\sigma$-algebra of Lebesgue measurable subsets of $\mathbb{R}^n$. Define a Borel measure $\beta_n$ on $\mathcal{B}_n$ by restricting the Lebesgue measure $\lambda_n$ to Borel sets. Since $\mathcal{B}_{m+n} = \mathcal{B}_m \otimes \mathcal{B}_n$ (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 $M \subset A$ 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 $M \subset A$, and $f : M \to \mathbb{C}$ be a function. Define $\tilde{f} : X \to \mathbb{C}$ by

$$\tilde{f} = \begin{cases} 0 & x \not\in M, \\ f(x) & x \in M. \end{cases}$$

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

$$\int g \, d(T_*\mu) = \int g \circ T \, d\mu.$$

Definition 2.8.5 (Polar coordinates). The polar coordinates for a point $p \in \mathbb{R}^n \setminus \{0\}$ are the pair $(r, \xi)$, with $r \in (0, \infty)$ and $\xi \in S^{n-1} \subset \mathbb{R}^n$, such that $p = r\xi$. 28
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

$$\rho(E) = \int_E r^{n-1} \, dr = \int_E x^{n-1} \, dx.$$

**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

$$\lambda_n (\{r\xi \mid r \in (0, \alpha) \text{ and } \xi \in B\}) = \lambda_n (\alpha \cdot \tilde{B}) = \alpha^n \lambda_n (\tilde{B})$$

$$= \frac{\alpha^n}{n} \sigma(B) = \left( \int_0^{\alpha} r^{n-1} \, dr \right) \cdot \sigma(B) = \rho((0, \alpha))\sigma(B).$$

For uniqueness, suppose $m(A \times B) = \rho(A)\sigma(B) = \rho(A)\tilde{\sigma}(B)$. Picking $A = (0, 1)$, we have $\sigma(B) = \tilde{\sigma}(B)$ 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,

$$\int_{\mathbb{R}^n} \chi_M \, d\lambda_n = \int_{\mathbb{R}^n \setminus \{0\}} \chi_N \, d\beta_n$$

$$= \int_{\mathbb{R}^n \setminus \{0\}} (\chi_N \circ \Phi^{-1}) \circ \Phi \, d(\beta_n|\mathbb{R}^n \setminus \{0\})$$

$$= \int_{(0, \infty) \times S^{n-1}} \chi_N (r\xi) \, dm(r, \xi)$$

$$= \int_0^\infty \left( \int_{S^{n-1}} \chi_M (r\xi) r^{n-1} \, d\sigma(\xi) \right) \, dr.$$

$\square$
**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.

**Definition 2.8.9 (Γ function).** The Γ 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 $\mathbb{B}^n = \{x \in \mathbb{R}^n | |x| < 1\}$, then

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

**Proof.** We compute

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

$$= \int_0^\infty \left( \int_{S_{n-1}} e^{-r^2} r^{n-1} d\sigma(\xi) \right) dr$$

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

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

Example 2.8.12. For the first few \(n\),
\[
\lambda_1(B^1) = 2, \quad \lambda_3(B^3) = \frac{4\pi}{3}, \quad \lambda_4(B^4) = \frac{\pi^2}{2}.
\]

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 \((C^1\text{-diffeomorphism})\). 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 \((\text{Jacobian})\). The Jacobian (determinant) is the determinant of the Jacobian matrix,
\[
J_T(x) = \det(DT(x)).
\]

Lemma 2.9.4 \((\text{Chain rule})\). 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 \((\text{Whitney decomposition})\). 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
\[
\text{diam}(Q_k) \leq \text{dist}(Q_k, \partial \Omega) \leq 4 \text{diam}(Q_k).
\]
2.9 The transformation formula

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

$$
\lambda_n(T(Q)) = \sum_{i=1}^N \lambda_n(T(Q_i)) \lesssim \sum_{i=1}^N \lambda_n(A_i(Q_i)) = \sum_{i=1}^N |\det(DT(x_i))| \cdot \lambda_n(Q_i) = \sum_{i=1}^N \int_{Q_i} |J_T(x_i)| \, d\lambda_n \approx \sum_{i=1}^N \int_{Q_i} |J_T| \, d\lambda_n = \int_Q |J_T| \, d\lambda_n.
$$

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 $g = (f \circ T)|J_T|$. \qed
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 \(\mu : \mathcal{A} \rightarrow \mathbb{R}\) such that

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

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

(iii) 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 \((X, \mathcal{A})\). We say that a measurable function \(f : X \rightarrow \mathbb{R}\) is \(\nu\)-integrable in the extended sense if

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

If \(f\) is \(\nu\)-integrable in the extended sense, then

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

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 disjoint \(E, F \in \mathcal{A}\) such that \(E \cup F = X\) and \(\mu(E) = \nu(F) = 0\).

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 \(A \subset P\). One can similarly define (totally) negative sets and total null sets.

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 \(\mu(A_n) \rightarrow \mu(A)\).

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

3. Measurable subsets and countable unions of positive sets are positive.

4. If \(\mu < +\infty\) and \(\mu(A) \neq -\infty\), then \(A\) contains a positive subset \(P\) with \(\mu(P) \geq \mu(A)\).
Let \( A \) be a signed measure on \((X, \mathcal{A})\). Then there exist 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 the symmetric differences \( P \triangle P' \) and \( N \triangle N' \) are total null.

**Proof.** Without loss of generality, suppose \( \mu \) does not take the value \(+\infty\), and let

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

Let \( A_n \in \mathcal{A} \) be measurable sets with \( \mu(A_n) \to s \). By the lemma, for each \( n \), there is a positive set \( P_n \subset A_n \) with \( \mu(P_n) \geq \mu(A_n) \). Then \( P = \bigcup_n P_n \) is positive and \( \mu(P) \geq \mu(P_n) \geq \mu(A_n) \) for each \( n \), so we have

\[ s \geq \mu(P) \geq \lim_{n \to \infty} \mu(A_n) = s. \]

Thus \( \mu(P) = s < +\infty \). The set \( N = X \setminus P \in \mathcal{A} \) is negative, as otherwise there exists \( E \subset N \) with \( \mu(E) > 0 \), and then

\[ \mu(P \cup E) = \mu(P) + \mu(E) > s, \]

a contradiction. Hence we have existence for the required 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 \( P \triangle P' \) is total null. The same proof works for \( N \triangle N' \). \(\)
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 $\mu^+ \perp \mu^-$. 

Proof. Let $X = P \cup N$ be a Hahn decomposition for $\mu$ and define 

$$\mu^+(A) = \mu(A \cap P) \quad \text{and} \quad \mu^-(A) = -\mu(A \cap N).$$

It is clear that this satisfies the required conditions, so we have proved existence.

For uniqueness, suppose that in addition to $\mu = \mu^+ - \mu^-$ as above, we have $\mu = \nu^+ - \nu^-$ for positive measures $\nu^+ \perp \nu^-$. Let $E, F$ be disjoint with $E \cup F = X$ and $\nu^+(E) = \nu^-(F) = 0$. For $A \in \mathcal{A}$ with $A \subset E$, we have 

$$\mu(A) = \mu(A \cap E) + \mu(A \cap F) = -\nu^-(A \cap E),$$

so $E$ is a negative set for $\mu$. Similarly, $F$ is a positive set for $\mu$, so $X = E \cup F$ is a Hahn decomposition for $\mu$. By the uniqueness of Hahn decomposition, $P \Delta F$ and $N \Delta E$ are total null sets for $\mu$. Hence for $A \in \mathcal{A}$, 

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

so $\nu^+ = \mu^+$, and similarly $\nu^- = \mu^-$. \qed

Definition 3.1.8 (Variations). The positive measures $\mu^+$ and $\mu^-$ are the positive and negative variations of $\mu$. The total variation of $\mu$ is the positive measure $|\mu| = \mu^+ + \mu^-$. 

Remark 3.1.9. For a given $A \in \mathcal{A}$, we have $|\mu(A)| \leq |\mu|(A)$, but equality might not hold. 

3.2 THE RADON-NIKODYM THEOREM

Definition 3.2.1 (Finite signed measure). A signed measure $\mu$ is said to be finite if $|\mu|$ is finite as a positive measure. Similarly, $\mu$ is $\sigma$-finite if $|\mu|$ is $\sigma$-finite. 

Notation. Let $(X, \mathcal{A})$ be a measurable space, $\nu$ be a positive measure on $X$, and $f : X \to \mathbb{R}$ be $\nu$-integrable in the extended sense. We write $d\mu = f \, d\nu$ to mean that $\mu$ is the signed measure defined by $\mu(A) = \int_A f \, d\nu$. 

Proposition 3.2.2. If $d\mu = f \, d\nu$, then $\{f = 0\}$ is a total null set for $\mu$. 

Proof. The sets $P = \{f \geq 0\}$ and $N = \{f < 0\}$ give a Hahn decomposition for $\mu$, as do $P' = \{f > 0\}$ and $N' = \{f \leq 0\}$. Hence $P \Delta P' = N \Delta N' = \{f = 0\}$ is total null. \qed

Proposition 3.2.3. If $d\mu = f \, d\nu$, then $d|\mu| = |f| \, d\nu$. 

Proof. The sets $P = \{f \geq 0\}$ and $N = \{f < 0\}$ give a Hahn decomposition for $\mu$, so the Jordan decomposition $\mu = \mu^+ - \mu^-$ satisfies $d\mu^+ = f^+ \, d\nu$ and $d\mu^- = f^- \, d\nu$. Then 

$$d|\mu| = d\mu^+ + d\mu^- = f^+ \, d\nu + f^- \, d\nu = |f| \, d\nu. \qed$$

Definition 3.2.4 (Absolute continuity). 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 $\nu(A) = 0$ for $A \in \mathcal{A}$ whenever $\mu(A) = 0$. 

37
Proposition 3.2.5.  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 3.2.6. 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 \( E \).

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

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

for each \( n \). This shows that \( \nu(N) = 0 \). If \( \mu(P) = 0 \), then \( \nu \perp \mu \). Otherwise, \( \mu(P) > 0 \), so \( \mu(P_n) > 0 \) for some \( n \). Since \( P_n \) is totally positive for \( \nu - (1/n)\mu \), we can take \( E = P_n \) and \( \epsilon = 1/n \).

Theorem 3.2.7 (Lebesgue-Radon-Nikodym). Let \( (X, \mathcal{A}) \) be a measurable space, \( \mu \) be a \( \sigma \)-finite positive measure, and \( \nu \) be a \( \sigma \)-finite signed measure. Then there are unique \( \sigma \)-finite signed measures \( \lambda, \rho \) with \( \nu = \lambda + \rho \) such that \( \lambda \perp \mu \) and \( \rho \ll \mu \). Moreover, there exists a function \( f : X \to \mathbb{R} \) which is \( \mu \)-integrable in the extended sense such that \( d\nu = f \, d\mu \), and \( f \) is unique up to equality \( \mu \)-a.e.

Proof. First suppose that \( \mu \) and \( \nu \) are finite positive measures, and let

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

We will construct \( f \in \mathcal{F} \) which is maximal in the sense that if \( g \in \mathcal{F} \), then \( \mu(\{ g \geq f \}) = 0 \). First, we note that \( \mathcal{F} \neq \emptyset \) as \( 0 \in \mathcal{F} \), so

\[
a = \sup \left\{ \int f \, d\mu \mid f \in \mathcal{F} \right\} > -\infty.
\]

Since each of these integrals is bounded above by \( \nu(X) < +\infty \), we have that \( a \) is finite. Pick \( f_n \in \mathcal{F} \) such that \( \int f_n \, d\mu \to a \) as \( n \to \infty \).

Define the functions \( g_n = \max(f_1, \ldots, f_n) \). We claim that \( g_n \in \mathcal{F} \) for each \( n \). To see this, it is enough to show that if \( h_1, h_2 \in \mathcal{F} \), then \( h = \max(h_1, h_2) \in \mathcal{F} \). Let \( A = \{ h_1 > h_2 \} \in \mathcal{A} \). For \( E \in \mathcal{A} \),

\[
\int_E h \, d\mu = \int_{E \cap A} h_1 \, d\mu + \int_{E \cap (X \setminus A)} h_2 \, d\mu \leq \nu(E \cap A) + \nu(E \cap (X \setminus A)) = \nu(E),
\]

so \( h \in \mathcal{F} \), as required.

By construction, \( g_n \uparrow \), so we can define \( f \) to be the pointwise limit. Then \( f \) is measurable, and to see that \( f \in \mathcal{F} \), we have by the monotone convergence theorem that

\[
\int_E f \, d\mu = \int_{E} \lim_{n \to \infty} g_n \, d\mu = \lim_{n \to \infty} \int_E g_n \, d\mu \leq \nu(E)
\]
for all \( E \in \mathcal{A} \). In the case \( E = X \), we obtain \( \int f \, d\mu = a < +\infty \), so \( f \) takes the value \(+\infty\) only on a set of measure zero. Redefining \( f \) by setting \( f(x) = 0 \) wherever it was \(+\infty\) before, we can suppose \( f \) takes values in \([0, \infty)\).

Consider the finite measure \( \rho \) defined by

\[
\rho(A) = \int_A f \, d\mu
\]

for \( A \in \mathcal{A} \), so \( d\rho = f \, d\mu \). 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 \( A \subseteq E \), 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 \( a \) as the supremum of all such integrals. Thus we have shown existence in the case that \( \mu \) and \( \nu \) are finite.

If \( \mu \) and \( \nu \) are \( \sigma \)-finite positive measures, then we can decompose \( X \) into countably many pairwise disjoint measurable pieces \( X_n \) on each of which both \( \mu \) and \( \nu \) are finite positive measures. Let \( \nu = \lambda_n + \rho_n \) be a decomposition on \( X_n \) with \( \lambda_n \perp \mu \) and \( \rho_n \ll \mu \) on \( X_n \), then define

\[
\lambda(A) = \sum_{n=1}^{\infty} \lambda_n (A \cap X_n) \quad \text{and} \quad \rho(A) = \sum_{n=1}^{\infty} \rho_n (A \cap X_n)
\]

for \( A \in \mathcal{A} \). The decomposition \( \nu = \lambda + \rho \) has the required properties.

Finally, in the general case where \( \nu \) need not be a positive measure, let \( \nu = \nu^+ - \nu^- \) be the Jordan decomposition of \( \nu \) and write \( \nu^\pm = \lambda^\pm - \rho^\pm \). Then for \( \lambda = \lambda^+ - \lambda^- \) and \( \rho = \rho^+ - \rho^- \), the required decomposition is \( \nu = \lambda + \rho \).

For uniqueness, first suppose \( \mu \) and \( \nu \) are finite. Given two decompositions \( \nu = \lambda_1 + \rho_1 = \lambda_2 + \rho_2 \), it must be that \( \lambda_1, \lambda_2, \rho_1, \rho_2 \) are all finite, so then \( \lambda = \lambda_1 - \lambda_2 \) is mutually singular with \( \mu \), but also \( \lambda = \rho_2 - \rho_1 \ll \mu \), so \( \lambda = 0 \), which implies \( \lambda_1 = \lambda_2 \) and \( \rho_1 = \rho_2 \). Then \( d(\rho_1 - \rho_2) = (f_1 - f_2) \, d\mu = 0 \), so \( f_1 - f_2 \) is zero \( \mu \)-a.e.

In the general case of \( \sigma \)-finite measures, we use the same idea of partitioning the space into subsets on which the measures are finite. \( \square \)

**Definition 3.2.8** (Lebesgue decomposition). A decomposition \( \nu = \lambda + \rho \) as in the theorem is a **Lebesgue decomposition** for \( \nu \).

**Definition 3.2.9** (Radon-Nikodym derivative). Let \( \mu \) be a \( \sigma \)-finite positive measure and \( \nu \) be a \( \sigma \)-finite signed measure with \( \nu \ll \mu \). The function \( f \) for which \( d\nu = f \, d\mu \) (unique up to modification on a set of measure zero) 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 \(\nu : \mathcal{A} \rightarrow \mathbb{C}\) such that

1. \(\nu(\emptyset) = 0\);
2. 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 measures \(\nu_r, \nu_i\) defined by \(\nu_r(A) = \text{Re} \nu(A)\) and \(\nu_i = \text{Im} \nu(A)\). 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** (Lebesgue-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 \(\nu\) has a unique Lebesgue decomposition \(\nu = \lambda + \rho\) with \(\lambda \perp \mu\) and \(\rho \ll \mu\). Moreover, there is a Radon-Nikodym derivative \(f \in L^1(\mu)\) for which \(d\rho = fd\mu\), and \(f\) is unique \(\mu\)-a.e.

**Proof.** This follows from applying Lebesgue-Radon-Nikodym to \(\nu_r\) and \(\nu_i\).

**Definition 3.3.3** (Total variation for complex measures). Let \(\nu\) be a complex measure on \((X, \mathcal{A})\). The total variation of \(\nu\) is
\[
|\nu|(A) = \sup \left\{ \sum_{n=1}^{\infty} |\nu(A_n)| \mid A_n \in \mathcal{A} \text{ pairwise disjoint and } \bigcup_{n=1}^{\infty} A_n = A \right\}.
\]

**Remark 3.3.4.** For signed measures, this agrees with our earlier definition of the total variation (Homework B1 Problem 2). However, for complex measures, we need not have \(|\nu| = |\nu_r| + |\nu_i|\).

**Proposition 3.3.5.** If \(\nu\) is a complex measure on \((X, \mathcal{A})\), then \(|\nu|\) is a finite positive measure.

**Proof.** The finite positive measure \(\mu = |\nu_r| + |\nu_i|\) satisfies \(\nu \ll \mu\), so there is a Radon-Nikodym derivative \(f \in L^1(\mu)\) 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 partition of \(A\) into countably many measurable sets. Then
\[
\sum_{n=1}^{\infty} |\nu(A_n)| = \sum_{n=1}^{\infty} \left| \int_{A_n} f \, d\mu \right| \leq \sum_{n=1}^{\infty} \int_{A_n} |f| \, d\mu = \int \left( \sum_{n=1}^{\infty} |f| \chi_{A_n} \right) \, d\mu = \int_A |f| \, d\mu,
\]
for \(A \in \mathcal{A}\).
where in the second to last equality, we use the monotone convergence theorem.

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

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

Let

$$B = \bigcup_{k=1}^{N} B_k, \quad A_0 = A \cap (X \setminus B), \quad A_i = A \cap B_i \quad (i = 1, \ldots, N).$$

Then $A_0, \ldots, A_N$ partition $A$ into pairwise disjoint measurable sets, and

$$\sum_{i=0}^{N} |\nu(A_i)| = \sum_{i=0}^{N} \left| \int_{A_i} (f - s) \, d\mu + \int_{A_i} s \, d\mu \right| \geq \sum_{i=0}^{N} \left| \int_{A_i} s \, d\mu \right| - \int |f - s| \, d\mu \geq \sum_{i=0}^{N} \int_{A_i} |s| \, d\mu - \epsilon \geq \int_{A} |f| \, d\mu - 2\epsilon,$$

where in the second to last inequality, we use the fact that $s$ is constant on each $A_i$. Since $\epsilon$ is arbitrary, it follows that

$$|\nu|(A) \geq \int_{A} |f| \, d\mu.$$
4 MORE ON $L^p$ SPACES

4.1 BOUNDED LINEAR MAPS AND DUAL SPACES

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

**Definition 4.1.1** (Bounded linear map). Let $X$ and $Y$ be normed vector spaces over $\mathbb{F}$ and let $T : X \to Y$ be a linear map. We say that $T$ is *bounded* if there exists a constant $C \geq 0$ such that $\|T(x)\| \leq C\|x\|$ for all $x \in X$.

**Theorem 4.1.2.** 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 4.1.3.** 1. If $X$ and $Y$ are finite-dimensional normed vector spaces, then any linear map $T : X \to Y$ is bounded.

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

**Definition 4.1.4** (Operator norm). Let $T : X \to Y$ be a bounded linear map. The *operator norm* of $T$ is $\|T\| = \inf\{C \mid \|T(x)\| \leq C\|x\| \text{ for all } x \in X\}$.

**Proposition 4.1.5.**

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

**Proposition 4.1.6.** For vector spaces $X$ and $Y$ over $\mathbb{F}$, the set $L(X,Y)$ of bounded linear operators $X \to Y$ is a vector space over $\mathbb{F}$, and the operator norm is a norm on $L(X,Y)$.

**Proposition 4.1.7.** If $T \in L(X,Y), S \in L(Y,Z)$, then $S \circ T \in L(X,Z)$ and $\|S \circ T\| \leq \|S\| \cdot \|T\|$.

**Proposition 4.1.8.** Let $X$ be a normed vector space and let $Y$ be a Banach space. Then $L(X,Y)$ is a Banach space.
4.1 Bounded linear maps and dual spaces

**Proof.** Let \((T_n)\) be a Cauchy sequence in \(L(X,Y)\), so for any \(\epsilon > 0\), there exists \(N\) such that for all \(n,m \geq N\), \(\|T_n - T_m\| < \epsilon\). Then for each \(x \in X\), the sequence \((T_n(x))\) is Cauchy in \(Y\), so it has a limit. Hence define a function \(T : X \rightarrow Y\) pointwise by

\[
T(x) = \lim_{n \to \infty} T_n(x).
\]

To see that \(T_n \rightarrow T\) in operator norm, fix \(\epsilon > 0\) and let \(x \in X\) with \(\|x\| \leq 1\). Then there exists \(N\) such that for all \(n \geq N\), we have

\[
\|(T_n - T)(x)\| \leq \|(T_n - T_N)(x)\| + \|(T_N - T)(x)\| \leq \|T_n - T_N\| + \|(T_N - T)(x)\| < 2\epsilon,
\]

hence \(\|T_n - T\| \leq 2\epsilon\).

**Definition 4.1.9** (Dual space). Let \(X\) be a normed vector space over \(\mathbb{F}\). The space \(L(X,\mathbb{F})\) of bounded linear functionals is the *dual space* of \(X\), denoted \(X^*\).

By construction, \(X^*\) is a Banach space.

**Theorem 4.1.10** (Hahn-Banach). Let \(X\) be a normed vector space over \(\mathbb{F}\) and let \(M \subset X\) be a linear subspace. Let \(f : M \rightarrow \mathbb{F}\) be a bounded linear functional. Then there exists a bounded linear functional \(F : X \rightarrow \mathbb{F}\) such that \(F|_M = f\) and \(\|F\| = \|f\|\).

**Proof.** First suppose \(\mathbb{F} = \mathbb{R}\). If \(\|f\| = 0\), then \(f = 0\) on \(M\), so we can take \(F = 0\) on \(X\). Otherwise, without loss of generality, \(\|f\| = 1\). Let \(Z\) be the set of pairs \((U,g)\), where \(U \subset X\) is a subspace with \(M \subset U\) and \(g : U \rightarrow \mathbb{R}\) is a bounded linear functional with \(g|_M = f\) and \(\|g\| = \|f\| = 1\); then \(Z\) is non-empty, as \((M,f) \in Z\). A partial order can be defined on \(Z\) by saying that \((U,g) \leq (U',g')\) if \(U \subset U'\) and \(g'|_U = g\). With this partial order, every chain in \(Z\) has an upper bound, so by Zorn’s lemma, \(Z\) has a maximal element \((V,F)\). We claim that \(V = X\).

Suppose otherwise, and let \(x_0 \in X\setminus V\), without loss of generality with \(\|x_0\| = 1\). Take \(W\) to be the span of \(V\) and \(x_0\). We wish to extend \(F\) to some \(G : W \rightarrow \mathbb{R}\) by taking \(G(x_0) = \alpha\) for some \(\alpha\) such that \(\|G\| = 1\) and extending linearly. To see that such an \(\alpha\) exists, note that we already have \(\|F\| = 1\) on \(V\), so we only require that \(\|G(v + tx_0)\| \leq \|v + tx_0\|\), or equivalently,

\[
-\|v + tx_0\| \leq F(v) + t\alpha \leq \|v + tx_0\|.
\]

We already have this inequality for \(t = 0\) independent of \(\alpha\), and for \(t \neq 0\), rearranging tells us that we must satisfy

\[
-\|u + x_0\| - F(u) \leq \alpha \leq \|u + x_0\| - F(u)
\]

for all \(u \in V\). It suffices to show that \(-\|u_1 + x_0\| - F(u_1) \leq \|u_2 + x_0\| - F(u_2)\) for all \(u_1, u_2 \in V\). Indeed, we have

\[
F(u_2) - F(u_1) = F(u_2 - u_1) \leq \|u_2 - u_1\| = \|(u_2 + x_0) - (u_1 + x_0)\| \leq \|u_2 + x_0\| + \|u_1 + x_0\|,
\]

so the required \(\alpha\) exists. We thus have a contradiction to the maximality of \((V,F)\).

If \(\mathbb{F} = \mathbb{C}\), 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 (\(\mathbb{C}\)-linear)
functional \( F : X \to \mathbb{C} \), it is immediate that \( u = \text{Re} \, F \) is an \( \mathbb{R} \)-linear functional. Conversely, if we have an \( \mathbb{R} \)-linear functional \( u : X \to \mathbb{R} \), we obtain the corresponding \( \mathbb{C} \)-linear functional
\[
F(x) = u(x) - iu(ix).
\]
Looking at operator norms, for all \( x \in X \),
\[
|u(x)| = |\text{Re} \, F(x)| \leq |F(x)| \leq \|F\| \|x\|
\]
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 \( \|u\| = \|F\| \).
Returning to the theorem, let \( f : M \to \mathbb{C} \) be a \( \mathbb{C} \)-linear functional and \( v = \text{Re} \, f : M \to \mathbb{R} \). Then \( \|f\| = \|v\| \), and by Hahn-Banach for \( \mathbb{R} \)-linear functionals, we can extend \( v \) to an \( \mathbb{R} \)-linear function \( u : X \to \mathbb{R} \) with \( \|u\| = \|v\| = \|f\| \). Define \( F : X \to \mathbb{C} \) from \( u \) as above. Then \( F \) is a \( \mathbb{C} \)-linear functional with \( \|F\| = \|u\| = \|f\| \), and \( F \) extends \( f \).

**Corollary 4.1.11.** 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.** The result is clear for \( x = 0 \). Otherwise, if \( \|f\| \leq 1 \), 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 other direction, let \( M = \mathbb{F}x \) 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 \( g \) to a functional \( f \in X^* \) with \( \|f\| = 1 \). Then \( |f(x)| = |g(x)| = \|x\| \).

**Remark 4.1.12.** In general, \( \|f\| = \sup\{|f(x)| \mid x \in X \text{ with } \|x\| \leq 1\} \), but the supremum is not generally attained as a maximum if \( X \) is infinite-dimensional.

**Definition 4.1.13** (Reflexive space). The bilinear form \( X \times X^* \to \mathbb{F} \) given by \( \langle x, f \rangle = f(x) \) defines an injection \( ev : X \to X^{**} \) by \( x \mapsto \langle x, \_ \rangle \). We say that \( X \) is reflexive if \( ev \) is an isomorphism.

Since the dual of any normed space is a Banach space, all reflexive spaces are Banach.

**Example 4.1.14.** Every finite-dimensional normed vector space is reflexive.

### 4.2 THE DUAL OF \( L^p \)

Let \( (X, \mathcal{A}, \mu) \) be a measure space, \( 1 \leq p \leq \infty \), and \( q \) be the conjugate exponent of \( p \). For \( g \in L^q(\mu) \), define
\[
\Phi_g : L^p(\mu) \to \mathbb{C}, \quad \Phi_g(f) = \int fg \, d\mu.
\]
By Hölder’s inequality, \( \Phi_g \) is a bounded linear functional on \( L^p(\mu) \).
Theorem 4.2.1. 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 $\Phi \in (L^p)^\ast$ 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 countably 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.

$$
\Phi(s) = \int s \, g \, d\mu
$$

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

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

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

$$
\Phi(h) = \lim_{n \to \infty} \Phi(s_n) = \lim_{n \to \infty} \int s_n g \, d\mu = \int h g \, d\mu,
$$

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

$$
\limsup_{n \to \infty} \left| \int s_n g \, d\mu - \int h g \, d\mu \right| \leq \limsup_{n \to \infty} \int |s_n - h||g| \, d\mu = \limsup_{n \to \infty} \|s_n - h\|_\infty \int |g| \, d\mu = 0.
$$

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
\( f_n = \overline{\nu}[g]^{q-1} \chi_{E_n} \in L^\infty \subset L^p \). We can now apply \( \Phi \) to get
\[
\int_{E_n} |g|^q \, d\mu = \left| \int_{E_n} \overline{\nu}[g]^{q-1} g \, d\mu \right|
\leq \| \Phi \| \| f_n \|_p
\leq \| \Phi \| \left( \int_{E_n} |g|^{(q-1)p} \right)^{1/p}
\leq \| \Phi \| \left( \int_{E_n} |g|^q \right)^{1/p},
\]
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 \( \| g \|_q \leq \| \Phi \| = \| \Phi \| \leq \| g \|_q \). 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 \( E_n \to x \) 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 \( \int f g \, d\mu = \int f \tilde{g} \, d\mu \) for all \( f \in L^p \), so equivalently, \( \int f : (g - \tilde{g}) \, d\mu \) 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. \( \square \)

**Remark 4.2.2.**
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 4.2.3 (Basic 5B-covering lemma).** Let \((X, d)\) be a metric space and let \( B \) be a collection of open balls in \( X \) with uniformly bounded radii. Then there exists a collection \( \tilde{B} \) of pairwise disjoint balls with
\[
\bigcup_{B \in B} B \subset \bigcup_{B \in \tilde{B}} 5B.
\]

**Proof.** See Homework B4 Problem 3. \( \square \)

### 4.3 THE HARDY-LITTLEWOOD MAXIMAL FUNCTIONS

**Definition 4.3.1 (Locally integrable function).** A measurable function \( f : \mathbb{R}^n \to \mathbb{C} \) is **locally integrable** if \( \int_B |f| \, d\lambda_n < \infty \) for all balls \( B \subset \mathbb{R}^n \), or equivalently, \( \int_K |f| \, d\lambda_n < \infty \) for all \( K \subset \mathbb{R}^n \) compact. The space of locally integrable functions is denoted \( L^1_{\text{loc}}(\mathbb{R}^n) \).

**Notation.** For \( f \in L^1_{\text{loc}}, \) write
\[
\int_B |f| \, d\lambda_n = \frac{1}{\lambda_n(B)} \int_B |f| \, d\lambda_n.
\]
If \( f \in L^1_{\text{loc}} \), we define the centered Hardy-Littlewood maximal function by

\[
(Hf)(x) = \sup_{r > 0} \int_{B(x, r)} |f| \, d\lambda_n
\]

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 4.3.2.** 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 \( \int_B |f| \geq \alpha \). Then \( (Mf)(y) > \alpha \) for all \( y \in B \), and so \( B \subset \{Mf > \alpha\} \).

**Remark 4.3.3.** 1. For \( f \in L^1_{\text{loc}} \), we have \( |f| \leq Hf \leq Mf \) almost everywhere, but \( Mf \) is typically not much larger than \( |f| \). For \( f \in L^p \) with \( 1 < p \leq \infty \), we get

\[
\|Mf\|_p \leq C(n, p)\|f\|_p.
\]

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 4.3.4 (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 \( B = B(a, r) \in B \), then \( \alpha < \int_B |f| \) means that
\[
\lambda_n(B) = c_n r^n \leq \frac{1}{\alpha} \int_B |f| \leq \frac{1}{\alpha} \|f\|_1 < \infty,
\]
so the balls of \( B \) have uniformly bounded radii. Applying the 5\( B \)-covering lemma, there is a subfamily \( \tilde{B} \subset B \) of pairwise disjoint balls with
\[
A \subset \bigcup_{B \in \mathcal{B}} B \subset \bigcup_{B \in \mathcal{B}} 5B.
\]
Note that \( \tilde{B} \) is countable, say \( \tilde{B} = \{ B_i \mid i \in I \subset \mathbb{N} \} \). Then
\[
\lambda_n(A) \leq \lambda_n \left( \bigcup_{i \in I} 5B_i \right)
\leq \sum_{i \in I} \lambda_n(5B_i)
= 5^n \sum_{i \in I} \lambda_n(B_i)
\leq \frac{5^n}{\alpha} \sum_{i \in I} \int_{B_i} |f|
= \frac{5^n}{\alpha} \int_{\bigcup B_i} |f|
\leq \frac{5^n}{\alpha} \|f\|_1.
\]

**Lemma 4.3.5.** 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 4.3.6** (\( L^p \)-boundedness of the maximal function). Let \( n \in \mathbb{N} \) and \( 1 < p \leq \infty \). Then there exists a constant \( C = C(p, n) = 2 \left( \frac{n^k}{p^{k+1}} \right)^{1/p} \) 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
\[
g = f \circ \chi_{\{|f| \leq \alpha/2\}}, \quad h = f \circ \chi_{\{|f| > \alpha/2\}}.
\]
4.3 The Hardy-Littlewood maximal functions

Then \( f = g + h \) and \( \|g\|_{\infty} \leq \alpha/2 \). We have

\[
\|h\|_1 = \int_{\{|f| > \alpha/2\}} |f| \leq \int_{\{|f| > \alpha/2\}} |f|^{p-1} \left( \frac{\|f\|_p}{\alpha/2} \right)^{p-1} \\
\leq \left( \frac{2}{\alpha} \right)^{p-1} \int |f|^p < \infty,
\]

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,

\[
\|Mf\|_p^p \leq p \int_{0}^{\alpha} \alpha^{p-2} \cdot \frac{5^n}{\alpha} \left( \int_{\{|f| > \alpha/2\}} |f| \right) d\alpha \\
= 2p \cdot 5^n \int_{0}^{\alpha} \alpha^{p-2} \int_{\mathbb{R}^n} \chi_{\{|f| > \alpha/2\}} |f| d\alpha d\lambda_n \\
= 2p \cdot 5^n \int_{\mathbb{R}^n} |f(x)| \left( \int_{0}^{\alpha} \alpha^{p-2} d\alpha \right) d\lambda_n(x) \\
= 2p \cdot 5^n \int_{\mathbb{R}^n} |f(x)| \left( \frac{\alpha^{p-1}}{p-1} \right) d\lambda_n(x) \\
= 2p \cdot \frac{5^n}{p-1} \int_{\mathbb{R}^n} |f(x)| \cdot 2^{p-1} |f(x)|^{p-1} d\lambda_n(x) \\
= \frac{2p \cdot 5^n}{p-1} \int_{\mathbb{R}^n} |f|^p = \frac{2p \cdot 5^n}{p-1} \|f\|_p^p.
\]
5 DIFFERENTIABILITY

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 \( \mathbb{R} \{0,1\} \).

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

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

and

\[
T f(x) = \limsup_{r \to 0^+} T_r f(x) \in [0, \infty].
\]

We must show that \( T f(x) = 0 \) 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, \( T g \equiv 0 \). Let \( h_k = f - g_k \). Then \( f = g_k + h_k \) with \( \|h_k\|_1 < 1/k \). We have \( T_r f \leq T_r g_k + T_r h_k \), and so \( T f \leq Th_k \). This gives

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

\[
\leq \int_{B(x,r)} |h_k| \, d\lambda_n + |h_k(x)|
\]

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

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

\[
\frac{\lambda_n(A(k, \alpha))}{\alpha/2} \leq \frac{5^n}{\alpha/2} \|h_k\|_1 + \frac{1}{\alpha/2} \|h_k\|_1 = \frac{12(5^n + 1)}{k \alpha} \to 0
\]
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.

\[\square\]

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

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

for almost every $x$.

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

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

\[\square\]

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

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

**Proof.**

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

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

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

\[\square\]
Let $E \subseteq \mathbb{R}^n$ 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 $E \subseteq \mathbb{R}^n$ 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 $r \to 0$.

**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,

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

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

\[
F_k \subset V_\varepsilon = \bigcup_{x \in F_k} B_k \subset \bigcup_{m=1}^{\infty} 5B_m.
\]

Then

\[
\lambda_n(F_k) \leq \lambda_n(V_\varepsilon) \leq \sum_{m=1}^{\infty} \lambda_n(5B_m)
= 5^n \sum_{m=1}^{\infty} \lambda_n(B_m)
\leq 5^n k \sum_{n=1}^{\infty} \mu(B_m) = 5^n k \mu \left( \bigcup_{m=1}^{\infty} B_m \right)
\leq 5^n k \mu(U_\varepsilon) \leq 5^n k \varepsilon.
\]

Letting \( \varepsilon \to 0 \) 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”.

**Definition 5.2.1** (Functions of bounded variation). A function \( F : \mathbb{R} \to \mathbb{C} \) has **bounded variation** (is a \( \text{BV-function} \)) 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 \( \text{BV} \)-functions on \( \mathbb{R} \) is denoted by \( \text{BV}(\mathbb{R}) \), or simply \( \text{BV} \) when understood.

If \( F \in \text{BV} \) 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. \( F \in \text{BV} \) if and only if \( \text{Re} \, F \in \text{BV} \) and \( \text{Im} \, F \in \text{BV} \).  
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 \( F \in \text{BV} \), then \( F(x^+) = \lim_{y \to x^+} F(y) \) and \( F(x^-) = \lim_{y \to x^-} F(y) \) exist for all \( x \in \mathbb{R} \). Furthermore, \( F(+\infty) \) and \( F(-\infty) \) also exist.  
4. If \( F \in \text{BV} \), then \( F \) has at most countably many discontinuities.  

**Proof.**  
1. trivial  
2. Suppose \( F = G - H \) for bounded increasing functions \( G, H \). Let \( x_0 < x_1 < \cdots < x_n \). Then  
   \[
   \sum_{k=1}^{n} |F(x_k) - F(x_{k-1})| 
   \leq \sum_{k=1}^{n} |G(x_k) - G(x_{k-1})| + \sum_{k=1}^{n} |H(x_k) - H(x_{k-1})| 
   = \sum_{k=1}^{n} G(x_k) - G(x_{k-1}) + \sum_{k=1}^{n} H(x_k) - H(x_{k-1}) 
   = G(x_n) - G(x_0) + H(x_n) - H(x_0) 
   \leq 2M_G + 2M_H, 
   \]
   where \( |G| \leq 2M_G \) and \( |H| \leq 2M_H \).  
Conversely, let \( F : \mathbb{R} \to \mathbb{R} \) be a BV-function and consider \( T_F : \mathbb{R} \to \mathbb{R} \). If \( x \leq y \), then  
   \[
   T(x) + |F(y) - F(x)| \leq T_F(y), 
   \]
so  
   \[
   T_F(x) - F(x) \leq T_F(y) - F(y), \quad T_F(x) + F(x) \leq T_F(y) + F(y). 
   \]
Thus \( T_F - F \) and \( T_F + F \) are both bounded increasing functions, and \( F \) is half their difference.  
3. This is true for all bounded increasing functions, hence for all BV-functions.  
4. This is true for all bounded increasing functions (standard exercise), hence for all BV-functions.  

**Definition 5.2.3.** A function \( F \in \text{BV} \) is called a normalized BV-function if \( F \) is right-continuous, i.e. \( F(x^+) = F(x) \) for all \( x \in \mathbb{R} \), and \( F(-\infty) = 0 \).  

The set of normalized BV-functions is denoted \( \text{NBV} \).  

**Lemma 5.2.4.** If \( F \in \text{NBV} \), then \( T_F \in \text{NBV} \).
Proof. Since $T_F$ is a bounded increasing function, $T_F \in BV$. Let $\epsilon > 0$ and $b \in \mathbb{R}$ be arbitrary. We can find $x_0 < \cdots < x_n \leq b$ such that

$$T_F(b) - \epsilon \leq \sum_{i=1}^{n} |F(x_i) - F(x_{i-1})| \leq T_F(b).$$

Then

$$T_F(x_k) - \epsilon \leq \sum_{i=1}^{k} |F(x_i) - F(x_{i-1})| \leq T_F(x_k)$$

for all $k \leq n$. To show that $T_F(-\infty) = 0$, we pick $b = 0$. Then for $k = 0$, we have

$$T_F(x_0) - \epsilon \leq 0,$$

so $0 \leq T_F(-\infty) \leq T_F(x_0) \leq \epsilon$. It remains to show that $T_F$ is right continuous.

To show that $T_F(x^+) = T_F(x)$ at $x \in \mathbb{R}$, let $b = x + 1$. Throwing the point $x$ into the sequence, we only increase the sum, but keep it below $T_F(b)$, so without loss of generality, $x$ is one of the points $x_0, \ldots, x_n$. Since $F(x^+) = F(x)$, we can find $h \in (0, 1)$ so that $|F(x+h) - F(x)| < \epsilon$. Without loss of generality, we can assume $x+h$ is in the sequence and is the next point after $x$. If $x_k = x$, then

$$T_F(x) - \epsilon \leq \sum_{i=1}^{k} |F(x_i) - F(x_{i-1})| \leq T_F(x),$$

so

$$\left| T_F(x) - \sum_{i=1}^{k} |F(x_i) - F(x_{i-1})| \right| \leq \epsilon.$$

Similarly,

$$\left| T_F(x+h) - \sum_{i=1}^{k+1} |F(x_i) - F(x_{i-1})| \right| \leq \epsilon.$$

The two sums differ precisely by $|F(x+h) - F(x)| < \epsilon$, so $|T_F(x+h) - T_F(x)| < 3\epsilon$. \qed

**Theorem 5.2.5** (Complex measures on $\mathbb{R}$). $F \in NBV$ if and only if there exists a complex Borel measure $\mu$ on $\mathbb{R}$ such that $F = F_\mu$. Moreover, for a given $F \in NBV$, the corresponding complex measure $\mu$ is unique.

Proof. First, we show that if $\mu$ is a complex Borel measure, then $F_\mu \in NBV$. Since $F_{\alpha \mu + \beta \nu} = \alpha F_\mu + \beta F_\nu$, the map $\mu \mapsto F_\mu$ is linear, so by splitting into real and imaginary parts, then applying Jordan decomposition, it suffices to suppose that $\mu$ is a finite positive Borel measure. To see that $F \in BV$, let $x_0 < \cdots < x_n$ be arbitrary. Then

$$F(x_k) - F(x_{k-1}) = \mu(-\infty, x_k] - \mu(-\infty, x_{k-1}] = \mu(x_{k-1}, x_k] \geq 0.$$ 

Hence

$$\sum_{k=1}^{n} |F(x_k) - F(x_{k-1})| = \mu(x_0, x_n] \leq \mu(\mathbb{R}) < \infty.$$
For right continuity, let \( x \in \mathbb{R} \) be arbitrary and \( x_n \uparrow x \). Then \( F(x_n) = \mu(-\infty, x_n] \to \mu(-\infty, x] = F(x) \) since the intervals are decreasing and have uniformly bounded measure. The proof that \( F(-\infty) = 0 \) proceeds similarly.

Next, we show that if \( F \in NBV \), then there exists a complex Borel measure \( \mu \) for which \( F_\mu = F \). Without loss of generality, assume that \( F \) is real-valued. Since \( T_F \in NBV \) by the lemma, we can write \( F = (T_F + F) - T_F \) as a difference of increasing functions, so without loss of generality, assume that \( F \) is increasing. We then attempt to define \( \mu \) by \( \mu(a, b] = F(b) - F(a) \). More specifically, we define a premeasure on the algebra generated by the \( h \)-intervals with this formula and use the Carathéodory extension theorem to get a measure. The proof that this has the required properties is in Homework B6.

Finally, uniqueness of \( \mu \) is in the same homework. \( \square \)

**Proposition 5.2.6.** \( F_\mu \) is continuous if and only if \( \mu \) has no atoms, i.e. points \( x \) for which \( \mu(\{x\}) \neq 0 \).

**Definition 5.2.7** (Absolutely continuous (AC) function). A function \( F : \mathbb{R} \to \mathbb{C} \) is absolutely continuous if for all \( \epsilon > 0 \), there exists \( \delta > 0 \) so that if \( (a_1, b_1), \ldots, (a_n, b_n) \subset \mathbb{R} \) are disjoint intervals, then

\[
\sum_{i=1}^{n} b_i - a_i < \delta \implies \sum_{i=1}^{n} |F(b_i) - F(a_i)| < \epsilon.
\]

Denote by \( AC \) or \( AC(\mathbb{R}) \) the set of all AC-functions on \( \mathbb{R} \).

**Proposition 5.2.8.** Let \( \mu \) be a complex Borel measure on \( \mathbb{R} \) and \( F = F_\mu \). Then \( \mu \ll \lambda_1 \) if and only if \( F \) is AC.

**Proof.** (\( \implies \)) If \( \mu \ll \lambda_1 \), then there exists \( f \in L^1(\lambda_1) \) such that \( d\mu = f \, d\lambda_1 \). Then \( d|\mu| = |f| \, d\lambda_1 \), hence \( |\mu| \ll \lambda_1 \), so for each \( \epsilon > 0 \), there exists \( \delta > 0 \) such that \( \lambda_1(A) < \delta \) implies \( |\mu|(A) < \epsilon \) for all Borel sets \( A \subset \mathbb{R} \). Suppose \( (a_1, b_1), \ldots, (a_n, b_n) \) are disjoint intervals. If \( \sum_i (b_i - a_i) < \delta \), then for \( A = \bigcup_i (a_i, b_i) \), we have \( \lambda_1(A) < \delta \), so

\[
\sum_{i=1}^{n} |F(b_i) - F(a_i)| = \sum_{i=1}^{n} |\mu(a_i, b_i)| \leq \sum_{i=1}^{n} |\mu|(a_i, b_i) = |\mu|(A) < \epsilon.
\]

Note that we used absolute continuity to conclude that \( |\mu|(\{b_i\}) = 0 \).

(\( \iff \)) That \( \mu \ll \lambda_1 \) is equivalent to saying that if \( E \subset \mathbb{R} \) is Borel and \( \lambda_1(E) = 0 \), then \( \mu(E) = 0 \). Let \( E \subset \mathbb{R} \) with \( \lambda_1(E) = 0 \) and \( \epsilon > 0 \) be arbitrary. Let \( \delta > 0 \) be as in the definition of absolute continuity of \( F \). Since \( |\mu| \) and \( \lambda_1 \) are outer regular, we can find open sets \( V_n \) containing \( E \) such that \( |\mu|(V_n) \to |\mu|(E) \) as \( n \to \infty \) and an open set \( U \supset E \) with \( \lambda_1(U) < \delta \). Define \( U_n = V_n \cap U \).

Then \( U_n \) is open, \( V_n \supset U_n \supset E \), and \( |\mu|(U_n) \to |\mu|(E) \). Furthermore, \( \lambda_1(U_n) < \delta \) for each \( n \in \mathbb{N} \). Then

\[
|\mu(U_n) - \mu(E)| = |\mu(U_n \backslash E)| = |\mu(U \backslash E)| = |\mu|(U_n) - |\mu|(E) \to 0,
\]

so \( \mu(U_n) \to \mu(E) \). Since \( U_n \) is open, we can write \( U_n = \bigcup_i (a_i, b_i) \) with the \( (a_i, b_i) \) pairwise disjoint. Moreover,

\[
\sum_{i=1}^{\infty} (b_i - a_i) = \lambda_1(U_n) < \delta.
\]

57
Proof. Let $F$ for a general the set of $x$. Define $G$ measure zero. redefining $F$ for almost every $x,x$. Then $F$ is continuous, so $\mu$ has no atoms. Hence $|\mu(U_n)| \leq \epsilon$ for all $n$, so $|\mu(E)| \leq \epsilon$.

\begin{remark}
AC and BV are not the same class of functions. For example, $F(x) = \sin x$ is in AC, since it is Lipschitz, but $F$ is not in BV.
\end{remark}

\begin{lemma}
Let $[a, b] \subset \mathbb{R}$ be a finite interval. Then $AC[a, b] \subset BV[a, b]$.
\end{lemma}

\begin{proof}
Let $F \in AC[a, b]$ and pick $\delta > 0$ for $\epsilon = 1$ as in the definition. To show that $F \in BV[a, b]$, let $a \leq x_0 < \cdots < x_n \leq b$. By adding points spaced $\delta/2$ apart, we may assume without loss of generality that $x_{k+1} - x_k < \delta$ for all $k$ and that the endpoints are in the collection. Greedily group the points starting from the left so that each group has size less than $\delta$, so the total number of groups is at most $N = \lceil 2(b - a)/\delta \rceil + 1$. Then

$$\sum_{k=0}^{n-1} |F(x_{k+1}) - F(x_k)| = \sum_{i=1}^{N} \sum_{\text{group of points}} |F(x_{k+1}) - F(x)| \leq N.$$ 

Hence $F \in BV[a, b]$.
\end{proof}

5.3 THE FUNDAMENTAL THEOREM OF CALCULUS

\begin{theorem}
Let $F : \mathbb{R} \to \mathbb{R}$ be increasing. Then $F'(x)$ exists almost everywhere.
\end{theorem}

\begin{proof}
Without loss of generality, suppose $F$ is bounded by fixing a finite interval $[-n, n]$ and redefining $F$ outside $[-n, n]$ by making $F$ constant on each outward ray. We can recover the result for a general $F$ by letting $n \to \infty$, noting that a countable union of sets of measure zero is still of measure zero.

Define $G(x) = F(x^+)$ for $x \in \mathbb{R}$. Then $G$ is increasing, bounded, and right continuous. Moreover, the set of $x \in \mathbb{R}$ where $H(x) = G(x) - F(x) \neq 0$ is countable. There exists a finite Borel measure $\mu$ on $\mathbb{R}$ such that $\mu([a, b]) = G(b) - G(a)$. Then

$$G(x + h) - G(x) = \begin{cases} \mu(x, x+h] & h > 0, \\ -\mu(x+h, x] & h < 0. \end{cases}$$

The sets $(x, x+r]$ and $(x-r, x]$ shrink nicely to $x$ as $r \to 0$. It follows that if $d\mu = f \, d\lambda_1 + d\nu$, then

$$\lim_{h \to 0^+} \frac{G(x+h) - G(x)}{h} = \lim_{h \to 0^+} \frac{\mu(x, x+h]}{\lambda_1(x, x+h]} = f(x)$$

for almost every $x$ and

$$\lim_{h \to 0^-} \frac{G(x+h) - G(x)}{h} = \lim_{h \to 0^-} \frac{\mu(x-h, x]}{\lambda_1(x-h, x]} = f(x)$$

58
for almost every $x$. Hence $G'(x)$ exists (and is $f(x)$) for almost every $x$. We claim that $H'(x)$ exists and $H'(x) = 0$ for almost every $x \in \mathbb{R}$. From this, it will follow that $F'(x) = G'(x) - H'(x)$ exists for almost every $x$ and $F'(x) = f(x)$ for such $x$. Let $E = \{x_n\}$ be an enumeration of the points where $H(x) \neq 0$. Define a measure $\rho = \sum_n H(x_n) \cdot \delta_{x_n}$. This is a finite Borel measure which is singular with respect to Lebesgue measure. Then

\[
\frac{|H(x+h) - H(x)|}{h} \leq \frac{H(x+h) + H(x)}{|h|} \\
\leq \frac{2\rho[x - |h|, x + |h|]}{|h|} \\
= 2 \cdot \frac{\rho[x - |h|, x + |h|]}{\lambda_1[x - |h|, x + |h|]} \rightarrow 0
\]

as $h \rightarrow 0$ for almost every $x$ by Lebesgue differentiation. This completes the proof.

\[\square\]

**Corollary 5.3.2.** Let $F \in \text{NBV}$. Then $F'(x)$ exists for almost every $x \in \mathbb{R}$. Moreover, if $\mu$ is the unique complex Borel measure such that $F\mu = F$ and $d\mu = f d\lambda_1 + dv$ is the Lebesgue decomposition of $\mu$, then $F'(x) = f(x)$ for almost every $x \in \mathbb{R}$. In particular, $F' \in L^1(\lambda_1)$. Finally, if $F \in AC \cap \text{NBV}$, then

\[F(x) = \int_{-\infty}^x F'(t) \, dt = \int F' \, d\lambda_1 \]

for each $x \in \mathbb{R}$.

**Proof.** Everything but the last part follows from the theorem if $F$ is increasing. In the general case, we split $F$ into its real and imaginary parts and write each as differences of increasing functions in $\text{NBV}$. If $F \in AC \cap \text{NBV}$, then $\mu \ll \lambda_1$, so $d\mu = f d\lambda_1 = F' d\lambda_1$. Hence for each $x \in \mathbb{R}$, we have

\[F(x) = F\mu(x) = \mu((-\infty, x]) = \int f \, d\lambda_1 = \int F' \, d\lambda_1. \]

\[\square\]

**Theorem 5.3.3** (Fundamental theorem of calculus for Lebesgue integrals). Let $[a,b] \subset \mathbb{R}$ be a finite interval and $F : [a,b] \rightarrow \mathbb{C}$ be a function. The following are equivalent:

1. $F \in AC[a,b]$;
2. $F$ is differentiable almost everywhere on $[a,b]$ with $F' \in L^1[a,b]$, and

\[F(x) - F(a) = \int_a^x F'(t) \, dt\]

for each $x \in [a,b]$;
3. there exists $f \in L^1[a,b]$ such that

\[F(x) - F(a) = \int_a^x f(t) \, dt\]

for $x \in [a,b]$. 

59
5.3 The fundamental theorem of calculus

Proof. By subtracting a constant, we may assume that $F(a) = 0$. We may extend $F$ to $\mathbb{R}$ by setting $F(x) = 0$ for $x < a$ and $F(x) = F(b)$ for $x > b$.

(1) $\implies$ (2) Since $F \in AC[a, b]$, we have $F \in BV[a, b]$, and so for the extended function $F \in AC \cap BV$, we have $F \in NBV \cap AC$. Then the corollary applies.

(2) $\implies$ (3) We take $f = F'$.

(3) $\implies$ (1) We extend $f$ to be 0 outside $[a, b]$. Then $f \in L^1(\mathbb{R})$, so we can define a complex Borel measure $d\mu = f \, d\lambda_1$ on $\mathbb{R}$ with $\mu \ll \lambda_1$. Then $F = F_\mu$.

$\square$
6 FUNCTIONAL ANALYSIS

6.1 LOCALLY COMPACT HAUSDORFF SPACES

Let $X$ be a locally compact Hausdorff space.

**Proposition 6.1.1.**

1. If $K \subset U \subset X$ with $K$ compact and $U$ open, then there exists an open set $V$ with $V \subset K \subset U$.

2. (Urysohn) If $K \subset U \subset X$ with $K$ compact and $U$ open, then there exists a continuous function $f : X \to [0, 1]$ such that $f = 1$ on $K$ and $	ext{supp } f \subset U$ is compact. We may write $K \prec f \prec U$.

3. (Tietze extension) If $K \subset X$ is compact and $f : K \to \mathbb{C}$ is continuous, then there exists $F : X \to \mathbb{C}$ continuous with compact support and $F|_{K} = f$.

4. (Partitions of unity) Let $K \subset X$ be compact and $\mathcal{U} = \{U_1, \ldots, U_n\}$ be an open cover of $K$. Then there exist continuous functions $h_1, \ldots, h_n : X \to [0, 1]$ such that $h_i \prec U_i$, $0 \leq h_1 + \cdots + h_n \leq 1$, and $h_1 + \cdots + h_n = 1$ on $K$. The functions $h_1, \ldots, h_n$ form a partition of unity on $K$ subordinate to $\mathcal{U}$.

*Proof of 4.* By a covering argument, we can find compact sets $F_1, \ldots, F_n$ such that $F_i \subset U_i$ for $i = 1, \ldots, n$ and $K \subset F_1 \cup \cdots \cup F_n$. By Urysohn, we may pick continuous functions $g_i$ such that $F_i \prec g_i \prec U_i$. Then $U = \{g_1 + \cdots + g_n > 0\} \supset K$ is open, so by Urysohn again, we can find $f$ with $K \prec f \prec U$. If $g_0 = 1 - f$, then $g_0|_{K} = 0$ and $g_0|_{X \setminus U} = 1$. Then $g = g_0 + \cdots + g_n > 0$ on $X$, so we can define $h_i = g_i / g$ for $i = 1, \ldots, n$. 

The set $C_c(X)$ of continuous functions $f : X \to \mathbb{C}$ with compact support is a vector space over $\mathbb{C}$ with uniform norm (or supremum norm)

$$
\|f\|_{\infty} = \sup_{x \in X} |f(x)| < \infty.
$$

In general, $C_c(X)$ is not a Banach space.

**Definition 6.1.2** (Vanishing at infinity). A continuous function $f : X \to \mathbb{C}$ vanishes at infinity if for all $\epsilon > 0$, there exists a compact set $K$ such that $|f(x)| < \epsilon$ for $x \in X \setminus K$.

The set of continuous functions vanishing at infinity is denoted $C_0(X)$.

**Theorem 6.1.3.** ($C_0(X), \| \cdot \|_{\infty}$) is a Banach space, and $C_c(X)$ is dense in $C_0(X)$.

**Remark 6.1.4.** The Alexandroff one-point compactification of $X$ is the space $\hat{X} = X \cup \{\infty\}$, with topology defined by saying that $U \subset \hat{X}$ is open if one of the following holds:

1. $U \subset X$ and $U$ is open in $X$;
2. $\infty \in U$ and $\hat{X} \setminus U \subset X$ is compact.

If $f \in C_0(X)$, then $f$ uniquely extends to $\hat{X}$ by setting $f(\infty) = 0$, and conversely, if $f : \hat{X} \to \mathbb{C}$ is continuous with $f(\infty) = 0$, then $f$ restricts to a function in $C_0(X)$.
**Definition 6.1.5** (Complex Radon measure). A *complex Radon measure* $\mu$ is a complex Borel measure whose total variation measure $|\mu|$ is a (positive) Radon measure.

**Theorem 6.1.6** (Homework B2 Problem 1). Let $\mathcal{M}(X)$ denote the set of complex Radon measures on $X$. For $\mu \in \mathcal{M}(X)$, define $\|\mu\| = |\mu|(X)$. Then $\mathcal{M}(X)$ is a vector space and $\|\| \|$ is a norm.

**Definition 6.1.7** (Positive linear functional). A linear functional $I : C_c(X) \to \mathbb{C}$ is *positive* if $I(f) \geq 0 \geq 0$ whenever $f \geq 0$.

**Theorem 6.1.8** (Riesz-Markov-Kakutani representation theorem). 1. There is an isometric isomorphism $\mathcal{M}(X) \to C_0(X)^*$ given by
$$\mu \mapsto I_\mu : C_0(X) \to \mathbb{C}, \quad I_\mu(f) = \int f \, d\mu.$$ 2. If $I : C_c(X) \to \mathbb{C}$ is a positive linear functional (not necessarily bounded), then there exists a unique positive Radon measure $\mu$ on $X$ such that $I(f) = \int f \, d\mu$ for all $f \in C_c(X)$.

*Proof of uniqueness for 2.* For uniqueness, by outer regularity, it suffices to show that $\mu(U)$ is determined by $I$ for all open sets $U \subset X$. Let $K \subset U$ be compact. Then there exists $f \in C_c(X)$ with $K \prec f \prec U$ by Urysohn, so $\chi_K \leq f \leq \chi_U$. This implies that
$$\mu(K) \leq \int f \, d\mu = I(f) \leq \mu(U).$$ Taking the supremum over all compact sets $K$, the inner regularity of $U$ tells us that
$$\mu(U) = \sup_{f \prec U} I(f).$$ For existence, the idea is to define a set function on open sets in this way, then use this on open covers to construct an outer measure. \qed

If $X$ is compact, so $C_0(X) = C(X)$. This gives us the following corollary.

**Corollary 6.1.9.** If $X$ is compact, then $C(X)^* \cong \mathcal{M}(X)$.

**Remark 6.1.10.** If $\mu$ is a positive Radon measure, then $I : f \mapsto \int f \, d\mu$ is a positive linear functional on $C_c(X)$.

### 6.2 Weak topologies

Let $X$ be a vector space over $F = \mathbb{R}$ or $F = \mathbb{C}$.

**Definition 6.2.1** (Seminorm). A *seminorm* on $X$ is a map $p : X \to [0, \infty)$ such that for all $x, y \in X$ and $\alpha \in F$, we have $p(\alpha x) = |\alpha| p(x)$ and $p(x + y) \leq p(x) + p(y)$.

**Example 6.2.2.** If $X$ is a Banach space and $f \in X^*$, then $p : x \mapsto |f(x)|$ is a seminorm on $X$.

**Proposition 6.2.3.** Let $X$ be a vector over $F$ and $\{p_\alpha\}_{\alpha \in A}$ be a family of seminorms on $X$. Call a set $U \subset X$ open if for all $x \in U$, there exist $\epsilon > 0$ and a finite set $F \subset A$ such that
$$N_{F, \epsilon}(x) = \{y \in X : p_\alpha(y - x) < \epsilon \text{ for all } \alpha \in F\} \subset U.$$ Then the open sets form a topology $\mathcal{T}$ on $X$. 62
Proof. It is certainly the case that \( \emptyset \) and \( X \) are open. Unions of open sets are clearly open. Let \( U, V \in T \), with associated \( F, G \) and \( \epsilon, \delta \). We can then take \( F \cup G \) and \( \min(\epsilon, \delta) \) to get that \( U \cap V \in T \).

**Remark 6.2.4.** Each set \( N_{F,\epsilon}(x) \) is an open set.

**Definition 6.2.5** (Weak and weak-\( \ast \) topologies). Let \( X \) be a Banach space. The weak topology on \( X \) is the topology induced by the family of seminorms \( \{ p_f \mid f \in X^* \} \) given by \( p_f(x) = |f(x)| \). The weak-\( \ast \) topology on \( X^* \) is the topology induced by the family of seminorms \( \{ q_x \mid x \in X \} \) given by \( q_x(f) = |f(x)| \).

**Remark 6.2.6.** Since \( X^* \) is a Banach space, it also has a weak topology. If \( X \) is reflexive, then the weak and weak-\( \ast \) topologies on \( X^* \) coincide.

**Lemma 6.2.7.** Let \( X \) be a Banach space. Every open set in the weak topology on \( X \) is also open in the norm topology. Similarly, every open subset of \( X^* \) in the weak-\( \ast \) topology is also open in the norm topology.

Proof. Let \( U \subset X \) be weakly open and \( x \in U \) be arbitrary. Then there exist \( \epsilon > 0 \) and \( f_1, \ldots, f_n \in X^* \) for which

\[
N_{F,\epsilon}(x) = \{ y \in X \mid |f_i(y - x)| < \epsilon \} \subset U.
\]

Let

\[
\delta = \epsilon \min(\|f_1\|, \ldots, \|f_n\|) + 1.
\]

Then \( B(x, \delta) \subset N_{F,\epsilon}(x) \subset U \), as required. The proof for the weak-\( \ast \) topology is similar.

**Remark 6.2.8.** The weak topology is more coarse than the norm topology, but is still Hausdorff. However, it is not metrizable if \( X \) has infinite dimension. Similar statements hold for the weak-\( \ast \) topology.

**Definition 6.2.9** (Directed set). A directed set is a set \( A \) with a partial order \( \leq \) such that for all \( \alpha, \beta \in A \), there exists \( \gamma \in A \) such that \( \alpha, \beta \in \gamma \).

**Example 6.2.10.**

1. With the usual order, \( \mathbb{N} \) is a directed set.

2. The set of finite subsets \( F \subset [0, 1] \) is a directed set with inclusion.

3. The family \( U_x \) of all open neighborhoods of \( x \in X \) in some topological space is a directed set with reverse inclusion.

**Definition 6.2.11** (Net). A net in a topological space \( X \) is a family \( \{ x_\alpha \}_{\alpha \in A} \) of points \( x_\alpha \in X \) indexed by a directed set \( A \).

**Definition 6.2.12** (Convergence of nets). If \( X \) is a topological space, \( \{ x_\alpha \}_{\alpha \in A} \) is a net, and \( x \in X \), we say that \( \{ x_\alpha \} \) converges to \( x \), written \( x_\alpha \to x \), if for every neighborhood \( U \subset X \), there exists \( \alpha_0 \in A \) such that \( x_\alpha \in U \) for all \( \alpha \geq \alpha_0 \).

**Example 6.2.13.** Let \( f \) be Riemann integrable on \([0, 1]\) and let \( A \) be the set of all finite partitions \( F \) of \([0, 1]\). Then the lower sums \( \{ L_F f \}_{F \in A} \) form a convergent net in \( \mathbb{R} \) which converges to the integral of \( f \).
### Proposition 6.2.14.  
1. If \( U \subset X \), then \( x \in U \) if and only if there is a net in \( U \) which converges to \( x \).

2. A map \( f : X \to Y \) of topological spaces is continuous if and only if for any convergent net in \( X \) with limit \( x \), the image net in \( Y \) converges to \( f(x) \).

### Remark 6.2.15.  
The topology on a space is uniquely determined by the convergent nets and their limits.

### Notation.  
Let \( X \) be a Banach space. If a net \( \{x_\alpha\}_{\alpha \in A} \) converges with respect to the weak topology to \( x \), we write \( x_\alpha \xrightarrow{w} x \) or \( x_\alpha \to x \).

### Proposition 6.2.16.  
\( x_\alpha \xrightarrow{w} x \) if and only if \( f(x_\alpha) \to f(x) \) for all \( f \in X^* \).

**Proof.**  
(\( \Rightarrow \)) Suppose \( x_\alpha \xrightarrow{w} x \) and let \( f \in X^* \) and \( \epsilon > 0 \) be arbitrary. Then there exists \( \alpha_0 \in A \) such that

\[
x_\alpha \in N_{(p_j)_j}(x) = \{ y \in X \mid |f(y) - x| < \epsilon \}
\]

for all \( \alpha \geq \alpha_0 \). Then \( |f(x_\alpha) - x| = |f(x_\alpha) - f(x)| < \epsilon \) for all \( \alpha \geq \alpha_0 \), so \( f(x_\alpha) \to f(x) \).

(\( \Leftarrow \)) Suppose \( f(x_\alpha) \to f(x) \) for all \( f \in X^* \). It suffices to check the convergence condition for any neighborhood of the form \( N_{(p_j)_j}(x) \) in the weak topology. If \( F = \{p_{j_1}, \ldots, p_{j_k}\} \), then \( f_i(x_\alpha) \to f_i(x) \), so there exist \( \alpha_i \in A \) such that for each \( i \), we have \( |f_i(x_\alpha) - x| < \epsilon \) for all \( \alpha \geq \alpha_i \). Since \( A \) is a directed set, there is an \( \alpha_0 \) with \( \alpha_0 \geq \alpha_i \) for all \( i = 1, \ldots, n \). We then have \( |f_i(x_\alpha) - x| < \epsilon \) for all \( i \) and \( \alpha \geq \alpha_0 \), so \( x_\alpha \xrightarrow{w} x \).

\( \Box \)

### Example 6.2.17.  
Let \( X_i \) for \( i \in I \) be topological spaces and \( = \prod_{i \in I} X_i \). If \( \pi_j : X \to X_j \) is the projection, then one can show that for a net \( \{x_\alpha\} \) in \( X \), we have \( x_\alpha \to x \) if and only if \( p_j(x_\alpha) \to p_j(x) \) for all \( j \).

### Example 6.2.18.  
Let \( (Z, \mathcal{A}, \mu) \) be a \( \sigma \)-finite measure space and \( 1 \leq p < \infty \). Then \( X = L^p(\mu) \) is a Banach space. If \( \{f_n\} \) is a sequence in \( X \), then \( f_n \xrightarrow{w} f \) if and only if \( \int f_n g d\mu = \int f g d\mu \) for all \( g \in L^q(\mu) \).

### Example 6.2.19.  
Let \( Z \) be a locally compact Hausdorff space and \( Y = \mathcal{M}(Z) \) be the set of all complex Radon measures on \( Z \). If \( X = C_0(Z) \), then by the Riesz representation theorem, \( X^* = Y \). Then \( \mu_n \xrightarrow{w} \mu \in \mathcal{M}(Z) \) if and only if \( \int f d\mu_n \to \int f d\mu \) for all \( f \in C_0(Z) \). Instead of weak-\( * \) convergence, this convergence of measures is sometimes called vague convergence.

### Example 6.2.20.  
Let \( Z = [0, 1] \) and \( \mu_n = (1/n) \sum_{k=1}^{n} \delta_{k/n} \). Then \( \mu_n \xrightarrow{w} \lambda_1 \), Lebesgue measure on \([0, 1] \).

### Theorem 6.2.21 (Banach-Alaoglu).  
Let \( X \) be a Banach space and \( B = \{ f \in X^* \mid \|f\| \leq 1 \} \) be the closed unit ball in \( X^* \). Then \( B \) is compact with respect to the weak-\( * \) topology.

**Proof.**  
For each \( x \in X \), let \( D_x = \{ z \in \mathbb{C} \mid \|z\| \leq \|x\| \} \) and \( P = \prod_x D_x \) with the product topology. By Tychonoff’s theorem, \( P \) is compact. Define a map \( I : B \to P \) by setting \( I(f) = (f(x))_{x \in X} \). Then \( I \) is a homeomorphism onto its image if \( B \) has the weak-\( * \) topology and \( I(B) \) has the subspace topology from \( P \), and \( I(B) \subset P \) is closed, hence compact. To see that \( I \) is a homeomorphism onto...
its image, \( I : B \to I(B) \) is certainly a bijection, so we must check continuity of \( I \) and \( I^{-1} \). It suffices to show that \( f_\alpha \overset{w^*}{\to} f \) in \( B \) if and only if \( I(f_\alpha) \to I(f) \) in \( I(B) \). Indeed, \( f_\alpha \overset{w^*}{\to} f \) if and only if \( f_\alpha(x) \to f(x) \) for all \( x \in X \), which happens if and only if \( I(f_\alpha) \to I(f) \) in \( I(B) \) (by the characterization of convergence of nets in the product topology). To see that \( I(B) \subset P \) is closed, consider a net in \( I(B) \), which has the form \( \{ I(f_\alpha) \}_{\alpha \in A} \), where \( \{ f_\alpha \}_{\alpha \in A} \) is a net in \( B \). Suppose \( I(f_\alpha) \to z = (z_x)_{x \in X} \). We must show that there exists \( f \in B \) such that \( z = I(f) \). Define \( f(x) = z_x \) for \( x \in X \). Then \( f \) is a linear functional on \( X \), as if \( a, b \in F \) and \( x, y \) are arbitrary, then

\[
f_\alpha(ax + by) = p_{ax + by}(I(f_\alpha)) \to p_{ax + by}(z) = az_x + by = af(x) + bf(y),
\]

and by linearity and a similar computation,

\[
f_\alpha(ax + by) = af_\alpha(x) + bf_\alpha(y) \to az_x + by = af(x) + bf(y).
\]

Since \( z \in P \), we have \( z_x = f(x) \in D_x \), so \( |f(x)| = |z_x| \leq \|x\| \) for all \( x \in X \), so \( \|f\| \leq 1 \).

**Corollary 6.2.22.** Let \( Z \) be a compact metric space and \( \{ \mu_n \} \) be a sequence of Borel probability measures on \( Z \). Then there is a subsequence \( \mu_{n_k} \) and a Borel probability measure \( \mu \) such that \( \mu_{n_k} \overset{w^*}{\rightharpoonup} \mu \).

**Proof (outline).** We regard \( \mu_n \) as complex measures with \( \|\mu_n\| = \mu_n(Z) = 1 \). By Banach-Alaoglu, there is a subsequence \( \mu_{n_k} \to \mu \) for some \( \mu \in \mathcal{M}(Z) \). (Here, we use the fact that \( C(Z) \) is separable and that the weak-* topology on the unit ball in \( C(Z)^* \) is metrizable.) One can show by integrating suitable non-negative functions that \( \mu \) is a positive measure, and \( \mu(X) = 1 \) is seen by integrating the function 1.

6.3 SOME THEOREMS IN FUNCTIONAL ANALYSIS

**Theorem 6.3.1** (Baire’s category theorem). Let \( X \) be a complete metric space.

1. If \( G_n \) for \( n \in \mathbb{N} \) are open and dense in \( X \), then \( \bigcap_n G_n \) is dense in \( X \).

2. If \( A = \bigcup_n A_n \) is a countable union of closed subsets \( A_n \) with empty interior, then \( A \) has empty interior.

**Proof.** The two statements are equivalent, so we will show the first one. Let \( D = \bigcup_1^n G_n \) and let \( B_0 = B(x, \epsilon) \) for \( x \in X \) and \( \epsilon > 0 \) be arbitrary. We inductively construct closed balls \( B_n = \overline{B(x_n, r_n)} \) such that \( 0 < r_n \leq \epsilon/2^n \) and \( B_{n+1} \subset B_n \cap G_1 \cap \cdots \cap G_{n+1} \subset B_0 \). Since \( G_1 \) is dense, \( G_1 \cap B_0 \neq \emptyset \), so we can pick \( x_1 \in G_1 \cap \text{Int } B_0 \) and \( r_1 \leq \epsilon/2 \) so that \( B_1 = \overline{B(x_1, r_1)} \subset G_1 \cap \text{Int } B_0 \). If \( B_n \) has been chosen then \( G_{n+1} \cap \text{Int } B_n \neq \emptyset \), so we can pick \( B_{n+1} \) in accordance with the given specifications. From the nesting of these closed balls, it follows that \( \{ x_n \} \) is a Cauchy sequence, so it has a limit \( x \) since \( X \) is complete, and \( x \in B_n \subset G_n \) for all \( n \).

**Remark 6.3.2.** Such sets \( A \) are often called sets of the first category or meagre.

**Example 6.3.3.** \( \mathbb{R}^n \) cannot be represented as a countable union of codimension 1 affine subspaces.

**Theorem 6.3.4** (Banach-Steinhaus). Let \( X \) be a Banach space, \( Y \) be a normed space, and \( \{ T_i \}_{i \in I} \) be a family of bounded linear operators \( T_i : X \to Y \). Then if \( \sup_i \| T_i(x) \| < \infty \) for each \( x \), we have \( \sup_i \| T_i \| < \infty \).
Proof. For \( n \in \mathbb{N} \), define
\[
A_n = \left\{ x \in X \mid \sup_{i \in I} \|T_i(x)\| \leq n \right\} = \bigcap_{i \in I} T_i^{-1}(\overline{B}(0, n)).
\]

Then \( \bigcup_n A_n = X \), so by Baire’s category theorem, one of the sets \( A_n \) has non-empty interior. Suppose that \( A_k \) has non-empty interior. Then there exists \( x_0 \in X \) and \( r > 0 \) such that \( \overline{B}(x_0, r) \subset A_k \). If \( x \in \overline{B}(0, r) \) is arbitrary, then \( x_0 + x \in \overline{B}(x_0, x) \), so
\[
\|T_i(x)\| = \|T_i(x + x_0) - T_i(x_0)\| \leq \|T_i(x + x_0)\| + \|T_i(x_0)\| \leq 2k.
\]

By homogeneity,
\[
\|T_i(x)\| \leq \frac{2k}{r} \|x\|
\]
for \( x \in X \) and \( i \in I \), so \( \|T_i\| \leq 2k/r \) for all \( i \in I \).

Corollary 6.3.5. Let \( X \) be a Banach space, \( Y \) be a normed space, and \( T_n : X \to Y \) be bounded linear operators for \( n \in \mathbb{N} \). If \( \{T_n(x)\} \) converges pointwise for each \( x \in X \), then \( \sup_n \|T_n\| < \infty \).

Corollary 6.3.6. Let \( X \) be a Banach space. Then every weakly convergent sequence in \( X \) is bounded in norm. Similarly, every sequence \( f_n \) in \( X^* \) converging in the weak-* topology is uniformly bounded in operator norm.

Proof. Each \( x \in X \) can be considered a bounded linear functional \( T_x \) on \( X^* \) with \( \|T_x\| = \|x\| \). If \( x_n \xrightarrow{w} x \), then \( f(x_n) \to f(x) \) for all \( f \in X^* \), so \( T_{x_n}(f) \to T_x(f) \). By Banach-Steinhaus, \( \sup_n \|T_{x_n}\| = \sup_n \|x_n\| < \infty \). A similar argument applies in \( X^* \), but more directly since we already have bounded linear functionals.

Theorem 6.3.7 (Open mapping theorem). Let \( X \) and \( Y \) be Banach spaces and \( T : X \to Y \) be a bounded linear operator. If \( T \) is surjective, then \( T \) is an open mapping.

Proof. We first claim that there exists \( C_0 > 0 \) such that for all \( y \in Y \) and \( \epsilon > 0 \), there exists \( x \in X \) with \( \|y - T(x)\| < \epsilon \) and \( \|x\| \leq C_0 \|y\| \). Since \( T \) is surjective, \( Y = \bigcup_k \overline{T(B_k)} \), where \( B_k = \overline{B}(0, k) \subset X \). By Baire’s category theorem, there exists \( k \) for which \( \overline{T(B_k)} \) has non-empty interior. Then there exists \( y_0 \in Y \) and \( r > 0 \) such that \( \overline{B}(y_0, r) \subset \overline{T(B_k)} \). Now let \( y \in Y \) be arbitrary with \( y \neq 0 \). Then \( y_0, y_1 = y_0 + (r/\|y\|)y \in \overline{T(B_k)} \), so there exist sequences \( \{z_n\} \) and \( \{z'_n\} \) in \( B_k \) such that \( T(z_n) \to y_0 \) and \( T(z'_n) \to y_1 \). If \( x'_n = z'_n - z_n \), then \( T(x'_n) \to (r/\|y\|)y \). Then \( x_n = (\|y\|/r)x'_n \in B_{2k\|y\|/r} \), which proves the claim.

Our second claim is that there is another constant \( C_1 > 0 \) such that for all \( y \in Y \), there exists \( x \in X \) with \( y = Tx \) and \( \|x\| \leq C_1 \|y\| \). Let \( y_0 = y \in Y \) be arbitrary, and without loss of generality suppose \( \|y\| = 1 \). By the first claim, there exists \( x_0 \in X \) such that \( \|y_0 - Tx_0\| \leq 1/2 \) and \( \|x_0\| \leq C_0 \). Let \( y_1 = y_0 - Tx_0 \). By induction, we can pick \( x_n \in X \) with \( \|y_n - Tx_n\| \leq 1/2^{n+1} \) and \( \|x_n\| \leq C_0/2^n \), and then set \( y_{n+1} = y_n - T(x_n) \). Define \( s_n = x_0 + x_1 + \cdots + x_n \). Then \( \{s_n\} \) is a Cauchy sequence, so \( s_n \to x \) for some \( x \in X \). By construction, we find that \( \|x\| \leq 2C_0 \) when \( \|y\| = 1 \) and
\[
y - T(x) = \lim_{n \to \infty} y - T(s_n) = \lim_{n \to \infty} (y_0 - T(x_0) - \cdots - T(x_n)) = \lim_{n \to \infty} y_{n+1} = 0.
\]
Finally, we prove that $T$ is an open mapping. Let $U \subset X$ be open and $y_0 \in T(U)$. There exists $x_0 \in U$ such that $y_0 = T(x_0)$. Since $U$ is open, there exists $\delta > 0$ such that $B_X(x_0, \delta) = x_0 + B_X(0, \delta) \subset U$. By the second claim, $B_Y(0, \delta/C_1) \subset T(B_X(0, \delta))$, so if $r = \delta/C_1$, then
\[ B_Y(y_0, r) = y_0 + B(0, r) = T(x_0 + B_X(0, \delta)) = T(B_X(x_0, \delta)). \]

\[ \square \]

Corollary 6.3.8. If $T$ is bijective, then $T^{-1}: Y \to X$ is a bounded linear operator.

Proof. Since $T$ is linear, $T^{-1}$ is linear. To see that $T^{-1}$ is bounded, it suffices to show that $T^{-1}$ is continuous. The preimage of an open set $U \subset X$ is $T(U) \subset Y$, which is open.

Theorem 6.3.9 (Closed graph theorem). Let $X$ and $Y$ be Banach spaces, and let $T : X \to Y$ be a linear map. Then $T$ is continuous if and only if graph $T \subset X \times Y$ is closed.

Proof. ($\implies$) Let $(z_n)$ be a sequence in graph $T$ converging to $z = (x, y) \in X \times Y$. If $z_n = (x_n, T(x_n))$, then $x_n \to x$ and $T(x_n) \to y$. Since $T$ is continuous, $T(x) = y$, so $z = (x, T(x)) \in$ graph $T$.

($\impliedby$) Suppose graph $T$ is closed. Then it is a Banach space with norm induced by the norm on $X \times Y$. Let $\pi_1$ and $\pi_2$ be the projection maps; these are bounded linear operators, and $R = \pi_1|_{\text{graph } T}$ is a bijection onto $X$. By the open mapping theorem, the inverse is a bounded linear operator, hence $T = \pi_2 \circ S$ is a bounded linear operator, so continuous.

\[ \square \]

6.4 HILBERT SPACES

Let $X$ be a vector space over $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$ with an inner product. We say that $X$ is a Hilbert space if $X$ is complete with the metric induced by the inner product.

Example 6.4.1. 1. $\mathbb{R}^n$ and $\mathbb{C}^n$ with the usual inner products.

2. If $f, g \in L^2(\mu)$, we can define
\[ \langle f, g \rangle = \int fg \, d\mu. \]

3. If $X$ has the counting measure $\mu$ on its power set, then we write $l^2(X)$ for $L^2(\mu)$.

Proposition 6.4.2. Let $X$ be a Hilbert space with inner product $\langle \ , \ \rangle$.

1. If $(x_n, y_n) \to (x, y)$, then $\langle x_n, y_n \rangle \to \langle x, y \rangle$.

2. $\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2$.

We say that $x, y \in X$ are orthogonal, written $x \perp y$, if $\langle x, y \rangle = 0$.

Proposition 6.4.3. If $x_1, \ldots, x_n \in X$ and $x_i \perp x_j$ for $i \neq j$, then $\|x_1 + \cdots + x_n\|^2 = \|x_1\|^2 + \cdots + \|x_n\|^2$. 

67
**Definition 6.4.4** (Convex set). A set \( K \subset X \) in a vector space \( X \) is **convex** if for all \( x, y \in K \) and \( \lambda \in [0,1] \), we have \((1 - \lambda)x + \lambda y \in K\).

**Theorem 6.4.5.** Let \( X \) be a Hilbert space, \( K \subset X \) be closed and convex, and \( z \in X \). Then there exists a unique \( x \in K \) such that \( \|x - z\| = \inf\{\|y - z\| \mid y \in K\} \).

**Proof.** For existence, suppose without loss of generality that \( z = 0 \). Let \( \{x_n\} \) be an infimizing sequence in \( K \), meaning that \( \|x_n\| \to d = \inf\{\|y\| \mid y \in K\} \). Then

\[
\|x_n - x_m\|^2 + 4 \left\| \frac{x_n + x_m}{2} \right\|^2 = 2\|x_n\|^2 + 2\|x_m\|^2,
\]

so

\[
\limsup_{n,m \to \infty} \|x_n - x_m\|^2 = \limsup_{n,m \to \infty} \left( 2\|x_n\|^2 + 2\|x_m\|^2 - 4 \left\| \frac{x_n + x_m}{2} \right\|^2 \right) \leq 2d^2 + 2d^2 - 4d^2 = 0,
\]

where we have used the fact that \((x_n + x_m)/2 \in K\) in the last non-trivial step. Thus \((x_n)\) is Cauchy, so \(x_n \to x\) with \(\|x\| = d\).

For uniqueness, suppose \( x, x' \in K \) and \( \|x\| = \|x'\| = d\). We use the same trick to get that \(\|x - x'\|^2 \leq 0\), so \(x = x'\). \(\square\)

Let \( M \subset X \) be a linear subspace. The **orthogonal complement** to \( M \) is \(\{x \in X \mid x \perp y \text{ for all } y \in M\}\).

**Theorem 6.4.6.** Let \( X \) be a Hilbert space and \( M \subset X \) be a closed subspace. Then \( X = M \oplus M^\perp\).

**Proof.** Let \( x \in X \) be arbitrary. Since \( M \) is closed and convex, there exists \( y \in M \) such that \(\|x - y\| = \inf\{\|x - y'\| \mid y' \in M\}\). We claim that \( x - y \in M^\perp\). For all \( t \in F \) and \( m \in M \), we have \( y + tm \in M \), so

\[
\|x - y\|^2 \leq \|(x - (y + tm))\|^2 = \|x - y\|^2 + 2\Re t(x - y, m) + |t|^2\|u\|^2.
\]

Since this is true for all \( t \in F \), we must have \( (x - y, m) = 0 \). \(\square\)

**Theorem 6.4.7** (Riesz representation theorem). Let \( X \) be a Hilbert space and \( f \in X^* \). Then there exists a unique \( y \in X \) such that \( f = f_y \), where \( f_y(x) = \langle x, y \rangle \). The map \( y \in X \mapsto f_y \in X^* \) is a conjugate linear isometric isomorphism between Banach spaces.

**Proof.** Let \( f \in X^* \) and \( M = \ker f \). Then \( M \) is a closed linear subspace of \( X \), so \( X = M \oplus M^\perp\). If \( M^\perp = 0 \), then \( X = M \) and \( f = 0 \), so we can pick \( y = 0 \). Otherwise, pick \( z \in M^\perp \) non-zero. Then \( z \not\in M \), so \( f(z) \neq 0 \). We pick \( y = \alpha z \) for some suitable \( \alpha \in F \). Given \( x \in X \), consider \( u = f(x)z - f(z)x \). Then \( f(u) = f(x)f(z) - f(z)f(x) = 0 \), so \( u \in M \). Hence \( u \perp z \), so

\[
0 = \langle f(x)z - f(z)x, z \rangle = f(x)\|z\|^2 - f(z)\langle x, z \rangle.
\]

Hence \( f(x) = f(z)/\|z\|^2 \langle x, z \rangle \), so we take \( \alpha = f(z)/\|z\|^2 \). \(\square\)

A set \( \{x_\alpha\}_{\alpha \in A} \) in a Hilbert space is an **orthonormal system** if \( \langle x_\alpha, x_\beta \rangle = 0 \) if \( \alpha \neq \beta \) and 1 if \( \alpha = \beta \).
Lemma 6.4.8. Let \( \{x_\alpha\}_{\alpha \in A} \) be an orthonormal system in a Hilbert space \( X \). Then the following are equivalent:

1. \( \{x_\alpha\} \) is a maximal orthonormal system;
2. \( U = \text{span}\{x_\alpha\} \) is dense in \( X \);
3. if \( x \in X \) and \( \langle x, x_\alpha \rangle = 0 \) for all \( \alpha \in A \), then \( x = 0 \).

Proof. (1) \( \Rightarrow \) (2) Suppose \( \{x_\alpha\}_{\alpha \in A} \) is maximal and let \( M = \overline{U} \). Then \( M \) is a linear subspace of \( X \), so \( X = M \oplus M^\perp \). If \( M^\perp \neq 0 \), then \( \{x_\alpha\} \) is not maximal, as we could adjoin a non-zero element of \( M^\perp \). Hence \( M^\perp = 0 \) and \( M = X \).

(2) \( \Rightarrow \) (3) Suppose \( x \in X \) and \( \langle x, x_\alpha \rangle = 0 \) for all \( \alpha \). Then there exist \( z_n \in U \) such that \( z_n \to x \).

Since \( x \perp U \), we have

\[
\|x\|^2 = \langle x, x \rangle = \lim_{n \to \infty} \langle x, z_n \rangle = 0
\]

(3) \( \Rightarrow \) (1) If not, then a vector that we could adjoin to \( \{x_\alpha\} \) would break (3).

\[\square\]

Definition 6.4.9 (Complete ONS). An orthonormal system satisfying any of these conditions is a complete orthonormal system or a Hilbert basis.

Lemma 6.4.10. Let \( X \) be an infinite-dimensional separable Hilbert space. Then \( X \) has a countable (necessarily infinite) Hilbert basis \( \{x_1, x_2, \ldots\} \).

Proof. Let \( \{z_n\} \) be a countable dense subset of \( X \) and \( \{y_n\} \) be a maximal linearly independent subset. Then \( U = \text{span}\{y_n\} \) is dense in \( X \). Running Gram-Schmidt (valid for a countable infinite set of vectors as well), we produce an orthonormal system \( \{x_n\} \) with span dense in \( X \). This is the required Hilbert basis.

\[\square\]

Theorem 6.4.11 (Parseval identities). Let \( \{x_n\} \) be a Hilbert basis of a separable Hilbert space \( X \) of infinite dimension. Then

1. \( x = \sum_n \langle x, x_n \rangle x_n \);
2. \( \langle x, y \rangle = \sum_n \langle x, x_n \rangle \langle y, x_n \rangle \);
3. \( \|x\|^2 = \sum_n \|x, x_n\|^2 \).

Proof. We first show that

\[
\sum_{n=1}^{\infty} |\langle x, x_n \rangle|^2 \leq \|x\|^2. \tag{*}
\]
(This is Bessel’s inequality, which holds for arbitrary orthonormal systems.) Write \( \alpha_n = \langle x, x_n \rangle \). Then

\[
0 \leq \left\| x - \sum_{n=1}^{N} \alpha_n x_n \right\|^2 \\
= \left\langle x - \sum_{n=1}^{N} \alpha_n x_n, x - \sum_{n=1}^{N} \alpha_n x_n \right\rangle \\
= \|x\|^2 - 2 \text{Re} \left( \sum_{n=1}^{N} \alpha_n x_n \right) + \left\| \sum_{n=1}^{N} \alpha_n x_n \right\|^2 \\
= \|x\|^2 - 2 \text{Re} \left( \sum_{n=1}^{N} |\alpha_n|^2 \right) + \sum_{n=1}^{N} |\alpha_n|^2 \\
= \|x\|^2 - \sum_{n=1}^{N} |\alpha_n|^2
\]

Hence \( \sum_n |\alpha_n|^2 \) converges and is bounded above by \( \|x\|^2 \).

Let the partial sums of \( \sum_n \alpha_n x_n \) be \( s_N \). We claim that \( (s_N) \) is a Cauchy sequence. We have, for \( M \leq N \),

\[
\limsup_{N,M \to \infty} \|s_N - s_M\|^2 = \limsup_{N,M \to 0} \left\| \sum_{n=M+1}^{N} \alpha_n x_n \right\|^2 \\
= \limsup_{N,M \to 0} \sum_{n=M+1}^{N} |\alpha_n|^2 = 0.
\]

Since \( (s_N) \) is a Cauchy sequence, it converges. let \( s = \sum_n \alpha_n x_n \). Then for all \( N \),

\[
\langle x - s, x_N \rangle = \alpha_N - \langle s, x_N \rangle \\
= \alpha_N - \lim_{N \to \infty} \langle \sum_{n=1}^{N} \alpha_n x_n, x_n \rangle \\
= \alpha_N - \alpha_N = 0,
\]

so by completeness, \( x - s = 0 \). This proves 1. The Parseval identities (2 and 3) follow directly. \( \square \)

**Theorem 6.4.12** (Weierstrass approximation theorem). Let \( [a, b] \subset \mathbb{R} \). Then for each \( f \in C[a, b] \) and \( \epsilon > 0 \), there exists a polynomial \( P \) such that \( \|f - P\|_{\infty} \leq \epsilon \).

**Proof.** Without loss of generality \( [a, b] \subset (0, 1) \). Pick \( f \in C[a, b] \) and extend \( f \) to \( \mathbb{R} \) with \( \text{supp} f \subset [0, 1] \). Define \( P_n(x) = (K_n * f)(x) = \int K_n(x-u) f(u) \, du \) for

\[
K_n(u) = \begin{cases} 
0 & |u| > 1, \\
c_n(1-u^2)^n & |u| \leq 1.
\end{cases}
\]
The constant $c_n$ is chosen so that $K_n$ integrates to 1. Then $|P_n(x) - f(x)|$ is uniformly small on $[a, b]$ for $n$ large, since $f$ is continuous.

**Theorem 6.4.13** (Stone-Weierstrass theorem). Let $X$ be a compact Hausdorff space and $A \subset C(X, \mathbb{R})$ be a subalgebra of $C(X, \mathbb{R})$ which is closed with respect to the uniform norm. If $A$ separates points in $X$ and for all $x \in X$, there exists $f \in A$ such that $f(x) \neq 0$, then $A = C(X, \mathbb{R})$.

**Proof.** We claim that if $f \in A$, then $|f| \in A$. Without loss of generality, $|f| \leq 1$. It follows from the Weierstrass approximation theorem that there exist polynomials $P_n$ with $P_n(0) = 0$ such that $P_n(x) \to |x|$ uniformly on $[-1, 1]$. Since $A$ is a subalgebra of $C(X, \mathbb{R})$ and $|f| \leq 1$, we have $P_n(f) \to |f|$ uniformly on $X$.

If $f, g \in A$, then $f \vee g = \max(f, g)$ and $f \wedge g = \min(f, g)$ are in $A$ (so $A$ is a lattice of functions), as

$$f \vee g = \frac{1}{2}(f + g) + \frac{1}{2}|f - g|, \quad f \wedge g = \frac{1}{2}(f + g) - \frac{1}{2}|f - g|.$$ 

It then follows, from the above and from $A$ separating points with non-zero function values at both points, that for all $x, y \in X$ with $x \neq y$ and for all $\alpha, \beta \in \mathbb{R}$, there exists $f \in A$ such that $f(x) = \alpha$ and $f(y) = \beta$.

Now let $f \in C(X, \mathbb{R})$ and $\epsilon > 0$ be arbitrary. For all $x, y \in X$, we can find $g_{xy} \in A$ such that $g_{xy}(x) = f(x)$ and $g_{xy}(y) = f(y)$. Let $U_{xy} = \{z \in X \mid f(z) - \epsilon < g_{xy}(z)\}$ and $V_{xy} = \{z \in X \mid g_{xy}(z) < f(z) + \epsilon\}$. Then $U_{xy}$ and $V_{xy}$ are open, and $x, y \in U_{xy}, V_{xy}$. Fix $y \in X$. Then $X = \bigcup_y U_{xy}$ is an open cover, so there is a finite subcover $X = U_{x_1y} \cup \cdots \cup U_{x_ny}$. Let $g_y = g_{x_1y} \vee \cdots \vee g_{x_ny} \in A$. Then $g_y > f - \epsilon$ by construction. Moreover, for $z \in V_y = V_{x_1y} \cap \cdots \cap V_{x_ny}$, we have $g_{x_i}(z) < f(z) + \epsilon$, so $g_y(z) < f(z) + \epsilon$ for $z \in V_y$. By compactness again, we can write $X = V_{y_1} \cup \cdots \cup V_{y_m}$ and take $g = g_{y_1} \wedge \cdots \wedge g_{y_m}$. This $g$ is in $A$, and $\|g - f\|_{\infty} < \epsilon$.

Pick $\epsilon = 1/n$ and get $g_n \in A$ with $\|f - g_n\| < 1/n$. Then $g_n \to f$ in $C(X, \mathbb{R})$, and by closure of $A$, we have $f \in A$.

**Remark 6.4.14.** If $A$ is not required to be closed, then we conclude that $A$ is dense in $C(X, \mathbb{R})$.

**Corollary 6.4.15.** Let $K \subset \mathbb{R}^n$ be compact. The set of all polynomial functions on $\mathbb{R}^n$ is dense in $C(K, \mathbb{R})$.

**Proof.** It is clear that the polynomial functions form a subalgebra. The coordinate functions are sufficient to separate points, and the constant function 1 is sufficient for non-vanishing. Hence Stone-Weierstrass applies.

**Theorem 6.4.16** (Complex Stone-Weierstrass). Let $X$ be a compact Hausdorff space and $A$ be a complex subalgebra of $C(X)$. If, in addition to the previous conditions, $A$ is closed under complex conjugation, then $A = C(X)$.

**Proof.** If $f \in A$, then $\text{Re} \ f$ and $\text{Im} \ f$ are in $A$. Then $A \cap C(X, \mathbb{R})$ satisfies the conditions of Stone-Weierstrass, so $C(X, \mathbb{R}) \subset A$. Since $A$ is a complex subalgebra, $A = C(X)$. 

71
7 FOURIER ANALYSIS

7.1 TRIGONOMETRIC SERIES

Let $\mathbb{T}$ be the unit circle. A function $f : \mathbb{T} \to \mathbb{C}$ is equivalently a function $f : \mathbb{R} \to \mathbb{C}$ which is $2\pi$-periodic, or also equivalently, a function $f : [-\pi, \pi) \to \mathbb{C}$.

For $1 \leq p < \infty$, we can regard $L^p(\mathbb{T})$ as the set of (equivalence classes) of $2\pi$-periodic Lebesgue-measurable functions $f : \mathbb{R} \to \mathbb{C}$ with $\|f\|_p = \left( \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(t)|^p \, dt \right)^{1/p} < \infty$.

We have $f \sim g$ if and only if $f = g$ almost everywhere. Equivalently we look at functions $f : \mathbb{T} \to \mathbb{C}$ with the arc length measure on $\mathbb{T}$. This is obtained by pushing Lebesgue measure on $[-\pi, \pi)$ forward via $t \mapsto e^{it}$.

The space $L^\infty(\mathbb{T})$ consists of $2\pi$-periodic Lebesgue measurable functions $f : \mathbb{R} \to \mathbb{C}$ with $\|f\|_\infty = \text{ess sup}\{|f(t)| \mid t \in \mathbb{R} = \inf\{m \mid \lambda_1\{|f| > m\} = 0\} < \infty$.

Let $C(\mathbb{T})$ be the space of continuous functions $F : \mathbb{T} \to \mathbb{C}$ (or space of $2\pi$-periodic continuous functions $f : \mathbb{R} \to \mathbb{C}$) equipped with the norm $\|f\|_\infty = \text{sup}\{f(t) \mid t \in \mathbb{R}\}$.

Each of these spaces is a Banach space. Moreover, we have the inclusions

$C(\mathbb{T}) \subset L^\infty(\mathbb{T}) \subset L^p(\mathbb{T}) \subset L^q(\mathbb{T}) \subset L^1(\mathbb{T})$

whenever $\infty \geq p \geq q \geq 1$, by Hölder’s inequality.

**Definition 7.1.1** (Trigonometric polynomial). A trigonometric polynomial is a function $f \in C(\mathbb{T})$ of the form

$$f(t) = a_0 + \sum_{n=1}^{N} (a_n \cos nt + b_n \sin nt),$$

where $a_0, a_n, b_n \in \mathbb{C}$.

Equivalently, we may write

$$f(t) = \sum_{n=-N}^{N} c_n e^{int}.$$

Note that $L^2(\mathbb{T})$ is a Hilbert space with

$$\langle f, g \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t)\overline{g(t)} \, dt,$$

and $C(\mathbb{T}) \subset L^2(\mathbb{T})$.

**Theorem 7.1.2.** The trigonometric polynomials are dense in $C(\mathbb{T})$. 

73
Proof. It is convenient to write \( P \) as the space of functions of the form
\[
z \in \mathbb{T} \mapsto \sum_{n=-N}^{N} c_n z^n.
\]
That these are dense in \( C(\mathbb{T}) \) follows from complex Stone-Weierstrass.

Corollary 7.1.3.  
(a) Trigonometric polynomials are dense in \( L^2(\mathbb{T}) \).

(b) Define \( u_n(t) = e^{int} \) for \( n \in \mathbb{Z} \) and \( t \in \mathbb{R} \). Then \( \{u_n\} \) is a Hilbert space basis for \( L^2(\mathbb{T}) \).

7.2 FOURIER SERIES

Definition 7.2.1 (Fourier series). Let \( f \in L^1(\mathbb{T}) \supset L^2(\mathbb{T}) \). The Fourier coefficients of \( f \) are
\[
\hat{f}(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t)u_n(t) \, dt = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t)e^{-int} \, dt.
\]
If \( f \in L^2(\mathbb{T}) \), then \( \hat{f}(n) = \langle f, u_n \rangle \).

The Fourier series of \( f \) is
\[
f \sim \sum_{n=-\infty}^{n} \hat{f}(n)e^{int}.
\]
(Here \( \sim \) is used to say that \( f \) has Fourier series given by the right hand side.) When \( f \in L^2(\mathbb{T}) \), we have actual equality.

Denote the partial sums by
\[
(s_N f)(t) = \sum_{n=-N}^{N} \hat{f}(n)e^{int}.
\]