Real Analysis

Armin Rainer

Fakultät für Mathematik, Universität Wien, Oskar-Morgenstern-Platz 1, A-1090 Wien, Austria

E-mail address: armin.rainer@univie.ac.at

Preface

These are lecture notes for the course *Reelle Analysis* held in Vienna in Spring 2014 and 2016 (two semester hours). The main sources are [**1**], [**3**], [**5**], [**6**], [**8**], [**10**], [**11**], [**12**], [**13**], and [**14**].

Contents

CHAPTER 1

Basic measure theory

1.1. σ-algebras and measures

Let X be a set. A collection $\mathfrak{S} \subseteq \mathfrak{P}(X)$ of subsets of X is called a **σ-algebra** if the following are satisfied:

- If $A \in \mathfrak{S}$, then $A^c = X \setminus A \in \mathfrak{S}$.
- If $\{A_i\}_{i=1}^{\infty}$ is a countable family of sets in $\mathfrak{S}$, then $\bigcup_{i=1}^{\infty} A_i \in \mathfrak{S}$.
- $X \in \mathfrak{S}$.

It is immediate from this definition that

- $\emptyset \in \mathfrak{S}$.
- If $\{A_i\}_{i=1}^{\infty}$ is a countable family of sets in $\mathfrak{S}$, then $\bigcap_{i=1}^{\infty} A_i \in \mathfrak{S}$.
- If $A_1, A_2 \in \mathfrak{S}$, then $A_1 \setminus A_2 \in \mathfrak{S}$.

Evidently, for any set X, the collections $\{\emptyset, X\}$ and $\mathfrak{P}(X)$ form σ-algebras, respectively. Given any family of subsets $\mathfrak{A} \subseteq \mathfrak{P}(X)$ the intersection of all σ-algebras containing $\mathfrak{A}$ is a σ-algebra. It is the smallest σ-algebra containing $\mathfrak{A}$ and is called the **σ-algebra generated by** $\mathfrak{A}$.

Let X be a topological space. The σ-algebra $\mathfrak{B}(X)$ generated by all open subsets in X is called the σ-algebra of **Borel sets** in X, or **Borel σ-algebra**. The Borel σ-algebra $\mathfrak{B}(\mathbb{R}^n)$ is generated by the open balls in $\mathbb{R}^n$. It contains all closed sets, but not all subsets of $\mathbb{R}^n$.

A **(positive) measure** μ on a σ-algebra $\mathfrak{S}$ is a mapping $\mu : \mathfrak{S} \to [0, \infty]$ with the following properties:

- $\mu(\emptyset) = 0$
- μ is **σ-additive**, i.e., if $\{A_i\}_{i=1}^{\infty}$ is a countable family of disjoint sets in $\mathfrak{S}$, then
$$\mu\Big(\bigcup_{i=1}^{\infty} A_i\Big) = \sum_{i=1}^{\infty} \mu(A_i).$$

Lemma 1.1. *Let μ be a measure on a σ-algebra $\mathfrak{S}$, and let $A_i \in \mathfrak{S}$. Then:*

(1) *μ is **finitely additive**, i.e., for finite families of disjoint sets $\{A_i\}_{i=1}^{m}$,*
$$\mu\Big(\bigcup_{i=1}^{m} A_i\Big) = \sum_{i=1}^{m} \mu(A_i),$$
(2) *μ is **monotone**, i.e., $\mu(A_1) \le \mu(A_2)$ if $A_1 \subseteq A_2$.*
(3) *If $A_1 \subseteq A_2 \subseteq \cdots$, then*
$$\lim_{j\to\infty} \mu(A_j) = \mu\Big(\bigcup_{i=1}^{\infty} A_i\Big).$$
(4) *If $A_1 \supseteq A_2 \supseteq \cdots$ and $\mu(A_1) < \infty$, then*
$$\lim_{j\to\infty} \mu(A_j) = \mu\Big(\bigcap_{i=1}^{\infty} A_i\Big).$$

PROOF. (1) follows immediately from the definition of measure.

(2) We have $A_2 = A_1 \cup (A_2 \setminus A_1)$ and so $\mu(A_2) = \mu(A_1) + \mu(A_2 \setminus A_1) \geq \mu(A_1)$.

(3) Setting $B_i := A_i \setminus A_{i-1}$, $i \geq 2$, and $B_1 := A_1$, we obtain a sequence of disjoint sets $B_i \in \mathfrak{S}$ so that $\bigcup_{i=1}^m A_i = \bigcup_{j=1}^m B_j$, for all $m \in \mathbb{N} \cup \{\infty\}$. Thus

$$\mu\Big(\bigcup_{i=1}^{\infty} A_i\Big) = \mu\Big(\bigcup_{j=1}^{\infty} B_j\Big) = \sum_{j=1}^{\infty} \mu(B_j) = \lim_{m\to\infty} \sum_{j=1}^{m} \mu(B_j)$$
$$= \lim_{m\to\infty} \mu\Big(\bigcup_{j=1}^{m} B_j\Big) = \lim_{m\to\infty} \mu(A_m).$$

(4) We have $\bigcap_{i=1}^{\infty} A_i = A_1 \setminus \bigcup_{j=1}^{\infty}(A_1 \setminus A_j)$, and thus, by (3),

$$\mu\Big(\bigcap_{i=1}^{\infty} A_i\Big) = \mu(A_1) - \mu\Big(\bigcup_{j=1}^{\infty}(A_1 \setminus A_j)\Big) = \mu(A_1) - \lim_{i\to\infty} \mu(A_1 \setminus A_i) = \lim_{i\to\infty} \mu(A_i). \quad \square$$

A **measure space** is a triple $(X, \mathfrak{S}, \mu)$ consisting of a set X, a σ-algebra $\mathfrak{S}$ on X, and a measure μ on $\mathfrak{S}$. The elements of $\mathfrak{S}$ are called (μ-)**measurable sets**. If $X' \in \mathfrak{S}$, then we may define the **measure subspace** $(X', \mathfrak{S}', \mu')$, where $\mathfrak{S}' := \{A : A \in \mathfrak{S} \text{ and } A \subseteq X'\} = \{A \cap X' : A \in \mathfrak{S}\}$ and $\mu' := \mu|_{\mathfrak{S}'}$.

A measure μ is called **finite** if $\mu(X) < \infty$, and **probability measure** if $\mu(X) = 1$. It is called **σ-finite** if there exists a sequence $X_i \in \mathfrak{S}$ such that $\mu(X_i) < \infty$ for all i and $X = \bigcup_{i=1}^{\infty} X_i$; note that the X_i can be chosen disjoint by setting $X_i' = X_i \setminus \bigcup_{k=1}^{i-1} X_k$. We say that μ has the **finite subset property** if for each $A \in \mathfrak{S}$ with $\mu(A) > 0$ there is $B \in \mathfrak{S}$ with $B \subseteq A$ and $0 < \mu(B) < \infty$. A σ-finite measure has the finite subset property; if $A \in \mathfrak{S}$ with $\mu(A) > 0$ then for some i we have $0 < \mu(A \cap X_i) < \infty$.

Example 1.2.

(1) For any set X we may take the σ-algebra $\mathfrak{P}(X)$ of all subsets and consider the **counting measure**
$$\mu(A) = \begin{cases} |A| & \text{if } A \text{ is finite} \\ \infty & \text{if } A \text{ is infinite} \end{cases}.$$
(2) If X is a topological space and μ is a measure on the Borel σ-algebra, then μ is called a **Borel measure**.
(3) Fix a point $x \in \mathbb{R}^n$. Then the **Dirac δ-measure** δ_x defined by
$$\delta_x(A) = \chi_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases}$$
is a measure defined on the Borel σ-algebra or even on $\mathfrak{P}(\mathbb{R}^n)$.

1.2. Monotone class theorem and uniqueness of measures

Let X be a set. A collection $\mathfrak{A} \subseteq \mathfrak{P}(X)$ of subsets of X is called an **algebra** if $X \in \mathfrak{A}$ and, for every $A, B \in \mathfrak{A}$, also $A^c \in \mathfrak{A}$ and $A \cup B \in \mathfrak{A}$.

A collection $\mathfrak{M} \subseteq \mathfrak{P}(X)$ of subsets of X is called an **monotone class** if, for $A_i \in \mathfrak{M}$, we have:

- If $A_1 \subseteq A_2 \subseteq \cdots$, then $\bigcup_{i=1}^{\infty} A_i \in \mathfrak{M}$.
- If $A_1 \supseteq A_2 \supseteq \cdots$, then $\bigcap_{i=1}^{\infty} A_i \in \mathfrak{M}$.

Clearly, $\mathfrak{P}(X)$ is a monotone class.

Theorem 1.3 (Monotone class theorem). *Let $\mathfrak{A}$ be an algebra of subsets of X. Then there exists a smallest monotone class $\mathfrak{M}$ that contains $\mathfrak{A}$, and $\mathfrak{M}$ is the σ-algebra generated by $\mathfrak{A}$.*

PROOF. Let $\mathfrak{M}$ be the intersection of all monotone classes that contain $\mathfrak{A}$. Then $\mathfrak{M}$ is a monotone class that contains $\mathfrak{A}$, and by definition it is the smallest.

In order to show that $\mathfrak{M}$ is the σ-algebra generated by $\mathfrak{A}$, it suffices to prove that $\mathfrak{M}$ is closed under complements and finite unions. Indeed, assuming this, we may conclude that, if $A_i \in \mathfrak{M}$ then $B_n := \bigcup_{i=1}^n A_i \in \mathfrak{M}$ and $B_1 \subseteq B_2 \subseteq \cdots$ and hence $\bigcup_{i=1}^\infty A_i = \bigcup_{n=1}^\infty B_n \in \mathfrak{M}$. Thus $\mathfrak{M}$ is a σ-algebra. Since any σ-algebra is a monotone class, $\mathfrak{M}$ is the smallest σ-algebra that contains $\mathfrak{A}$, i.e., the σ-algebra generated by $\mathfrak{A}$.

Let us show that $\mathfrak{M}$ is closed under finite unions. Fix $A \in \mathfrak{M}$ and consider $\mathfrak{C}(A) := \{B \in \mathfrak{M} : A \cup B \in \mathfrak{M}\}$. Let $B_i \in \mathfrak{C}(A)$ so that $B_1 \subseteq B_2 \subseteq \cdots$. Then $(A \cup B_1) \subseteq (A \cup B_2) \subseteq \cdots$ is a sequence in $\mathfrak{M}$, hence $A \cup \bigcup_{i=1}^\infty B_i = \bigcup_{i=1}^\infty (A \cup B_i) \in \mathfrak{M}$, and so $\bigcup_{i=1}^\infty B_i \in \mathfrak{C}(A)$. Similarly, the intersection of a decreasing sequence of sets in $\mathfrak{C}(A)$ belongs to $\mathfrak{C}(A)$. Thus $\mathfrak{C}(A)$ is a monotone class.

If $A \in \mathfrak{A}$, then $\mathfrak{A} \subseteq \mathfrak{C}(A) \subseteq \mathfrak{M}$, since $\mathfrak{A}$ is an algebra, and thus $\mathfrak{C}(A) = \mathfrak{M}$. If $A \in \mathfrak{M}$ is arbitrary, then $\mathfrak{A} \subseteq \mathfrak{C}(A)$, for if $B \in \mathfrak{A}$ then $\mathfrak{C}(B) = \mathfrak{M}$, by the previous sentence, and hence $A \cup B \in \mathfrak{M}$. Thus $\mathfrak{C}(A) = \mathfrak{M}$ for each $A \in \mathfrak{M}$, that means that $\mathfrak{M}$ is closed under finite unions.

In order to prove that $\mathfrak{M}$ is closed under complements, we consider $\mathfrak{C} := \{B \in \mathfrak{M} : B^c \in \mathfrak{M}\}$. Since $\mathfrak{A}$ is an algebra, $\mathfrak{A} \subseteq \mathfrak{C}$. If $B_i \in \mathfrak{C}$ so that $B_1 \subseteq B_2 \subseteq \cdots$, then $B_i^c \in \mathfrak{M}$ and $B_1^c \supseteq B_2^c \supseteq \cdots$, and hence $\left(\bigcup_{i=1}^\infty B_i\right)^c = \bigcap_{i=1}^\infty B_i^c \in \mathfrak{M}$. Similarly, the intersection of a decreasing sequence of sets in $\mathfrak{C}$ belongs to $\mathfrak{C}$. It follows that $\mathfrak{C} = \mathfrak{M}$. The proof is complete. $\square$

Theorem 1.4 (Uniqueness of measures). *Let $\mathfrak{A}$ be an algebra of subsets of X and let $\mathfrak{S}$ be the σ-algebra generated by $\mathfrak{A}$. Let μ_1 and μ_2 be measures on $\mathfrak{S}$ that coincide on $\mathfrak{A}$. Suppose that there is a sequence of sets $A_i \in \mathfrak{A}$ so that $\mu_1(A_i) = \mu_2(A_i) < \infty$, $i \geq 1$, and $\bigcup_{i=1}^\infty A_i = X$. Then $\mu_1 = \mu_2$ on $\mathfrak{S}$.*

PROOF. First we assume that $\mu_1(X) < \infty$. Lemma 1.1 implies that $\mathfrak{M} := \{A \in \mathfrak{S} : \mu_1(A) = \mu_2(A)\}$ is a monotone class;

$$\mu_1\Big(\bigcup_{i=1}^\infty A_i\Big) = \lim_{j\to\infty} \mu_1(A_j) = \lim_{j\to\infty} \mu_2(A_j) = \mu_2\Big(\bigcup_{i=1}^\infty A_i\Big) \quad \text{if } A_i \subseteq A_{i+1}$$
$$\mu_1\Big(\bigcap_{i=1}^\infty A_i\Big) = \lim_{j\to\infty} \mu_1(A_j) = \lim_{j\to\infty} \mu_2(A_j) = \mu_2\Big(\bigcap_{i=1}^\infty A_i\Big) \quad \text{if } A_i \supseteq A_{i+1}.$$

By Theorem 1.3, we can conclude that $\mathfrak{M} = \mathfrak{S}$ which gives the assertion.

For the case $\mu_1(X) = \infty$, note that, for each $A \in \mathfrak{A}$, $A \cap \mathfrak{S}$ is the σ-algebra (on A) generated by $A \cap \mathfrak{A}$ (exercise!). Thus $\mu_1(A \cap B) = \mu_2(A \cap B)$ for all $B \in \mathfrak{S}$ if $\mu_1(A) < \infty$, by the finite case. By assumption, $X = \bigcup_{i=1}^\infty A_i$ for sets $A_i \in \mathfrak{A}$ so that $\mu_1(A_i) = \mu_2(A_i) < \infty$. Without loss of generality we may assume that the A_i are disjoint. Then, for $B \in \mathfrak{S}$,

$$\mu_1(B) = \mu_1\Big(\bigcup_{i=1}^\infty (A_i \cap B)\Big) = \sum_{i=1}^\infty \mu_1(A_i \cap B) = \sum_{i=1}^\infty \mu_2(A_i \cap B) = \mu_2(B). \qquad \square$$

An **elementary family** $\mathfrak{E}$ is a collection of subsets of X satisfying

- $\emptyset \in \mathfrak{E}$,
- if $E, F \in \mathfrak{E}$ then $E \cap F \in \mathfrak{E}$,

- if $E \in \mathfrak{E}$, then E^c is a finite disjoint union of elements in $\mathfrak{E}$.

Proposition 1.5. *The collection $\mathfrak{A}$ of finite disjoint unions of elements in an elementary family $\mathfrak{E}$ forms an algebra.*

Proof. Suppose that $A, B \in \mathfrak{E}$ and $B^c = \bigcup_{i=1}^n C_i$, where $C_i \in \mathfrak{E}$ are disjoint. Then $A \setminus B = \bigcup_{i=1}^n (A \cap C_i) \in \mathfrak{A}$ and $A \cup B = (A \setminus B) \cup B \in \mathfrak{A}$, since these unions are disjoint. By induction, we can conclude that if $A_1, \dots, A_n \in \mathfrak{E}$ then $\bigcup_{i=1}^n A_i \in \mathfrak{A}$. For, by inductive hypothesis we may assume that $A_1, \dots, A_{n-1}$ are disjoint, and then $\bigcup_{i=1}^n A_i = A_n \cup \bigcup_{i=1}^{n-1} (A_i \setminus A_n) \in \mathfrak{A}$. Thus if $A, B \in \mathfrak{A}$ then $A \cup B \in \mathfrak{A}$.

Let us show that $\mathfrak{A}$ is stable under complements. Let $A_1, \dots, A_n \in \mathfrak{E}$ and $A_i^c = \bigcup_{j=1}^{m_i} B_{ij}$ with $B_{ij} \in \mathfrak{E}$ disjoint for all i, j. Then

$$\Big(\bigcup_{i=1}^n A_i\Big)^c = \bigcap_{i=1}^n \bigcup_{j=1}^{m_i} B_{ij} = \bigcup_{\substack{1 \le j_i \le m_i \\ 1 \le i \le n}} B_{1j_1} \cap \cdots \cap B_{nj_n}$$

which belongs to $\mathfrak{A}$. □

1.3. Outer measures and Caratheodory's construction

An **outer measure** on a set X is a mapping $\mu : \mathfrak{P}(X) \to [0, \infty]$ satisfying:

- $\mu(\emptyset) = 0$.
- μ is monotone, i.e., $\mu(A) \le \mu(B)$ if $A \subseteq B$.
- μ is **σ-subadditive**, i.e., for any countable family $\{A_i\}_{i=1}^\infty$ of sets $A_i \subseteq X$,
$$\mu\Big(\bigcup_{i=1}^\infty A_i\Big) \le \sum_{i=1}^\infty \mu(A_i).$$

Theorem 1.6 (Caratheodory)**.** *Let μ be an outer measure on X. Set*

$$\mathfrak{S} := \{E \in \mathfrak{P}(X) : \mu(A) = \mu(A \cap E) + \mu(A \setminus E) \text{ for every } A \subseteq X\}.$$

Then $\mathfrak{S}$ is a σ-algebra and $(X, \mathfrak{S}, \mu|_{\mathfrak{S}})$ is a measure space.

Proof. Clearly, $X \in \mathfrak{S}$. If $E \in \mathfrak{S}$ then $E^c \in \mathfrak{S}$, since, for every $A \subseteq X$,

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

Next we claim that, for $E, F \in \mathfrak{S}$, also $E \cup F \in \mathfrak{S}$. Indeed, for every $A \subseteq X$,

$$\begin{aligned}
&\mu(A \cap (E \cup F)) + \mu(A \setminus (E \cup F)) \\
&\qquad = \mu(A \cap (E \cup F) \cap E) + \mu((A \cap (E \cup F)) \setminus E) + \mu(A \setminus (E \cup F)) \\
&\qquad = \mu(A \cap E) + \mu((A \setminus E) \cap F) + \mu((A \setminus E) \setminus F) \\
&\qquad = \mu(A \cap E) + \mu(A \setminus E) \\
&\qquad = \mu(A).
\end{aligned}$$

(The first and last equality hold, because $E \in \mathfrak{S}$, the third, because $F \in \mathfrak{S}$.) Let $\{E_i\}_{i=1}^\infty$ be a sequence of sets in $\mathfrak{S}$, and set $E := \bigcup_{i=1}^\infty E_i$ and $E_{\le n} := \bigcup_{i=1}^n E_i$. By induction on n, each $E_{\le n} \in \mathfrak{S}$. Set $F_n := E_{\le n} \setminus E_{\le n-1} = E_n \setminus E_{\le n-1}$, $n \ge 2$, and $F_1 = E_1$. For any $n \ge 2$ and $A \subseteq X$, we have

$$\begin{aligned}
\mu(A \cap E_{\le n}) &= \mu(A \cap E_{\le n} \cap E_{\le n-1}) + \mu(A \cap E_{\le n} \setminus E_{\le n-1}) \\
&= \mu(A \cap E_{\le n-1}) + \mu(A \cap F_n),
\end{aligned}$$

and, by induction, $\mu(A \cap E_{\leq n}) = \sum_{i=1}^{n} \mu(A \cap F_i)$ for each $n \geq 1$. This, together with σ-subadditivity, implies

$$\mu(A \cap E) = \mu\Big(A \cap \bigcup_{i=1}^{\infty} F_i\Big) \leq \sum_{i=1}^{\infty} \mu(A \cap F_i)$$
$$= \lim_{n\to\infty} \sum_{i=1}^{n} \mu(A \cap F_i) = \lim_{n\to\infty} \mu(A \cap E_{\leq n}).$$

Using monotonicity, we find

$$\mu(A \setminus E) = \mu\Big(A \setminus \bigcup_{i=1}^{\infty} E_{\leq i}\Big) \leq \inf_{i\geq 1} \mu(A \setminus E_{\leq i}) = \lim_{i\to\infty} \mu(A \setminus E_{\leq i}),$$

since the sequence $\mu(A \setminus E_{\leq i})$ is non-increasing and bounded from below by $\mu(A \setminus E)$. Thus,

$$\mu(A \cap E) + \mu(A \setminus E) \leq \lim_{n\to\infty} (\mu(A \cap E_{\leq n}) + \mu(A \setminus E_{\leq n}) = \mu(A).$$

This shows that $E \in \mathfrak{S}$, since the converse inequality is trivially satisfied by subadditivity. So $\mathfrak{S}$ is a σ-algebra.

In order to see that $(X, \mathfrak{S}, \mu|_{\mathfrak{S}})$ is a measure space, we need to show that μ is σ-additive on $\mathfrak{S}$. Let $\{E_i\}_{i=1}^{\infty}$ be a sequence of disjoint sets in $\mathfrak{S}$, and define E and $E_{\leq n}$ as above. Then

$$\mu(E_{\leq n}) = \mu(E_{\leq n} \cap E_n) + \mu(E_{\leq n} \setminus E_n) = \mu(E_n) + \mu(E_{\leq n-1}),$$

and, by induction, $\mu(E_{\leq n}) = \sum_{i=1}^{n} \mu(E_i)$ for each $n \geq 1$. Thus,

$$\mu(E) \geq \mu(E_{\leq n}) = \sum_{i=1}^{n} \mu(E_i)$$

for all n, and hence $\mu(E) \geq \sum_{i=1}^{\infty} \mu(E_i)$, which implies $\mu(E) = \sum_{i=1}^{\infty} \mu(E_i)$, as μ is σ-subadditive. □

1.4. Complete measures

Let $(X, \mathfrak{S}, \mu)$ be a measure space. Sets $E \in \mathfrak{S}$ with $\mu(E) = 0$ are called **μ-null sets**. If a statement about points $x \in X$ is true except for x in some null set, we say that it holds **μ-almost everywhere**, or **μ-a.e.** The measure μ is called **complete** if all subsets of null sets are measurable, i.e., $E \in \mathfrak{S}$, $\mu(E) = 0$, and $F \subseteq E$ implies $F \in \mathfrak{S}$.

Theorem 1.7 (Completion)**.** *Let $(X, \mathfrak{S}, \mu)$ be a measure space. Define*

$$\overline{\mathfrak{S}} := \{E \subseteq X :\ \exists A, B \in \mathfrak{S},\ A \subseteq E \subseteq B,\ \mu(B \setminus A) = 0\},$$

and set $\mu(E) := \mu(A)$ in this situation. Then $\overline{\mathfrak{S}}$ is a σ-algebra and μ is a measure on $\overline{\mathfrak{S}}$.

The measure space $(X, \overline{\mathfrak{S}}, \mu)$ is complete. The σ-algebra $\overline{\mathfrak{S}}$ is called the **μ-completion** of $\mathfrak{S}$.

Proof. Let us check that $\overline{\mathfrak{S}}$ is a σ-algebra. Clearly, $\mathfrak{S} \subseteq \overline{\mathfrak{S}}$. If $E \in \overline{\mathfrak{S}}$, then $A \subseteq E \subseteq B$ and hence $B^c \subseteq E^c \subseteq A^c$, and $A^c \setminus B^c = A^c \cap B = B \setminus A$ has measure 0, that is $E^c \in \overline{\mathfrak{S}}$. Suppose that $A_i, B_i \in \mathfrak{S}$ with $A_i \subseteq E_i \subseteq B_i$ and $\mu(B_i \setminus A_i) = 0$ for all i. Then $\bigcup_{i=1}^{\infty} A_i \subseteq \bigcup_{i=1}^{\infty} E_i \subseteq \bigcup_{i=1}^{\infty} B_i$ and

$$\Big(\bigcup_{i=1}^{\infty} B_i\Big) \setminus \Big(\bigcup_{i=1}^{\infty} A_i\Big) = \bigcup_{i=1}^{\infty} (B_i \setminus \bigcup_{i=1}^{\infty} A_i) \subseteq \bigcup_{i=1}^{\infty} (B_i \setminus A_i)$$

has measure zero. Hence $\bigcup_{i=1}^{\infty} E_i \in \overline{\mathfrak{S}}$ and $\overline{\mathfrak{S}}$ is a σ-algebra.

Next we show that μ is well-defined on $\overline{\mathfrak{S}}$. If $A, B, A', B' \in \mathfrak{S}$ satisfy

$$A \subseteq E \subseteq B,\ \mu(B \setminus A) = 0,\ A' \subseteq E \subseteq B',\ \mu(B' \setminus A') = 0,$$

then $A \setminus A' \subseteq E \setminus A' \subseteq B' \setminus A'$ and hence $\mu(A \setminus A') = 0$. Therefore $\mu(A) = \mu(A \cap A')$. Similarly, we find $\mu(A') = \mu(A \cap A')$, and thus $\mu(A) = \mu(A')$.

σ-additivity of μ on $\overline{\mathfrak{S}}$ follows from σ-additivity on $\mathfrak{S}$; if the sets E_i above are disjoint then so are A_i. $\square$

CHAPTER 2

Lebesgue measure on $\mathbb{R}^n$

2.1. Construction of the Lebesgue measure

A **box** I in $\mathbb{R}^n$ is given by the product of n compact intervals

$$I = [a, b] := [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_n, b_n],$$

where $a = (a_1, \dots, a_n)$, $b = (b_1, \dots, b_n)$, and $a_i \leq b_i$, $i = 1, \dots, n$, are real numbers. The **volume** $|I|$ of I is defined by

$$|I| = (b_1 - a_1) \cdots (b_n - a_n).$$

A box is called a **cube** if all its sides have the same length. A union of boxes is said to be **almost disjoint** if the interiors of the boxes are disjoint; the interior of a box I is denoted by

$$\mathring{I} = (a, b) := (a_1, b_1) \times (a_2, b_2) \times \cdots \times (a_n, b_n).$$

We denote by $\operatorname{dist}(E_1, E_2) := \inf\{|x_1 - x_2| : x_1 \in E_1, x_2 \in E_2\}$ the **distance** of two subsets $E_1, E_2 \subseteq \mathbb{R}^n$.

Theorem 2.1 (Lebesgue measure). *Let $\lambda^* : \mathfrak{P}(\mathbb{R}^n) \to [0, \infty]$ be defined by*

$$\lambda^*(E) := \inf\Big\{ \sum_{i=1}^{\infty} |Q_i| : \{Q_i\}_{i=1}^{\infty} \text{ is a countable cover of } E \text{ by cubes} \Big\},$$

and set

$$\mathfrak{L}(\mathbb{R}^n) := \{E \in \mathfrak{P}(\mathbb{R}^n) : \lambda^*(A) = \lambda^*(A \cap E) + \lambda^*(A \setminus E) \text{ for every } A \subseteq \mathbb{R}^n\}.$$

Then:

(1) *λ^* is an outer measure; the so-called* ***Lebesgue outer measure****.*
(2) *If $\operatorname{dist}(E_1, E_2) > 0$, then $\lambda^*(E_1 \cup E_2) = \lambda^*(E_1) + \lambda^*(E_2)$.*
(3) *$\mathfrak{L}(\mathbb{R}^n)$ is a σ-algebra that contains the Borel σ-algebra $\mathfrak{B}(\mathbb{R}^n)$.*

Proof. (1) Evidently, $\lambda^*(\emptyset) = 0$ and λ^* is monotone. In order to show that λ^* is σ-subadditive, let $E = \bigcup_{i=1}^{\infty} E_i$. We may assume that each $\lambda^*(E_i) < \infty$ for all i; otherwise there is nothing to prove. For given $\epsilon > 0$ and each j, there exists a cover $E_j \subseteq \bigcup_{k=1}^{\infty} Q_{j,k}$ by cubes so that

$$\sum_{k=1}^{\infty} |Q_{j,k}| \leq \lambda^*(E_j) + \frac{\epsilon}{2^j}.$$

Then $\{Q_{j,k}\}_{j,k=1}^{\infty}$ is a cover of E by cubes, and hence

$$\lambda^*(E) \leq \sum_{j=1}^{\infty} \sum_{k=1}^{\infty} |Q_{j,k}| \leq \sum_{j=1}^{\infty} \lambda^*(E_j) + \epsilon$$

which implies the assertion as ϵ was arbitrary.

(2) Choose $\operatorname{dist}(E_1, E_2) > \delta > 0$ and fix $\epsilon > 0$. There exists a cover $\{Q_j\}_{j=1}^\infty$ by cubes of $E := E_1 \cup E_2$ so that

$$\sum_{j=1}^\infty |Q_j| \le \lambda^*(E) + \epsilon.$$

We may assume that each Q_j has diameter less than δ, after possibly subdividing Q_j. Then each Q_j can intersect at most one of E_1 or E_2, and setting $J_i := \{j : Q_j \cap E_i \neq \emptyset\}$, $i = 1, 2$, we have $J_1 \cap J_2 = \emptyset$, and $E_i \subseteq \bigcup_{j \in J_i}^\infty Q_j$, $i = 1, 2$. Thus,

$$\lambda^*(E_1) + \lambda^*(E_2) \le \sum_{j \in J_1}^\infty |Q_j| + \sum_{j \in J_2}^\infty |Q_j| \le \sum_{j=1}^\infty |Q_j| \le \lambda^*(E) + \epsilon,$$

which implies (2), as ϵ was arbitrary; the converse inequality holds by (1).

(3) That $\mathfrak{L}(\mathbb{R}^n)$ is a σ-algebra follows from Theorem 1.6. In order to show that $\mathfrak{B}(\mathbb{R}^n) \subseteq \mathfrak{L}(\mathbb{R}^n)$ it suffices to prove that $\mathfrak{L}(\mathbb{R}^n)$ contains all closed subsets of $\mathbb{R}^n$. Let $F \subseteq \mathbb{R}^n$ be closed, and let A be any subset of $\mathbb{R}^n$. By (1), it is enough to show that

$$\lambda^*(A) \ge \lambda^*(A \cap F) + \lambda^*(A \setminus F),$$

and so we may assume that $\lambda^*(A) < \infty$. We set

$$A_0 := \{x \in A : \operatorname{dist}(x, F) \ge 1\},$$
$$A_i := \{x \in A : (i+1)^{-1} \le \operatorname{dist}(x, F) < i^{-1}\}, \quad i \ge 1.$$

Then any two sets A_{2j} and A_{2k} with even indices have positive distance; the same applies to sets A_{2j+1} with odd indices. By (2), for each $m \in \mathbb{N}$,

$$\sum_{i=0}^m \lambda^*(A_{2i}) = \lambda^*\Big(\bigcup_{i=0}^m A_{2i}\Big) \le \lambda^*(A),$$
$$\sum_{i=0}^m \lambda^*(A_{2i+1}) = \lambda^*\Big(\bigcup_{i=0}^m A_{2i+1}\Big) \le \lambda^*(A),$$

and therefore $\sum_{i=0}^\infty \lambda^*(A_i) < \infty$. Using $A \setminus F = \bigcup_{i=0}^\infty A_i$ and (1), we find

$$\begin{aligned}\lambda^*(A \cap F) + \lambda^*(A \setminus F) &\le \lambda^*(A \cap F) + \lambda^*\Big(\bigcup_{i=0}^m A_i\Big) + \sum_{i=m+1}^\infty \lambda^*(A_i)\\ &= \lambda^*\Big((A \cap F) \cup \bigcup_{i=0}^m A_i\Big) + \sum_{i=m+1}^\infty \lambda^*(A_i) \qquad \text{(by (2))}\\ &\le \lambda^*(A) + \sum_{i=m+1}^\infty \lambda^*(A_i),\end{aligned}$$

which implies the required inequality, since $\sum_{i=m+1}^\infty \lambda^*(A_i) \to 0$ as $m \to \infty$. □

Theorems 1.6 and 2.1 imply that the restriction of the Lebesgue outer measure λ^* to the σ-algebra $\mathfrak{L}(\mathbb{R}^n)$ is a measure. We call it the **Lebesgue measure**, and we denote it by λ or by λ^n, when the dimension n is important. The elements of $\mathfrak{L}(\mathbb{R}^n)$ are called the **(Lebesgue) measurable sets** in $\mathbb{R}^n$.

The Lebesgue measure is complete. Indeed, if $E \subseteq F$ and $\lambda(F) = 0$, then $\lambda^*(E) = 0$, and hence

$$\lambda^*(A) \le \lambda^*(A \cap E) + \lambda^*(A \setminus E) \le \lambda^*(E) + \lambda^*(A) = \lambda^*(A),$$

for any $A \subseteq \mathbb{R}^n$. But a Lebesgue null set need not be a Borel set; see Example 3.5. In fact, we shall see in Corollary 2.10 that the Lebesgue measure is the completion of the Borel measure $\lambda^*|_{\mathfrak{B}(\mathbb{R}^n)}$.

Example 2.2. One point sets are null sets. Indeed, for $x \in \mathbb{R}^n$,

$$0 \le \lambda^*(\{x\}) \le |\prod_{i=1}^n [x_i - \tfrac{1}{k}, x_i + \tfrac{1}{k}]| = (\tfrac{2}{k})^n$$

for all $k \ge 1$. It follows that finite sets and countable sets are null sets.

Example 2.3 (The Cantor set)**.** Consider the interval $C_0 = [0,1]$ and let C_1 be the set obtained by deleting the middle third open interval from $[0,1]$, i.e., $C_1 = [0,1/3] \cup [2/3,1]$. Next delete each middle third open interval of each subinterval in C_1, i.e., $C_2 = [0,1/3^2] \cup [2/3^2,1/3] \cup [2/3,7/3^2] \cup [8/3^2,1]$. Continuing this procedure we obtain a sequence $C_0 \supseteq C_1 \supseteq \cdots$ of compact sets. The intersection $C := \bigcap_{k=0}^\infty C_k$ is called the **Cantor set**. The Cantor set is a null set. Each C_k is a disjoint union of 2^k closed intervals, each of length 3^{-k}. Since $C \subseteq C_k$ for all k, $\lambda(C) \le (2/3)^k$ for all k, and thus $\lambda(C) = 0$.

The Cantor set is uncountable. To see this observe that

$$C = \Big\{x \in [0,1] : x = \sum_{j=1}^\infty \frac{a_j}{3^j},\ a_j \in \{0,2\}\Big\}$$

and consider the function $f : C \to [0,1]$ defined by

$$x = \sum_{j=1}^\infty \frac{a_j}{3^j} \mapsto f(x) = \sum_{j=1}^\infty \frac{b_j}{2^j}, \text{ where } b_j = \frac{a_j}{2}. \tag{2.1}$$

The function f is clearly surjective and thus C is uncountable.

Proposition 2.4. *We have $\lambda([a,b]) = |[a,b]| = (b_1-a_1)\cdots(b_n-a_n)$. In particular, degenerate boxes (where $a_i = b_i$ for at least one i) are null sets.*

Proof. Clearly, $\lambda([a,b]) \ge |[a,b]|$. Consider a grid in $\mathbb{R}^n$ of cubes Q of side length $1/k$. Let $\mathfrak{C}_1$ be the collection of all Q contained in $[a,b]$, and let $\mathfrak{C}_2$ be the collection of all Q intersecting $[a,b]$ as well as $[a,b]^c$. Then the number of cubes in $\mathfrak{C}_2$ is bounded by k^{n-1} times a constant C independent of k, and thus $\sum_{Q\in\mathfrak{C}_2} |Q| \le C/k$. Then, as $\bigcup_{Q\in\mathfrak{C}_1} Q \subseteq [a,b]$,

$$\sum_{Q\in\mathfrak{C}_1\cup\mathfrak{C}_2} |Q| \le |[a,b]| + C/k,$$

for all k, and therefore $\lambda([a,b]) \le |[a,b]|$. □

Lemma 2.5. *If $E = \bigcup_{i=1}^\infty Q_i$ is an almost disjoint union of cubes, then $\lambda(E) = \sum_{i=1}^\infty |Q_i|$.*

Proof. Let $\epsilon > 0$. For each Q_i choose a cube $\tilde{Q}_i$ contained in the interior of Q_i and such that $|Q_i| \le |\tilde{Q}_i| + \epsilon/2^i$. Then the cubes $\tilde{Q}_i$ are disjoint, and hence

$$\sum_{i=1}^\infty |Q_i| \ge \lambda(E) \ge \lambda\Big(\bigcup_{i=1}^\infty \tilde{Q}_i\Big) = \sum_{i=1}^\infty |\tilde{Q}_i| \ge \sum_{i=1}^\infty |Q_i| - \epsilon.$$

The statement follows, as ϵ was arbitrary. □

Lemma 2.6. *Every open set $U \subseteq \mathbb{R}^n$ is a countable almost disjoint union of cubes.*

Proof. Consider the collection $\mathfrak{C}_0$ of cubes of side length 1 defined by the lattice $\mathbb{Z}^n$. Set

$$\mathfrak{U}_0 := \{Q \in \mathfrak{C}_0 : Q \subseteq U\} \quad \text{and}$$
$$\mathfrak{V}_0 := \{Q \in \mathfrak{C}_0 : Q \cap U \ne \emptyset \text{ and } Q \cap U^c \ne \emptyset\}.$$

Let $\mathfrak{C}_1$ be the collection of cubes that we obtain by subdividing each cube in $\mathfrak{V}_0$ into 2^n cubes of side length $1/2$, and set

$$\begin{aligned}\mathfrak{U}_1 &:= \{Q \in \mathfrak{C}_1 : Q \subseteq U\} \quad \text{and} \\ \mathfrak{V}_1 &:= \{Q \in \mathfrak{C}_1 : Q \cap U \neq \emptyset \text{ and } Q \cap U^c \neq \emptyset\}.\end{aligned}$$

Continue this procedure. Then $U = \bigcup_{Q \in \mathfrak{U}} Q$, where $\mathfrak{U} := \bigcup_{i=0}^{\infty} \mathfrak{U}_i$, is a countable almost disjoint union of cubes. □

2.2. Radon measures on $\mathbb{R}^n$

Let X be a topological space. A measure μ on a σ-algebra $\mathfrak{S} \supseteq \mathfrak{B}(X)$ is called **outer regular** if

$$\mu(E) = \inf\{\mu(U) : E \subseteq U, U \text{ open}\}, \quad E \in \mathfrak{S},$$

and **inner regular** if

$$\mu(E) = \sup\{\mu(K) : K \subseteq E, K \text{ compact}\}, \quad E \in \mathfrak{S}.$$

If μ is both outer and inner regular, it is called **regular**.

A **Radon measure** on $\mathbb{R}^n$ is a Borel measure that is finite on compact sets. More generally, a Radon measure on a locally compact Hausdorff space X is a Borel measure that is finite on compact sets, outer regular on Borel sets, and inner regular on open sets. The next theorem shows that on $\mathbb{R}^n$ finiteness on compact sets implies regularity. By the Riesz representation theorem (e.g. [**5**]), the Radon measures on a locally compact Hausdorff space X correspond to the positive linear functionals on the space $C_c(X)$ of continuous functions with compact support.

We denote by $B_r(x) := \{y \in \mathbb{R}^n : |x-y| < r\}$ the open ball centered at $x \in \mathbb{R}^n$ of radius r with respect to the Euclidean norm $|x| := (x_1^2 + \cdots + x_n^2)^{1/2}$.

Theorem 2.7. *Each Radon measure μ on $\mathbb{R}^n$ is σ-finite and regular. For each Borel set A and each $\epsilon > 0$ there is an open set U and a closed set F so that*

$$F \subseteq A \subseteq U, \quad \text{and} \quad \mu(U \setminus F) \leq \epsilon. \tag{2.2}$$

PROOF. Evidently, μ is σ-finite.

Let us prove (2.2). First we assume that μ is finite. Let $\mathfrak{A}$ be the set of all Borel sets A that satisfy (2.2). We claim that $\mathfrak{A}$ is a σ-algebra. If $A \in \mathfrak{A}$, then for given $\epsilon > 0$ there are U and F satisfying (2.2), and thus $U^c \subseteq A^c \subseteq F^c$ and $\mu(F^c \setminus U^c) = \mu(U \setminus F) \leq \epsilon$, i.e., $A^c \in \mathfrak{A}$. Suppose that $A_i \in \mathfrak{A}$, $i \geq 1$, and $\epsilon > 0$. So there are open U_i and closed F_i so that $F_i \subseteq A_i \subseteq U_i$ and $\mu(U_i \setminus F_i) \leq \epsilon/2^{i+1}$. Then $U := \bigcup_{i=1}^{\infty} U_i$ is open and $F := \bigcup_{i=1}^{m} F_i$ is closed for finite m. Since μ is finite,

$$\mu\Big(\bigcup_{i=m+1}^{\infty} F_i \setminus F\Big) \leq \mu\Big(\bigcup_{i=1}^{\infty} F_i \setminus F\Big) = \mu\Big(\bigcup_{i=1}^{\infty} F_i\Big) - \mu\Big(\bigcup_{i=1}^{m} F_i\Big) \leq \epsilon/2,$$

for sufficiently large m, by Lemma 1.1. Since $U \setminus F \subseteq (U \setminus \bigcup_{i=1}^{\infty} F_i) \cup (\bigcup_{i=m+1}^{\infty} F_i \setminus F)$,

$$\mu(U \setminus F) \leq \sum_{i=1}^{\infty} \mu(U_i \setminus F_i) + \mu\Big(\bigcup_{i=m+1}^{\infty} F_i \setminus F\Big) \leq \epsilon.$$

Thus $\mathfrak{A}$ is a σ-algebra.

Every closed set $F \subseteq \mathbb{R}^n$ belongs to $\mathfrak{A}$, since the sets $U_k := \{x : \operatorname{dist}(x, F) < 1/k\}$ are open and satisfy $\mu(U_k \setminus F) \to 0$ as $k \to \infty$, by Lemma 1.1. It follows that $\mathfrak{A} = \mathfrak{B}(\mathbb{R}^n)$ and hence (2.2).

Assume that μ is not finite. Let A be a Borel set and let $\epsilon > 0$ be given. Since $\nu_i(E) := \mu(E \cap B_i(0))$ is a finite Radon measure on $\mathbb{R}^n$, by the above, there exists a closed set $C_i \subseteq (B_i(0) \setminus A)$ with $\nu_i((B_i(0) \setminus A) \setminus C_i) = \mu((B_i(0) \setminus A) \setminus C_i) \le \epsilon/2^i$. Then $U := \bigcup_{i=1}^{\infty} (B_i(0) \setminus C_i)$ is open,

$$A = \bigcup_{i=1}^{\infty} B_i(0) \cap A \subseteq \bigcup_{i=1}^{\infty} (B_i(0) \setminus C_i) = U$$

and

$$\mu(U \setminus A) \le \sum_{i=1}^{\infty} \mu((B_i(0) \setminus C_i) \setminus A) \le \epsilon.$$

Similarly, there exists a closed set $F_i \subseteq A_i := A \cap \{x : i \le |x| < i+1\}$ with $\mu(A_i \setminus F_i) \le \epsilon/2^{i+1}$,

$$F := \bigcup_{i=0}^{\infty} F_i \subseteq \bigcup_{i=0}^{\infty} A_i = A,$$

and

$$\mu(A \setminus F) \le \sum_{i=0}^{\infty} \mu(A_i \setminus F_i) \le \epsilon.$$

It remains to show that F is closed. If $x \in \overline{F}$ and $F \ni x_k \to x$, then $|x_k| \to |x|$ and so $x_k \in F_j \cup F_{j+1}$ for some j and for all sufficiently large k. Consequently, $x \in F_j \cup F_{j+1} \subseteq F$, since $F_j \cup F_{j+1}$ is closed. Thus (2.2) is proved.

Finally, we show that μ is regular. Let A be a Borel set, and let $\epsilon > 0$. Outer regularity is clear if $\mu(A) = \infty$ and follows from (2.2) if $\mu(A) < \infty$: there exists an open set $U \supseteq A$ so that $\mu(A) + \epsilon \ge \mu(A) + \mu(U \setminus A) = \mu(U)$. Next we show

$$\mu(A) = \sup\{\mu(F) : F \subseteq A, F \text{ closed}\}. \tag{2.3}$$

It follows from (2.2) if $\mu(A) < \infty$: there is a closed set $F \subseteq A$ so that $\mu(A) - \epsilon \le \mu(A) - \mu(A \setminus F) = \mu(F)$. If $\mu(A) = \infty$, write $A = \bigcup_{i=0}^{\infty} A_i$ where A_i is as above. Since μ is finite on compact sets, $\mu(A_i) < \infty$, and, again by (2.2), there exist closed $F_i \subseteq A_i$ with $\mu(F_i) \ge \mu(A_i) - 1/2^{i+1}$. By Lemma 1.1,

$$\lim_{k\to\infty} \mu\Big(\bigcup_{i=0}^{k} F_i\Big) = \mu\Big(\bigcup_{i=0}^{\infty} F_i\Big) = \sum_{i=0}^{\infty} \mu(F_i) \ge \mu(A) - 1 = \infty,$$

which shows (2.3), since $\bigcup_{i=0}^{k} F_i$ is closed. We finally have

$$\sup\{\mu(K) : K \subseteq A, K \text{ compact}\} = \sup\{\mu(F) : F \subseteq A, F \text{ closed}\},$$

since for any closed $F \subseteq \mathbb{R}^n$ the sets $K_k := F \cap \overline{B_k(0)}$ are compact and $\mu(F) = \lim_{k\to\infty} \mu(K_k)$. □

2.3. Properties of the Lebesgue measure

Proposition 2.8. *The Lebesgue outer measure is **Borel regular**, i.e., for each $E \subseteq \mathbb{R}^n$ there exists a Borel set $B \supseteq E$ such that $\lambda^*(E) = \lambda^*(B)$.*

Proof. If $\lambda^*(E) = \infty$ take $B = \mathbb{R}^n$. Suppose that $\lambda^*(E) < \infty$. For each $k \ge 1$ choose a countable collection $\mathfrak{C}_k$ of cubes so that

$$E \subseteq \bigcup_{Q\in\mathfrak{C}_k} Q =: B_k \quad \text{and} \quad \sum_{Q\in\mathfrak{C}_k} |Q| \le \lambda^*(E) + 1/k.$$

Then $B := \bigcap_{k=1}^{\infty} B_k$ is a Borel set that contains E and satisfies

$$\lambda^*(B) \le \lambda^*(B_k) \le \sum_{Q\in\mathfrak{C}_k} |Q| \le \lambda^*(E) + 1/k,$$

for all k, hence $\lambda^*(E) = \lambda^*(B)$. □

Theorem 2.9 (Regularity)**.** *The Lebesgue measure λ on $\mathbb{R}^n$ is σ-finite and regular. Its restriction to $\mathfrak{B}(\mathbb{R}^n)$ is a Radon measure.*

Proof. Clearly, λ is finite on compact sets and hence a Radon measure when restricted to $\mathfrak{B}(\mathbb{R}^n)$. Thus λ is σ-finite. By Theorem 2.7,

$$\lambda(B) = \inf\{\lambda(U) : B \subseteq U, U \text{ open}\} = \sup\{\lambda(K) : K \subseteq B, K \text{ compact}\}$$

for each Borel set B. If $E \subseteq \mathbb{R}^n$ is arbitrary, then, by Proposition 2.8, there is a Borel set $B \supseteq E$ with $\lambda^*(E) = \lambda^*(B)$, and thus

$$\lambda^*(E) = \lambda^*(B) = \inf\{\lambda(U) : B \subseteq U, U \text{ open}\} \geq \inf\{\lambda(U) : E \subseteq U, U \text{ open}\},$$

which shows that λ is outer regular.

To see that λ is inner regular let $E \subseteq \mathbb{R}^n$ be measurable, and suppose first that E is contained in a cube Q. Let $\epsilon > 0$. Then $\lambda(Q \setminus E) < \infty$ and, as λ is outer regular, there exists an open $U \supseteq Q \setminus E$ so that $\lambda(U) \leq \lambda(Q \setminus E) + \epsilon$. The set $K := Q \setminus U \subseteq E$ is compact and satisfies

$$\lambda(E) = \lambda(Q) - \lambda(Q \setminus E) \leq \lambda(Q) - \lambda(U) + \epsilon \leq \lambda(Q) - \lambda(Q \cap U) + \epsilon = \lambda(K) + \epsilon.$$

If E is not contained in a cube, for each $k \geq 1$, there is a compact $K_k \subseteq E \cap [-k,k]^n$ so that $\lambda(K_k) \geq \lambda(E \cap [-k,k]^n) - 1/k$. Hence $\lambda(K_k) \to \lambda(E)$ as $k \to \infty$ and hence λ is inner regular. □

Corollary 2.10 (Characterization of Lebesgue measurability)**.** *A set $E \subseteq \mathbb{R}^n$ is Lebesgue measurable if and only if there are an F_σ-set A and a G_δ-set B satisfying $A \subseteq E \subseteq B$ and $\lambda(B \setminus A) = 0$.*

An F_σ**-set** is a countable union of closed sets, and a G_δ**-set** is a countable intersection of open sets. The corollary implies, in view of Theorem 1.7, that the Lebesgue σ-algebra $\mathfrak{L}(\mathbb{R}^n)$ is the completion of the Borel σ-algebra $\mathfrak{B}(\mathbb{R}^n)$.

Proof. Assume that E is Lebesgue measurable. Theorem 2.9 implies that there exist open sets G_i and closed sets F_i satisfying $F_i \subseteq E \subseteq G_i$ and $\lambda(G_i \setminus F_i) \leq 1/i$. The sets $F = \bigcup_{i=1}^\infty F_i$ and $G = \bigcap_{i=1}^\infty G_i$ are as required.

Conversely, if there exist such F and G, then for any $A \subseteq \mathbb{R}^n$, we have $A \cap F \subseteq A \cap E \subseteq A \cap G$, $A \setminus G \subseteq A \setminus E \subseteq A \setminus F$,

$$\lambda^*((A \cap G) \setminus (A \cap F)) = \lambda^*(A \cap (G \setminus F)) \leq \lambda^*(G \setminus F) = 0,$$

and similarly $\lambda^*((A \setminus F) \setminus (A \setminus G)) = 0$. This implies $\lambda^*(A \cap E) = \lambda^*(A \cap F)$ and $\lambda^*(A \setminus E) = \lambda^*(A \setminus F)$, and thus

$$\lambda^*(A \cap E) + \lambda^*(A \setminus E) = \lambda^*(A \cap F) + \lambda^*(A \setminus F) = \lambda^*(A),$$

since F is measurable. □

Theorem 2.11 (Uniqueness of Lebesgue measure I)**.** *The Lebesgue measure λ is the unique measure on the Borel σ-algebra $\mathfrak{B}(\mathbb{R}^n)$ satisfying $\lambda([a,b]) = |[a,b]|$.*

Proof. By Proposition 2.4, $\lambda([a,b]) = |[a,b]|$. Suppose there is a second measure μ on $\mathfrak{B}(\mathbb{R}^n)$ with this property. We claim that λ and μ coincide on the collection $\mathfrak{A}$ of all finite disjoint unions of sets of the form $F \cap G$, where F is closed and G is open, and that $\mathfrak{A}$ is an algebra. The statement of the Theorem is then a consequence of Theorem 1.4, since the σ-algebra generated by $\mathfrak{A}$ is the Borel σ-algebra.

That $\mathfrak{A}$ is an algebra follows from Proposition 1.5, since the collection of sets of the form $F \cap G$, where F is closed and G is open, is an elementary family, in fact,

$$(F_1 \cap G_1) \cap (F_2 \cap G_2) = (F_1 \cap F_2) \cap (G_1 \cap G_2),$$

$$(F \cap G)^c = (F \cap G^c) \cup (F^c \cap G) \cup (F^c \cap G^c).$$

If F is closed and G is open, set $G_k := \{x \in \mathbb{R}^n : \operatorname{dist}(x, F) < 1/k\}$. Then G_k is open, $G_k \supseteq G_{k+1}$, and $F = \bigcap_{k=1}^{\infty} G_k$. If $\mu(G) < \infty$ then, by Lemma 1.1,

$$\mu\Big(\bigcap_{k=1}^{\infty}(G_k \cap G)\Big) = \lim_{j\to\infty} \mu(G_j \cap G) = \lim_{j\to\infty} \lambda(G_j \cap G) = \lambda\Big(\bigcap_{k=1}^{\infty}(G_k \cap G)\Big),$$

since λ and μ coincide on open sets, by Lemmas 2.5 and 2.6. Thus $\mu(F \cap G) = \lambda(F \cap G)$ if $\mu(G) < \infty$. If $\mu(G) = \infty$, then

$$\mu(F \cap G \cap (-k, k)^n) = \lambda(F \cap G \cap (-k, k)^n)$$

and letting $k \to \infty$ we find again $\mu(F \cap G) = \lambda(F \cap G)$. By σ-additivity, μ and λ coincide on $\mathfrak{A}$. $\square$

Corollary 2.12. *A Borel regular outer measure μ on $\mathbb{R}^n$ so that all Borel sets are μ-measurable and so that $\mu([a, b]) = |[a, b]|$ coincides with the Lebesgue outer measure.*

Proof. By Theorem 2.11, μ and λ^* coincide on all Borel sets. Let $E \subseteq \mathbb{R}^n$ be arbitrary. As μ and λ^* are Borel regular, there exist Borel sets $B_1, B_2 \supseteq E$ so that $\mu(B_1) = \mu(E)$ and $\lambda^*(B_2) = \lambda^*(E)$. Then, as $B_1 \cap B_2 \supseteq E$, we have $\mu(E) = \mu(B_1) \geq \mu(B_1 \cap B_2) \geq \mu(E)$, thus $\mu(E) = \mu(B_1 \cap B_2)$, and analogously $\lambda^*(E) = \lambda^*(B_1 \cap B_2)$. Therefore $\mu(E) = \lambda^*(E)$. $\square$

Proposition 2.13 (Translation invariance)**.** *The Lebesgue measure λ on $\mathbb{R}^n$ is translation invariant, i.e., if E is measurable and $y \in \mathbb{R}^n$, then the set $E + y := \{x + y : x \in E\}$ is measurable and $\lambda(E + y) = \lambda(E)$.*

Proof. The assertion is clearly true in the case that E is a cube. Consequently, for arbitrary $E \subseteq \mathbb{R}^n$ we have $\lambda^*(E + y) = \lambda^*(E)$. If E is measurable and $A \subseteq \mathbb{R}^n$ is arbitrary, then

$$\begin{aligned}
&\lambda^*(A \cap (E + y)) + \lambda^*(A \setminus (E + y)) \\
&\quad = \lambda^*(((A - y) \cap E) + y) + \lambda^*(((A - y) \setminus E) + y) \\
&\quad = \lambda^*((A - y) \cap E) + \lambda^*((A - y) \setminus E) \\
&\quad = \lambda^*(A - y) \\
&\quad = \lambda^*(A),
\end{aligned}$$

and so $E + y$ is measurable. $\square$

For further invariance properties, see Lemma 3.32 and Theorem 3.33.

Theorem 2.14 (Uniqueness of Lebesgue measure II)**.** *If μ is a translation invariant Radon measure on $\mathbb{R}^n$, then there is a constant $C > 0$ such that $\mu(E) = C\lambda(E)$ for all Borel sets E.*

Proof. Set $\mu([0, 1)^n) =: C < \infty$. Consider the grid of dyadic cubes of the form $[a_1, b_1) \times \cdots \times [a_n, b_n)$ defined by the lattice $2^{-k}\mathbb{Z}^n$. Since these cubes are all translates of each other,

$$2^{kn}\mu(Q) = \mu([0, 1)^n) = C\lambda([0, 1)^n) = C2^{kn}\lambda(Q),$$

for each such cube Q. We may infer that μ vanishes on degenerate boxes, and so $\mu(Q) = C\lambda(Q)$ for each closed dyadic cube $Q = [a_1, b_1] \times \cdots \times [a_n, b_n]$. Then $\mu(E) = C\lambda(E)$ for each open set E, by Lemmas 2.5 and 2.6, and thus for each Borel set E, by regularity of μ and λ, see Theorems 2.7 and 2.9. $\square$

Proposition 2.15 (Approximation by cubes)**.** *Let $E \subseteq \mathbb{R}^n$ be measurable with $\lambda(E) < \infty$. For each $\epsilon > 0$ there exist cubes $Q_1, \dots, Q_m$ such that $\lambda(E \triangle \bigcup_{i=1}^m Q_i) < \epsilon$, where $E \triangle F := (E \setminus F) \cup (F \setminus E) = (E \cup F) \setminus (E \cap F)$ is the **symmetric difference**.*

Proof. Let $\epsilon > 0$ be fixed and let Q_i be cubes such that $E \subseteq \bigcup_{i=1}^\infty Q_i$ and $\sum_{i=1}^\infty |Q_i| \le \lambda(E) + \epsilon/2$. Since $\lambda(E) < \infty$ the infinite sum converges and there exists m such that $\sum_{i=m+1}^\infty |Q_i| < \epsilon/2$. Then,

$$\begin{aligned}
\lambda\Big(E \triangle \bigcup_{i=1}^m Q_i\Big) &= \lambda\Big(E \setminus \bigcup_{i=1}^m Q_i\Big) + \lambda\Big(\bigcup_{i=1}^m Q_i \setminus E\Big) \\
&\le \lambda\Big(\bigcup_{i=m+1}^\infty Q_i\Big) + \lambda\Big(\bigcup_{i=1}^m Q_i \setminus E\Big) \\
&\le \sum_{i=m+1}^\infty |Q_i| + \sum_{i=1}^\infty |Q_i| - \lambda(E) < \epsilon. \qquad \square
\end{aligned}$$

2.4. Non-measurable sets

Every set of positive measure in $\mathbb{R}$ has non-measurable subsets.

Theorem 2.16 (Existence of non-measurable sets)**.** *Let $E \subseteq \mathbb{R}$. If every subset of E is Lebesgue measurable, then $\lambda(E) = 0$.*

Proof. On $\mathbb{R}$ consider the equivalence relation $x \sim y :\Leftrightarrow x - y \in \mathbb{Q}$. The axiom of choice allows us to choose exactly one element in each equivalence class and to gather these elements in one set N; such a set is called a **Vitali set**.

For $q \in \mathbb{Q}$ consider the translates $N+q$ which are pairwise disjoint; otherwise we have $x+q_1 = y+q_2$ and thus $x-y \in \mathbb{Q}$, but x and y belong to different equivalence classes, a contradiction. Fix $p \in \mathbb{Q}$ and set $E_p := E \cap (N + p)$. By assumption, E_p is measurable. Let $K \subseteq E_p$ be compact and set $L := \bigcup_{q \in \mathbb{Q} \cap [0,1]} K + q$. Then $\lambda(L) < \infty$, since L is bounded, and, since the sets $K + q$ are disjoint, $\lambda(L) = \sum_{q \in \mathbb{Q} \cap [0,1]} \lambda(K)$. Thus $\lambda(K) = 0$. Since K was arbitrary, we may conclude that $\lambda(E_p) = 0$, by regularity of λ. Consequently, $\lambda(E) = 0$, because $E = \bigcup_{p \in \mathbb{Q}} E_p$. $\square$

In the previous proof the axiom of choice plays an essential role. In fact, Solovay constructed a model in which all axioms of Zermelo–Frankel set theory, except the axiom of choice, hold and in which every subset of $\mathbb{R}$ is Lebesgue measurable.

There exists a *finitely* additive translation-invariant set function assigning boxes their volume that is defined on *all* subsets of $\mathbb{R}$, respectively $\mathbb{R}^2$, but not in higher dimensions. In fact, any ball in $\mathbb{R}^3$ can be decomposed into finitely many disjoint subsets, which can then be reassembled using only rotations and translations to form two copies of the original ball; this results is called the **Banach–Tarski paradox**.

CHAPTER 3

Integration

3.1. Measurable functions

A set X equipped with a σ-algebra $\mathfrak{S} \subseteq \mathfrak{P}(X)$ is called a **measurable space** $(X, \mathfrak{S})$. A mapping $f : X \to Y$ between measurable spaces $(X, \mathfrak{S})$ and $(Y, \mathfrak{T})$ is called $(\mathfrak{S}, \mathfrak{T})$**-measurable** if $f^{-1}(E) \in \mathfrak{S}$ for every $E \in \mathfrak{T}$.

It is obvious by definition that the composition of measurable mappings is measurable, more precisely, if $f : X \to Y$ is $(\mathfrak{S}, \mathfrak{T})$-measurable and $g : Y \to Z$ is $(\mathfrak{T}, \mathfrak{U})$-measurable then $g \circ f$ is $(\mathfrak{S}, \mathfrak{U})$-measurable.

Lemma 3.1. *If $\mathfrak{T}$ is generated by $\mathfrak{A}$, then a mapping $f : X \to Y$ is $(\mathfrak{S}, \mathfrak{T})$-measurable if and only if $f^{-1}(E) \in \mathfrak{S}$ for every $E \in \mathfrak{A}$.*

Proof. This follows from the fact that $\{E \subseteq Y : f^{-1}(E) \in \mathfrak{S}\}$ is a σ-algebra on Y containing $\mathfrak{A}$, and hence containing $\mathfrak{T}$. $\square$

If follows that any continuous mapping $f : X \to Y$ between topological spaces X and Y is $(\mathfrak{B}(X), \mathfrak{B}(Y))$-measurable.

If f is a real or complex valued function on a measurable space $(X, \mathfrak{S})$ then we say that f is $\mathfrak{S}$**-measurable** if f is $(\mathfrak{S}, \mathfrak{B}(\mathbb{R}))$- or $(\mathfrak{S}, \mathfrak{B}(\mathbb{C}))$-measurable. For instance, $f : \mathbb{R}^n \to \mathbb{C}$ is **Lebesgue measurable** if it is $(\mathfrak{L}(\mathbb{R}^n), \mathfrak{B}(\mathbb{C}))$-measurable, and it is **Borel measurable** or also a **Borel function** if it is $(\mathfrak{B}(\mathbb{R}^n), \mathfrak{B}(\mathbb{C}))$-measurable.

Note that if $f, g : \mathbb{R} \to \mathbb{R}$ are Lebesgue measurable, then $g \circ f$ need not be Lebesgue measurable.

The **characteristic function** $\chi_A : X \to \mathbb{R}$ of a subset $A \subseteq X$,

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

is $\mathfrak{S}$-measurable if and only if A is $\mathfrak{S}$-measurable.

Proposition 3.2. *Let X be a measurable space.*

(1) *If $f_1, f_2 : X \to \mathbb{R}$ are measurable, then $f = (f_1, f_2) : X \to \mathbb{R}^2$ is measurable.*
(2) *A complex valued function $f : X \to \mathbb{C}$ is measurable if and only if* $\operatorname{Re} f$ *and* $\operatorname{Im} f$ *are measurable. In this case $|f|$ is measurable.*
(3) *If $f, g : X \to \mathbb{C}$ are measurable, then so are $f + g$ and fg.*

Proof. (1) Every open subset $U \subseteq \mathbb{R}^2$ is a countable union of cubes $U = \bigcup_{i=1}^\infty Q_i$, by Lemma 2.6. Then $f^{-1}(U) = f^{-1}(\bigcup_{i=1}^\infty Q_i) = \bigcup_{i=1}^\infty f^{-1}(Q_i)$ is measurable, since each $f^{-1}(Q_i) = f_1^{-1}(I_{i,1}) \cap f_2^{-1}(I_{i,2})$ is measurable, where $Q_i = I_{i,1} \times I_{i,2}$ and $I_{i,1}$, $I_{i,2}$ are compact intervals.

(2) follows from (1) and the fact if $f : X \to \mathbb{C}$ is measurable then the composite $g \circ f$ for any continuous mapping g is measurable. This also implies (3). $\square$

The **extended real line** is the set $[-\infty,\infty] = \mathbb{R} \cup \{\pm\infty\}$ with the topology generated by the open sets of $\mathbb{R}$ and all intervals $[-\infty, a)$ and $(a, \infty]$. Then $\mathfrak{B}([-\infty,\infty]) = \{E \subseteq [-\infty,\infty] : E \cap \mathbb{R} \in \mathfrak{B}(\mathbb{R})\}$. A function $f : X \to [-\infty,\infty]$ on a measurable space $(X, \mathfrak{S})$ is said to be $\mathfrak{S}$-measurable if it is $(\mathfrak{S}, \mathcal{B}([-\infty,\infty]))$-measurable.

Proposition 3.3. *Let $(X, \mathfrak{S})$ be a measurable space. A function $f : X \to [-\infty,\infty]$ is $\mathfrak{S}$-measurable if and only if $f^{-1}((a,\infty]) \in \mathfrak{S}$ for all $a \in \mathbb{R}$.*

Proof. By Lemma 3.1, it suffices to show that $\{(a,\infty] : a \in \mathbb{R}\}$ generates $\mathcal{B}([-\infty,\infty])$. This follows from

$$[-\infty, a) = \bigcup_{i=1}^{\infty} [-\infty, a - \tfrac{1}{i}] = \bigcup_{i=1}^{\infty} (a - \tfrac{1}{i}, \infty]^c$$

and from $(a,b) = [-\infty, b) \cap (a, \infty]$. $\square$

It follows that every upper or lower semicontinuous function is Borel measurable. Recall that a function $f : X \to [-\infty,\infty]$ on a topological space X is upper (or lower) semicontinuous if $\{x : f(x) < a\}$ (or $\{x : f(x) > a\}$) is open for all $a \in \mathbb{R}$.

Theorem 3.4 (Pointwise limits of measurable functions)**.** *Let $f_n : X \to [-\infty,\infty]$, $n \in \mathbb{N}$, be a sequence of measurable functions on a measurable space $(X, \mathfrak{S})$. Then*

$$\inf_{n\in\mathbb{N}} f_n, \quad \sup_{n\in\mathbb{N}} f_n, \quad \liminf_{n\to\infty} f_n, \quad \limsup_{n\to\infty} f_n$$

are measurable. Thus, the limit of any pointwise convergent sequence of complex valued measurable functions is measurable.

Proof. Let $g := \sup_{n\in\mathbb{N}} f_n$. Then $g^{-1}((a,\infty]) = \bigcup_{n\in\mathbb{N}} f_n^{-1}((a,\infty])$ and thus g is measurable, by Proposition 3.3. The result for the infimum is analogous (note that $\inf_n f_n = -\sup_n(-f_n)$). Since

$$\limsup_{n\to\infty} f_n = \inf_{n\in\mathbb{N}} \sup_{m\geq n} f_m \quad \text{and} \quad \liminf_{n\to\infty} f_n = \sup_{n\in\mathbb{N}} \inf_{m\geq n} f_m$$

the result follows. $\square$

Thus, if $f, g : X \to [-\infty,\infty]$ are measurable, then so are the functions $\min\{f, g\}$ and $\max\{f, g\}$. In particular, this is true for $f^+ := \max\{f, 0\}$ and $f^- := -\min\{f, 0\}$, the **positive** and **negative part** of f. Note that

$$f = f^+ - f^- \quad \text{and} \quad |f| = f^+ + f^-.$$

For a complex valued function $f : X \to \mathbb{C}$ we have its **polar decomposition**,

$$f = |f| \operatorname{sgn} f, \qquad \text{where} \qquad \operatorname{sgn} z := \begin{cases} z/|z| & z \neq 0 \\ 0 & z = 0 \end{cases}.$$

If f is measurable, then so is $|f|$ and $\operatorname{sgn} f$. Indeed, $|\ | : \mathbb{C} \to \mathbb{R}$ is continuous, and the preimage $\operatorname{sgn}^{-1}(U)$ of an open set $U \subseteq \mathbb{C}$ is either open or of the form $V \cup \{0\}$, where V is open, and hence sgn is Borel.

Example 3.5 (The Cantor function)**.** Consider $\mathbb{R}$ with the Lebesgue measure λ. Let C be the Cantor set from Example 2.3. The Cantor set is a closed null set, in particular, C is Borel. Let $f : C \to [0,1]$ be the function defined in (2.1). It is easy to see that $x, y \in C$ implies $f(x) < f(y)$ unless x and y are the endpoints of one of the intervals removed from $[0,1]$ to obtain C. In the latter case $f(x) = k/2^\ell$ for some integers k, ℓ, and $f(x)$ and $f(y)$ are the two expansions in base 2 of this number. Thus, we can extend f to a function $f : [0,1] \to [0,1]$ by setting

$f|_{(a,b)} \equiv f(a) = f(b)$ on each connected component (a, b) of $[0, 1] \setminus C$. Then f is still nondecreasing and it is continuous, since its range is all of $[0, 1]$. This is called the **Cantor function**.

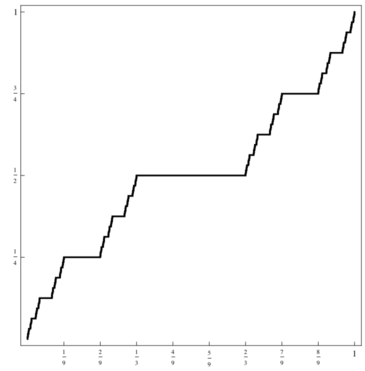


FIGURE 1. The Cantor function. (Generated with *Mathematica* and based on the code provided in [**15**, p.173].)

As a by-product we obtain the existence of Lebesgue null sets which are not Borel as follows. The function $g(x) = x+f(x)$ is strictly increasing and continuous, thus a homeomorphism onto its image. The image $g(C)$ has positive measure and so, by Theorem 2.16, there is a non-measurable subset $F \subseteq g(C)$. If we set $E = g^{-1}(F)$, then $E \subseteq C$ and hence E is a null set. But E is not Borel. Indeed, if E were Borel, then so were F, since g^{-1} is continuous.

3.2. Approximation by simple functions

Let $(X, \mathfrak{S})$ be a measurable space. A **simple function** is a complex valued measurable function on X with finite image. A simple function is representable in the form

$$s = \sum_{i=1}^{N} a_i \chi_{E_i},$$

where all $E_i \in \mathfrak{S}$ and $a_i \in \mathbb{C}$. In fact, setting $E_i := \{x : s(x) = a_i\}$, where $s(X) = \{a_1, \dots, a_N\}$, yields such a representation with the additional property that all a_i are distinct and all E_i are disjoint; we call this particular representation **canonical**.

Simple functions will be for the Lebesgue integral what **step functions** (where E_i are just boxes in $X = \mathbb{R}^n$) are for the Riemann integral.

Theorem 3.6 (Approximation by simple functions)**.** *Let $f : X \to [0, \infty]$ be measurable. There exist simple functions s_i on X such that*

(1) $0 \le s_1 \le s_2 \le \dots \le f$
(2) $\lim_{i\to\infty} s_i(x) = f(x)$ *for every $x \in X$.*

PROOF. To each integer $m \ge 1$ and each $t > 0$ there corresponds a unique integer $k = k(m, t)$ that satisfies $k/2^m \le t < (k + 1)/2^m$. Define

$$g_m(t) := \begin{cases} k(m,t)/2^m & \text{if } 0 \le t < m \\ m & \text{if } m \le t \le \infty. \end{cases}$$

We have

$$t - 2^{-m} < g_m(t) \le t \quad \text{if } 0 \le t < m.$$

Thus $\lim_{m\to\infty} g_m(t) = t$ for every $t \in [0,\infty]$, and clearly $0 \le g_1 \le g_2 \le \cdots \le t$. Then $s_m := g_m \circ f$ are simple functions with the required properties. □

Corollary 3.7. *Let $f : X \to [-\infty,\infty]$ or $f : X \to \mathbb{C}$ be measurable. There exist simple functions s_i on X such that*

(1) $0 \le |s_1| \le |s_2| \le \cdots \le |f|$
(2) $\lim_{i\to\infty} s_i(x) = f(x)$ *for every* $x \in X$.

Proof. Consider first the case $f : X \to [-\infty,\infty]$. By Theorem 3.6 applied to f^+ and f^-, there are simple functions $0 \le s_1^+ \le s_2^+ \le \cdots \le f^+$ and $0 \le s_1^- \le s_2^- \le \cdots \le f^-$ so that $\lim_{i\to\infty} s_i^{\pm}(x) = f^{\pm}(x)$ for every $x \in X$. Then $s_i := s_i^+ - s_i^-$ is as required. The case $f : X \to \mathbb{C}$ is an easy consequence. □

Given a measure μ on $(X,\mathfrak{S})$, one often wants to ignore μ-null sets. In this respect we have for complete measures:

Proposition 3.8. *Assume that μ is complete, and that f, g, f_i are functions with values in $[-\infty,\infty]$ or in $\mathbb{C}$.*

(1) *If f is measurable and $f = g$ μ-a.e., then g is measurable.*
(2) *If f_i are measurable and $f_i \to f$ μ-a.e., then f is measurable.*

Proof. We may assume that all functions have values in the extended real line.

(1) Since μ is complete, the sets $E = \{x : f(x) \neq g(x)\}$ and $g^{-1}((a,\infty]) \cap E$ are measurable, and thus $g^{-1}((a,\infty]) = (f^{-1}((a,\infty]) \cap E^c) \cup (g^{-1}((a,\infty]) \cap E)$ is measurable.

(2) Let $E = \{x : f_i(x) \to f(x)\}$. Then $f_i\chi_E \to f\chi_E$ and $\mu(E^c) = 0$. By Theorem 3.4, $f\chi_E$ is measurable, and so $f^{-1}((a,\infty]) = (f^{-1}((a,\infty]) \cap E^c) \cup ((f\chi_E)^{-1}((a,\infty]) \cap E)$ is measurable. □

If the measure is not complete we still have:

Proposition 3.9. *Let $(X,\mathfrak{S},\mu)$ be a measure space and $(X,\overline{\mathfrak{S}},\overline{\mu})$ its completion. If f is a $\overline{\mathfrak{S}}$-measurable function on X, then there is a $\mathfrak{S}$-measurable function g such that $f = g$ $\overline{\mu}$-a.e.*

Proof. This is immediate from the definition of the completion $\overline{\mu}$, if $f = \chi_E$ with $E \in \overline{\mathfrak{S}}$ and hence if f is a $\overline{\mathfrak{S}}$-measurable simple function. By Corollary 3.7, there is a sequence of $\overline{\mathfrak{S}}$-measurable simple functions s_i converging pointwise to f. For each i, there is a $\mathfrak{S}$-measurable function g_i so that $s_i = g_i$ except on a set $E_i \in \overline{\mathfrak{S}}$ with $\overline{\mu}(E_i) = 0$. Choose a set $F \in \mathfrak{S}$ with $\mu(F) = 0$ and $F \supseteq \bigcup_{i=1}^{\infty} E_i$; it exists by the definition of $\overline{\mathfrak{S}}$. Then $g = \lim_{i\to\infty} g_i\chi_{F^c}$ is as required, by Theorem 3.4. □

3.3. Integration on a measure space

Let us fix the arithmetic in $[0,\infty]$. We define

$$a + \infty = \infty + a = \infty \qquad \text{if } a \in [0,\infty]$$

$$a \cdot \infty = \infty \cdot a = \begin{cases} \infty & \text{if } a \in (0,\infty] \\ 0 & \text{if } a = 0. \end{cases}$$

Then addition and multiplication in $[0,\infty]$ are commutative, associative, and distributive. The cancellation laws have to be treated with some care; $a + c = b + c$ implies $a = b$ only if $c \in [0,\infty)$, and $ac = bc$ implies $a = b$ only if $c \in (0,\infty)$.

Lemma 3.10. *If $f, g : X \to [0, \infty]$ are measurable, then so are $f + g$ and fg.*

Proof. By Theorem 3.6 there exist simple functions $0 \le s_1 \le s_2 \le \cdots \le f$ and $0 \le t_1 \le t_2 \le \cdots \le g$ such that $s_i(x) \to f(x)$ and $t_i(x) \to g(x)$ for all x. Then $s_i(x) + t_i(x) \to f(x) + g(x)$ and $s_i(x)t_i(x) \to f(x)g(x)$, and Theorem 3.4 implies the statement. $\square$

Throughout this section let $(X, \mathfrak{S}, \mu)$ be a fixed measure space. We will define the integral in three steps:

- for positive simple functions,
- for positive functions,
- for complex valued functions.

Step 1. Integrating positive simple functions. The **(Lebesgue) integral** $\int s \, d\mu$ with respect to the measure μ of a simple function $s : X \to [0, \infty)$ with canonical representation $s = \sum_{i=1}^{N} a_i \chi_{E_i}$ is defined by

$$\int s \, d\mu := \sum_{i=1}^{N} a_i \mu(E_i),$$

where we use the convention $0 \cdot \infty = 0$. If $E \in \mathfrak{S}$, then $s\chi_E$ is a simple function, and we define

$$\int_E s \, d\mu := \int s\chi_E \, d\mu = \sum_{i=1}^{N} a_i \mu(E_i \cap E).$$

Lemma 3.11. *Let $s : X \to [0, \infty)$ be a simple function and let $s = \sum_{i=1}^{N} a_i \chi_{E_i}$ be any representation as a linear combination of characteristic functions. Then*

$$\int s \, d\mu = \sum_{i=1}^{N} a_i \mu(E_i).$$

Proof. There exists a refinement $\{F_1, \ldots, F_M\}$ of $\bigcup_{i=1}^{N} E_i$ such that

$$F_j \in \mathfrak{S} \text{ are disjoint}, \quad \bigcup_{i=1}^{N} E_i = \bigcup_{j=1}^{M} F_j, \quad \text{and} \quad E_i = \bigcup_{F_j \subseteq E_i} F_j.$$

It suffices to take

$$\{F_1, \ldots, F_M\} = \{G_1 \cap \cdots \cap G_N : G_i \in \{E_i, (E_i)^c\}\} \setminus \{(E_1)^c \cap \cdots \cap (E_N)^c\}.$$

If we set $b_j := \sum_{F_j \subseteq E_i} a_i$ then $s = \sum_{j=1}^{M} b_j \chi_{F_j}$. The numbers b_j may not be distinct and some may be zero. If $b \in \{b_j\}$ is non-zero, set $H_b := \bigcup_{b_j = b} F_j$. Clearly, the sets H_b are pairwise disjoint and satisfy $\mu(H_b) = \sum_{b_j = b} \mu(F_j)$. We have $s = \sum b\chi_{H_b}$ where the sum is over the non-zero values in $\{b_j\}$, and then

$$\int s \, d\mu = \sum b\mu(H_b) = \sum_{j=1}^{M} b_j \mu(F_j) = \sum_{j=1}^{M} \sum_{F_j \subseteq E_i} a_i \mu(F_j) = \sum_{i=1}^{N} a_i \mu(E_i). \qquad \square$$

Lemma 3.12. *Let s and t be positive simple functions on X, and $E, F, E_i \in \mathfrak{S}$.*

(1) *For $a \in [0, \infty)$ we have $\int as \, d\mu = a \int s \, d\mu$.*
(2) *$\int (s + t) \, d\mu = \int s \, d\mu + \int t \, d\mu$.*
(3) *If $s \le t$, then $\int s \, d\mu \le \int t \, d\mu$.*
(4) *If $E \subseteq F$, then $\int_E s \, d\mu \le \int_F s \, d\mu$.*
(5) *The mapping $E \mapsto \int_E s \, d\mu$ is a measure on $\mathfrak{S}$.*
(6) *If $\mu(E) = 0$ then $\int_E s \, d\mu = 0$.*

Proof. (1) is obvious. Let $s = \sum_{i=1}^N a_i \chi_{E_i}$ and $t = \sum_{j=1}^M b_j \chi_{F_j}$ be canonical representations. Then $E_i = \bigcup_{j=1}^M (E_i \cap F_j)$ and $F_j = \bigcup_{i=1}^N (E_i \cap F_j)$ and these unions are disjoint. Thus, by finite additivity of μ,

$$\int s\, d\mu + \int t\, d\mu = \sum_{i=1}^N \sum_{j=1}^M (a_i + b_j)\mu(E_i \cap F_j) = \int (s+t)\, d\mu,$$

which shows (2). If $s \le t$, then $a_i \le b_j$ whenever $E_i \cap F_j \ne \emptyset$, and hence

$$\int s\, d\mu = \sum_{i=1}^N \sum_{j=1}^M a_i \mu(E_i \cap F_j) \le \sum_{i=1}^N \sum_{j=1}^M b_j \mu(E_i \cap F_j) = \int t\, d\mu,$$

that is (3). (4) follows from (3), or from monotonicity of μ. For (5), if $F_1, F_2, \ldots \in \mathfrak{S}$ are disjoint, then

$$\int_{\bigcup_{j=1}^\infty F_j} s\, d\mu = \sum_{i=1}^N a_i \mu(E_i \cap \bigcup_{j=1}^\infty F_j) = \sum_{j=1}^\infty \sum_{i=1}^N a_i \mu(E_i \cap F_j) = \sum_{j=1}^\infty \int_{F_j} s\, d\mu.$$

(6) follows from the definition. □

Step 2. Integrating positive functions. The **(Lebesgue) integral** $\int f\, d\mu$ with respect to the measure μ of a positive measurable function $f : X \to [0,\infty]$ is defined by

$$\int f\, d\mu := \sup \Big\{ \int s\, d\mu : s \text{ simple and } 0 \le s \le f \Big\} \in [0,\infty].$$

If $E \in \mathfrak{S}$, we define

$$\int_E f\, d\mu := \int f\chi_E\, d\mu = \sup \Big\{ \int_E s\, d\mu : s \text{ simple and } 0 \le s \le f \Big\}.$$

For simple f this definition coincides with the earlier one, by Lemma 3.12, (3).

Lemma 3.13. *For measurable functions $f, g : X \to [0,\infty]$ we have*

$$\int af\, d\mu = a \int f\, d\mu, \qquad \textit{for } a \in [0,\infty),$$

and

$$\int f\, d\mu \le \int g\, d\mu, \qquad \textit{if } f \le g.$$

Proof. This is clear from the definition. □

Note that this implies $\int_E f\, d\mu \le \int_F f\, d\mu$ if $E \subseteq F$.

Theorem 3.14 (Monotone convergence theorem or Beppo Levi's theorem)**.** *Let f_i be measurable functions on X satisfying*

(1) $0 \le f_1 \le f_2 \le \cdots \le \infty$
(2) $\lim_{i\to\infty} f_i(x) = f(x)$ *for all $x \in X$.*

Then f is measurable, and

$$\lim_{i\to\infty} \int f_i\, d\mu = \int f\, d\mu.$$

Proof. By Theorem 3.4, f is measurable. Since $f_i \le f_{i+1} \le f$ for all i, we have $\int f_i\, d\mu \le \int f_{i+1}\, d\mu \le \int f\, d\mu$, by Lemma 3.13, and hence $\lim_{i\to\infty} \int f_i\, d\mu$ exists (possibly equal to ∞) and satisfies

$$\lim_{i\to\infty} \int f_i\, d\mu \le \int f\, d\mu.$$

Let s be a simple function satisfying $0 \leq s \leq f$, and let $a \in (0,1)$. Set

$$E_i := \{x : f_i(x) \geq as(x)\}.$$

Then $E_i \in \mathfrak{S}$, $E_1 \subseteq E_2 \subseteq \cdots$, $X = \bigcup_{i=1}^\infty E_i$, and, by Lemma 3.13,

$$\int f_i \, d\mu \geq \int_{E_i} f_i \, d\mu \geq a \int_{E_i} s \, d\mu.$$

Since $E \mapsto \int_E s \, d\mu$ is a measure, by Lemma 3.12, $\lim_{i\to\infty} \int_{E_i} s \, d\mu = \int s \, d\mu$, by Lemma 1.1, and so

$$\lim_{i\to\infty} \int f_i \, d\mu \geq a \int s \, d\mu,$$

and, as this holds for every $a < 1$, it remains true for $a = 1$. Taking the supremum over all simple functions s satisfying $0 \leq s \leq f$, we get

$$\lim_{i\to\infty} \int f_i \, d\mu \geq \int f \, d\mu.$$

The proof is complete. $\square$

Corollary 3.15. *Let $f_i : X \to [0,\infty]$ be measurable functions, and $f = \sum_{i=1}^\infty f_i$. Then*

$$\int f \, d\mu = \sum_{i=1}^\infty \int f_i \, d\mu.$$

Proof. First we prove the statement for the sum of two functions f and g. By Theorem 3.6, there exist simple functions $0 \leq s_1 \leq s_2 \leq \cdots \leq f$ and $0 \leq t_1 \leq t_2 \leq \cdots \leq g$ with $s_i(x) \to f(x)$ and $t_i(x) \to g(x)$ for all x. Then $s_i + t_i$ is an increasing sequence of simple functions that converges pointwise to $f + g$, and Theorem 3.14 together with Lemma 3.12 imply

$$\int (f+g)\, d\mu = \lim_{i\to\infty} \int (s_i + t_i)\, d\mu = \lim_{i\to\infty} \int s_i \, d\mu + \lim_{i\to\infty} \int t_i \, d\mu = \int f \, d\mu + \int g \, d\mu.$$

By induction, we obtain $\int \sum_{i=1}^n f \, d\mu = \sum_{i=1}^n \int f_i \, d\mu$ for finite n, and applying Theorem 3.14 to $F_n := \sum_{i=1}^n f_i$, implies the result for infinite sums. $\square$

Corollary 3.16. *Let $f : X \to [0,\infty]$ be measurable. Then $\nu(E) = \int_E f \, d\mu$ is a measure on $\mathfrak{S}$. If $g : X \to [0,\infty]$ is measurable, then*

$$\int g \, d\nu = \int gf \, d\mu.$$

Proof. Let $E_i \in \mathfrak{S}$ be pairwise disjoint. By Corollary 3.15,

$$\nu\Big(\bigcup_{i=1}^\infty E_i\Big) = \int \sum_{i=1}^\infty \chi_{E_i} f \, d\mu = \sum_{i=1}^\infty \int \chi_{E_i} f \, d\mu = \sum_{i=1}^\infty \nu(E_i),$$

so ν is a measure on $\mathfrak{S}$. By definition, $\int g \, d\nu = \int gf \, d\mu$ holds for $g = \chi_E$, $E \in \mathfrak{S}$, and hence for each positive simple function,

$$\int \sum_{i=1}^N a_i \chi_{E_i} \, d\nu = \sum_{i=1}^N a_i \int \chi_{E_i} \, d\nu = \sum_{i=1}^N a_i \int \chi_{E_i} f \, d\mu = \int \sum_{i=1}^N a_i \chi_{E_i} f \, d\mu.$$

The general case follows from Theorem 3.6 and the monotone convergence theorem 3.14. $\square$

Corollary 3.17 (Fatou's lemma)**.** *For measurable functions $f_i : X \to [0,\infty]$,*

$$\int \liminf_{i\to\infty} f_i \, d\mu \leq \liminf_{i\to\infty} \int f_i \, d\mu.$$

PROOF. Set $g_j := \inf_{i \geq j} f_i$. Then $g_j \leq g_{j+1}$ and $g_j \leq f_i$ for all $i \geq j$. Thus, $\int g_j \, d\mu \leq \inf_{i \geq j} \int f_i \, d\mu$. Since $\lim_{j \to \infty} g_j = \liminf_{i \to \infty} f_i$, the monotone convergence theorem 3.14 implies that

$$\int \liminf_{i \to \infty} f_i \, d\mu = \lim_{j \to \infty} \int g_j \, d\mu \leq \liminf_{i \to \infty} \int f_i \, d\mu. \qquad \square$$

Proposition 3.18. *For a measurable function $f : X \to [0, \infty]$, $\int f \, d\mu = 0$ if and only if $f = 0$ μ-a.e.*

PROOF. This is clearly true if f is a simple function; if $f = \sum_{i=1}^{N} a_i \chi_{E_i}$ is the canonical representation then $a_i \geq 0$, and $\int f \, d\mu = 0$ if and only if for each i either $a_i = 0$ or $\mu(E_i) = 0$. In general, if $f = 0$ μ-a.e. and s is a simple function with $0 \leq s \leq f$, then $s = 0$ μ-a.e. and thus $\int f \, d\mu = \sup_{s \leq f} \int s \, d\mu = 0$. Conversely, if $f \neq 0$ μ-a.e., then there is an integer $k \geq 1$ so that $\mu(\{x : f(x) > 1/k\}) > 0$, since $\{x : f(x) > 0\} = \bigcup_{k=1}^{\infty} \{x : f(x) > 1/k\}$. But then $f > k^{-1} \chi_{\{x : f(x) > 1/k\}}$ and therefore $\int f \, d\mu \geq k^{-1} \mu(\{x : f(x) > 1/k\}) > 0$. $\square$

Corollary 3.19. *Let $f_i, f : X \to [0, \infty]$ be measurable functions so that $f_i(x) \nearrow f(x)$ for μ-a.e. $x \in X$, then $\lim_{i \to \infty} \int f_i \, d\mu = \int f \, d\mu$.*

PROOF. There is a measurable set E with $\mu(E^c) = 0$ and such that $f_i(x) \nearrow f(x)$ for each $x \in E$. Then $f - f\chi_E = 0$ a.e. and $f_i - f_i \chi_E = 0$ a.e. and by the monotone convergence theorem 3.14 and Proposition 3.18,

$$\lim_{i \to \infty} \int f_i \, d\mu = \lim_{i \to \infty} \int f_i \chi_E \, d\mu = \int f \chi_E \, d\mu = \int f \, d\mu. \qquad \square$$

Step 3. Integrating complex valued functions. We define

$$L^1(\mu) := \Big\{ f : X \to \mathbb{C} \text{ measurable } : \int |f| \, d\mu < \infty \Big\}.$$

If f is measurable, then so is $|f|$, by Proposition 3.2, any hence the integral is defined. The members of $L^1(\mu)$ are called **(Lebesgue) integrable** functions with respect to the measure μ.

For $f \in L^1(\mu)$, $f = u + iv$, and $E \in \mathfrak{S}$, we define the **(Lebesgue) integral** over E with respect to the measure μ by

$$\int_E f \, d\mu := \Big(\int_E u^+ \, d\mu - \int_E u^- \, d\mu \Big) + i \Big(\int_E v^+ \, d\mu - \int_E v^- \, d\mu \Big).$$

The measurability of f guarantees the measurability of $u^\pm, v^\pm$, which are all positive functions. So all integrals on the right-hand side exist. As $u^\pm \leq |u| \leq |f|$ and $v^\pm \leq |v| \leq |f|$ all four integrals are finite, and thus $\int_E f \, d\mu \in \mathbb{C}$.

If $f : X \to [-\infty, \infty]$ is measurable, we define

$$\int_E f \, d\mu := \int_E f^+ \, d\mu - \int_E f^- \, d\mu$$

provided that at least one integral on the right-hand side is finite; then $\int_E f \, d\mu \in [-\infty, \infty]$.

Proposition 3.20. *Let $f, g \in L^1(\mu)$. Then*

(1) ***Linearity.*** *If $a, b \in \mathbb{C}$, then $af + bg \in L^1(\mu)$ and*

$$\int (af + bg) \, d\mu = a \int f \, d\mu + b \int g \, d\mu.$$

(2) ***Monotony.*** *If $f \le g$, then*

$$\int f\,d\mu \le \int g\,d\mu.$$

(3) ***Triangle inequality.***

$$\Big|\int f\,d\mu\Big| \le \int |f|\,d\mu.$$

(4) ***σ-additivity.*** *If $E_i \in \mathfrak{S}$ are disjoint, then*

$$\int_{\bigcup_{i=1}^{\infty} E_i} f\,d\mu = \sum_{i=1}^{\infty} \int_{E_i} f\,d\mu.$$

Proof. (1) By Proposition 3.2, $af + bg$ is measurable, and, by the properties of the integral for positive functions,

$$\int |af + bg|\,d\mu \le \int |a||f| + |b||g|\,d\mu = |a| \int |f|\,d\mu + |b| \int |g|\,d\mu < \infty.$$

Hence $af + bg \in L^1(\mu)$. Next we show

$$\int f + g\,d\mu = \int f\,d\mu + \int g\,d\mu. \tag{3.1}$$

To this end we may assume without loss of generality that f and g are real valued. Setting $h = f + g$ we have

$$h^+ - h^- = f^+ - f^- + g^+ - g^-$$

or equivalently

$$h^+ + f^- + g^- = f^+ + g^+ + h^-$$

and thus

$$\int h^+\,d\mu + \int f^-\,d\mu + \int g^-\,d\mu = \int f^+\,d\mu + \int g^+\,d\mu + \int h^-\,d\mu.$$

Each of these integrals is finite, so (3.1) follows. Let us show

$$\int af\,d\mu = a \int f\,d\mu. \tag{3.2}$$

If $a \ge 0$ this follows easily from Lemma 3.13. For $a = -1$ we have, writing $f = u + iv$,

$$\begin{aligned}\int -f\,d\mu &= \Big(\int (-u)^+\,d\mu - \int (-u)^-\,d\mu\Big) + i\Big(\int (-v)^+\,d\mu - \int (-v)^-\,d\mu\Big)\\ &= \Big(\int u^-\,d\mu - \int u^+\,d\mu\Big) + i\Big(\int v^-\,d\mu - \int v^+\,d\mu\Big)\\ &= -\int f\,d\mu,\end{aligned}$$

for $a = i$,

$$\begin{aligned}\int if\,d\mu = \int (iu - v)\,d\mu &= i\int u\,d\mu - \int v\,d\mu\\ &= i\Big(\int u\,d\mu + i\int v\,d\mu\Big) = i\int f\,d\mu.\end{aligned}$$

Combining these cases with (3.1) implies (3.2), and (1) follows.

(2) By assumption $f^+ - f^- \le g^+ - g^-$, or equivalently $f^+ + g^- \le g^+ + f^-$, thus $\int (f^+ + g^-)\,d\mu \le \int (g^+ + f^-)\,d\mu$, and (1) implies the assertion.

(3) Since $\int f\,d\mu \in \mathbb{C}$, there exists $a \in \mathbb{C}$, $|a| = 1$, so that $a \int f\,d\mu = |\int f\,d\mu|$. Then

$$\Big|\int f\,d\mu\Big| = a\int f\,d\mu = \int af\,d\mu = \int \mathrm{Re}(af)\,d\mu \le \int |af|\,d\mu = \int |f|\,d\mu.$$

(4) follows from the definition and from Corollary 3.16. $\square$

Proposition 3.21. *Let $f, g \in L^1(\mu)$. Then $\int_E f\,d\mu = \int_E g\,d\mu$ for all $E \in \mathfrak{S}$ if and only if $f = g$ μ-a.e.*

Proof. By Proposition 3.18, $f = g$ μ-a.e. if and only if $\int |f-g|\,d\mu = 0$. If $\int |f-g|\,d\mu = 0$, then for any $E \in \mathfrak{S}$,

$$\Big|\int_E f\,d\mu - \int_E g\,d\mu\Big| \le \int_E |f-g|\,d\mu \le \int |f-g|\,d\mu = 0,$$

whence $\int_E f\,d\mu = \int_E g\,d\mu$. Conversely, if $u = \mathrm{Re}(f-g)$ and $v = \mathrm{Im}(f-g)$ and $f \ne g$ μ-a.e., then at least one of u^+, u^-, v^+, v^- must be nonzero on a set of positive measure. If $E = \{x : u^+(x) > 0\}$ has positive measure, then $\mathrm{Re}(\int_E f\,d\mu - \int_E g\,d\mu) = \int_E u^+\,d\mu > 0$, since $u^- = 0$ on E. The other cases work analogously. $\square$

This proposition implies that regarding integration it makes no difference if we modify functions on null sets.

Theorem 3.22 (Dominated convergence theorem)**.** *Let $f_i : X \to \mathbb{C}$ be measurable functions such that $f_i \to f$ μ-a.e. If there is a function $g \in L^1(\mu)$ such that $|f_i| \le g$ μ-a.e. for all i, then $f \in L^1(\mu)$ and*

$$\lim_{i\to\infty} \int |f_i - f|\,d\mu = 0 \quad \text{and} \quad \int f\,d\mu = \lim_{i\to\infty} \int f_i\,d\mu.$$

Proof. The function f is measurable (maybe after redefinition on a null set), by Theorem 3.4. Since $|f| \le g$ μ-a.e., $f \in L^1(\mu)$. Since $|f_i - f| \le 2g$ μ-a.e., hence $2g - |f_i - f| \ge 0$ μ-a.e., Fatou's lemma 3.17 implies

$$\begin{aligned}\int 2g\,d\mu &\le \liminf_{i\to\infty} \int (2g - |f_i - f|)\,d\mu \\ &= \int 2g\,d\mu + \liminf_{i\to\infty} \Big(-\int |f_i - f|\,d\mu\Big) \\ &= \int 2g\,d\mu - \limsup_{i\to\infty} \int |f_i - f|\,d\mu.\end{aligned}$$

As $\int 2g\,d\mu$ is finite, we may conclude $\limsup_{i\to\infty} \int |f_i - f|\,d\mu \le 0$ and thus $\lim_{i\to\infty} \int |f_i - f|\,d\mu = 0$. Finally,

$$\Big|\int f\,d\mu - \lim_{i\to\infty}\int f_i\,d\mu\Big| = \lim_{i\to\infty}\Big|\int (f - f_i)\,d\mu\Big| \le \lim_{i\to\infty}\int |f - f_i|\,d\mu = 0$$

shows that $\int f\,d\mu = \lim_{i\to\infty} \int f_i\,d\mu$. $\square$

Corollary 3.23. *If f_i is a sequence in $L^1(\mu)$ such that $\sum_{i=1}^\infty \int |f_i|\,d\mu < \infty$, then $\sum_{i=1}^\infty f_i$ converges μ-a.e. to a function in $L^1(\mu)$, and $\int \sum_{i=1}^\infty f_i\,d\mu = \sum_{i=1}^\infty \int f_i\,d\mu$.*

Proof. Corollary 3.15 implies $\int \sum_{i=1}^\infty |f_i|\,d\mu = \sum_{i=1}^\infty \int |f_i|\,d\mu < \infty$, and so $g := \sum_{i=1}^\infty |f_i| \in L^1(\mu)$. Then $\sum_{i=1}^\infty |f_i(x)|$ is finite for μ-a.e. x, and for these x the series $\sum_{i=1}^\infty f_i(x)$ converges. The dominated convergence theorem 3.22 applied to the partial sums gives $\int \sum_{i=1}^\infty f_i\,d\mu = \sum_{i=1}^\infty \int f_i\,d\mu$. $\square$

3.4. Fubini's theorem

Let $(X, \mathfrak{S})$ and $(Y, \mathfrak{T})$ be two measurable spaces. On the cartesian product $X \times Y$ we consider the σ-algebra $\mathfrak{S} \otimes \mathfrak{T}$ generated by all **measurable rectangles**, that is by the set $\mathfrak{E} := \{E \times F : E \in \mathfrak{S}, F \in \mathfrak{T}\}$. Since

$$(A \times B) \cap (E \times F) = (A \cap E) \times (B \cap F) \text{ and } (A \times B)^c = (X \times B^c) \cup (A^c \times B),$$

$\mathfrak{E}$ is an elementary family.

For a set $E \subseteq X \times Y$ we denote by $E_x = \{y : (x, y) \in E\}$ and $E^y = \{x : (x, y) \in E\}$ its respective sections.

Lemma 3.24. *If $E \in \mathfrak{S} \otimes \mathfrak{T}$ then $E_x \in \mathfrak{T}$ and $E^y \in \mathfrak{S}$ for each $x \in X$ and $y \in Y$. We say that every set in $\mathfrak{S} \otimes \mathfrak{T}$ has the **section property**.*

Proof. We set $\mathfrak{R} := \{E \in \mathfrak{S} \otimes \mathfrak{T} : E_x \in \mathfrak{T} \text{ for all } x \in X\}$ and show that $\mathfrak{R}$ is a σ-algebra containing all measurable rectangles. This implies the statement; the proof for E^y is analogous.

If $E = A \times B$ is a measurable rectangle, then $E_x = B$ if $x \in A$ and $E_x = \emptyset$ if $x \in A^c$, so $E \in \mathfrak{R}$. That $\mathfrak{R}$ is a σ-algebra follows from the identities $(E^c)_x = (E_x)^c$ and $(\bigcup_{i=1}^\infty E_i)_x = \bigcup_{i=1}^\infty (E_i)_x$. ☐

With a function f on $X \times Y$ we associate functions f_x on Y given by $f_x(y) := f(x, y)$ and functions f^y on X given by $f^y(x) := f(x, y)$.

Lemma 3.25. *Let f be a $\mathfrak{S} \otimes \mathfrak{T}$-measurable function on $X \times Y$. Then f_x is $\mathfrak{T}$-measurable for all $x \in X$, and f^y is $\mathfrak{S}$-measurable for all $y \in Y$.*

Proof. This follows from Lemma 3.24, since $(f_x)^{-1}(E) = (f^{-1}(E))_x$ and $(f^y)^{-1}(E) = (f^{-1}(E))^y$. ☐

Theorem 3.26 (Product measure)**.** *Let $(X, \mathfrak{S}, \mu)$ and $(Y, \mathfrak{T}, \nu)$ be σ-finite measure spaces. If $E \in \mathfrak{S} \otimes \mathfrak{T}$, then the functions $x \mapsto \nu(E_x)$ and $y \mapsto \mu(E^y)$ are measurable on X and Y, respectively, and*

$$(\mu \otimes \nu)(E) := \int_X \nu(E_x)\, d\mu(x) = \int_Y \mu(E^y)\, d\nu(y) \tag{3.3}$$

*is a σ-finite measure on $\mathfrak{S} \otimes \mathfrak{T}$. It is called the **product** of the measures μ and ν.*

Proof. First assume that μ and ν are finite. Let $\mathfrak{R}$ be the collection of all $E \in \mathfrak{S} \otimes \mathfrak{T}$ for which $x \mapsto \nu(E_x)$ and $y \mapsto \mu(E^y)$ are measurable and (3.3) holds. If $E = A \times B$ is a measurable rectangle, then $\nu(E_x) = \nu(B)\chi_A(x)$ and $\mu(E^y) = \mu(A)\chi_B(y)$ are obviously measurable, and

$$\int_X \nu(E_x)\, d\mu(x) = \mu(A)\nu(B) = \int_Y \mu(E^y)\, d\nu(y),$$

hence $E \in \mathfrak{R}$. Since the measurable rectangles form an elementary family, the collection of finite disjoint unions of measurable rectangles forms an algebra, by Proposition 1.5. By the monotone class theorem 1.3, we may conclude $\mathfrak{R} = \mathfrak{S} \otimes \mathfrak{T}$ if we show that $\mathfrak{R}$ is a monotone class.

Let $E_1 \subseteq E_2 \subseteq \cdots$, $E_i \in \mathfrak{R}$, and set $E = \bigcup_{i=1}^\infty E_i$. Then $f_i(x) := \nu((E_i)_x)$ and $g_i(y) := \mu((E_i)^y)$ are measurable functions satisfying $f_i \le f_{i+1}$, $g_i \le g_{i+1}$, $f_i(x) \to \nu(E_x)$, and $g_i(y) \to \mu(E^y)$ for all x and y, by Lemma 1.1. By the monotone convergence theorem 3.14,

$$\int_X \nu(E_x)\, d\mu(x) = \lim_{i\to\infty} \int_X \nu((E_i)_x)\, d\mu(x)$$

$$= \lim_{i\to\infty} \int_Y \mu((E_i)^y)\, d\nu(y) = \int_Y \mu(E^y)\, d\nu(y),$$

thus $E \in \mathfrak{R}$. If $E_1 \supseteq E_2 \supseteq \cdots$, $E_i \in \mathfrak{R}$, then we may conclude in a similar way that $\bigcap_{i=1}^\infty E_i \in \mathfrak{R}$, using the dominated convergence theorem 3.22. So $\mathfrak{R}$ is a monotone class.

If μ and ν are σ-finite, we can write $X \times Y$ as an increasing union of measurable rectangles $X_i \times Y_i$ with $\mu(X_i) < \infty$ and $\nu(Y_i) < \infty$. For $E \in \mathfrak{S} \otimes \mathfrak{T}$, we may apply the preceding argument to each $E \cap (X_i \times Y_i)$,

$$\begin{aligned}\int \chi_{X_i}(x)\nu(E_x \cap Y_i)\, d\mu(x) &= \int_{X_i} \nu(E_x \cap Y_i)\, d\mu(x)\\ &= \int_{Y_i} \mu(E^y \cap X_i)\, d\nu(y) = \int \chi_{Y_i}(y)\mu(E^y \cap X_i)\, d\nu(y)\end{aligned}$$

and conclude (3.3) from the monotone convergence theorem 3.14.

Let us prove that $(\mu \otimes \nu)(E) := \int_X \nu(E_x)\, d\mu(x)$ is a σ-finite measure on $\mathfrak{S} \otimes \mathfrak{T}$. σ-additivity follows from Corollary 3.15: If $E_i \in \mathfrak{S} \otimes \mathfrak{T}$ are disjoint, then $(E_i)_x \in \mathfrak{T}$ are disjoint, so, for $E = \bigcup_{i=1}^\infty E_i$,

$$\nu(E_x) = \nu\Big(\Big(\bigcup_{i=1}^\infty E_i\Big)_x\Big) = \nu\Big(\bigcup_{i=1}^\infty (E_i)_x\Big) = \sum_{i=1}^\infty \nu((E_i)_x)$$

and thus

$$(\mu \otimes \nu)(E) = \int_X \nu(E_x)\, d\mu(x) = \int_X \sum_{i=1}^\infty \nu((E_i)_x)\, d\mu(x) = \sum_{i=1}^\infty (\mu \otimes \nu)(E_i).$$

Clearly, the measure $\mu \otimes \nu$ is σ-finite; indeed $(\mu \otimes \nu)(X_i \times Y_i) = \mu(X_i)\nu(Y_i) < \infty$. □

Theorem 3.27 (Fubini's theorem)**.** *Let $(X, \mathfrak{S}, \mu)$ and $(Y, \mathfrak{T}, \nu)$ be σ-finite measure spaces, and let f be an $(\mathfrak{S} \otimes \mathfrak{T})$-measurable function on $X \times Y$.*

(1) *If $0 \le f \le \infty$, then the functions*

$$\begin{aligned}\varphi &: X \to [0, \infty],\ \ \varphi(x) := \int_Y f_x\, d\nu,\\ \psi &: Y \to [0, \infty],\ \ \psi(y) := \int_X f^y\, d\mu\end{aligned}$$

are measurable, and

$$\int_{X\times Y} f\, d(\mu \otimes \nu) = \int_X \varphi\, d\mu = \int_Y \psi\, d\nu. \tag{3.4}$$

(2) *If f is complex valued and*

$$\int_X \varphi^*\, d\mu < \infty,\ \ \text{where } \varphi^*(x) := \int_Y |f|_x\, d\nu,$$

then $f \in L^1(\mu \otimes \nu)$.

(3) *If $f \in L^1(\mu \otimes \nu)$, then $f_x \in L^1(\nu)$ for μ-a.e. $x \in X$, $f^y \in L^1(\mu)$ for ν-a.e. $y \in Y$, the a.e. defined functions φ and ψ are in $L^1(\mu)$ and $L^1(\nu)$, respectively, and* (3.4) *holds.*

The identity (3.4) may be written in the form

$$\int_{X\times Y} f\, d(\mu \otimes \nu) = \int_X \Big(\int_Y f(x, y)\, d\nu(y)\Big)\, d\mu(x) = \int_Y \Big(\int_X f(x, y)\, d\mu(x)\Big)\, d\nu(y).$$

The left most integral is called a **double integral**, the other two are called **iterated integrals**. The assertion in (1) is often referred to as **Tonelli's theorem**.

Proof. (1) The definitions of φ and ψ are meaningful by Lemma 3.25. Theorem 3.26 implies (1) in the case that $f = \chi_E$ for $E \in \mathfrak{S} \otimes \mathfrak{T}$, and thus (1) holds for all positive simple functions s. In the general case, there exists a sequence of simple functions $0 \le s_1 \le s_2 \le \cdots$ such that $s_i(x,y) \to f(x,y)$ for all $(x,y) \in X \times Y$, by Theorem 3.6. Then, if

$$\varphi_i(x) := \int_Y (s_i)_x \, d\nu, \tag{3.5}$$

we have

$$\int_X \varphi_i \, d\mu = \int_{X\times Y} s_i \, d(\mu \otimes \nu). \tag{3.6}$$

The monotone convergence theorem 3.14, applied to (3.5), implies that $\varphi_i(x) \to \varphi(x)$ for all $x \in X$. Clearly, $\varphi_i \le \varphi_{i+1}$. Thus we may again apply the monotone convergence theorem to both sides of (3.6), and we obtain the first equality in (3.4). The other half of (3.4) follows similarly.

(2) follows by applying (1) to $|f|$.

(3) It is no restriction to assume that $f \in L^1(\mu \otimes \nu)$ is real valued. Then (1) applies to f^+ and f^-; set $\varphi^\pm(x) := \int_Y (f^\pm)_x \, d\nu$. As $f^\pm \le |f|$ we may conclude that $\varphi^\pm \in L^1(\mu)$. Thanks to $f_x = (f^+)_x - (f^-)_x$ we have $f_x \in L^1(\nu)$ for every x satisfying $\varphi^\pm(x) < \infty$. Since $\varphi^\pm \in L^1(\mu)$, this happens for μ-a.e. x; at any such x we have $\varphi(x) = \varphi^+(x) - \varphi^-(x)$. Thus $\varphi \in L^1(\mu)$. Now (3.4) holds for $f^\pm$ and $\varphi^\pm$ in place of f and φ. Subtracting the respective equalities yield the first equality of (3.4). The other half follows analogously. $\square$

The following example shows that the theorem is not true if one of the measure spaces is not σ-finite.

Example 3.28. If $X = Y = [0,1]$, μ the Lebesgue measure, ν the counting measure, and $f(x,y) = 1$ for $x = y$ and $f(x,y) = 0$ otherwise, then

$$\int_X f(x,y)\, d\mu(x) = 0 \quad \text{and} \quad \int_Y f(x,y)\, d\nu(y) = 1$$

for all $x, y \in [0,1]$ so that

$$\int_X \Big(\int_Y f(x,y)\, d\nu(y) \Big)\, d\mu(x) = 1 \ne 0 = \int_Y \Big(\int_X f(x,y)\, d\mu(x) \Big)\, d\nu(y).$$

The function $f = \chi_{\{x=y\}}$ is $(\mathfrak{L}([0,1]) \otimes \mathfrak{P}([0,1]))$-measurable, since $\{x = y\} = \bigcap_{n=1}^\infty Q_n$ where $Q_n = ([\frac{0}{n}, \frac{1}{n}] \times [\frac{0}{n}, \frac{1}{n}]) \cup \cdots \cup ([\frac{n-1}{n}, \frac{n}{n}] \times [\frac{n-1}{n}, \frac{n}{n}])$.

The product measure $\mu \otimes \nu$ rarely is complete, even if μ and ν are complete. If $A \in \mathfrak{S}$ is non-empty with $\mu(A) = 0$ and $B \subseteq Y$ so that $B \notin \mathfrak{T}$, then $A \times B \subseteq A \times Y$ and $(\mu \otimes \nu)(A \times Y) = 0$, but $A \times B \notin \mathfrak{S} \otimes \mathfrak{T}$, by Lemma 3.24. This applies in particular to the Lebesgue measure: $\lambda^1 \otimes \lambda^1 \ne \lambda^2$. However the following is true.

Theorem 3.29. *λ^{m+n} is the completion of $\lambda^m \otimes \lambda^n$, for $m, n \ge 1$.*

Proof. First we show that

$$\mathfrak{B}(\mathbb{R}^{m+n}) \subseteq \mathfrak{L}(\mathbb{R}^m) \otimes \mathfrak{L}(\mathbb{R}^n) \subseteq \mathfrak{L}(\mathbb{R}^{m+n}).$$

The first inclusion follows from the fact that each cube in $\mathbb{R}^{m+n}$ belongs to $\mathfrak{L}(\mathbb{R}^m) \otimes \mathfrak{L}(\mathbb{R}^n)$ and $\mathfrak{B}(\mathbb{R}^{m+n})$ is the σ-algebra generated by the cubes in $\mathbb{R}^{m+n}$; see Lemma 2.6. Suppose that $E \in \mathfrak{L}(\mathbb{R}^m)$ and $F \in \mathfrak{L}(\mathbb{R}^n)$. Then $E \times \mathbb{R}^n$ and $\mathbb{R}^m \times F$ belong to $\mathfrak{L}(\mathbb{R}^{m+n})$, by Corollary 2.10, and thus $E \times F = (E \times \mathbb{R}^n) \cap (\mathbb{R}^m \times F)$ belongs to $\mathfrak{L}(\mathbb{R}^{m+n})$, which implies the second inclusion.

Both λ^{m+n} and $\lambda^m \otimes \lambda^n$ coincide on boxes and hence on $\mathfrak{B}(\mathbb{R}^{m+n})$, by Theorem 2.11. If $A \in \mathfrak{L}(\mathbb{R}^m) \otimes \mathfrak{L}(\mathbb{R}^n)$, then $A \in \mathfrak{L}(\mathbb{R}^{m+n})$ and so there exist $B_1, B_2 \in$

$\mathfrak{B}(\mathbb{R}^{m+n})$ such that $B_1 \supseteq A \supseteq B_2$ and $\lambda^{m+n}(B_1 \setminus B_2) = 0$, by Corollary 2.10. Consequently,

$$(\lambda^m \otimes \lambda^n)(A \setminus B_2) \le (\lambda^m \otimes \lambda^n)(B_1 \setminus B_2) = \lambda^{m+n}(B_1 \setminus B_2) = 0,$$

and thus $(\lambda^m \otimes \lambda^n)(A) = (\lambda^m \otimes \lambda^n)(B_2) = \lambda^{m+n}(B_2) = \lambda^{m+n}(A)$. So λ^{m+n} and $\lambda^m \otimes \lambda^n$ coincide on $\mathfrak{L}(\mathbb{R}^m) \otimes \mathfrak{L}(\mathbb{R}^n)$ which implies the statement. $\square$

Theorem 3.30 (Fubini's theorem for complete measures)**.** *Let $(X, \mathfrak{S}, \mu)$ and $(Y, \mathfrak{T}, \nu)$ be complete σ-finite measure spaces, and let $\overline{\mathfrak{S} \otimes \mathfrak{T}}$ be the completion of $\mathfrak{S} \otimes \mathfrak{T}$ with respect to $\mu \otimes \nu$. Let f be an $\overline{\mathfrak{S} \otimes \mathfrak{T}}$-measurable function on $X \times Y$. Then all conclusions of Theorem 3.27 hold, except that the $\mathfrak{T}$-measurability of f_x can be asserted only for μ-a.e. $x \in X$ so that $\varphi(x)$ is only defined μ-a.e., and similarly for f^y and ψ.*

Proof. By Proposition 3.9, $f = g + h$, where $h = 0$ $\overline{\mu \otimes \nu}$-a.e. and g is $(\mathfrak{S} \otimes \mathfrak{T})$-measurable. We claim that for μ-a.e. $x \in X$ we have $h(x, y) = 0$ for ν-a.e. $y \in Y$ and h_x is $\mathfrak{T}$-measurable for μ-a.e. $x \in X$. Similarly, for h^y.

Indeed, $A := \{(x, y) \in X \times Y : h(x, y) \neq 0\}$ is a $\overline{\mu \otimes \nu}$-null set. So there exists $B \in \mathfrak{S} \otimes \mathfrak{T}$ such that $A \subseteq B$ and $(\mu \otimes \nu)(B) = 0$. By Theorem 3.26, $\int_X \nu(B_x)\, d\mu(x) = (\mu \otimes \nu)(B) = 0$. By Proposition 3.21, $\mu(E) = 0$, where $E := \{x \in X : \nu(B_x) > 0\}$. If $x \notin E$, then $\nu(B_x) = 0$ and, as $(Y, \mathfrak{T}, \nu)$ is complete, each subset of $A_x (\subseteq B_x)$ belongs to $\mathfrak{T}$. If $y \notin A_x$, then $h_x(y) = 0$. It follows that, for every $x \notin E$, h_x is $\mathfrak{T}$-measurable and $h_x(y) = 0$ ν-a.e. The claim is proved.

Apply Theorem 3.27 to g. By the claim, $f_x = g_x$ ν-a.e. for μ-a.e. x and $f^y = g^y$ μ-a.e. for ν-a.e. y. Thus the two iterated integrals and the double integral of f are the same as those of g. $\square$

3.5. Transformation of measures and integrals

Let $(X, \mathfrak{S})$ and $(Y, \mathfrak{T})$ be measurable spaces and let $f : X \to Y$ be $(\mathfrak{S}, \mathfrak{T})$-measurable. Given a measure μ on $(X, \mathfrak{S})$ we may define the **push-forward** $f_*\mu$ on $(Y, \mathfrak{T})$ by

$$f_*\mu(E) := \mu(f^{-1}(E)), \quad E \in \mathfrak{T}.$$

It is easy to check that $f_*\mu$ is a measure.

Proposition 3.31. *Let $g : Y \to \mathbb{C}$ be $\mathfrak{T}$-measurable. Then $g \circ f \in L^1(\mu)$ if and only if $g \in L^1(f_*\mu)$, and*

$$\int_Y g\, d(f_*\mu) = \int_X g \circ f\, d\mu.$$

Proof. For $E \in \mathfrak{T}$ and $g = \chi_E$ the formula follows from $\chi_E \circ f = \chi_{f^{-1}(E)}$. So it holds for simple functions and hence for positive functions, by Theorem 3.6 and the monotone convergence theorem 3.14. In particular, the equality holds for $|g|$ instead of g, and so $g \circ f \in L^1(\mu)$ if and only if $g \in L^1(f_*\mu)$. That it is also valid for complex valued g follows immediately. $\square$

In the following we focus on the Lebesgue measure λ.

Lemma 3.32. *Let $A : \mathbb{R}^n \to \mathbb{R}^n$ be linear invertible, and let E be measurable. Then $A(E)$ is measurable and $\lambda(A(E)) = |\det A|\, \lambda(E)$. In particular, λ is invariant under orthogonal transformations.*

Proof. It suffices to prove the statement for Borel sets E. Then null sets are invariant under A and A^{-1}, and hence so are Lebesgue measurable sets.

If E is a Borel set then so is $A(E)$, since $\chi_{A(E)} = \chi_E \circ A^{-1}$ and since χ_E and A^{-1} and hence $\chi_E \circ A^{-1}$ are Borel mappings.

We shall use translation invariance, see Proposition 2.13, and dilation invariance of λ^1 on Borel sets, i.e., if $a \in \mathbb{R}\backslash\{0\}$ and $E \in \mathfrak{B}(\mathbb{R})$ then $aE = \{ax : x \in E\} \in \mathfrak{B}(\mathbb{R})$ and $\lambda^1(aE) = |a|\lambda^1(E)$. The collection of intervals in $\mathbb{R}$ is invariant under dilations, and hence so is $\mathfrak{B}(\mathbb{R})$. Then $\mu_a(E) := \lambda^1(aE)/|a|$ defines a Borel measure that coincides with λ^1 on boxes, and thus on all Borel sets, by Theorem 2.11.

Suppose that A, and thus also A^{-1}, is upper triangular with all diagonal entries equal to 1. Then,

$$\begin{aligned}\lambda(A(E)) &= \int_{\mathbb{R}^n} \chi_{A(E)}(x)\,dx = \int_{\mathbb{R}^n} \chi_E(A^{-1}(x))\,dx \\ &= \int_{\mathbb{R}^{n-1}} \int_{\mathbb{R}} \chi_E(x_1 + f_1(x_{\geq 2}), x_2 + f_2(x_{\geq 3}), \dots, x_n)\,dx_1\,dx_{\geq 2} \\ &= \int_{\mathbb{R}^{n-1}} \int_{\mathbb{R}} \chi_E(x_1, x_2 + f_2(x_{\geq 3}), \dots, x_n)\,dx_1\,dx_{\geq 2},\end{aligned}$$

using Fubini's theorem 3.27 and translation invariance of λ^1. Repeating this procedure for the other variables, we find

$$\lambda(A(E)) = \int_{\mathbb{R}^n} \chi_E(x)\,dx = \lambda(E).$$

Similarly, the assertion holds for lower triangular matrices with all diagonal entries equal to 1. If $A = \operatorname{diag}(a_1, \dots, a_n)$ is diagonal, then Fubini's theorem 3.27 and dilation invariance of λ^1 analogously imply

$$\lambda(A(E)) = |a_1 \cdots a_n|\lambda(E).$$

An arbitrary square matrix A admits a decomposition $A = LDU$, where L (U) is an lower (upper) triangular matrix with all diagonal entries equal 1 and D is diagonal. Thus the result follows. □

Theorem 3.33 (Transformation formula)**.** *Let $U, V \subseteq \mathbb{R}^n$ be open and let $f \in C^1(U, V)$ be bijective. If g is a measurable function on V, then $g \circ f$ is measurable on U. If $g \geq 0$ or $g \in L^1(V)$, then*

$$\int_U g(f(x))|J_f(x)|\,dx = \int_V g(y)\,dy,$$

where $J_f = \det(\partial f/\partial x)$ is the Jacobi determinant of f. In particular, for measurable $E \subseteq U$, $f(E)$ is measurable, and

$$\lambda(f(E)) = \int_E |J_f(x)|\,dx.$$

Proof. It is sufficient to consider Borel measurable functions and sets. Since f and f^{-1} are continuous, there are no measurability problems in this case. If g is Lebesgue measurable and B is a Borel set in $\mathbb{C}$, then $g^{-1}(B) = E \cup N$, where E is Borel and N is a null set. Moreover, $f^{-1}(E)$ is Borel and $f^{-1}(N)$ is a null set (by the result for Borel sets), and thus $(g \circ f)^{-1}(B)$ is Lebesgue measurable, i.e., $g \circ f$ is Lebesgue measurable.

We use the norm $|x|_\infty = \max_{1\leq i\leq n} |x_i|$ for $x \in \mathbb{R}^n$ and the matrix norm $\|A\| = \max_{1\leq i\leq n} \sum_{j=1}^n |A_{ij}|$; then $|Ax|_\infty \leq \|A\||x|_\infty$. Let $Q = \{x : |x - a|_\infty \leq h\}$ be a cube contained in U. By the mean value theorem, $f(x) - f(a) = f'(z)(x - a)$ for some z on the segment between x and a, and hence, for $x \in Q$,

$$|f(x) - f(a)|_\infty \leq \sup_{z\in Q} \|f'(z)\|\,|x - a|_\infty \leq \sup_{z\in Q} \|f'(z)\|\,h.$$

So $f(Q)$ is contained in a cube of side length $\sup_{z\in Q}\|f'(z)\|$ times the side length of Q, thus
$$\lambda(f(Q)) \le \big(\sup_{z\in Q}\|f'(z)\|\big)^n \lambda(Q).$$
If $A:\mathbb{R}^n\to\mathbb{R}^n$ is linear invertible, we find, by Lemma 3.32,
$$\lambda(f(Q)) = |\det A|\,\lambda(A^{-1}f(Q)) \le |\det A|\big(\sup_{z\in Q}\|A^{-1}f'(z)\|\big)^n\lambda(Q).$$
Since f' is uniformly continuous on Q, for each $\epsilon>0$ there exists $\delta>0$ so that for $x,y\in Q$ with $|x-y|_\infty\le\delta$,
$$\|f'(x)^{-1}f'(y)\| = \|f'(x)^{-1}f'(y) - f'(x)^{-1}f'(x) + \mathrm{Id}\,\| \le 1+\epsilon.$$
By decomposing Q into subcubes $Q_1,\dots,Q_N$ with side length $\le\delta$ and centers $x_1,\dots,x_N$, we may conclude
$$\begin{aligned}\lambda(f(Q)) &\le \sum_{i=1}^N \lambda(f(Q_i))\\ &\le \sum_{i=1}^N |J_f(x_i)|\Big(\sup_{z\in Q_i}\Big\|f'(x_i)^{-1}f'(z)\Big\|\Big)^n\lambda(Q_i)\\ &\le (1+\epsilon)^n\sum_{i=1}^N |J_f(x_i)|\lambda(Q_i).\end{aligned}$$
Note that $\sum_{i=1}^N |J_f(x_i)|\chi_{Q_i}$ is a simple function which tends uniformly on Q to $x\mapsto|J_f(x)|$ as $\delta\to0$, by continuity of $x\mapsto J_f(x)$. Letting δ and ϵ approach 0 implies
$$\lambda(f(Q)) \le \int_Q |J_f(x)|\,dx.$$

We shall show that this estimate holds with Q replaced by any Borel set in U. If $\Omega\subseteq U$ is open, then $\Omega=\bigcup_{i=1}^\infty Q_i$ is a almost disjoint union of cubes Q_i, by Lemma 2.6, and thus
$$\lambda(f(\Omega)) \le \sum_{i=1}^\infty \lambda(f(Q_i)) \le \sum_{i=1}^\infty \int_{Q_i}|J_f(x)|\,dx = \int_\Omega |J_f(x)|\,dx.$$
If $E\subseteq U$ is a Borel set of finite measure, then by outer regularity, Theorem 2.9, there exists a sequence $U\supseteq\Omega_i\supseteq\Omega_{i+1}\supseteq E$ of open sets Ω_i of finite measure so that $\lambda(\bigcap_{i=1}^\infty\Omega_i\setminus E)=0$. By Lemma 1.1 and the dominated convergence theorem 3.22,
$$\lambda(f(E)) \le \lambda\Big(f\Big(\bigcap_{i=1}^\infty\Omega_i\Big)\Big) \le \lim_{i\to\infty}\lambda(f(\Omega_i)) \le \lim_{i\to\infty}\int_{\Omega_i}|J_f(x)|\,dx = \int_E |J_f(x)|\,dx.$$
Since λ is σ-finite, the estimate holds for all Borel sets E.

We may infer that
$$\int_{f(U)} g(y)\,dy \le \int_U g(f(x))|J_f(x)|\,dx,$$
first for positive simple g and, by Theorem 3.6 and the monotone convergence theorem 3.14, for positive measurable g. Applying this to f^{-1} and $(g\circ f)|J_f|$ instead of f and g, we get
$$\int_U (g\circ f)(x)|J_f(x)|\,dx \le \int_{f(U)} g(x)|J_f(f^{-1}(x))||J_{f^{-1}}(x)|\,dx = \int_{f(U)} g(y)\,dy.$$
So the assertion is shown for $g\ge0$, and the case $g\in L^1(V)$ follows easily. The second statement is the special case, where $g=\chi_{f(E)}$. □

Let $S^{n-1} = \{x \in \mathbb{R}^n : |x| = 1\}$ denote the unit sphere in $\mathbb{R}^n$. The mapping

$$\varphi : \mathbb{R}^n \setminus \{0\} \to (0,\infty) \times S^{n-1} : x \mapsto (|x|, x/|x|)$$

defines a diffeomorphism with inverse $(r,y) \mapsto ry$; we call $(r,y) = \varphi(x)$ the **polar coordinates** of x. Let ρ be the measure on $(0,\infty)$ defined by $\rho(E) = \int_E r^{n-1}\,dr$.

Theorem 3.34 (Polar coordinates)**.** *There is a unique Borel measure σ on S^{n-1} such that $\varphi_*\lambda = \rho \otimes \sigma$. If f is Borel measurable on $\mathbb{R}^n$ and $f \ge 0$ or $f \in L^1(\lambda)$, then*

$$\int_{\mathbb{R}^n} f(x)\,dx = \int_{(0,\infty)} \int_{S^{n-1}} f(ry) r^{n-1}\,d\sigma(y)\,dr.$$

Proof. By Proposition 3.31 and Fubini's theorem 3.27, it suffices to show that there is a unique Borel measure σ on S^{n-1} such that $\varphi_*\lambda = \rho \otimes \sigma$. For Borel sets E in S^{n-1} we define

$$\sigma(E) := n\lambda(\varphi^{-1}((0,1] \times E)),$$

which is a Borel measure on S^{n-1}, since the mapping $E \mapsto \varphi^{-1}((0,1] \times E)$ maps Borel sets to Borel sets and commutes with unions, intersections, and complements. For $a > 0$, we have by Lemma 3.32,

$$\begin{aligned}\varphi_*\lambda((0,a] \times E) &= \lambda(\varphi^{-1}((0,a] \times E)) = a^n \lambda(\varphi^{-1}((0,1] \times E)) \\ &= \frac{a^n}{n}\sigma(E) = \rho((0,a])\sigma(E) = (\rho \otimes \sigma)((0,a] \times E).\end{aligned}$$

As an immediate consequence, $\varphi_*\lambda = \rho \otimes \sigma$ holds on sets of the form $(a,b] \times E$. For $N \in \mathbb{N}$ and a fixed Borel set $E \subseteq S^{n-1}$, the collection $\mathfrak{A}_{N,E}$ of finite disjoint unions of sets of the form $(a,b] \times E$, where $b \le N$, forms an algebra on $(0,N] \times E$, by Proposition 1.5, that generates the σ-algebra $\mathfrak{S}_{N,E} = \{A \times E : A \in \mathfrak{B}((0,N])\}$. By Theorem 1.4, $\varphi_*\lambda = \rho \otimes \sigma$ holds on $\mathfrak{S}_{N,E}$, and since all Borel rectangles in $(0,\infty) \times S^{n-1}$ are disjoint countable unions of sets in $\bigcup_{N \in \mathbb{N}, E \in \mathfrak{B}(S^{n-1})} \mathfrak{S}_{N,E}$, we have $\varphi_*\lambda = \rho \otimes \sigma$ on all Borel set, again by Theorem 1.4. □

The formula of the previous theorem can be extended to Lebesgue measurable functions by considering the completion of σ. If $f(x) = g(|x|)$ it gives

$$\int_{\mathbb{R}^n} f(x)\,dx = \sigma(S^{n-1}) \int_{(0,\infty)} g(r) r^{n-1}\,dr. \tag{3.7}$$

Example 3.35 (Integral of a Gaussian function)**.** We have

$$\int_{\mathbb{R}^n} e^{-a|x|^2}\,dx = \Big(\frac{\pi}{a}\Big)^{n/2}, \quad a > 0.$$

If we denote the integral on the left by I_n, then $I_n = (I_1)^n$ by Fubini's theorem 3.27. By (3.7),

$$I_2 = 2\pi \int_{(0,\infty)} r e^{-ar^2}\,dr = -\frac{\pi}{a} e^{-ar^2}\Big|_0^\infty = \frac{\pi}{a}.$$

Thus $I_1 = (\pi/a)^{1/2}$ and $I_n = (\pi/a)^{n/2}$.

Example 3.36 (Volume and surface area of the unit ball)**.** If $B^n := \{x \in \mathbb{R}^n : |x| \le 1\}$ denotes the closed unit ball in $\mathbb{R}^n$, then

$$\sigma(S^{n-1}) = \frac{2\pi^{n/2}}{\Gamma(n/2)} \quad \text{and} \quad \lambda(B^n) = \frac{\pi^{n/2}}{\Gamma(n/2+1)}.$$

By Example 3.35, (3.7), and Theorem 3.33,

$$\pi^{n/2} = \int_{\mathbb{R}^n} e^{-|x|^2}\,dx = \sigma(S^{n-1}) \int_{(0,\infty)} r^{n-1} e^{-r^2}\,dr$$

$$= \frac{\sigma(S^{n-1})}{2} \int_{(0,\infty)} t^{n/2-1} e^{-t}\, dt = \frac{\sigma(S^{n-1})}{2} \Gamma(n/2),$$

and by the definition of σ,

$$\lambda(B^n) = \frac{\sigma(S^{n-1})}{n} = \frac{\pi^{n/2}}{n/2 \cdot \Gamma(n/2)} = \frac{\pi^{n/2}}{\Gamma(n/2+1)}.$$

3.6. Integrals depending on parameters

We study continuity and differentiability of functions of the form

$$F(y) = \int_X f(x,y)\, d\mu(x), \quad y \in Y.$$

Theorem 3.37 (Continuity of integrals depending on parameters)**.** *Let $(X, \mathfrak{S}, \mu)$ be a measure space, let Y be a metric space, and let $f : X \times Y \to \mathbb{C}$ be a function. Assume that:*

1. *For each fixed $y \in Y$ the function $X \ni x \to f(x,y)$ is measurable.*
2. *For each fixed $x \in X$ the function $Y \ni y \to f(x,y)$ is continuous at y_0.*
3. *There is a positive function $g \in L^1(\mu)$ so that $|f(x,y)| \le g(x)$ for all $(x,y) \in X \times Y$.*

Then the function $F : Y \to \mathbb{C}$ given by

$$F(y) = \int_X f(x,y)\, d\mu(x), \quad y \in Y,$$

is well-defined and continuous at y_0.

Proof. The function F is well-defined by (1) and (3). Let $y_k \in Y$ by a sequence converging to y_0, and consider the sequence of functions $f_k : X \to \mathbb{C}$ given by

$$f_k(x) := f(x, y_k).$$

By (2), $f_k(x) \to f(x,y_0)$ for every $x \in X$, and, by (3), $|f_k| \le g$ for all k. The dominated convergence theorem 3.22 implies

$$\lim_{k\to\infty} F(y_k) = \lim_{k\to\infty} \int_X f_k\, d\mu = \int_X f(x,y_0)\, d\mu(x) = F(y_0). \qquad \square$$

Theorem 3.38 (Differentiability of integrals depending on parameters)**.** *Let $(X, \mathfrak{S}, \mu)$ be a measure space, let Y be open in $\mathbb{R}^n$, and let $f : X \times Y \to \mathbb{C}$ be a function. Assume that:*

1. *For each fixed $x \in X$ the function $Y \ni y \to f(x,y)$ is C^1.*
2. *For each fixed $y \in Y$ the function $X \ni x \to f(x,y)$ is in $L^1(\mu)$, and $X \ni x \to \frac{\partial}{\partial y_i} f(x,y)$, $i = 1, \dots, n$, is measurable.*
3. *There is a positive function $g \in L^1(\mu)$ so that $|\frac{\partial}{\partial y_i} f(x,y)| \le g(x)$ for all $(x,y) \in X \times Y$.*

Then the function $F : Y \to \mathbb{C}$ given by

$$F(y) = \int_X f(x,y)\, d\mu(x), \quad y \in Y,$$

is well-defined and C^1 with

$$\frac{\partial}{\partial y_i} F(y) = \int_X \frac{\partial}{\partial y_i} f(x,y)\, d\mu(x).$$

PROOF. The function F is well-defined by (2). Let $y_0 \in Y$ and let the open ball $B_r(y_0)$ be contained in Y. Let $h_k \in \mathbb{R} \setminus \{0\}$ with $h_k \to 0$ and such that $y_k := y_0 + h_k e_i \in B_r(y_0)$, where e_i is the ith standard unit vector in $\mathbb{R}^n$. Set

$$\varphi_k(x) := \frac{f(x, y_k) - f(x, y_0)}{h_k}.$$

Then each φ_k is in $L^1(\mu)$, and, for all $x \in X$,

$$\lim_{k\to\infty} \varphi_k(x) = \frac{\partial}{\partial y_i} f(x, y_0).$$

By (3) and the mean value theorem, $|\varphi_k| \leq g$. The dominated convergence theorem 3.22 implies that $x \mapsto \frac{\partial}{\partial y_i} f(x, y_0)$ is in $L^1(\mu)$ and we have

$$\lim_{k\to\infty} \int_X \varphi_k \, d\mu = \int_X \frac{\partial}{\partial y_i} f(x, y_0) \, d\mu(x).$$

Since

$$\int_X \varphi_k \, d\mu = \frac{1}{h_k}\Big(\int_X f(x, y_k) \, d\mu(x) - \int_X f(x, y_0) \, d\mu(x) \Big) = \frac{F(y_k) - F(y_0)}{h_k},$$

we see that $\frac{\partial}{\partial y_i} F(y_0)$ exists and equals $\int_X \frac{\partial}{\partial y_i} f(x, y_0) \, d\mu(x)$. The continuity of $\frac{\partial}{\partial y_i} F$ follows from Theorem 3.37. □

Theorem 3.39 (Holomorphy of integrals depending on parameters)**.** *Let $(X, \mathfrak{S}, \mu)$ be a measure space, let Y be open in $\mathbb{C}$, and let $f : X \times Y \to \mathbb{C}$ be a function. Assume that:*

1. *For each fixed $x \in X$ the function $Y \ni y \to f(x, y)$ is holomorphic.*
2. *For each fixed $y \in Y$ the function $X \ni x \to f(x, y)$ is measurable.*
3. *There is a positive function $g \in L^1(\mu)$ so that $|f(x, y)| \leq g(x)$ for all $(x, y) \in X \times Y$.*

Then the function $F : Y \to \mathbb{C}$ given by

$$F(y) = \int_X f(x, y) \, d\mu(x), \quad y \in Y,$$

is well-defined and holomorphic with

$$F'(y) = \int_X \partial_y f(x, y) \, d\mu(x).$$

PROOF. The function F is well-defined by (2) and (3). Let $y_0 \in Y$ and let $\overline{B_r(y_0)}$ be contained in Y. For all $y \in B_r(y_0)$ and all $x \in X$, we have

$$\partial_y f(x, y) = \frac{1}{2\pi i} \int_{\partial B_r(y_0)} \frac{f(x, z)}{(z - y)^2} \, dz$$

and thus, if we write $y = y_1 + iy_2$ and use (3), for all $y \in B_{r/2}(y_0)$ and all $x \in X$

$$|\partial_{y_i} f(x, y)| \leq r \max_{z \in \partial B_r(y_0)} \frac{|f(x, z)|}{|z - y|^2} \leq \frac{4g(x)}{r}.$$

By Theorem 3.38, F is C^1 in $B_{r/2}(y_0)$ and satisfies the Cauchy–Riemann equations

$$\partial_{y_1} F(y) + i\partial_{y_2} F(y) = \int_X \partial_{y_1} f(x, y) + i\partial_{y_1} f(x, y) \, d\mu(x) = 0.$$

The proof is complete. □

3.7. Relation to the Riemann integral

Let $[a,b]$ be a compact interval and let $f : [a,b] \to \mathbb{R}$ be bounded. For each **partition** P of $[a,b]$, i.e., a finite sequence $P = (t_i)_{i=0}^n$ with $a = t_0 < t_1 < \cdots < t_n = b$, define

$$U_P f := \sum_{i=1}^n \sup_{t_{i-1} \le t \le t_i} f(t)\,(t_i - t_{i-1}),$$
$$L_P f := \sum_{i=1}^n \inf_{t_{i-1} \le t \le t_i} f(t)\,(t_i - t_{i-1}),$$

and set

$$\overline{I}_a^b(f) := \inf_P U_P f, \quad \underline{I}_a^b(f) := \sup_P L_P f,$$

where P varies over all partitions of $[a,b]$. If $\overline{I}_a^b(f) = \underline{I}_a^b(f)$ then their common value is the **Riemann integral** $\int_a^b f(x)\,dx$, and f is called **Riemann integrable**.

Theorem 3.40. *Let $f : [a,b] \to \mathbb{R}$ be bounded. Then:*

(1) *If f is Riemann integrable, then f is Lebesgue measurable and thus integrable (since bounded), and $\int_a^b f(x)dx = \int_{[a,b]} f\,d\lambda$.*

(2) *f is Riemann integrable if and only if*

$$\lambda(\{t \in [a,b] : f \textit{ is discontinuous at } t\}) = 0.$$

The second result is **Lebesgue's criterion for Riemann integrability**.

Proof. (1) Without loss of generality assume that $f \ge 0$. For each partition P of $[a,b]$ set

$$G_P := \sum_{i=1}^n \sup_{t_{i-1} \le t \le t_i} f(t)\,\chi_{(t_{i-1},t_i]},$$
$$g_P := \sum_{i=1}^n \inf_{t_{i-1} \le t \le t_i} f(t)\,\chi_{(t_{i-1},t_i]},$$

such that $U_P f = \int G_P\,d\lambda$ and $L_P f = \int g_P\,d\lambda$. If f is Riemann integrable, there exists a sequence of partitions P_k whose mesh size (that is $\max_i(t_i - t_{i-1})$) tends to 0, such that $P_k \subseteq P_{k+1}$, and so that $U_{P_k} f$ and $L_{P_k} f$ converge to $\int_a^b f(x)\,dx$. Then $G_{P_k} \ge G_{P_{k+1}} \ge f \ge g_{P_{k+1}} \ge g_{P_k}$, and $G := \lim_{k\to\infty} G_{P_k}$, $g := \lim_{k\to\infty} g_{P_k}$ satisfy $g \le f \le G$. By the dominated convergence theorem 3.22,

$$\int g\,d\lambda = \int_a^b f(x)\,dx = \int G\,d\lambda,$$

and thus $\int (G-g)\,d\lambda = 0$. By Proposition 3.21, $G = g = f$ a.e. Since G is measurable, by Proposition 3.4, so is f, by Proposition 3.8 (as λ is complete), and we have

$$\int_a^b f(x)\,dx = \int G\,d\lambda = \int_{[a,b]} f\,d\lambda.$$

(2) Assume that f is Riemann integrable. By the first part of the proof, the set

$$E := \{t \in [a,b] : g(t) \ne G(t)\} \cup \bigcup_{k=1}^\infty P_k$$

has measure zero. We will show that the set of discontinuities of f lies in E. Fix $t_0 \in [a,b] \setminus E$ and $\epsilon > 0$. Then $g(t_0) = G(t_0)$ and hence $G_{P_k}(t_0) - g_{P_k}(t_0) \le \epsilon$ for

k sufficiently large. Since $t_0 \notin P_k$, G_{P_k} and g_{P_k} are constant near t_0. Thus there is $\delta > 0$ so that for $|t - t_0| \le \delta$,

$$f(t) - f(t_0) \le G_{P_k}(t) - g_{P_k}(t_0) = G_{P_k}(t_0) - g_{P_k}(t_0) \le \epsilon,$$

and similarly $f(t) - f(t_0) \ge \epsilon$. This implies that f is continuous at t_0.

Conversely, let f be continuous except on a set E of measure zero. By Theorem 2.9, given $\epsilon > 0$ we may find open intervals I_i so that $E \subseteq \bigcup_{i=1}^\infty I_i$ and $\sum_{i=1}^\infty |I_i| \le \epsilon/(4M)$, where $M = \sup_{t\in[a,b]} f(t)$. If f is continuous at t, then there is an open interval $J_t \ni t$ such that $|f(s) - f(r)| \le \epsilon/2(b-a)$ for $s, r \in J_t \cap [a, b]$. The open cover $\{I_i\} \cup \{J_t : t \in [a, b] \setminus E\}$ of $[a, b]$ has a finite subcover; let $P = (t_i)_{i=0}^n$ be the partition of $[a, b]$ given by the endpoints (inside $[a, b]$) of the intervals in this subcover. Let $L = \{\ell : (t_{\ell-1}, t_\ell) \subseteq I_i \text{ for some } i\}$. Then

$$\begin{aligned}
U_P f - L_P f &= \sum_{i=1}^n \sup_{t_{i-1}\le s,t\le t_i} (f(t) - f(s)) (t_i - t_{i-1}) \\
&\le \sum_{i\in L} 2M (t_i - t_{i-1}) + \sum_{i\notin L} \frac{\epsilon}{2(b-a)} (t_i - t_{i-1}) \\
&\le 2M \frac{\epsilon}{4M} + \frac{\epsilon}{2(b-a)} (b - a) = \epsilon.
\end{aligned}$$

This implies that f is Riemann integrable. □

The proper Riemann integral is thus subsumed in the Lebesgue integral. The latter allows for integration of a wider class of functions. For instance, $\chi_{\mathbb{Q}\cap[0,1]}$ is discontinuous everywhere and hence not Riemann integrable. It is however Lebesgue integrable with $\int \chi_{\mathbb{Q}\cap[0,1]}\, d\lambda = 0$.

For improper Riemann integrals the situation is different. The functions $f = \sum_{k=1}^\infty \frac{(-1)^k}{k} \chi_{(k,k+1]}$ or $g(x) = \sin(x)/x$ have improper Riemann integrals over $[1, \infty)$ (to see this for g use partial integration and the majorant criterion), but they are not Lebesgue integrable. A Lebesgue integrable function on $[a, \infty)$ that is Riemann integrable on $[a, b]$, for each $b > a$, has absolutely convergent improper Riemann integral and

$$\int_{[a,\infty)} f\, d\lambda = \lim_{b\to\infty} \int_a^b f(x)\, dx. \tag{3.8}$$

Indeed, for each $b > a$, $\int_a^b |f(x)|\, dx = \int_{[a,b]} |f|\, d\lambda \le \int_{[a,\infty)} |f|\, d\lambda$ and hence $\lim_{b\to\infty} \int_a^b |f(x)|\, dx$ exists. Moreover, choose a sequence $b_k \nearrow \infty$ and set $f_k := f\chi_{[a,b_k]}$. Then the dominated convergence theorem 3.22 implies (3.8).

3.8. Hausdorff measure

In this section we consider the d-dimensional Hausdorff measure in $\mathbb{R}^n$. It allows for a definition of d-dimensional area in an intrinsic way, i.e., without reference to parameterizations. Moreover, it makes sense in any metric space and even for non-integer d.

For $d \ge 0$ let us set

$$\omega_d := \frac{\pi^{d/2}}{\Gamma(d/2 + 1)},$$

where $\Gamma(t) := \int_0^\infty s^{t-1} e^{-s}\, ds$ is the Gamma function. If $d \ge 1$ is an integer, then ω_d is the d-dimensional Lebesgue measure of the unit ball in $\mathbb{R}^d$; see Example 3.36.

Let $E \subseteq \mathbb{R}^n$ be any subset. The **d-dimensional Hausdorff measure** of E is given by

$$\mathcal{H}^d(E) := \lim_{\epsilon \to 0+} \mathcal{H}^d_\epsilon(E), \tag{3.9}$$

where for $0 < \epsilon \leq \infty$,

$$\mathcal{H}^d_\epsilon(E) := \frac{\omega_d}{2^d} \inf \Big\{ \sum_i (\mathrm{diam}(E_i))^d : \mathrm{diam}(E_i) < \epsilon,\ E \subseteq \bigcup_i E_i \Big\}$$

for countable covers $\{E_i\}_i$ of E and with the convention $\mathrm{diam}(\emptyset) = 0$. Note that the limit in (3.9) exits (finite or infinite), since $\epsilon \mapsto \mathcal{H}^d_\epsilon(E)$ is decreasing, and that $\mathcal{H}^0$ is the counting measure. It is possible to restrict the E_i in the definition to closed (or open) and convex sets such that $E_i \cap E \neq \emptyset$, but further restrictions produce other outer measures, e.g., using only balls yields the so-called **spherical Hausdorff measure**.

The definition of Hausdorff measure extends to any metric space. It depends on the metric but not on the ambient space, i.e., $\mathcal{H}^d_X(E) = \mathcal{H}^d_Y(E)$ whenever $E \subseteq X$ and the metric space X is isometrically embedded in the metric space Y.

Proposition 3.41. *Let $d \geq 0$, $n \in \mathbb{N}$.*

(1) *$\mathcal{H}^d$ is an outer measure on $\mathbb{R}^n$ and a measure on $\mathfrak{B}(\mathbb{R}^n)$.*
(2) *For each $E \subseteq \mathbb{R}^n$, $z \in \mathbb{R}^n$, and $a > 0$,*

$$\mathcal{H}^d(E + z) = \mathcal{H}^d(E), \quad \mathcal{H}^d(aE) = a^d \mathcal{H}^d(E).$$

(3) *$\mathcal{H}^d = 0$ if $d > n$.*
(4) *If $d > d' \geq 0$, then $\mathcal{H}^d(E) > 0$ implies $\mathcal{H}^{d'}(E) = \infty$.*
(5) *If $f : \mathbb{R}^n \to \mathbb{R}^m$ is a Lipschitz function with Lipschitz constant* $\mathrm{Lip}(f)$, *then*

$$\mathcal{H}^d(f(E)) \leq \mathrm{Lip}(f)^d\, \mathcal{H}^d(E).$$

Proof. (1) Let us show that $\mathcal{H}^d$ is σ-subadditive; monotony is obvious. It is easy to see that each $\mathcal{H}^d_\epsilon$ is σ-subadditive. Thus $\mathcal{H}^d$ is σ-subadditive, since the supremum of σ-subadditive set functions is σ-subadditive. So $\mathcal{H}^d$ is an outer measure on $\mathbb{R}^n$.

Suppose that $\delta = \mathrm{dist}(E_1, E_2) > 0$ and $\epsilon \leq \delta$. Then any set of diameter $< \epsilon$ intersecting $E_1 \cup E_2$ is intersecting only one of the sets E_1, E_2. Hence, $\mathcal{H}^d_\epsilon(E_1 \cup E_2) \geq \mathcal{H}^d_\epsilon(E_1) + \mathcal{H}^d_\epsilon(E_2)$. Since $\mathcal{H}^d_\epsilon$ is σ-subadditive, we obtain $\mathcal{H}^d_\epsilon(E_1 \cup E_2) = \mathcal{H}^d_\epsilon(E_1) + \mathcal{H}^d_\epsilon(E_2)$, and letting $\epsilon \to 0$, $\mathcal{H}^d(E_1 \cup E_2) = \mathcal{H}^d(E_1) + \mathcal{H}^d(E_2)$. The proof of Theorem 2.1(3) shows that all closed sets, and hence all Borel sets, are $\mathcal{H}^d$-measurable.

(2) This follows from $\mathrm{diam}(E+z)^d = \mathrm{diam}(E)^d$ and $\mathrm{diam}(aE)^d = a^d \mathrm{diam}(E)^d$.

(3) Let $d > n$. Any cube Q of side length 1 can be covered by k^n closed cubes of side length $1/k$. Thus, $\mathcal{H}^d_\epsilon(Q) \leq \omega_d(\sqrt{n}/2)^d k^{n-d}$ for $\epsilon > \sqrt{n}/k$. Letting $k \to \infty$ implies $\mathcal{H}^d(Q) = 0$. The assertion now follows from translation invariance and σ-subadditivity.

(4) We have $(\mathrm{diam}(E_i)/\epsilon)^d \leq (\mathrm{diam}(E_i)/\epsilon)^{d'}$ if $\mathrm{diam}(E_i) < \epsilon$. Thus for $0 < \epsilon < \infty$,

$$\frac{2^d}{\omega_d} \mathcal{H}^d_\epsilon(E) \leq \epsilon^{d-d'} \frac{2^{d'}}{\omega_{d'}} \mathcal{H}^{d'}_\epsilon(E)$$

which implies the statement.

(5) follows from $\mathrm{diam}(f(E)) \leq \mathrm{Lip}(f)\ \mathrm{diam}(E)$. □

Note that $\mathcal{H}^d$ is not σ-finite if $d < n$.

The **Hausdorff dimension** of a subset $E \subseteq \mathbb{R}^n$ is defined by

$$\dim_{\mathcal{H}} E := \inf\{d \geq 0 : \mathcal{H}^d(E) = 0\}.$$

Then, by Proposition 3.41,

$$\mathcal{H}^d(E) = \begin{cases} \infty & \text{if } d < \dim_{\mathcal{H}} E, \\ 0 & \text{if } d > \dim_{\mathcal{H}} E. \end{cases}$$

Finite sets have Hausdorff dimension 0. But there also exist compact uncountable sets with Hausdorff dimension 0.

Example 3.42 (Hausdorff dimension of the Cantor set)**.** Let $C = \bigcap_{k=0}^\infty C_k$ be the Cantor set; see Example 2.3. Recall that C_k is a disjoint union of 2^k closed intervals with length 3^{-k}. Thus

$$\mathcal{H}^d_{3^{-k}}(C) \leq \frac{\omega_d}{2^d} \frac{2^k}{3^{kd}}.$$

This bound remains bounded as $k \to \infty$ provided that $2/3^d \leq 1$. So for the choice

$$d = \frac{\log 2}{\log 3} \tag{3.10}$$

we have $\mathcal{H}^d(C) = \lim_{k\to\infty} \mathcal{H}^d_{3^{-k}}(C) < \infty$ and hence $\dim_{\mathcal{H}} C \leq d$.

To conclude that the Hausdorff dimension of the Cantor set C is $d = \log 2/\log 3$, we need to show that $\mathcal{H}^d(C) > 0$. To this end we prove that $\sum_j \operatorname{diam}(I_j)^d \geq 1/4$ whenever $\{I_j\}$ is a cover of C by open intervals. Since C is compact, we may assume that $I_1, \ldots, I_n$ cover C. As the interior of C is empty, we may also assume that the endpoints of each I_j lie outside of C (making the I_j slightly larger if necessary). Let $\delta > 0$ be the distance between C and the set of all endpoints of intervals I_j, and choose a positive integer k such that $3^{-k} < \delta$. Then each connected component $C_{k,i}$ of C_k is contained in some I_j.

We assert that, for each open interval I and each fixed ℓ,

$$\sum_{C_{\ell,i} \subseteq I} \operatorname{diam}(C_{\ell,i})^d \leq 4 \operatorname{diam}(I)^d. \tag{3.11}$$

This will imply the strived for inequality,

$$4 \sum_j \operatorname{diam}(I_j)^d \geq \sum_j \sum_{C_{k,i} \subseteq I_j} \operatorname{diam}(C_{k,i})^d \geq \sum_{i=1}^{2^k} \operatorname{diam}(C_{k,i})^d = 1,$$

since $\operatorname{diam}(C_{k,i})^d = 3^{-kd} = 2^{-k}$ by (3.10). Let us show (3.11). If m denotes the least integer for which I contains some $C_{m,i}$, then $m \leq \ell$. There are at most 4 connected components $C_{m,i_1}, \ldots, C_{m,i_p}$ of C_m which intersect I; otherwise m would not be minimal. Thus,

$$\sum_{C_{\ell,i} \subseteq I} \operatorname{diam}(C_{\ell,i})^d \leq \sum_{q=1}^p \sum_{C_{\ell,i} \subseteq C_{m,i_q}} \operatorname{diam}(C_{\ell,i})^d = \sum_{q=1}^p \operatorname{diam}(C_{m,i_q})^d \leq 4 \operatorname{diam}(I)^d.$$

because $\sum_{C_{\ell,i} \subseteq C_{m,i_q}} \operatorname{diam}(C_{\ell,i})^d = 2^{\ell-m} 3^{-\ell d} = 2^{\ell-m} 2^{-\ell} = \operatorname{diam}(C_{m,i_q})^d$.

Theorem 3.43 (Isodiametric inequality)**.** *For every Lebesgue measurable set $E \subseteq \mathbb{R}^n$,*

$$\lambda^n(E) \leq \omega_n \Big(\frac{\operatorname{diam}(E)}{2}\Big)^n. \tag{3.12}$$

Proof. For $v \in S^{n-1}$ let π_v be the hyperplane perpendicular to v, and for $w \in \pi_v$ set

$$E_{v,w} := \{t \in \mathbb{R} : w + tv \in E\}.$$

Consider the symmetriced set

$$S_v(E) := \{w + tv : w \in \pi_v,\ 2|t| \le \lambda^1(E_{v,w})\}.$$

By Fubini's theorem 3.27, we may conclude that the mapping $\pi_v \ni w \mapsto \lambda^1(E_{v,w})$ is $\mathfrak{L}(\mathbb{R}^{n-1})$-measurable where $\pi_v \cong \mathbb{R}^{n-1}$, and hence $S_v(E)$ is Lebesgue measurable and $\lambda^n(S_v(E)) = \lambda^n(E)$. We have $\operatorname{diam}(S_v(E)) \le \operatorname{diam}(E)$ thanks to the easy inequality $\lambda^1(I) + \lambda^1(J) \le 2\sup\{|t-s| : t \in I,\ s \in J\}$ for $I, J \in \mathfrak{B}(\mathbb{R})$. If E is symmetric with respect to a direction orthogonal to v, then so is $S_v(E)$.

Define iteratively $E_0 := \overline{E}$ and $E_i := S_{e_i}(E_{i-1})$, where $e_1, \ldots, e_n$ denote the standard unit vectors in $\mathbb{R}^n$. Then E_n is Lebesgue measurable, satisfies $\lambda^n(E_n) = \lambda^n(E_0)$, $\operatorname{diam}(E_n) \le \operatorname{diam}(E)$, and is invariant under the mapping $x \mapsto -x$. Hence E_n is contained in the closed ball with radius $\operatorname{diam}(E)/2$, and

$$\lambda^n(E) \le \lambda^n(E_0) = \lambda^n(E_n) \le \omega_n \Big(\frac{\operatorname{diam}(E)}{2}\Big)^n. \qquad \square$$

This argument is called **Steiner symmetrization**.

Theorem 3.44. *For every Borel set $E \subseteq \mathbb{R}^n$ and every $\epsilon \in (0,\infty]$,*

$$\lambda^n(E) = \mathcal{H}^n_\epsilon(E) = \mathcal{H}^n(E).$$

Proof. Let us prove $\lambda^n(E) \le \mathcal{H}^n_\epsilon(E)$. Let $(E_i)_i$ be a cover of E by closed sets with $\operatorname{diam}(E_i) < \epsilon$. Then, by the isodiametric inequality (3.12),

$$\lambda^n(E) \le \sum_i \lambda^n(E_i) \le \frac{\omega_n}{2^n} \sum_i (\operatorname{diam}(E_i))^n.$$

We may conclude $\lambda^n(E) \le \mathcal{H}^n_\epsilon(E)$, since the cover $(E_i)_i$ was arbitrary.

Note that $\mathcal{H}^n$ is finite on bounded sets; use the argument in the proof of Proposition 3.41(3). Hence $\mathcal{H}^n$ is a translation invariant Radon measure on $\mathbb{R}^n$. By Theorem 2.14, there is a constant $C > 0$ such that $\lambda^n(E) = C\,\mathcal{H}^n(E)$ for all Borel sets $E \subseteq \mathbb{R}^n$.

It remains to show that $C = 1$. If B is the unit ball in $\mathbb{R}^n$, then

$$\lambda^n(B) \le \mathcal{H}^n_\epsilon(B) \le \mathcal{H}^n(B) = C^{-1}\,\lambda^n(B),$$

whence $C \le 1$. On the other hand, for all ϵ,

$$\mathcal{H}^n_\epsilon(B) \le \lambda^n(B) = C\,\mathcal{H}^n(B)$$

and thus $C \ge 1$. In order to see the inequality $\mathcal{H}^n_\epsilon(B) \le \lambda^n(B)$ note that it is possible to find a collection of disjoint closed balls $\overline{B}_1, \overline{B}_2, \ldots$ with $\operatorname{diam}(\overline{B}_i) < \epsilon$ such that $\bigcup_{i=1}^\infty \overline{B}_i \subseteq B$ and $\lambda^n(B \setminus \bigcup_{i=1}^\infty \overline{B}_i) = 0$; this is a consequence of the Besicovitch–Vitali covering theorem, cf. [**3**]. Then

$$\mathcal{H}^n_\epsilon\Big(\bigcup_{i=1}^\infty \overline{B}_i\Big) \le \frac{\omega_n}{2^n} \sum_{i=1}^\infty (\operatorname{diam}(\overline{B}_i))^n = \sum_{i=1}^\infty \lambda^n(\overline{B}_i) = \lambda^n\Big(\bigcup_{i=1}^\infty \overline{B}_i\Big) = \lambda^n(B).$$

We may conclude that $\mathcal{H}^n_\epsilon(B) \le \lambda^n(B)$, since a λ^n-null set is also a $\mathcal{H}^n_\epsilon$-null set. In fact, for every cube $Q \subseteq \mathbb{R}^n$, we have $\omega_n(\operatorname{diam}(Q)/2)^n = \omega_n(\sqrt{n}/2)^n \lambda^n(Q)$ and thus

$$\begin{aligned}\mathcal{H}^n_\epsilon(E) &\le \frac{\omega_n}{2^n} \inf\Big\{\sum_i (\operatorname{diam}(Q_i))^n : Q_i \text{ cubes},\ \operatorname{diam}(Q_i) < \epsilon,\ E \subseteq \bigcup_i Q_i\Big\} \\ &= \omega_n \Big(\frac{\sqrt{n}}{2}\Big)^n \lambda^n(E). \qquad \square\end{aligned}$$

CHAPTER 4

L^p-spaces

Let $(X, \mathfrak{S}, \mu)$ be a measure space.

4.1. Definition of L^p-spaces

For $1 \leq p < \infty$, we set

$$L^p(\mu) := \{f : X \to \mathbb{C} : f \text{ is measurable and } |f|^p \in L^1(\mu)\}.$$

We shall also use the notation $L^p(X)$ if there is no ambiguity. Note that

$$|f+g|^p \leq 2^p \max(|f|, |g|)^p = 2^p \max(|f|^p, |g|^p) \leq 2^p(|f|^p + |g|^p)$$

which implies that $L^p(\mu)$ is a vector space. For $f \in L^p(X)$ we define

$$\|f\|_p := \Big(\int |f|^p \, d\mu\Big)^{1/p}.$$

For $p = \infty$, we set

$$L^\infty(\mu) := \{f : X \to \mathbb{C} : f \text{ is measurable and }$$
$$\exists M \in \mathbb{R} : |f(x)| \leq M \text{ for } \mu\text{-a.e. } x \in X\}.$$

For $f \in L^\infty(X)$ we define the **essential supremum**

$$\|f\|_\infty := \inf\{M : |f(x)| \leq M \text{ for } \mu\text{-a.e. } x \in X\}.$$

We shall see below that $\|f\|_p$, $1 \leq p \leq \infty$, defines a norm on (equivalence classes of functions in) $L^p(\mu)$; it is called the **L^p-norm**; we will also use $\| \; \|_{L^p(\mu)}$ or $\| \; \|_{L^p(X)}$.

If A is a nonempty set, we denote by $l^p(A)$ the space $L^p(\mu)$, where μ is the counting measure on $(A, \mathfrak{P}(A))$.

By Proposition 3.21, for a measurable function f, $\|f\|_p = 0$ if and only if $f = 0$ μ-a.e. So $\| \; \|_p$ is not a norm on $L^p(\mu)$ as defined above. For this reason we *redefine* $L^p(\mu)$: The equivalence relation $f \sim g :\Longleftrightarrow f = g$ μ-a.e. partitions $L^p(\mu)$ into equivalence classes. The L^p-norm is constant on every equivalence class. Henceforth we use the symbol $L^p(\mu)$ for the vector space of equivalence classes of measurable functions whose L^p-norm is finite.

For the sake of simplicity, we will nevertheless speak of *L^p-functions*. However, one should keep in mind that it makes no sense to ask for the value of an L^p-function at some particular point.

4.2. Inequalities

Recall that a real valued function φ defined on an open interval (a, b) is called convex if, for $x, y \in (a, b)$,

$$\varphi((1-t)x + ty) \leq (1-t)\varphi(x) + t\varphi(y), \quad 0 < t < 1,$$

and strictly convex if the inequality is strict. Setting $z = (1-t)x + ty$ we obtain

$$\frac{\varphi(z) - \varphi(x)}{z - x} \leq \frac{\varphi(y) - \varphi(x)}{y - x} \leq \frac{\varphi(y) - \varphi(z)}{y - z}, \quad x < z < y. \tag{4.1}$$

with strict inequalities if φ is strictly convex. The inequalities in (4.1) imply that the one-sided derivatives $\varphi'_\pm(x)$ of φ exist in $\mathbb{R}$ at every $x \in (a,b)$; indeed, the difference quotients $\delta(x,y) := \frac{\varphi(y)-\varphi(x)}{y-x}$ satisfy $\delta(x,y) \geq \delta(x,z)$ for $x < z < y$ and are bounded from below by $\delta(w,x)$ for some $w < x$, thus $\varphi'_+(x) = \lim_{y\to x^+} \delta(x,y)$. As a consequence φ is continuous.

Theorem 4.1 (Jensen's inequality)**.** *Let $(X, \mathfrak{S}, \mu)$ be a measure space with $\mu(X) = 1$. If $f \in L^1(\mu)$ is real valued and $f(X) \subseteq (a,b)$ ($a = -\infty$ and $b = \infty$ are allowed), and if $\varphi : (a,b) \to \mathbb{R}$ is convex, then*

$$\varphi\Big(\int f\,d\mu\Big) \leq \int \varphi \circ f\,d\mu.$$

Proof. Since $\mu(X) = 1$, we have $a < z := \int f\,d\mu < b$. By (4.1),

$$\alpha := \sup_{x<z} \frac{\varphi(z)-\varphi(x)}{z-x} \leq \frac{\varphi(y)-\varphi(z)}{y-z}, \qquad \text{for all } y \in (z,b),$$

and therefore

$$\varphi(w) \geq \varphi(z) + \alpha(w-z), \qquad \text{for all } w \in (a,b).$$

In particular, $\varphi(f(x)) \geq \varphi(z) + \alpha(f(x)-z)$ for all $x \in X$. Since φ is continuous, $\varphi \circ f$ is measurable, and integrating the last inequality yields

$$\int \varphi \circ f\,d\mu \geq \varphi(z) + \alpha(\int f\,d\mu - z) = \varphi\Big(\int f\,d\mu\Big). \qquad \square$$

A pair of positive real numbers p and q are called **conjugate exponents** if

$$\frac{1}{p} + \frac{1}{q} = 1;$$

we regard also 1 and ∞ to be conjugate.

Theorem 4.2 (Hölder's inequality)**.** *Let p and q be conjugate exponents, $1 \leq p \leq \infty$. Let $f \in L^p(\mu)$ and $g \in L^q(\mu)$. Then $fg \in L^1(\mu)$, and*

$$\|fg\|_1 \leq \|f\|_p \|g\|_q.$$

If $p = q = 2$ this is also called **Schwarz inequality**.

Proof. For $p = 1$ this follows easily from the definition of the integral. Let us assume that $1 < p < \infty$. Set $A := \{x \in X : |g(x)| > 0\}$ and $\nu(E) := \int_E |g|^q\,d\mu$, for $E \in \mathfrak{S}$. Since $g \in L^q(\mu)$, we have $\nu(A) = \nu(X) = \|g\|_q^q < \infty$. By Corollary 3.16, $\gamma := \nu/\nu(A)$ is a probability measure on A. By Jensen's inequality 4.1 and since $(1-q)p = -q$,

$$\begin{aligned}
\frac{1}{\nu(A)^p}\|fg\|_1^p &= \Big|\int_A |f||g|^{1-q}\frac{|g|^q}{\nu(A)}\,d\mu\Big|^p = \Big|\int_A |f||g|^{1-q}\,d\gamma\Big|^p \\
&\leq \int_A (|f||g|^{1-q})^p\,d\gamma = \int_A |f|^p|g|^{-q}\frac{|g|^q}{\nu(A)}\,d\mu \\
&= \frac{1}{\nu(A)}\int_A |f|^p\,d\mu = \frac{1}{\nu(A)}\|f\|_p^p
\end{aligned}$$

and hence $\|fg\|_1 \leq \|f\|_p \nu(A)^{1-1/p} = \|f\|_p\|g\|_q$. $\square$

Corollary 4.3. *If $f_i \in L^{p_i}(\mu)$ and $\sum_{i=1}^n 1/p_i = 1/p$ for $p, p_i \in [1,\infty]$, then*

$$\Big\|\prod_{i=1}^n f_i\Big\|_p \leq \prod_{i=1}^n \|f_i\|_{p_i}.$$

PROOF. If $p = \infty$ then $p_i = \infty$ for all i and the inequality is obvious. So assume that $p < \infty$. If $p_i = \infty$ for some i, the result can be reduced to that case that all $p_i < \infty$. So let us make this assumption. If $n = 2$, we have, as $1 = 1/(p_1/p) + 1/(p_2/p)$,

$$\int |f_1 f_2|^p \, d\mu \le \Big(\int |f_1|^{p_1} \, d\mu \Big)^{p/p_1} \Big(\int |f_2|^{p_2} \, d\mu \Big)^{p/p_2} = \|f_1\|_{p_1}^p \|f_2\|_{p_2}^p.$$

In the general case, define q by $1/q = \sum_{i=2}^n 1/p_i$, and use induction:

$$\Big\| \prod_{i=1}^n f_i \Big\|_p \le \|f_1\|_{p_1} \Big\| \prod_{i=2}^n f_i \Big\|_q \le \prod_{i=1}^n \|f_i\|_{p_i}. \qquad \square$$

Proposition 4.4. *Let p and q be conjugate exponents, $1 \le p \le \infty$. If $p = \infty$ we assume that μ has the finite subset property. Then for every $f \in L^p(\mu)$,*

$$\|f\|_p = \sup_{\substack{g \in L^q(\mu) \\ \|g\|_q \le 1}} \int |fg| \, d\mu = \sup_{\substack{g \in L^q(\mu) \\ \|g\|_q \le 1}} \Big| \int fg \, d\mu \Big|.$$

PROOF. The identities are clear if $f = 0$. So let us assume that $\|f\|_p > 0$. By Hölder's inequality 4.2, for each $g \in L^q(\mu)$ with $\|g\|_q \le 1$,

$$\Big| \int fg \, d\mu \Big| \le \int |fg| \, d\mu \le \|f\|_p,$$

hence $\sup \big| \int fg \, d\mu \big| \le \sup \int |fg| \, d\mu \le \|f\|_p$.

It remains to prove that $\|f\|_p \le \sup \big| \int fg \, d\mu \big|$. Consider first the case that $p \in [1, \infty)$. Set $h(x) := |f(x)|^{p-2}\overline{f(x)}$ if $f(x) \ne 0$ and $h(x) := 0$ if $f(x) = 0$, then $fh = |f|^p$. If $p > 1$, then $|h|^q = |f|^p$ and hence $g := \|f\|_p^{-p/q} h$ satisfies $\|g\|_q = 1$ and $\int fg \, d\mu = \|f\|_p$. If $p = 1$, then $\|h\|_\infty = 1$ and $\int fh \, d\mu = \|f\|_1$.

If $p = \infty$, choose $0 < m < \|f\|_\infty$ and set $A_m := \{x \in X : |f(x)| \ge m\}$. Then $\mu(A_m) > 0$. Since μ has the finite subset property, there exists $B_m \subseteq A_m$ with $0 < \mu(B_m) < \infty$. The function $\varphi := \chi_{\{|f(x)|=0\}} + \chi_{\{|f(x)|>0\}} f|f|^{-1}$ is measurable and satisfies $|\varphi| = 1$ and $f = \varphi|f|$. Thus $g_m := \chi_{B_m}/(\varphi\mu(B_m))$ satisfies $\|g_m\|_1 = 1$ and $\int f g_m \, d\mu = \frac{1}{\mu(B_m)} \int_{B_m} |f| \, d\mu \ge m$, and thus

$$\sup \Big\{ \Big| \int fg \, d\mu \Big| : g \in L^1(\mu), \|g\|_1 \le 1 \Big\} \ge m.$$

Letting $m \to \|f\|_\infty$ finishes the proof. $\square$

Theorem 4.5 (Minkowski's integral inequality)**.** *Let $(X, \mathfrak{S}, \mu)$ and $(Y, \mathfrak{T}, \nu)$ be σ-finite measure spaces, let $f : X \times Y \to [0, \infty]$ be $(\mathfrak{S} \otimes \mathfrak{T})$-measurable, and let $1 \le p \le \infty$. Then*

$$\Big\| \int_Y f(\ , y) \, d\nu(y) \Big\|_{L^p(\mu)} \le \int_Y \|f(\ , y)\|_{L^p(\mu)} \, d\nu(y).$$

PROOF. It follows from Fubini's theorem 3.27 that the function $h(x) := \int_Y f(x, y) \, d\nu(y)$, $x \in X$, is measurable. Furthermore, by Proposition 4.4,

$$\begin{aligned}
&\Big\| \int_Y f(\ , y) \, d\nu(y) \Big\|_{L^p(\mu)} = \|h\|_{L^p(\mu)} \\
&= \sup \Big\{ \int |hg| \, d\mu : g \in L^q(\mu), \|g\|_{L^q(\mu)} \le 1 \Big\} \\
&= \sup \Big\{ \int_X \int_Y f(x, y) |g(x)| \, d\nu(y) \, d\mu(x) : g \in L^q(\mu), \|g\|_{L^q(\mu)} \le 1 \Big\}
\end{aligned}$$

$$
\begin{aligned}
&= \sup\Big\{ \int_Y \int_X f(x,y)|g(x)|\, d\mu(x)\, d\nu(y) : g \in L^q(\mu), \|g\|_{L^q(\mu)} \le 1 \Big\} \\
&\le \int_Y \sup\Big\{ \int_X f(x,y)|g(x)|\, d\mu(x) : g \in L^q(\mu), \|g\|_{L^q(\mu)} \le 1 \Big\}\, d\nu(y) \\
&= \int_Y \|f(\ ,y)\|_{L^p(\mu)}\, d\nu(y). \qquad \square
\end{aligned}
$$

Corollary 4.6 (Minkowski's inequality)**.** *Let $1 \le p \le \infty$. For $f_1, f_2 \in L^p(\mu)$,*

$$\|f_1 + f_2\|_p \le \|f_1\|_p + \|f_2\|_p.$$

It follows that $\| \ \|_p$ is a norm on $L^p(\mu)$.

Proof. As $\int |f_1 + f_2|^p \, d\mu \le \int ||f_1| + |f_2||^p \, d\mu$, we may assume without loss of generality that f_1, f_2 are nonnegative. Then Minkowski's inequality follows from Minkowski's integral inequality 4.5 if we let Y be the two point set $\{1, 2\}$ with the counting measure.

Note that in this case the use of Fubini's theorem in the proof of Theorem 4.5 reduces to linearity of the integral, and hence it is not necessary to assume σ-finiteness: if $f(x, 1) = f_1(x)$ and $f(x, 2) = f_2(x)$, then

$$
\begin{aligned}
&\int_X \int_Y f(x,y)|g(x)|\, d\nu(y)\, d\mu(x) \\
&= \int_X f_1(x)|g(x)| + f_2(x)|g(x)|\, d\mu(x) \\
&= \int_X f_1(x)|g(x)|\, d\mu(x) + \int_X f_2(x)|g(x)|\, d\mu(x) \\
&= \int_Y \int_X f(x,y)|g(x)|\, d\mu(x)\, d\nu(y). \qquad \square
\end{aligned}
$$

In general $L^p(\mu) \not\subseteq L^q(\mu)$ for all $p \ne q$; consider x^{-a}, $a > 0$, on $(0, \infty)$ with the Lebesgue measure. However, we have the following results; see also Section 7.3 on interpolation of L^p-spaces.

Proposition 4.7 (Inclusion relations)**.** *If $1 \le p < q < r \le \infty$, then*

$$L^p(\mu) \cap L^r(\mu) \subseteq L^q(\mu) \subseteq L^p(\mu) + L^r(\mu),$$

and

$$\|f\|_q \le \|f\|_p^t \|f\|_r^{1-t}, \qquad \textit{where} \qquad \frac{1}{q} = \frac{t}{p} + \frac{1-t}{r}.$$

Proof. Let us first prove $L^q(\mu) \subseteq L^p(\mu) + L^r(\mu)$. For $f \in L^q(\mu)$ set $E := \{x : |f(x)| > 1\}$ and decompose $f = f\chi_E + f\chi_{E^c}$. This shows the asserted inclusion, since $|f\chi_E|^p = |f|^p\chi_E \le |f|^q\chi_E$, thus $f\chi_E \in L^p(\mu)$, and $|f\chi_{E^c}|^r = |f|^r\chi_{E^c} \le |f|^q\chi_{E^c}$, thus $f\chi_{E^c} \in L^r(\mu)$; for $r = \infty$ we clearly have $\|f\chi_{E^c}\|_\infty \le 1$.

Now we turn to the other inclusion. Consider first the case $r < \infty$. By assumption, $p/(tq)$ and $r/((1-t)q)$ are conjugate, and so by Hölder's inequality 4.2,

$$
\begin{aligned}
\int |f|^q \, d\mu &= \int |f|^{tq} |f|^{(1-t)q} \, d\mu \le \| |f|^{tq} \|_{p/(tq)} \| |f|^{(1-t)q} \|_{r/((1-t)q)} \\
&= \Big(\int |f|^p \, d\mu \Big)^{tq/p} \Big(\int |f|^r \, d\mu \Big)^{(1-t)q/r} = \|f\|_p^{tq} \|f\|_r^{(1-t)q}.
\end{aligned}
$$

If $r = \infty$, then $t = p/q$ and

$$\int |f|^q \, d\mu = \int |f|^{q-p} |f|^p \, d\mu \le \|f\|_\infty^{q-p} \|f\|_p^p$$

which implies the assertion. □

Corollary 4.8. *If A is any set and $1 \le p < q \le \infty$, then $l^p(A) \subseteq l^q(A)$ and*

$$\|f\|_q \le \|f\|_p.$$

Proof. Obviously, $\|f\|_\infty \le \|f\|_p$. For $q < \infty$, setting $r = \infty$ and $t = p/q$ in Proposition 4.7 implies

$$\|f\|_q \le \|f\|_p^{p/q} \|f\|_\infty^{1-p/q} \le \|f\|_p^{p/q} \|f\|_p^{1-p/q} = \|f\|_p.$$ □

Proposition 4.9. *If $\mu(X) < \infty$ and $1 \le p < q \le \infty$, then $L^q(\mu) \subseteq L^p(\mu)$ and*

$$\|f\|_p \le \mu(X)^{1/p-1/q} \|f\|_q.$$

Proof. By Hölders inequality 4.2, for $1/r + 1/r' = 1$,

$$\||f|^p\|_1 \le \|1\|_{r'} \||f|^p\|_r = \mu(X)^{1/r'} \Big(\int_X |f|^{pr} \, d\mu \Big)^{1/r}$$

and thus

$$\|f\|_p \le \mu(X)^{1/pr'} \|f\|_{pr}.$$

Setting $r = q/p$ gives the assertion. □

4.3. Completeness

Let $1 \le p \le \infty$. The normed space $(L^p(\mu), \| \, \|_p)$ comes with a natural notion of convergence. A sequence (f_i) in $L^p(\mu)$ is called **(strongly) convergent** if there exists an element $f \in L^p(\mu)$ such that $\|f_i - f\|_p \to 0$ as $i \to \infty$. A sequence (f_i) in $L^p(\mu)$ is a **Cauchy sequence** if for all $\epsilon > 0$ there is $k \in \mathbb{N}$ so that $\|f_i - f_j\|_p < \epsilon$ if $i, j \ge k$. Recall that a normed space is **complete** if each Cauchy sequence is convergent.

Theorem 4.10 (Riesz–Fischer)**.** *Let $1 \le p \le \infty$. The space $L^p(\mu)$ is complete and hence a Banach space.*

Proof. Let $1 \le p < \infty$. Let (f_i) be a Cauchy sequence in $L^p(\mu)$. Choose i_1 such that $\|f_{i_1} - f_j\|_p < 1/2$ for $j \ge i_1$, choose $i_2 > i_1$ such that $\|f_{i_2} - f_j\|_p < 1/2^2$ for $j \ge i_2$, etc. In this way we obtain a subsequence (f_{i_k}) such that $\|f_{i_k} - f_{i_{k+1}}\|_p < 1/2^k$ for all $k \ge 1$. Let us define

$$F := |f_{i_1}| + \sum_{k=1}^{\infty} |f_{i_{k+1}} - f_{i_k}|.$$

Then F is an element of $L^p(\mu)$, by the monotone convergence theorem 3.14, since, for all $m \ge 1$,

$$\Big\| |f_{i_1}| + \sum_{k=1}^{m} |f_{i_{k+1}} - f_{i_k}| \Big\|_p \le \|f_{i_1}\|_p + 1.$$

In particular, $F(x)$ is finite for μ-a.e. x, and for such x the series $f_{i_1}(x) + \sum_{k=1}^{\infty} f_{i_{k+1}}(x) - f_{i_k}(x)$ is absolutely convergent, and thus the sequence of partial sums

$$f_{i_1}(x) + \sum_{k=1}^{m} f_{i_{k+1}}(x) - f_{i_k}(x) = f_{i_{m+1}}(x)$$

converges to some number $f(x)$. Since $|f_{i_k}(x)| \le F(x)$ and $F \in L^p(\mu)$, the dominated convergence theorem 3.22 implies that $f \in L^p(\mu)$, and in turn that

$\|f_{i_k} - f\|_p \to 0$ as $k \to \infty$, since $|f_{i_k} - f|^p \to 0$ and $|f_{i_k} - f|^p \le (2F)^p$ μ-a.e. That $\|f_i - f\|_p \to 0$ as $i \to \infty$ follows from

$$\|f_i - f\|_p \le \|f_i - f_{i_k}\|_p + \|f_{i_k} - f\|_p.$$

Let (f_i) be a Cauchy sequence in $L^\infty(\mu)$. The sets $E_i = \{x : |f_i(x)| > \|f_i\|_\infty\}$ and $E_{jk} = \{x : |f_j(x) - f_k(x)| > \|f_j - f_k\|_\infty\}$ and thus also their union E for all $i, j, k \in \mathbb{N}$ have measure zero. On E^c the sequence f_i converges uniformly to a bounded function f. Extending f by 0 on E we obtain a measurable bounded function satisfying $\|f_i - f\|_\infty \to 0$. (In more details: clearly, f_i converges pointwise to a function f on E^c. To see uniform convergence, let, for given $\epsilon > 0$, k be such that $\sup_{x\in E^c} |f_i(x) - f_j(x)| < \epsilon/2$ for $i, j \ge k$, and for $x \in E^c$ choose $i_x \ge k$ such that $|f(x) - f_{i_x}(x)| < \epsilon/2$. Then $|f(x) - f_j(x)| \le |f(x) - f_{i_x}(x)| + |f_{i_x}(x) - f_j(x)| < \epsilon$ for $j \ge k$, independently of x. In particular, $|f(x)| \le |f(x) - f_k(x)| + |f_k(x)| \le \epsilon + \sup_{x\in E^c} |f_k(x)|$ for all $x \in E^c$, i.e., f is bounded.) □

Corollary 4.11. *Let $1 \le p \le \infty$. Any Cauchy sequence in $L^p(\mu)$ has a subsequence that converges pointwise μ-a.e.*

Proof. This was shown in the proof of Theorem 4.10; see also Proposition 4.24 and Theorem 4.25. □

Corollary 4.12. *$L^2(\mu)$ is a Hilbert space with inner product $\langle f, g\rangle = \int_X f\overline{g}\, d\mu$.*

Proof. $\langle f, g\rangle$ is well-defined by Hölder's inequality 4.2 and it is easy to see that it defines an inner product on $L^2(\mu)$. Since $\|f\|_2 = \langle f, f\rangle^{1/2}$, the completeness follows from Theorem 4.10. □

4.4. Convolution and approximation by smooth functions

We will see in this section that L^p-functions on open subsets of $\mathbb{R}^n$ can be approximated by *nicer* functions if $1 \le p < \infty$. We start with the following proposition.

Proposition 4.13. *Let S denote the class of all simple functions s on X satisfying $\mu(\{x : s(x) \ne 0\}) < \infty$. If $1 \le p < \infty$, then S is dense in $L^p(\mu)$.*

Proof. Clearly, $S \subseteq L^p(\mu)$. Let $f \in L^p(\mu)$, $f \ge 0$. By Theorem 3.6, there exist simple functions $0 \le s_1 \le s_2 \le \cdots \le f$ so that $s_i(x) \to f(x)$ for μ-a.e. x. Thanks to $s_i \le f$ we have $\mu(\{x : s_i(x) \ne 0\}) < \infty$, i.e., $s_i \in S$. Since $|f - s_i|^p \le f^p$, the dominated convergence theorem 3.22 implies that $\|f - s_i\|_p \to 0$. The general complex case follows immediately. □

For the rest of the section let X be an open subset of $\mathbb{R}^n$ equipped with the Lebesgue measure λ; we shall write $L^p(X)$ instead of $L^p(\lambda)$ and $\int_X f\, dx$ instead of $\int_X f\, d\lambda$.

Theorem 4.14 (Approximation by continuous functions)**.** *For $1 \le p < \infty$, the class $C_c(X)$ of continuous functions with compact support in X is dense in $L^p(X)$.*

Proof. By Proposition 4.13, it suffices to show that, for each measurable $E \subseteq X$ with $\lambda(E) < \infty$, χ_E is the L^p-limit of a sequence of functions in $C_c(X)$. Since λ is regular, see Theorem 2.9, for given $\epsilon > 0$ there exist an open set U and a compact set K such that $K \subseteq E \subseteq U \subseteq X$, $\lambda(E) < \lambda(K) + \epsilon$, and $\lambda(U) < \lambda(E) + \epsilon$.

Let L be a compact neighborhood of K contained in U. If f is a continuous function on $\mathbb{R}$ so that $0 \leq f \leq 1$ and $f|_{\{t\leq 0\}} \equiv 1$ and $f|_{\{t>1/2\}} \equiv 0$, then

$$g(x) := f\Big(1 - \frac{\operatorname{dist}(x, L^c)}{\operatorname{dist}(K, L^c)}\Big)$$

is a continuous function with support in L and 1 on K. So $\chi_K \leq g \leq \chi_U$ and hence

$$\chi_K - \chi_E \leq g - \chi_E \leq \chi_U - \chi_E$$

which implies

$$(g - \chi_E)^+ \leq \chi_U - \chi_E \quad \text{ and } \quad (g - \chi_E)^- \leq \chi_E - \chi_K.$$

Therefore, using $(a + b)^p \leq (2 \max(a, b))^p \leq 2^p(a^p + b^p)$ for $a, b \geq 0$,

$$\begin{aligned}\int_X |g - \chi_E|^p \, dx &= \int_X ((g - \chi_E)^+ + (g - \chi_E)^-)^p \, dx \\ &\leq 2^p \int_X ((g - \chi_E)^+)^p + ((g - \chi_E)^-)^p \, dx \\ &\leq 2^{p+1}\epsilon.\end{aligned}$$

This finishes the proof, since ϵ was arbitrary. □

Note that $C_c(X)$ is not dense in $L^\infty(X)$. If f is a bounded and continuous function on X then

$$\|f\|_\infty = \sup_{x\in X} |f(x)|. \tag{4.2}$$

Clearly, $\|f\|_\infty \leq \sup_{x\in X} |f(x)|$. Conversely, for any $\epsilon > 0$ there exists a nonempty open subset $U \subseteq X$ such that $|f(y)| \geq \sup_{x\in X} |f(x)| - \epsilon$ for all $y \in U$. So the supremum of $|f(x)|$ on the complement of any null set is $\geq \sup_{x\in X} |f(x)| - \epsilon$, since this complement has nonempty intersection with U. As $\epsilon > 0$ was arbitrary we obtain (4.2). Consequently, any limit of functions in $C_c(X)$ with respect to $\| \; \|_\infty$ must be continuous, but there are elements in $L^\infty(X)$ with no continuous representative.

Let f and g be complex valued functions on $\mathbb{R}^n$. We formally define their **convolution** $f * g$ by

$$(f * g)(x) := \int_{\mathbb{R}^n} f(x - y)g(y) \, dy.$$

One has to be careful to make sure that the definition makes sense. The integral is well-defined for all $x \in \mathbb{R}^n$, if we require that $f \in L^p(\mathbb{R}^n)$ and $g \in L^q(\mathbb{R}^n)$ for p, q conjugate exponents, by Hölder's inequality 4.2. But actually more is true:

Theorem 4.15 (Young's inequality)**.** *Let $1 \leq p, q, r \leq \infty$ be such that $1/p + 1/q = 1/r + 1$. If $f \in L^p(\mathbb{R}^n)$ and $g \in L^q(\mathbb{R}^n)$, then $f * g \in L^r(\mathbb{R}^n)$ and*

$$\|f * g\|_r \leq \|f\|_p \|g\|_q. \tag{4.3}$$

Proof. We may assume without loss of generality that f and g are Borel functions, since there exist Borel functions which coincide with f and g a.e., by Proposition 3.9. Then the mapping $(x, y) \mapsto f(x - y)g(y)$ is also a Borel function, since $(x, y) \mapsto x - y$ and $(x, y) \mapsto y$ are Borel.

The case $r = \infty$ follows easily from Hölder's inequality 4.2:

$$|(f * g)(x)| \leq \int_{\mathbb{R}^n} |f(x - y)g(y)| \, dy \leq \|f\|_p \|g\|_q,$$

where we used translation invariance of the integral.

So assume $r < \infty$. Set $h(x) = (f * g)(x) = \int_{\mathbb{R}^n} f(x-y)g(y)\,dy$; we shall see in the course of the proof that $h(x)$ is defined and finite for a.e. x.

Set $s = p(1 - 1/q)$ and let q' be the conjugate exponent of q. By Hölder's inequality 4.2,

$$\begin{aligned}|h(x)| &\le \int_{\mathbb{R}^n} |f(x-y)g(y)|\,dy = \int_{\mathbb{R}^n} |f(x-y)|^{1-s}|f(x-y)|^s|g(y)|\,dy \\ &\le \Big(\int_{\mathbb{R}^n} |f(x-y)|^{(1-s)q}|g(y)|^q\,dy\Big)^{1/q}\Big(\int_{\mathbb{R}^n} |f(y)|^{sq'}\,dy\Big)^{1/q'},\end{aligned}$$

where we used translation invariance of the integral. Since $sq' = p$, we have

$$|h(x)|^q \le \Big(\int_{\mathbb{R}^n} |f(x-y)|^{(1-s)q}|g(y)|^q\,dy\Big)\|f\|_p^{sq}.$$

Note that $1/p + 1/q = 1/r + 1$ implies that $r \ge q$; in fact, $r = pq/(p+q-pq)$ and $p \ge p+q-pq$. So $t := r/q \ge 1$ and we can apply Minkowski's integral inequality 4.5:

$$\||h|^q\|_t \le \||g|^q\|_1\||f|^{(1-s)q}\|_t\|f\|_p^{sq} = \|g\|_q^q\|f\|_{t(1-s)q}^{(1-s)q}\|f\|_p^{sq}$$

and hence

$$\|h\|_r \le \|g\|_q\|f\|_{r(1-s)}^{1-s}\|f\|_p^s$$

which is (4.3), since $(1-s)r = p$. □

In particular, the convolution of $f, g \in L^1(\mathbb{R}^n)$ is a function $f * g \in L^1(\mathbb{R}^n)$ satisfying

$$\|f * g\|_1 \le \|f\|_1\|g\|_1,$$

and, for $f \in L^1(\mathbb{R}^n)$ and $g \in L^p(\mathbb{R}^n)$, $f * g \in L^p(\mathbb{R}^n)$ with

$$\|f * g\|_p \le \|f\|_1\|g\|_p. \tag{4.4}$$

Assuming that all integrals in question exist, the convolution is commutative, $f * g = g * f$, by Theorem 3.33, associative, $(f * g) * h = f * (g * h)$, by Fubini's theorem 3.27, and satisfies

$$\operatorname{supp}(f * g) \subseteq \overline{\operatorname{supp} f + \operatorname{supp} g}; \tag{4.5}$$

indeed, if $x \notin \operatorname{supp} f + \operatorname{supp} g$ then for all $y \in \operatorname{supp} g$ we have $x - y \notin \operatorname{supp} f$, and hence $f(x-y)g(y) = 0$ for all y.

We denote by $L^1_{\mathrm{loc}}(\mathbb{R}^n)$ the set of **locally integrable** functions, i.e., measurable functions $f : \mathbb{R}^n \to \mathbb{C}$ such that $\int_K |f(x)|\,dx < \infty$ for all bounded measurable subsets $K \subseteq \mathbb{R}^n$, and $C^k_c(\mathbb{R}^n)$ is the class of k times continuously differentiable functions on $\mathbb{R}^n$ with compact support.

Lemma 4.16. *If $\varphi \in C^k_c(\mathbb{R}^n)$ and $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$, then $\varphi * f \in C^k(\mathbb{R}^n)$, and*

$$\partial^\alpha(\varphi * f) = (\partial^\alpha\varphi) * f.$$

Proof. Clearly, $\varphi * f$ is well-defined. The lemma then follows from Theorem 3.38. □

For a function f on $\mathbb{R}^n$ and $y \in \mathbb{R}^n$ we consider the **translation**

$$T_y f(x) := f(x-y), \quad x \in \mathbb{R}^n. \tag{4.6}$$

Note that $\|T_y f\|_p = \|f\|_p$, for $1 \le p \le \infty$.

Lemma 4.17. *For $1 \le p < \infty$, translation is continuous in the L^p-norm, i.e., if $f \in L^p(\mathbb{R}^n)$ and $z \in \mathbb{R}^n$, then $\lim_{y\to 0}\|T_{y+z}f - T_z f\|_p = 0$.*

PROOF. It suffices to assume that $z = 0$, since $T_{y+z} = T_y T_z$. If $g \in C_c(\mathbb{R}^n)$, then the support of $T_y g$ is contained in a fixed compact set K for all $|y| \le 1$, and thus

$$\int_{\mathbb{R}^n} |T_y g(x) - g(x)|^p \, dx \le \|T_y g - g\|_\infty^p \, \lambda(K) \to 0, \quad \text{as } y \to 0,$$

since g is uniformly continuous. If $f \in L^p(\mathbb{R}^n)$ and $\epsilon > 0$, then there exists $g \in C_c(\mathbb{R}^n)$ with $\|g - f\|_p \le \epsilon/3$, by Theorem 4.14, and so

$$\|T_y f - f\|_p \le \|T_y f - T_y g\|_p + \|T_y g - g\|_p + \|g - f\|_p \le \epsilon,$$

for y sufficiently small. □

For any function φ on $\mathbb{R}^n$ and $\epsilon > 0$ we set

$$\varphi_\epsilon(x) = \epsilon^{-n} \varphi(x/\epsilon), \quad x \in \mathbb{R}^n. \tag{4.7}$$

If $\varphi \in L^1(\mathbb{R}^n)$, then $\int_{\mathbb{R}^n} \varphi_\epsilon(x)\, dx$ is independent of ϵ, by Theorem 3.33, and, for every $r > 0$ we have $\lim_{\epsilon \to 0} \int_{|x| \ge r} \varphi_\epsilon(x)\, dx = 0$, indeed

$$\int_{|x| \ge r} \varphi_\epsilon(x)\, dx = \int_{|x| \ge r} \epsilon^{-n} \varphi(x/\epsilon)\, dx = \int_{|x| \ge r/\epsilon} \varphi(x)\, dx.$$

Proposition 4.18. *Let $\varphi \in L^1(\mathbb{R}^n)$ with $\int_{\mathbb{R}^n} \varphi(x)\, dx = a$, and let $1 \le p < \infty$. If $f \in L^p$, then $\|f * \varphi_\epsilon - af\|_p \to 0$ as $\epsilon \to 0$.*

PROOF. By Theorem 3.33,

$$\begin{aligned} f * \varphi_\epsilon(x) - af(x) &= \int_{\mathbb{R}^n} (f(x-y) - f(x)) \varphi_\epsilon(y)\, dy \\ &= \int_{\mathbb{R}^n} (f(x - \epsilon z) - f(x)) \varphi(z)\, dz \\ &= \int_{\mathbb{R}^n} (T_{\epsilon z} f(x) - f(x)) \varphi(z)\, dz, \end{aligned}$$

and by Minkowski's integral inequality 4.5,

$$\|f * \varphi_\epsilon - af\|_p = \int_{\mathbb{R}^n} \|T_{\epsilon z} f - f\|_p |\varphi(z)|\, dz.$$

Now $\|T_{\epsilon z} f - f\|_p \to 0$ as $\epsilon \to 0$, by Lemma 4.17, and as $\|T_{\epsilon z} f - f\|_p \le 2\|f\|_p$, the assertion follows from the dominated convergence theorem 3.22. □

If $\int_{\mathbb{R}^n} \varphi\, dx = 1$ we say that the family $\{\varphi_\epsilon\}_{0 < \epsilon \le 1}$ is an **approximate identity**. A **mollifier** is a nonnegative function $\varphi \in C_c^\infty(\mathbb{R}^n)$ satisfying $\|\varphi\|_1 = 1$.

Example 4.19. Consider the function

$$\psi(x) := \begin{cases} \exp \frac{1}{|x|^2 - 1} & |x| < 1 \\ 0 & |x| \ge 1 \end{cases}.$$

Then $\varphi = (\int \psi\, dx)^{-1} \psi$ is a mollifier.

Theorem 4.20 (Approximation by smooth functions)**.** *For $1 \le p < \infty$, $C_c^\infty(X)$ is dense in $L^p(X)$.*

PROOF. Let $f \in L^p(X)$ and let $\delta > 0$. We may assume that $f \in L^p(\mathbb{R}^n)$ by setting $f \equiv 0$ on X^c. By Theorem 4.14, there exists $g \in C_c(\mathbb{R}^n)$ so that

$$\|f - g\|_p \le \delta/2.$$

Let φ be a mollifier and let φ_ϵ be defined by (4.7). By Lemma 4.16 and (4.5), $g_\epsilon := \varphi_\epsilon * g \in C_c^\infty(\mathbb{R}^n)$. By Proposition 4.18, $\|g_\epsilon - g\|_p \le \delta/2$ for sufficiently small ϵ. Thus,

$$\|g_\epsilon - f\|_p \le \|g_\epsilon - g\|_p + \|g - f\|_p \le \delta$$

which implies the assertion. □

Lemma 4.21 (Smooth Urysohn lemma)**.** *If $K \subseteq \mathbb{R}^n$ is compact and U is an open set containing K, then there exists $f \in C_c^\infty(\mathbb{R}^n)$ such that $0 \le f \le 1$, $f|_K \equiv 1$, and* $\operatorname{supp} f \subseteq U$.

Proof. Let $\delta := \operatorname{dist}(K, U^c)$, $V := \{x : \operatorname{dist}(x, K) < \delta/3\}$, and let φ be a mollifier with $\operatorname{supp}\varphi \subseteq B_{\delta/3}(0)$. Then $f := \chi_V * \varphi$ is as required. □

Finally, we will show that, for $1 \le p < \infty$, $L^p(X)$ is **separable**, i.e., it contains a countable dense subset.

Lemma 4.22. *If $1 \le p < \infty$, then the set of step functions is dense in $L^p(\mathbb{R}^n)$.*

Proof. By Proposition 4.13, simple functions s so that $\lambda(\{x : s(x) \neq 0\}) < \infty$ are dense in $L^p(\mathbb{R}^n)$. Such s are finite linear combinations of characteristic functions of sets E with $\lambda(E) < \infty$. So it suffices to show that for given $\epsilon > 0$ there exists a step function f so that $\|\chi_E - f\|_p \le \epsilon$. By Proposition 2.15, there exist almost disjoint cubes $Q_1, \ldots, Q_m$ such that $\lambda\big(E \triangle \bigcup_{i=1}^m Q_i\big) < \epsilon$, and thus $f = \sum_{i=1}^m \chi_{Q_i}$ satisfies

$$\int |\chi_E - f|^p \, d\lambda \le \lambda\Big(E \triangle \bigcup_{i=1}^m Q_i\Big) < \epsilon.$$ □

Theorem 4.23 (Separability)**.** *For $1 \le p < \infty$, $L^p(\mathbb{R}^n)$ is separable.*

Proof. Let $f \in L^p(\mathbb{R}^n)$ and let $\epsilon > 0$. By Lemma 4.22, there is a step function s satisfying $\|f - s\|_p \le \epsilon/2$. We may conclude that there is a step function t satisfying $\|f - t\|_p \le \epsilon$ and such that the real and imaginary parts of the coefficients and the coordinates of the boxes appearing in the canonical form of t are all rational numbers. So the set of step functions with rational real and imaginary parts of the coefficients and rational coordinates of the boxes appearing in its canonical form is dense in $L^p(\mathbb{R}^n)$. □

4.5. Modes of convergence

Let $(X, \mathfrak{S}, \mu)$ be a measure space. A sequence f_i of measurable complex valued functions on X is said to be **Cauchy in measure** if

$$\forall \epsilon > 0 \quad \mu(\{x : |f_i(x) - f_j(x)| \ge \epsilon\}) \to 0 \qquad \text{as } i, j \to \infty,$$

and we say that f_i **converges in measure** to f if

$$\forall \epsilon > 0 \quad \mu(\{x : |f_i(x) - f(x)| \ge \epsilon\}) \to 0 \qquad \text{as } i \to \infty.$$

Proposition 4.24. *If $f_i \to f$ in $L^1(\mu)$, then $f_i \to f$ in measure.*

The converse is not true.

Proof. If $E_{i,\epsilon} := \{x : |f_i(x) - f(x)| \ge \epsilon\}$, then

$$\int |f_i - f| \, d\mu \ge \int_{E_{i,\epsilon}} |f_i - f| \, d\mu \ge \epsilon \mu(E_{i,\epsilon})$$

goes to 0 as $i \to \infty$. □

Theorem 4.25. *If f_i is Cauchy in measure, then there is a measurable function f such that $f_i \to f$ in measure, and there is a subsequence of f_i that converges to f μ-a.e. If also $f_i \to g$ in measure, then $f = g$ μ-a.e.*

Proof. The sequence f_i has a subsequence h_j satisfying

$$\mu(\{x : |h_j(x) - h_{j+1}(x)| \geq 1/2^j\}) \leq 1/2^j.$$

Set $E_j := \{x : |h_j(x) - h_{j+1}(x)| \geq 1/2^j\}$ and $F_k := \bigcup_{j=k}^{\infty} E_j$. Then $\mu(F_k) \leq 2^{1-k}$. If $x \notin F_k$, then for all $i \geq j \geq k$,

$$|h_i(x) - h_j(x)| \leq \sum_{\ell=j}^{i-1} |h_{\ell+1}(x) - h_\ell(x)| \leq \sum_{\ell=j}^{i-1} 2^{-\ell} \leq 2^{1-j}. \tag{4.8}$$

It follows that h_j is pointwise Cauchy on $(F_k)^c$. For $F = \bigcap_{k=1}^{\infty} F_k$, we have $\mu(F) = 0$, and we define $f(x) := \lim_{j\to\infty} h_j(x)$ for $x \notin F$ and $f(x) := 0$ for $x \in F$. Then f is measurable and $h_j \to f$ μ-a.e. For $x \notin F_k$ and $j \geq k$, we have $|h_j(x) - f(x)| \leq 2^{1-j}$, by (4.8), and hence $h_j \to f$ in measure, since $\mu(F_k) \to 0$ as $k \to \infty$. It follows that $f_i \to f$ in measure, since

$$\{x : |f_i(x) - f(x)| \geq \epsilon\} \subseteq \{x : |f_i(x) - h_j(x)| \geq \epsilon/2\} \cup \{x : |h_j(x) - f(x)| \geq \epsilon/2\}.$$

If $f_i \to g$ in measure, then

$$\{x : |f(x) - g(x)| \geq \epsilon\} \subseteq \{x : |f(x) - f_i(x)| \geq \epsilon/2\} \cup \{x : |f_i(x) - g(x)| \geq \epsilon/2\}$$

implies $f = g$ μ-a.e. □

Convergence a.e. does not imply convergence in measure. However, this implication holds on a finite measure space, actually more is true:

Theorem 4.26 (Egorov's theorem)**.** *Let $\mu(X) < \infty$ and let $f_1, f_2, \ldots$ and f be measurable complex valued functions on X such that $f_i \to f$ μ-a.e. Then for every $\epsilon > 0$ there is a set $E \subseteq X$ such that $\mu(E) < \epsilon$ and $f_i \to f$ uniformly on E^c.*

Proof. Without loss of generality assume that $f_i(x) \to f(x)$ for every $x \in X$. For $k, \ell \in \mathbb{N}$ define

$$E_{k,\ell} := \bigcup_{i\geq k} \{x : |f_i(x) - f(x)| \geq 1/\ell\}.$$

Clearly, $E_{k,\ell} \supseteq E_{k+1,\ell}$ and $\bigcap_{k=1}^{\infty} E_{k,\ell} = \emptyset$, thus $\lim_{k\to\infty} \mu(E_{k,\ell}) = 0$, by Lemma 1.1. So, given $\epsilon > 0$, we find a subsequence k_ℓ such that $\mu(E_{k_\ell,\ell}) < \epsilon/2^\ell$. For $E = \bigcup_{\ell=1}^{\infty} E_{k_\ell,\ell}$, we have $\mu(E) < \epsilon$, and $|f_i(x) - f(x)| < 1/\ell$ if $i > k_\ell$ and $x \notin E$. It follows that $f_i \to f$ uniformly on E^c. □

Let us call the type of convergence in the conclusion of Egorov's theorem **almost uniform convergence**. The following diagram summarizes different modes of convergence $f_i \to f$ of a sequence of measurable complex valued functions on a measure space $(X, \mathfrak{S}, \mu)$.

$$\begin{tikzcd}
 & \text{uniform convergence} \arrow[d, Rightarrow] \\
 & \text{pointwise convergence} \arrow[d, Rightarrow] \\
f_i \to f \ \mu\text{-a.e.},\ |f_i| \le g \in L^1(\mu) \arrow[r, Rightarrow] \arrow[d, Rightarrow] & \mu\text{-a.e. convergence} \\
L^1 - \text{convergence} \arrow[d, Rightarrow] & \\
\text{convergence in measure} \arrow[d, Rightarrow] & \text{almost uniform convergence} \arrow[l, Rightarrow] \arrow[uu, Rightarrow] \\
\mu\text{-a.e. convergence for a subsequence} & \mu(X) < \infty,\ \mu\text{-a.e. convergence} \arrow[u, Rightarrow]
\end{tikzcd}$$

Theorem 4.27 (Lusin's theorem)**.** *Let f be a Lebesgue measurable complex valued function defined on a Lebesgue measurable set $E \subseteq \mathbb{R}^n$ with $\lambda(E) < \infty$. Then for every $\epsilon > 0$ there exists a compact set $K \subseteq E$ such that $\lambda(E \setminus K) \le \epsilon$ and such that $f|_K$ is continuous.*

PROOF. Assume without loss of generality that f is real valued and defined on $\mathbb{R}^n$ by setting $f \equiv 0$ in E^c. For each positive integer i, let $\{B_{ij}\}_{j=1}^{\infty}$ be a collection of disjoint Borel sets so that $\mathbb{R} = \bigcup_{j=1}^{\infty} B_{ij}$ and $\operatorname{diam} B_{ij} < 1/i$. Set $E_{ij} := E \cap f^{-1}(B_{ij})$. By regularity of λ, Theorem 2.9, there are compact sets $K_{ij} \subseteq E_{ij}$ satisfying $\lambda(E_{ij} \setminus K_{ij}) < \epsilon/2^{i+j}$. Since $E = \bigcup_{j=1}^{\infty} E_{ij}$,

$$\lambda\Big(E \setminus \bigcup_{j=1}^{\infty} K_{ij}\Big) \le \lambda\Big(\bigcup_{j=1}^{\infty} (E_{ij} \setminus K_{ij})\Big) < \epsilon/2^i.$$

By Lemma 1.1, $\lim_{k\to\infty} \lambda(E \setminus \bigcup_{j=1}^{k} K_{ij}) = \lambda(E \setminus \bigcup_{j=1}^{\infty} K_{ij})$, and so there are integers k_i such that $\lambda(E \setminus \bigcup_{j=1}^{k_i} K_{ij}) < \epsilon/2^i$. The sets $L_i := \bigcup_{j=1}^{k_i} K_{ij}$ are compact. Choose $b_{ij} \in B_{ij}$ and define $g_i : L_i \to \mathbb{R}$ be setting $g_i|_{K_{ij}} = b_{ij}$; the sets $K_{i,1}, \dots, K_{i,k_i}$ are compact and disjoint, so their mutual distance is positive, and g_i is continuous. As $\operatorname{diam} B_{ij} < 1/i$, we have $|f(x) - g_i(x)| < 1/i$ for all $x \in L_i$. Then the set $K := \bigcap_{i=1}^{\infty} L_i$ is compact, we have

$$\lambda(E \setminus K) \le \sum_{i=1}^{\infty} \lambda(E \setminus L_i) < \epsilon,$$

and $g_i \to f$ uniformly on K. It follows that $f|_K$ is continuous. $\square$

This does not mean that f is continuous at every $x \in K$; consider e.g. $\chi_{\mathbb{Q}\cap[0,1]}$.

4.6. The distribution function

Let $f : X \to \mathbb{C}$ be a measurable function on a measure space $(X, \mathfrak{S}, \mu)$. The **distribution function** d_f of f is defined by

$$d_f(\alpha) := \mu(\{x \in X : |f(x)| > \alpha\}), \quad \alpha \ge 0.$$

It follows from the definition that d_f is decreasing. Let us set $E_{f,\alpha} := \{x \in X : |f(x)| > \alpha\}$.

Lemma 4.28. *Let $(X, \mathfrak{S}, \mu)$ be a measure space and let $f, g : X \to \mathbb{C}$ be measurable functions. Then for all $\alpha, \beta > 0$:*

(1) *If $|f| \leq |g|$ μ-a.e., then $d_f \leq d_g$.*
(2) *$d_{cf}(\alpha) = d_f(\alpha/|c|)$ for every $c \in \mathbb{C} \setminus \{0\}$.*
(3) *$d_{f+g}(\alpha + \beta) \leq d_f(\alpha) + d_g(\beta)$.*
(4) *$d_{fg}(\alpha\beta) \leq d_f(\alpha) + d_g(\beta)$.*

Proof. (1) If $|f| \leq |g|$ μ-a.e., then $d_f(\alpha) = \mu(E_{f,\alpha}) \leq \mu(E_{g,\alpha}) = d_g(\alpha)$.

(2) We have $E_{cf,\alpha} = E_{f,\alpha/|c|}$.

(3) & (4) If $|f(x) + g(x))| > \alpha + \beta$ then $|f(x)| > \alpha$ or $|g(x)| > \beta$. Similarly if $|f(x)g(x))| > \alpha\beta$. $\square$

The distribution function d_f does not provide information about the behavior of f near any given point. However, the L^p-norm ($p < \infty$) of f can be computed if we only know d_f.

Proposition 4.29. *Let $(X, \mathfrak{S}, \mu)$ be a σ-finite measure space. If f is a measurable function on X and $0 < p < \infty$, then*

$$\int |f|^p \, d\mu = p \int_0^\infty \alpha^{p-1} d_f(\alpha) \, d\alpha. \tag{4.9}$$

Proof. By Fubini's theorem 3.27,

$$\begin{aligned} p \int_0^\infty \alpha^{p-1} d_f(\alpha) \, d\alpha &= p \int_0^\infty \alpha^{p-1} \int_X \chi_{E_{f,\alpha}} \, d\mu \, d\alpha \\ &= \int_X \int_0^{|f(x)|} p\alpha^{p-1} \, d\alpha \, d\mu \\ &= \int_X |f(x)|^p \, d\mu. \end{aligned}$$ $\square$

Remark 4.30. This result holds without the assumption of σ-finiteness; cf. [**5**, 6.24].

Let $(X, \mathfrak{S}, \mu)$ be a measure space, and let $1 \leq p < \infty$. The **weak Lebesgue space** $L^{p,\infty}(\mu)$ is defined as the set of all measurable functions f such that

$$\begin{aligned} \|f\|_{p,\infty} &:= \inf\Big\{C > 0 : d_f(\lambda) \leq (C/\alpha)^p \text{ for all } \alpha > 0\Big\} \\ &= \sup_{\alpha > 0} \alpha \, d_f(\alpha)^{1/p} < \infty. \end{aligned} \tag{4.10}$$

By definition $L^{\infty,\infty}(\mu) := L^\infty(\mu)$. As usual two functions in $L^{p,\infty}(\mu)$ are considered equal if they are equal μ-a.e.

By Lemma 4.28, we obtain that

$$\|cf\|_{p,\infty} = |c| \|f\|_{p,\infty},$$

for each $c \in \mathbb{C} \setminus \{0\}$, and

$$\|f + g\|_{p,\infty} \leq 2(\|f\|_{p,\infty} + \|g\|_{p,\infty}).$$

Moreover, $\|f\|_{p,\infty} = 0$ implies that $f = 0$ μ-a.e. That means that $L^{p,\infty}(\mu)$ is a **quasinormed space**. One can show that it is complete.

Proposition 4.31 (Chebyshev's inequality)**.** *Let $1 \le p < \infty$. If $f \in L^p(\mu)$ then $f \in L^{p,\infty}(\mu)$ and*

$$\|f\|_{p,\infty} \le \|f\|_p. \tag{4.11}$$

Proof. We have for all $\alpha > 0$,

$$\|f\|_p^p = \int |f|^p \, d\mu \ge \int_{E_{f,\alpha}} |f|^p \, d\mu \ge \alpha^p \mu(E_{f,\alpha}) = \alpha^p d_f(\alpha). \qquad \square$$

The inclusion $L^p(\mu) \subseteq L^{p,\infty}(\mu)$ is strict. For example, the function $f(x) = |x|^{-1/p}$ is in $L^{p,\infty}(\mathbb{R})$ but not in $L^p(\mathbb{R})$ (with the Lebesgue measure).

Proposition 4.32. *Let $(X, \mathfrak{S}, \mu)$ be a finite measure space. If $1 \le q < p < \infty$ then $L^{p,\infty}(\mu) \subseteq L^q(\mu)$ and*

$$\|f\|_q \le \Big(\frac{p}{p-q}\Big)^{1/q} \mu(X)^{1/q-1/p} \|f\|_{p,\infty}, \quad f \in L^{p,\infty}(\mu). \tag{4.12}$$

Proof. Let $f \in L^{p,\infty}(\mu)$. Then $d_f(\alpha) \le \min\{\mu(X), \alpha^{-p}\|f\|_{p,\infty}^p\}$, by (4.10). Thus, for $A := \mu(X)^{-1/p}\|f\|_{p,\infty}$, using Proposition 4.29,

$$\begin{aligned}
\|f\|_q^q &= q\int_0^\infty \alpha^{q-1} d_f(\alpha)\, d\alpha \\
&\le q\int_0^A \alpha^{q-1}\mu(X)\, d\alpha + q\int_A^\infty \alpha^{q-p-1}\|f\|_{p,\infty}^p \, d\alpha \\
&= A^q\,\mu(X) + \frac{q}{p-q}A^{q-p}\,\|f\|_{p,\infty}^p \\
&= \mu(X)^{1-q/p}\|f\|_{p,\infty}^q + \frac{q}{p-q}\mu(X)^{1-q/p}\|f\|_{p,\infty}^q \\
&= \frac{p}{p-q}\mu(X)^{1-q/p}\|f\|_{p,\infty}^q. \qquad \square
\end{aligned}$$

Proposition 4.33. *If $1 \le p < q < r < \infty$, then*

$$L^{p,\infty}(\mu) \cap L^{r,\infty}(\mu) \subseteq L^{q,\infty}(\mu)$$

and

$$\|f\|_{q,\infty} \le \|f\|_{p,\infty}^t \|f\|_{r,\infty}^{1-t}, \qquad \textit{where} \quad \frac{1}{q} = \frac{t}{p} + \frac{1-t}{r}.$$

Proof. Since $tq/p + (1-t)q/r = 1$, for all $\alpha > 0$,

$$\alpha^q d_f(\alpha) = (\alpha d_f(\alpha)^{1/p})^{tq}(\alpha d_f(\alpha)^{1/r})^{(1-t)q} \le \|f\|_{p,\infty}^{tq}\|f\|_{r,\infty}^{(1-t)q}. \qquad \square$$

CHAPTER 5

Absolute continuity of measures

5.1. Complex measures

Let $(X, \mathfrak{S})$ be a measurable space. A **complex measure** is a mapping $\nu : \mathfrak{S} \to \mathbb{C}$ satisfying

$$\nu\Big(\bigcup_{i=1}^{\infty} E_i\Big) = \sum_{i=1}^{\infty} \nu(E_i),$$

if $E_i \in \mathfrak{S}$ are pairwise disjoint. Note that setting $E_i = \emptyset$ for all i yields $\nu(\emptyset) = 0$. A positive measure is a complex measure only if it is finite. The above series is independent of the order of its terms, i.e., it converges unconditionally and hence absolutely.

Complex measures arise naturally. For instance, let μ be a positive measure on X and let $f \in L^1(\mu)$. Then $\nu(E) = \int_E f\, d\mu$ is a complex measure; cf. the proof of Corollary 3.16 and use the dominated convergence theorem 3.22.

For a complex measure ν one defines its **total variation** by

$$|\nu|(E) := \sup\Big\{ \sum_{i=1}^{\infty} |\nu(E_i)| : E = \bigcup_{i=1}^{\infty} E_i, E_i \in \mathfrak{S} \text{ disjoint}\Big\}.$$

By definition we have

$$|\nu(E)| \le |\nu|(E)$$

and, if ν is a positive measure, then $|\nu|(E) = \nu(E)$.

Theorem 5.1. *The total variation $|\nu|$ of a complex measure ν is a finite positive measure.*

The total variation $|\nu|$ is the smallest positive measure that dominates ν, i.e., if μ is a positive measure such that $|\nu(E)| \le \mu(E)$ for all $E \in \mathfrak{S}$, then $|\nu|(E) \le \mu(E)$ for all $E \in \mathfrak{S}$. The fact that $|\nu|$ is finite implies that every complex measure is bounded: $|\nu(E)| \le |\nu|(E) \le |\nu|(X)$.

Proof. Let $E_i \in \mathfrak{S}$ be disjoint and $E = \bigcup_{i=1}^{\infty} E_i$. In order to see that $|\nu|$ is a positive measure we need to show

$$|\nu|(E) = \sum_{i=1}^{\infty} |\nu|(E_i). \tag{5.1}$$

If $|\nu|(E_i) = \infty$ for some i, then clearly $|\nu|(E) = \infty$; so let us assume that $|\nu|(E_i) < \infty$ for all i. Let $\epsilon > 0$. For each i, there are disjoint $E_{ij} \in \mathfrak{S}$ so that $E_i = \bigcup_{j=1}^{\infty} E_{ij}$ and $|\nu|(E_i) \le \sum_{j=1}^{\infty} |\nu(E_{ij})| + \epsilon/2^i$. Then

$$\sum_{i=1}^{\infty} |\nu|(E_i) \le \sum_{i,j=1}^{\infty} |\nu(E_{ij})| + \epsilon \le |\nu|(E) + \epsilon,$$

since $E = \bigcup_{i,j=1}^{\infty} E_{ij}$ is a disjoint union. This implies $\sum_{i=1}^{\infty} |\nu|(E_i) \le |\nu|(E)$.

Conversely, if $F_j \in \mathfrak{S}$ are disjoint and $E = \bigcup_{j=1}^\infty F_j$, then

$$\begin{aligned}\sum_{j=1}^\infty |\nu(F_j)| &= \sum_{j=1}^\infty \Big| \sum_{i=1}^\infty \nu(F_j \cap E_i) \Big| \\ &\le \sum_{j=1}^\infty \sum_{i=1}^\infty |\nu(F_j \cap E_i)| \\ &= \sum_{i=1}^\infty \sum_{j=1}^\infty |\nu(F_j \cap E_i)| \\ &\le \sum_{i=1}^\infty |\nu|(E_i),\end{aligned}$$

and taking the supremum over all such partitions $\{F_j\}$ we may conclude that $|\nu|(E) \le \sum_{i=1}^\infty |\nu|(E_i)$. Thus, we proved (5.1) and $|\nu|$ is a positive measure.

It remains to show that $|\nu|(X) < \infty$. Since $|\nu|(E) \le |\operatorname{Re}\nu|(E) + |\operatorname{Im}\nu|(E)$, we may assume that ν is real valued. That $|\nu|(X) < \infty$ will follow from the claim that, if $E \in \mathfrak{S}$ and $|\nu|(E) = \infty$, then $E = A \cup B$ with disjoint $A, B \in \mathfrak{S}$ and

$$|\nu(A)| \ge 1 \quad \text{and} \quad |\nu|(B) = \infty.$$

Indeed, this assertion can be applied recursively (starting with $E = X$) to obtain disjoint sets $A_1, A_2, \ldots \in \mathfrak{S}$ with $|\nu(A_i)| \ge 1$ for all i. This leads to a contradiction, since $\nu(\bigcup_{i=1}^\infty A_i) = \sum_{i=1}^\infty \nu(A_i)$, but this series cannot converge.

Let us prove the claim. Suppose that $|\nu|(E) = \infty$. Then there exist disjoint sets $E_i \in \mathfrak{S}$ with $E = \bigcup_{i=1}^\infty E_i$ so that

$$\sum_{i=1}^\infty |\nu(E_i)| \ge 2 + |\nu(E)|.$$

Set $E_+ := \bigcup_{\nu(E_i)\ge 0} E_i$ and $E_- := \bigcup_{\nu(E_i)<0} E_i$. Then the previous inequality becomes

$$|\nu(E_+)| + |\nu(E_-)| \ge 2 + \big||\nu(E_+)| - |\nu(E_-)|\big|$$

and thus $|\nu(E_\pm)| \ge 1$. Since $E = E_+ \cup E_-$ and so $|\nu|(E) = |\nu|(E_+) + |\nu|(E_-)$, $|\nu|(E_+) = \infty$ or $|\nu|(E_-) = \infty$ (or both). $\square$

A **real measure** $\nu : \mathfrak{S} \to \mathbb{R}$ (often called a **signed measure**) can be decomposed into **positive** and **negative variations**,

$$\nu = \nu^+ - \nu^- \quad \text{where} \quad \nu^\pm := \frac{|\nu| \pm \nu}{2}.$$

By Theorem 5.1, $\nu^\pm$ are finite positive measures. This is known as the **Jordan decomposition**. If $\nu = \nu_1 - \nu_2$ is any other decomposition into positive measures, then $\nu_1 \ge \nu^+$ and $\nu_2 \ge \nu^-$; see the remarks after Theorem 5.7.

If ν is a real measure and f is $|\nu|$-integrable, then the integral of f with respect to ν is defined by

$$\int f\, d\nu := \int f\, d\nu^+ - \int f\, d\nu^-.$$

This definition can evidently be extended to any complex measure ν by applying it to the real and imaginary part of ν.

One can show that the set of all complex measures on a measurable space X equipped with the norm $\|\nu\| = |\nu|(X)$ forms a Banach space.

5.2. Absolute continuity and decomposition of measures

Let $(X, \mathfrak{S})$ be a measurable space, and let μ be a positive measure on $\mathfrak{S}$. In the following we assume that ν, ν_1, ν_2, etc., are further either positive or complex measures on $\mathfrak{S}$.

We say that ν is **absolutely continuous** with respect to μ, and write $\nu \ll \mu$, if, for each $E \in \mathfrak{S}$, $\mu(E) = 0$ implies $\nu(E) = 0$. For instance, the measure $\nu(E) = \int_E f\, d\mu$, where $f \in L^1(\mu)$, satisfies $\nu \ll \mu$; we shall see below that every measure absolutely continuous with respect to μ is of this form.

Two measures ν_1 and ν_2 on $\mathfrak{S}$ are called **mutually singular**, and we write $\nu_1 \perp \nu_2$, if they are supported on disjoint sets, i.e., there exist disjoint $E_1, E_2 \in \mathfrak{S}$ such that $\nu_i(E) = 0$ if $E \cap E_i = \emptyset$, $i = 1, 2$. For instance, the Lebesgue measure and the Dirac measure on $\mathbb{R}^n$ are mutually singular.

Lemma 5.2.

(1) *If $\nu_i \ll \mu$, $i = 1, 2$, then $\nu_1 + \nu_2 \ll \mu$.*
(2) *If $\nu_i \perp \mu$, $i = 1, 2$, then $\nu_1 + \nu_2 \perp \mu$.*
(3) *If $\nu_1 \ll \mu$ and $\nu_2 \perp \mu$, then $\nu_1 \perp \nu_2$.*
(4) *If $\nu \ll \mu$ and $\nu \perp \mu$, then $\nu = 0$.*
(5) *If $\nu \ll \mu$, then $|\nu| \ll \mu$.*

Proof. (1) is obvious.

(2) There exist $E_1, E_2, E \in \mathfrak{S}$ such that $E_i \cap E = \emptyset$ and ν_i is supported on E_i, $i = 1, 2$, and μ is supported on E. Then $\nu_1 + \nu_2$ is supported on $E_1 \cup E_2$ and $(E_1 \cup E_2) \cap E = \emptyset$.

(3) There exists $E_2 \in \mathfrak{S}$ so that ν_2 is supported on E_2 and $\mu(E_2) = 0$. Since $\nu_1 \ll \mu$, $\nu_1(E_2) = 0$ and hence ν_1 has support in E_2^c.

(4) By (3), $\nu \perp \nu$ and hence $\nu = 0$.

(5) Suppose that $\mu(E) = 0$ and let $E = \bigcup_{i=1}^{\infty} E_i$ for disjoint $E_i \in \mathfrak{S}$. Then $\mu(E_i) = 0$ for all i. Since $\nu \ll \mu$ we have $\nu(E_i) = 0$ for all i, and thus $\sum_i |\nu(E_i)| = 0$. This implies $|\nu|(E) = 0$. $\square$

Theorem 5.3 (Lebesgue–Radon–Nikodym theorem)**.** *Let μ and ν be positive finite measures on a measurable space $(X, \mathfrak{S})$. Then we have*

(1) *There is a unique pair of positive measures ν_a and ν_s on $\mathfrak{S}$ such that*
$$\nu = \nu_a + \nu_s, \quad \nu_a \ll \mu, \quad \nu_s \perp \mu, \quad \nu_a \perp \nu_s.$$
(2) *There is a unique $f \in L^1(\mu)$ such that*
$$\nu_a(E) = \int_E f\, d\mu, \quad E \in \mathfrak{S}.$$

The decomposition $\nu = \nu_a + \nu_s$ is called the **Lebesgue decomposition** of ν with respect to μ. Part (2) is known as the **Radon–Nikodym theorem**. The function f in (2) is called the Radon–Nikodym derivative of ν_a with respect to μ; one writes $d\nu_a = f\, d\mu$ or $f = d\nu_a / d\mu$.

Proof. To see uniqueness in (1) let ν_a' and ν_s' be another pair satisfying (1). Then $\nu_a - \nu_a' = \nu_s' - \nu_s$, $\nu_a - \nu_a' \ll \mu$, and $\nu_s' - \nu_s \perp \mu$, and thus $\nu_a - \nu_a' = \nu_s' - \nu_s = 0$, by Lemma 5.2. Uniqueness in (2) follows from Proposition 3.21.

Set $\varphi = \nu + \mu$. Then φ is a positive finite measure on $\mathfrak{S}$, and we have
$$\int_X f\, d\varphi = \int_X f\, d\nu + \int_X f\, d\mu$$

which is obvious for characteristic functions of sets in $\mathfrak{S}$, hence for simple functions, and thus also for arbitrary measurable functions. If $f \in L^2(\varphi)$, then

$$\Big|\int_X f\,d\nu\Big| \le \int_X |f|\,d\nu \le \int_X |f|\,d\varphi \le \varphi(X)^{1/2}\|f\|_{L^2(\varphi)},$$

by Hölder's inequality 4.2. We may infer that $f \mapsto \int_X f\,d\nu$ is a bounded linear functional on $L^2(\varphi)$. By Corollary 4.12 and Theorem A.8, there exists $g \in L^2(\varphi)$ such that, for all $f \in L^2(\varphi)$,

$$\int_X f\,d\nu = \int_X fg\,d\varphi.$$

In particular, for all $E \in \mathfrak{S}$,

$$\nu(E) = \int_X \chi_E g\,d\varphi = \int_E g\,d\varphi.$$

It follows that $g(x) \ge 0$ for φ-a.e. x, and since

$$\mu(E) = \varphi(E) - \nu(E) = \int_E (1-g)\,d\varphi,$$

we also have $g(x) \le 1$ for φ-a.e. x. Without loss of generality we may assume that $0 \le g(x) \le 1$ for all x. We obtain, for $f \in L^2(\varphi)$,

$$\int_X (1-g)f\,d\nu = \int_X f\,d\nu - \int_X fg\,d\nu = \int_X fg\,d\varphi - \int_X fg\,d\nu = \int_X fg\,d\mu. \tag{5.2}$$

Set $A := \{x : 0 \le g(x) < 1\}$ and $B := \{x : g(x) = 1\}$, and define

$$\nu_a(E) := \nu(A \cap E), \quad \nu_s(E) := \nu(B \cap E), \quad E \in \mathfrak{S}.$$

Taking $f = \chi_B$ in (5.2) we find $0 = \int_B (1-g)\,d\nu = \int_B g\,d\mu = \mu(B)$, and hence $\nu_s \perp \mu$. Since g is bounded and φ is finite, $f = (1 + g + g^2 + \cdots + g^k)\chi_E \in L^2(\varphi)$, for $E \in \mathfrak{S}$, and inserting f in (5.2) gives

$$\int_E (1-g^{k+1})\,d\nu = \int_E g(1 + g + g^2 + \cdots + g^k)\,d\mu.$$

For $x \in B$, $1 - g^{k+1}(x) = 0$, and for $x \in A$, $g^{k+1}(x) \searrow 0$ as $k \to \infty$, and therefore the left side converges to $\nu_a(E)$, by the monotone convergence theorem 3.14. The integrand of the right side converges monotonically to a positive measurable function h, and, by the monotone convergence theorem 3.14, we find that, for $E \in \mathfrak{S}$,

$$\nu_a(E) = \int_E h\,d\mu.$$

For $E = X$ we see that $h \in L^1(\mu)$, since $\nu_a(X) < \infty$. So we have proved (2). In particular, $\nu_a \ll \mu$ which completes the proof of (1). $\square$

Corollary 5.4 (Lebesgue–Radon–Nikodym theorem)**.** *We have the following extensions:*

1. *Theorem 5.3 remains true if μ is a positive σ-finite measure and ν is a complex measure (where ν_a and ν_s now are complex measures).*
2. *If μ and ν are positive σ-finite measures, then Theorem 5.3 still holds with the restriction that the function f is no longer in $L^1(\mu)$.*

PROOF. If μ is σ-finite, then $\bigcup_{i=1}^\infty X_i = X$ for disjoint $X_i \in \mathfrak{S}$ with $\mu(X_i) < \infty$.

(1) Suppose first that ν is positive with $\nu(X) < \infty$. Then we may apply Theorem 5.3 to each X_i. The Lebesgue decompositions of the restrictions of ν to X_i add up to a Lebesgue decomposition of ν. We obtain L^1-functions f_i on X_i with respect to the restriction of μ to X_i. Then $f := \sum_{i=1}^\infty f_i\chi_{X_i}$ satisfies

$\nu_a(E) = \int_E f\,d\mu$ and is $L^1(\mu)$, since $\nu(X) < \infty$. If ν is complex valued, we apply this to positive and negative variations of the real and imaginary part of ν.

(2) This follows in the same way as (1); we can assume that also $\nu(X_i) < \infty$. The function f satisfies $\int_{X_i} f\,d\mu < \infty$ for each i. □

The result fails if we go beyond σ-finiteness. For example, on $X = \mathbb{R}$ consider the σ-algebra $\mathfrak{L}(\mathbb{R})$ of Lebesgue measurable sets and let μ be the counting measure and $\nu = \lambda$ the Lebesgue measure on $\mathfrak{L}(\mathbb{R})$. Then $\nu \ll \mu$, but there is no function f satisfying $d\nu = f\,d\mu$. If there were such f, then $f(x_0) > 0$ for some $x_0 \in \mathbb{R}$ and $0 < f(x_0) = \int_{\{x_0\}} f\,d\mu = \nu(\{x_0\}) = 0$.

Proposition 5.5 (Characterization of absolute continuity)**.** *Let μ and ν be measures on a measurable space $(X, \mathfrak{S})$, μ positive and ν complex. Then the following are equivalent:*

(1) $\nu \ll \mu$.
(2) *For each $\epsilon > 0$ there is a $\delta > 0$ so that $|\nu(E)| < \epsilon$ for all $E \in \mathfrak{S}$ with $\mu(E) < \delta$.*

Proof. Clearly, (2) implies (1). Assume that (2) does not hold. Then there is $\epsilon > 0$ and there are $E_i \in \mathfrak{S}$ so that $\mu(E_i) < 2^{-i}$ and $|\nu(E_i)| \geq \epsilon$. Let us set $F_k := \bigcup_{i=k}^\infty E_i$ and $F = \bigcap_{k=1}^\infty F_k$. Then $\mu(F_k) \leq 2^{-k+1}$ and $\mu(F) = \lim_{k\to\infty} \mu(F_k) = 0$, by Lemma 1.1. Similarly, $|\nu|(F) = \lim_{k\to\infty} |\nu|(F_k) \geq \epsilon > 0$. Thus we do not have $|\nu| \ll \mu$, and hence (1) does not hold, by Lemma 5.2. □

Theorem 5.6 (Polar decomposition)**.** *Let ν be a complex measure on a measurable space $(X, \mathfrak{S})$. Then there exists a measurable function f on X satisfying $|f(x)| = 1$ for all $x \in X$, and such that*

$$d\nu = f\,d|\nu|.$$

Proof. The Radon–Nikodym theorem 5.3 implies that there is a function $f \in L^1(|\nu|)$ so that $d\nu = f\,d|\nu|$. Let us show that $|f(x)| = 1$ for all $x \in X$.

Set $E_a := \{x : |f(x)| < a\}$ and let $E_a = \bigcup_{i=1}^\infty E_{ai}$ be a partition of E_a. Then

$$\sum_{i=1}^\infty |\nu(E_{ai})| = \sum_{i=1}^\infty \Big| \int_{E_{ai}} f\,d|\nu| \Big| \leq \sum_{i=1}^\infty a|\nu|(E_{ai}) = a|\nu|(E_a),$$

and hence $|\nu|(E_a) \leq a|\nu|(E_a)$. This implies that $|\nu|(E_a) = 0$ if $a < 1$, and therefore $|f| \geq 1$ $|\nu|$-a.e.

On the other hand, whenever $|\nu|(E) > 0$,

$$\Big| \frac{1}{|\nu|(E)} \int_E f\,d|\nu| \Big| = \frac{|\nu(E)|}{|\nu|(E)} \leq 1.$$

We will show that this implies that $|f| \leq 1$ $|\nu|$-a.e. Take an open disk $B_r(c)$ in the complement of the closed unit disk $\overline{B_1(0)}$ in $\mathbb{C}$. It suffices to show that $E := f^{-1}(B_r(c))$ is a $|\nu|$-null set, since $\overline{B_1(0)}^c$ is a countable union of such disks. If $|\nu|(E) > 0$ then

$$\Big| \frac{1}{|\nu|(E)} \int_E f\,d|\nu| - c \Big| = \Big| \frac{1}{|\nu|(E)} \int_E (f - c)\,d|\nu| \Big| \leq r,$$

a contradiction.

By redefining f on the set $\{x : |f(x)| \neq 1\}$, the statement follows. □

Theorem 5.7 (Hahn decomposition)**.** *Let ν be a signed measure on a measurable space $(X, \mathfrak{S})$. Then there exist disjoint sets $P, N \in \mathfrak{S}$ such that $X = P \cup N$ and*

$$\nu^+(E) = \nu(P \cap E) \quad \textit{and} \quad \nu^-(E) = -\nu(N \cap E), \quad E \in \mathfrak{S}.$$

Proof. By Theorem 5.6, $d\nu = f\, d|\nu|$ for a measurable function f with $|f| = 1$. Since ν is real valued, so is f; this is true a.e. and everywhere after redefining f. Thus $f(X) = \{\pm 1\}$. Set $P := \{x : f(x) = 1\}$ and $N := \{x : f(x) = -1\}$. Note that

$$\frac{1 + f(x)}{2} = \begin{cases} f(x) & x \in P \\ 0 & x \in N \end{cases},$$

and since $\nu^+ = (|\nu| + \nu)/2$, we have for $E \in \mathfrak{S}$,

$$\nu^+(E) = \frac{1}{2} \int_E (1 + f)\, d|\nu| = \int_{P \cap E} f\, d|\nu| = \nu(P \cap E).$$

That $\nu^-(E) = -\nu(N \cap E)$ follows from $\nu = \nu^+ - \nu^-$ and from $\nu(E) = \nu(P \cap E) + \nu(N \cap E)$. $\square$

As a corollary we obtain that the Jordan decomposition is minimal in the following sense: if $\nu = \nu_1 - \nu_2$ for positive measures ν_1 and ν_2 then $\nu_1 \geq \nu^+$ and $\nu_2 \geq \nu^-$. In fact, as $\nu \leq \nu_1$ we have $\nu^+(E) = \nu(P \cap E) \leq \nu_1(P \cap E) \leq \nu_1(E)$.

CHAPTER 6

Differentiation and integration

6.1. The Lebesgue differentiation theorem

Recall that $L^1_{\mathrm{loc}}(\mathbb{R}^n)$ is the set of measurable functions $f : \mathbb{R}^n \to \mathbb{C}$ such that $\int_K |f(x)|\,dx < \infty$ for all bounded measurable subsets $K \subseteq \mathbb{R}^n$.

For $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$, $x \in \mathbb{R}^n$, and $r > 0$ we consider the **average** $A_r f(x)$ of f over the open ball $B_r(x)$,

$$A_r f(x) := \frac{1}{\lambda(B_r(x))} \int_{B_r(x)} f(y)\,dy = ⨍_{B_r(x)} f(y)\,dy.$$

We shall use the notation $⨍_E f\,dx = \lambda(E)^{-1} \int_E f\,dx$ whenever E is bounded and measurable, $\lambda(E) > 0$, and $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$.

Lemma 6.1. *The mapping* $(0,\infty) \times \mathbb{R}^n \ni (r,x) \mapsto A_r f(x) \in \mathbb{C}$ *is continuous.*

Proof. The functions $\chi_{B_r(x)}$ converge pointwise to $\chi_{B_{r_0}(x_0)}$ on the set $\mathbb{R}^n \setminus \{x : |x - x_0| = r_0\}$ as (r,x) tends to (r_0,x_0). Thus, $\chi_{B_r(x)} \to \chi_{B_{r_0}(x_0)}$ λ-a.e. on $\mathbb{R}^n$. Moreover, $|\chi_{B_r(x)}| \le \chi_{B_{r_0+1}(x_0)}$ if $r < r_0+1/2$ and $|x-x_0| < 1/2$. By the dominated convergence theorem 3.22, we have

$$\int_{B_r(x)} f(y)\,dy \to \int_{B_{r_0}(x_0)} f(y)\,dy,$$

and since $\lambda(B_r(x)) = \lambda(B_1(0))r^n \to \lambda(B_1(0))r_0^n = \lambda(B_{r_0}(x_0))$, the statement follows. $\square$

For $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$ we may define the **Hardy–Littlewood maximal function** Mf by

$$Mf(x) := \sup_{r>0} A_r|f|(x) = \sup_{r>0} ⨍_{B_r(x)} |f(y)|\,dy.$$

Then Mf is measurable, since $(Mf)^{-1}((a,\infty)) = \bigcup_{r>0}(A_r|f|)^{-1}((a,\infty))$ is open, by Lemma 6.1.

Lemma 6.2. *Let $\mathcal{C}$ be a collection of open balls in $\mathbb{R}^n$, and $U = \bigcup \mathcal{C}$. If $c < \lambda(U)$, then there are finitely many disjoint $B_1,\dots,B_k \in \mathcal{C}$ so that $\sum_{j=1}^k \lambda(B_j) > 3^{-n}c$.*

Proof. By Theorem 2.9, there is a compact set $K \subseteq U$ with $\lambda(K) > c$. The set K is covered by finitely many balls $A_1,\dots,A_\ell \in \mathcal{C}$. Let B_1 be one of the balls A_i with maximal radius. Let B_2 be a ball of maximal radius among the balls A_i disjoint from B_1. Let B_3 be a ball of maximal radius among the balls A_i disjoint from B_1 and B_2, etc., until the collection of A_i is exhausted. If $A_i \notin \{B_1,\dots,B_k\}$ then $A_i \cap B_j \neq \emptyset$ for some j, and if j is the smallest integer with that property, then the radius of A_i is at most that of B_j. Consequently, $A_i \subseteq B_j^*$, where B_j^* is the open ball concentric with B_j whose radius is three times that of B_j. Then

$B_1^*, \dots, B_k^*$ cover K and so

$$c < \lambda(K) \le \sum_{j=1}^{k} \lambda(B_j^*) = 3^n \sum_{j=1}^{k} \lambda(B_j). \qquad \square$$

Theorem 6.3 (M is weak type $(1,1)$)**.** *For each $f \in L^1(\mathbb{R}^n)$ and each $a > 0$, we have*

$$\lambda(\{x : Mf(x) > a\}) \le \frac{C}{a} \int_{\mathbb{R}^n} |f(x)|\, dx,$$

where C is a constant depending only on n.

Proof. Set $E_a := \{x : Mf(x) > a\}$ and let $x \in E_a$. Then there exists $r_x > 0$ so that $A_{r_x}|f|(x) > a$. The collection of balls $\{B_{r_x}(x)\}_{x \in E_a}$ covers E_a, and by Lemma 6.2, given $c < \lambda(E_a)$ there exist $x_1, \dots, x_k \in E_a$ so that the balls $B_j = B_{r_{x_j}}(x_j)$ are disjoint and $\sum_{j=1}^k \lambda(B_j) > 3^{-n} c$. Thus,

$$c < 3^n \sum_{j=1}^{k} \lambda(B_j) \le \frac{3^n}{a} \sum_{j=1}^{k} \int_{B_j} |f(x)|\, dx \le \frac{3^n}{a} \int_{\mathbb{R}^n} |f(x)|\, dx.$$

Letting $c \to \lambda(E_a)$ yields the result. $\square$

A sublinear mapping T (i.e. $|T(f+g)| \le |Tf| + |Tg|$ and $|T(cf)| = c|Tf|$ for $c > 0$) is called **weak type** (p, q) for $1 \le p \le \infty$ and $1 \le q < \infty$ if T maps $L^p(\mu)$ into $L^{q,\infty}(\mu)$ and $\|Tf\|_{q,\infty} \le C\|f\|_p$ for all $f \in L^p(\mu)$.

Theorem 6.3 means that the Hardy–Littlewood maximal operator M satisfies $\|Mf\|_{1,\infty} \le C\|f\|_1$ for $f \in L^1(\mathbb{R}^n)$, so it is weak type $(1,1)$; see also Corollary 7.11.

Proposition 6.4. *If $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$ then $\lim_{r\to 0} A_r f(x) = f(x)$ for λ-a.e. $x \in \mathbb{R}^n$, i.e.,*

$$\lim_{r \to 0} \fint_{B_r(x)} (f(y) - f(x))\, dy = 0 \quad \text{for } \lambda\text{-a.e. } x \in \mathbb{R}^n. \tag{6.1}$$

Proof. It suffices to show that, for each $N \in \mathbb{N}$, we have $\lim_{r\to 0} A_r f(x) = f(x)$ for λ-a.e. $x \in B_N(0)$. As, for $x \in B_N(0)$ and $r \le 1$, the values of $A_r f(x)$ depend only on the values of $f(y)$ for $y \in B_{N+1}(0)$, we may replace f by $\chi_{B_{N+1}(0)} f$ and hence assume that $f \in L^1(\mathbb{R}^n)$.

Let $\epsilon > 0$. By Theorem 4.14, there is a continuous function g with $\|f - g\|_1 \le \epsilon$. By continuity of g, for each $x \in \mathbb{R}^n$,

$$|A_r g(x) - g(x)| \le \fint_{B_r(x)} |g(y) - g(x)|\, dy \le \sup_{y \in B_r(x)} |g(y) - g(x)| \to 0$$

as $r \to 0$. Now

$$|A_r f(x) - f(x)| \le A_r|f - g|(x) + |A_r g(x) - g(x)| + |g(x) - f(x)|,$$

and taking $\limsup_{r\to 0} = \lim_{\epsilon \to 0} \sup_{0<r<\epsilon}$ on both sides we find

$$\limsup_{r \to 0} |A_r f(x) - f(x)| \le M(f - g)(x) + |g(x) - f(x)|.$$

This implies that

$$E_a := \{x : \limsup_{r\to 0} |A_r f(x) - f(x)| > a\}$$

satisfies

$$E_a \subseteq \{x : M(f - g)(x) > a/2\} \cup \{x : |g(x) - f(x)| > a/2\}.$$

It follows from Theorem 6.3 and Chebyshev's inequality 4.31 that

$$\lambda(E_a) \le \frac{2(C+1)}{a}\int_{\mathbb{R}^n} |f(x)-g(x)|\,dx \le \frac{2(C+1)}{a}\epsilon.$$

As $\epsilon > 0$ was arbitrary, $\lambda(E_a) = 0$.

Since $\lim_{r\to 0} A_r f(x) = f(x)$ if and only if $\limsup_{r\to 0} |A_r f(x) - f(x)| = 0$, we have $\lim_{r\to 0} A_r f(x) = f(x)$ if $x \notin \bigcup_{k=1}^{\infty} E_{1/k}$. This implies the assertion. □

We will show in the next theorem that (6.1) remains true if we replace the integrand by its absolute value. A **Lebesgue point** of a function $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$ is a point $x \in \mathbb{R}^n$ so that

$$\lim_{r\to 0} ⨍_{B_r(x)} |f(y) - f(x)|\,dy = 0.$$

Let L_f denote the set of all Lebesgue points of f.

Theorem 6.5. *If $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$ then $\lambda((L_f)^c) = 0$.*

Proof. Let $c \in \mathbb{C}$. Applying (6.1) to $x \mapsto |f(x) - c|$ shows that

$$\lim_{r\to 0} ⨍_{B_r(x)} |f(y) - c|\,dy = |f(x) - c|$$

except on a null set E_c. Let D be a countable dense subset of $\mathbb{C}$. Then $E = \bigcup_{c\in D} E_c$ is a null set. Assume $x \notin E$. For each $\epsilon > 0$ there is $c \in D$ so that $|f(x) - c| < \epsilon$, and thus

$$\limsup_{r\to 0} ⨍_{B_r(x)} |f(y) - f(x)|\,dy \le \limsup_{r\to 0} ⨍_{B_r(x)} |f(y) - c|\,dy + \epsilon = |f(x) - c| + \epsilon < 2\epsilon.$$

Since ϵ was arbitrary, the proof is complete. □

We shall now establish Theorem 6.5 for families of sets more general than $\{B_r(x)\}_r$. A family of Borel sets $\{E_r\}_{r>0}$ is said to **shrink nicely** to x if

- $E_r \subseteq B_r(x)$ for all $r > 0$,
- there is $a > 0$ so that $\lambda(E_r) > a\lambda(B_r(x))$ for all $r > 0$.

The sets E_r need not contain x.

Theorem 6.6 (Lebesgue differentiation theorem). *Let $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$. Then, for each $x \in L_f$ and each family $\{E_r\}_{r>0}$ that shrinks nicely to x,*

$$\lim_{r\to 0} ⨍_{E_r} |f(y) - f(x)|\,dy = 0 \quad \text{and} \quad \lim_{r\to 0} ⨍_{E_r} f(y)\,dy = f(x).$$

Proof. Since $\{E_r\}_{r>0}$ shrinks nicely to x,

$$\frac{1}{\lambda(E_r)}\int_{E_r} |f(y) - f(x)|\,dy \le \frac{1}{a\lambda(B_r(x))}\int_{B_r(x)} |f(y) - f(x)|\,dy \to 0$$

as $r \to 0$, by Theorem 6.5. The second equality may be written in the form

$$\lim_{r\to 0} ⨍_{E_r} (f(y) - f(x))\,dy = 0$$

and thus is a consequence of the first. □

Corollary 6.7 (Antiderivatives). *If $f \in L^1(\mathbb{R})$ and $F(x) = \int_{-\infty}^{x} f(t)\,dt$, $x \in \mathbb{R}$, then $F'(x) = f(x)$ on every Lebesgue point of f.*

PROOF. For $E_r = [x, x+r)$, Theorem 6.6 shows that, for $x \in L_f$,

$$\lim_{r\to 0} \frac{F(x+r)-F(x)}{r} = \lim_{r\to 0} \frac{1}{r}\int_x^{x+r} f(y)\,dy = \lim_{r\to 0} \fint_{E_r} f(y)\,dy = f(x),$$

so the right derivative of F at x exists and equals $f(x)$. Similarly for the left derivative. □

6.2. Derivatives of measures

The Radon–Nikodym theorem provides an abstract notion of derivative of a complex measure with respect to a positive measure. On the measurable space $(\mathbb{R}^n, \mathfrak{B}(\mathbb{R}^n))$ we can define a pointwise derivative of a complex measure with respect to Lebesgue measure which coincides λ-a.e. with the Radon–Nikodym derivative.

Theorem 6.8. *Let μ be a complex Borel measure on $\mathbb{R}^n$ with Lebesgue decomposition $d\mu = d\nu + f\,d\lambda$. Then for λ-a.e. $x \in \mathbb{R}^n$,*

$$\lim_{r\to 0} \frac{\mu(E_r)}{\lambda(E_r)} = f(x),$$

for every family $\{E_r\}_{r>0}$ that shrinks nicely to x.

PROOF. By the Radon–Nikodym theorem 5.3, $f \in L^1(\mathbb{R}^n)$. So, by Theorem 6.6, it suffices to show that, for λ-a.e. $x \in \mathbb{R}^n$,

$$\lim_{r\to 0} \frac{\nu(E_r)}{\lambda(E_r)} = 0$$

for every family $\{E_r\}_{r>0}$ that shrinks nicely to x. We may assume without loss of generality that ν is positive and $E_r = B_r(x)$, thanks to

$$\Big|\frac{\nu(E_r)}{\lambda(E_r)}\Big| \le \frac{|\nu|(E_r)}{\lambda(E_r)} \le \frac{|\nu|(B_r(x))}{a\lambda(B_r(x))}.$$

Let A be a Borel set such that $\nu(A) = \lambda(A^c) = 0$, and set

$$F_k := \Big\{x \in A : \limsup_{r\to 0} \frac{\nu(B_r(x))}{\lambda(B_r(x))} > \frac{1}{k}\Big\}.$$

To complete the proof it is enough to show that $\lambda(F_k) = 0$ for all k.

Since ν is finite (because μ is finite), ν is regular, by Theorem 2.7. Hence, for given $\epsilon > 0$ there is an open set $U \supseteq A$ so that $\nu(U) < \epsilon$. By definition of F_k, if $x \in F_k$ then there is a ball $B_x := B_{r_x}(x) \subseteq U$ such that $\nu(B_x) > k^{-1}\lambda(B_x)$. Set $V := \bigcup_{x\in F_k} B_x$ and choose $c < \lambda(V)$. By Lemma 6.2, there exist $x_1, \dots, x_j$ so that $B_{x_1}, \dots, B_{x_j}$ are disjoint and

$$c < 3^n \sum_{i=1}^{j} \lambda(B_{x_i}) \le 3^n k \sum_{i=1}^{j} \nu(B_{x_i}) \le 3^n k\nu(V) \le 3^n k\nu(U) < 3^n k\epsilon.$$

Letting $c \to \lambda(V)$ we may conclude that $\lambda(F_k) = 0$. □

For a complex Borel measure μ on $\mathbb{R}^n$ we call

$$(D\mu)(x) = \lim_{r\to 0} \frac{\mu(B_r(x))}{\lambda(B_r(x))}$$

the **derivative** of μ at $x \in \mathbb{R}^n$, provided that the limit exists. Theorem 6.8 tells us that the derivative of a complex Borel measure exists $|$*Leb*-a.e. and equals the Radon–Nikodym derivative of the absolutely continuous part of μ with respect to λ.

6.3. The fundamental theorem of calculus

A function $f : [a, b] \to \mathbb{C}$, $a, b \in \mathbb{R}$, is said to be **absolutely continuous** on $[a, b]$, we write $f \in AC([a, b])$, if for each $\epsilon > 0$ there is a $\delta > 0$ so that for any $n \in \mathbb{N}$ and any disjoint collection of subintervals $(a_i, b_i) \subseteq [a, b]$

$$\sum_{i=1}^{n} (b_i - a_i) < \delta \quad \Longrightarrow \quad \sum_{i=1}^{n} |f(b_i) - f(a_i)| < \epsilon. \tag{6.2}$$

Obviously, $f \in AC([a, b])$ is uniformly continuous on $[a, b]$. Note that $AC([a, b])$ forms a vector space.

Lemma 6.9. *Let $I = [a, b]$ and let $f \in L^1(I)$. Then*

$$F(x) := F(a) + \int_a^x f(t)\, dt, \quad x \in I,$$

is absolutely continuous on I.

Proof. Let μ be the measure on I defined by $d\mu = f\, d\lambda$. Since $\mu \ll \lambda$ and hence $|\mu| \ll \lambda$ by Lemma 5.2, for each $\epsilon > 0$ there is a $\delta > 0$ so that $|\mu|(E) < \epsilon$ if $\lambda(E) < \delta$, by Proposition 5.5. It follows that F is absolutely continuous on I, as $F(y) - F(x) = \mu((x, y))$ for $a \le x < y \le b$. $\square$

Proposition 6.10. *For a continuous nondecreasing function $f : I = [a, b] \to \mathbb{R}$ the following are equivalent:*

1. *$f \in AC(I)$.*
2. *f maps sets of measure zero to sets of measure zero.*
3. *f is differentiable a.e. on I, $f' \in L^1(I)$, and*
$$f(x) - f(a) = \int_a^x f'(t)\, dt, \quad x \in I.$$

Property (2) is called the **Lusin (N)-property**.

Proof. (1) $\Rightarrow$ (2) Let $E \subseteq I$ be measurable and $\lambda(E) = 0$. Without loss of generality assume that $E \subseteq (a, b)$. Let $\epsilon > 0$. Then there is $\delta > 0$ such that (6.2) holds. There exists an open set V with $E \subseteq V \subseteq I$ and $\lambda(V) < \delta$, by Theorem 2.9. Let (a_i, b_i) denote the connected components of V. Then $\lambda(V) = \sum(b_i - a_i) < \delta$ and thus $\sum(f(b_i) - f(a_i)) < \epsilon$, by (6.2), where we first consider partial sums and then proceed to the limit. Since $f(E) \subseteq \bigcup[f(a_i), f(b_i)]$ and the latter is a Borel set of measure bounded by ϵ, we we may conclude that $\lambda(f(E)) = 0$ (as λ is complete).

(2) $\Rightarrow$ (3) We define

$$g(x) := x + f(x), \quad x \in I.$$

Then g has the Lusin (N)-property, since, if f maps an interval J of length ℓ to an interval of length ℓ', then $g(J)$ is an interval of length $\ell + \ell'$. We claim that g maps measurable sets $E \subseteq I$ to measurable sets. Indeed, by Corollary 2.10, $E = E_0 \cup E_1$ where $\lambda(E_0) = 0$ and E_1 is a F_σ-set. In particular, E_1 is a countable union of compact sets and, as g is continuous, so is $g(E_1)$. Since g has the Lusin (N)-property, $\lambda(g(E_0)) = 0$ and we may conclude that $g(E) = g(E_0) \cup g(E_1)$ is measurable.

We define

$$\mu(E) := \lambda(g(E)), \quad E \subseteq I \text{ measurable}.$$

Then μ is a positive bounded measure on the Lebesgue measurable sets $E \subseteq I$, since g is injective and so σ-additivity of λ transfers to μ. Moreover, $\mu \ll \lambda$, since g has the Lusin (N)-property. By the Radon–Nikodym theorem 5.3, there exists $h \in$

$L^1(I)$ such that $d\mu = h\,d\lambda$. Consequently, for $E = [a, x]$ we find $g(E) = [g(a), g(x)]$ and

$$g(x) - g(a) = \lambda(g(E)) = \mu(E) = \int_E h\,d\lambda = \int_a^x h(t)\,dt,$$

which gives

$$f(x) - f(a) = \int_a^x (h(t) - 1)\,dt, \quad x \in I.$$

By Corollary 6.7, $f' = h - 1$ a.e., and (3) is shown.

(3) $\Rightarrow$ (1) follows from Lemma 6.9. $\square$

To any function $f : I = [a, b] \to \mathbb{C}$ we associate the **total variation function**

$$T_f(x) := \sup\Big\{\sum_{i=1}^n |f(x_i) - f(x_{i-1})| : n \in \mathbb{N}, a = x_0 < \cdots < x_n = x\Big\}, \quad x \in I.$$

In general $0 \le T_f(x) \le T_f(y) \le \infty$ if $x < y$. We say that f is of **bounded variation**, and write $f \in BV(I)$, if $T_f(b) < \infty$; $V_a^b(f) = T_f(b)$ is called the **total variation** of f.

Proposition 6.11. *An absolutely continuous function $f : I = [a, b] \to \mathbb{R}$ has bounded variation. The functions T_f, $T_f + f$, and $T_f - f$ are nondecreasing and absolutely continuous on I.*

Proof. For $\epsilon = 1$ there is a $\delta > 0$ such that (6.2) holds. Set $n := \lfloor 2(b - a)/\delta \rfloor$ and divide $[a, b]$ into n intervals $[x_{i-1}, x_i]$ of equal length $(b-a)/n$. Since $(b-a)/n < \delta$, (6.2) implies that $V_{x_{i-1}}^{x_i}(f) \le 1$ and therefore

$$V_a^b(f) = \sum_{i=1}^n V_{x_{i-1}}^{x_i}(f) \le n < \infty,$$

whence f has bounded variation on I.

If $a = x_0 < \cdots < x_n = x < y \le b$ then

$$T_f(y) \ge |f(y) - f(x)| + \sum_{i=1}^n |f(x_i) - f(x_{i-1})|$$

and hence $T_f(y) \ge |f(y) - f(x)| + T_f(x)$ and, in particular,

$$T_f(y) \ge f(y) - f(x) + T_f(x) \quad \text{and} \quad T_f(y) \ge f(x) - f(y) + T_f(x).$$

Thus T_f, $T_f + f$, and $T_f - f$ are nondecreasing.

It remains to show that T_f is absolutely continuous on I. For $a \le x < y \le b$,

$$T_f(y) - T_f(x) = \sup\Big\{\sum_{i=1}^n |f(x_i) - f(x_{i-1})| : n \in \mathbb{N}, x = x_0 < \cdots < x_n = y\Big\}. \tag{6.3}$$

For $\epsilon > 0$ there is a $\delta > 0$ such that (6.2) holds. Let (a_j, b_j) be disjoint subintervals of I so that $\sum_{j=1}^N (b_j - a_j) < \delta$. Applying (6.3) to each (a_j, b_j), we get

$$\sum_{j=1}^N (T_f(b_j) - T_f(a_j)) \le \epsilon,$$

by (6.2). Thus T_f is absolutely continuous on I. $\square$

Example 6.12 (Cantor function)**.** The Cantor function f from Example 3.5 is not absolutely continuous. In fact, $f(C) = [0,1]$ and so the Lusin (N)-property fails. f is differentiable a.e., $f' = 0$ on $[0,1] \setminus C$, but $1 = f(1) - f(0) \neq \int_0^1 f'(t)\,dt = 0$. However, f has bounded variation with $V_0^1(f) = 1$.

Theorem 6.13 (Fundamental theorem of calculus)**.** *For a function $f : I = [a,b] \to \mathbb{C}$ the following are equivalent:*

(1) $f \in AC(I)$.
(2) $f(x) = f(a) + \int_a^x g(t)\,dt$ *for some* $g \in L^1(I)$.
(3) f *is differentiable a.e. in* I, $f' \in L^1(I)$, *and* $f(x) = f(a) + \int_a^x f'(t)\,dt$.

Proof. (2) $\Rightarrow$ (1) is Lemma 6.9 and (3) $\Rightarrow$ (2) is trivial.

(1) $\Rightarrow$ (3) Without loss of generality assume that f is real valued. Write

$$f = \frac{T_f + f}{2} - \frac{T_f - f}{2}.$$

By Proposition 6.11, the functions $f_\pm := (T_f \pm f)/2$ are nondecreasing and absolutely continuous, and by Proposition 6.10, $f_\pm$ satisfy (3). It follows that $f = f_+ - f_-$ satisfies (3). □

Corollary 6.14 (Integration by parts)**.** *If $f, g \in AC([a,b])$ then $fg \in AC([a,b])$, and*

$$\int_a^b f'(x)g(x)\,dx = f(b)g(b) - f(a)g(a) - \int_a^b f(x)g'(x)\,dx.$$

Proof. Let $\epsilon > 0$. Then there is $\delta > 0$ so that for any finite disjoint collection of subintervals $(a_i, b_i) \subseteq [a,b]$ with $\sum_{i=1}^n (b_i - a_i) < \delta$ we have

$$\sum_{i=1}^n |f(b_i) - f(a_i)| < \epsilon \quad \text{and} \quad \sum_{i=1}^n |g(b_i) - g(a_i)| < \epsilon.$$

Let $C := \max\{\|f\|_\infty, \|g\|_\infty\}$. Then

$$|f(b_i)g(b_i) - f(a_i)g(a_i)| \le |f(b_i)||g(b_i) - g(a_i)| + |g(a_i)||f(b_i) - f(a_i)|$$

and thus

$$\sum_{i=1}^n |f(b_i)g(b_i) - f(a_i)g(a_i)| \le 2C\epsilon.$$

Hence $fg \in AC([a,b])$. By Theorem 6.13,

$$f(b)g(b) - f(a)g(a) = \int_a^b (fg)'(x)\,dx$$

and, as f, g, and fg are differentiable a.e., the desired formula follows from the product rule. □

6.4. Rademacher's theorem

Let $A \subseteq \mathbb{R}^n$. Recall that a mapping $f : A \to \mathbb{R}^m$ is said to be **Lipschitz** if

$$\mathrm{Lip}(f) := \sup_{\substack{x,y \in A \\ x \neq y}} \frac{|f(x) - f(y)|}{|x - y|} < \infty.$$

We say that f is **locally Lipschitz** if the restriction $f|_K$ to every compact subset $K \subseteq A$ is Lipschitz.

Theorem 6.15 (Lipschitz extensions). *Let $A \subseteq \mathbb{R}^n$ and let $f : A \to \mathbb{R}^m$ be Lipschitz. Then there exists a Lipschitz extension $\tilde{f} : \mathbb{R}^n \to \mathbb{R}^m$ of f with $\operatorname{Lip}(\tilde{f}) \le \sqrt{m}\operatorname{Lip}(f)$.*

Proof. If $m = 1$ we may define

$$\tilde{f}(x) := \inf_{a \in A} \big(f(a) + \operatorname{Lip}(f)|x - a|\big).$$

Indeed, if $x \in A$ then for all $a \in A$,

$$\tilde{f}(x) \le f(x) \le f(a) + \operatorname{Lip}(f)|x - a|$$

and thus $\tilde{f}(x) = f(x)$. For $x, y \in \mathbb{R}^n$,

$$\tilde{f}(x) \le \inf_{a \in A} \big(f(a) + \operatorname{Lip}(f)(|y - a| + |x - y|)\big) = \tilde{f}(y) + \operatorname{Lip}(f)|x - y|,$$

and symmetrically $\tilde{f}(y) \le \tilde{f}(x) + \operatorname{Lip}(f)|x - y|$.

If $f = (f_1, \ldots, f_m) : A \to \mathbb{R}^m$, then $\tilde{f} := (\tilde{f}_1, \ldots, \tilde{f}_m)$ is as required, since

$$|\tilde{f}(x) - \tilde{f}(y)|^2 = \sum_{i=1}^m |\tilde{f}_i(x) - \tilde{f}_i(y)|^2 \le m \operatorname{Lip}(f)^2 |x - y|^2. \qquad \square$$

Actually, by **Kirszbraun's theorem** there is an extension $\tilde{f}$ with $\operatorname{Lip}(\tilde{f}) = \operatorname{Lip}(f)$; cf. [**4**].

We shall now prove **Rademacher's theorem** that a Lipschitz function $f : \mathbb{R}^n \to \mathbb{R}^m$ is differentiable a.e. That is at a.e. $x \in \mathbb{R}^n$ there exists a linear mapping $T : \mathbb{R}^n \to \mathbb{R}^m$ such that

$$\lim_{y \to x} \frac{|f(y) - f(x) - T(x - y)|}{|x - y|} = 0.$$

If such a linear mapping exists, it is obviously unique. We denote it by $df(x)$ and call it the **derivative** of f at x.

Theorem 6.16 (Rademacher). *Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be locally Lipschitz. Then f is differentiable a.e.*

Proof. We may assume without loss of generality that $m = 1$ and that f is Lipschitz, by Theorem 6.15, since differentiability is a local property.

For $v \in \mathbb{R}^n$ with $|v| = 1$, we consider the directional derivative of f at x,

$$d_v f(x) := \lim_{t \to 0} \frac{f(x + tv) - f(x)}{t}$$

provided this limit exists. We claim that $d_v f(x)$ exists for a.e. $x \in \mathbb{R}^n$.

We work with the **Dini derivatives** $\overline{d}_v f(x)$ and $\underline{d}_v f(x)$. Since f is continuous,

$$\overline{d}_v f(x) := \limsup_{t \to 0} \frac{f(x + tv) - f(x)}{t} = \lim_{k \to \infty} \sup_{\substack{0 < |t| < 1/k \\ t \in \mathbb{Q}}} \frac{f(x + tv) - f(x)}{t}$$

is Borel measurable, by Theorem 3.4; the same holds for

$$\underline{d}_v f(x) := \liminf_{t \to 0} \frac{f(x + tv) - f(x)}{t}.$$

Consequently, the set

$$E_v := \{x \in \mathbb{R}^n : d_v f(x) \text{ fails to exist}\} = \{x \in \mathbb{R}^n : \underline{d}_v f(x) < \overline{d}_v f(x)\}$$

is a Borel set; note that $\underline{d}_v f(x), \overline{d}_v f(x) \in \mathbb{R}$ since f is Lipschitz. For fixed $x, v \in \mathbb{R}^n$ with $|v| = 1$ then function $\mathbb{R} \ni t \mapsto f(x + tv)$ is Lipschitz, hence absolutely continuous, and thus differentiable at a.e. $t \in \mathbb{R}$, by Theorem 6.13. So $\mathcal{H}^1(E_v \cap L) = 0$ for each line L whose direction is v. By Fubini's theorem 3.27, E_v is a null set.

If we take the standard unit vectors in $\mathbb{R}^n$ for v, we may conclude that the gradient

$$\nabla f(x) := (\partial_1 f(x), \ldots, \partial_n f(x))$$

exists for a.e. $x \in \mathbb{R}^n$.

We next claim that $d_v f(x) = \nabla f(x) \cdot v$ for a.e. $x \in \mathbb{R}^n$. Let $\varphi \in C_c^\infty(\mathbb{R}^n)$. We have

$$\int_{\mathbb{R}^n} \Big(\frac{f(x+tv) - f(x)}{t}\Big)\varphi(x)\, dx = -\int_{\mathbb{R}^n} f(x)\Big(\frac{\varphi(x) - \varphi(x-tv)}{t}\Big)\, dx.$$

As $|\frac{f(x+v/k)-f(x)}{1/k}| \leq \operatorname{Lip}(f)$, the dominated convergence theorem 3.22 yields

$$\begin{aligned}
\int_{\mathbb{R}^n} d_v f(x)\varphi(x)\, dx &= -\int_{\mathbb{R}^n} f(x) d_v\varphi(x)\, dx \\
&= -\int_{\mathbb{R}^n} f(x)\nabla\varphi(x) \cdot v\, dx \\
&= -\sum_{i=1}^n \int_{\mathbb{R}^n} f(x)\partial_i\varphi(x) v_i\, dx \\
&= \sum_{i=1}^n \int_{\mathbb{R}^n} \partial_i f(x)\varphi(x) v_i\, dx \\
&= \int_{\mathbb{R}^n} \varphi(x)\nabla f(x) \cdot v\, dx,
\end{aligned}$$

where we used Fubini's theorem 3.27, the absolute continuity of f on lines, and Corollary 6.14. Since the equality holds for every $\varphi \in C_c^\infty(\mathbb{R}^n)$, we have $d_v f(x) = \nabla f(x) \cdot v$ for a.e. $x \in \mathbb{R}^n$; cf. Proposition 3.21.

Choose a countable dense subset $\{v_1, v_2, \ldots\}$ of S^{n-1}. Set

$$E_k := \{x \in \mathbb{R}^n : d_{v_k} f(x) \text{ and } \nabla f(x) \text{ exist and satisfy } d_{v_k} f(x) = \nabla f(x) \cdot v_k\}$$

and $E := \bigcap_{k=1}^\infty E_k$. Then $\lambda(E^c) = 0$.

Let us show that f is differentiable at every $x \in E$. Fix $x \in E$. For $v \in S^{n-1}$ and $t \in \mathbb{R} \setminus \{0\}$ consider

$$Q(x, v, t) := \frac{f(x+tv) - f(x)}{t} - \nabla f(x) \cdot v.$$

Then, for $w \in S^{n-1}$,

$$\begin{aligned}
|Q(x,v,t) - Q(x,w,t)| &\leq \frac{|f(x+tv) - f(x+tw)|}{|t|} + |\nabla f(x) \cdot (v-w)| \\
&\leq \operatorname{Lip}(f)|v-w| + |\nabla f(x)||v-w| \\
&\leq (\sqrt{n}+1)\operatorname{Lip}(f)|v-w|. \qquad (6.4)
\end{aligned}$$

Fix $\epsilon > 0$ and choose an integer N sufficiently large such that if $v \in S^{n-1}$ then

$$|v - v_k| \leq \frac{\epsilon}{2(\sqrt{n}+1)\operatorname{Lip}(f)} \qquad (6.5)$$

for some $k \in \{1, \ldots, N\}$. Since $Q(x, v_k, t) \to 0$ as $t \to 0$, there exists $\delta > 0$ such that

$$|Q(x, v_k, t)| < \epsilon/2 \quad \text{for } 0 < |t| < \delta,\ k = 1, \ldots, N. \qquad (6.6)$$

By (6.4), (6.5), and (6.6), for each $v \in S^{n-1}$ there exists $k \in \{1, \ldots, N\}$ such that

$$|Q(x,v,t)| \leq |Q(x,v_k,t)| + |Q(x,v,t) - Q(x,v_k,t)| < \epsilon$$

if $0 < |t| < \delta$; the same δ works for all $v \in S^{n-1}$. Let $y \in \mathbb{R}^n$, $y \neq x$. Then $y = x + tv$ with $v = (y-x)/|y-x|$ and $t = |x-y|$, and therefore

$$\frac{|f(y) - f(x) - \nabla f(x) \cdot (y-x)|}{|y-x|} = |Q\big(x, \tfrac{y-x}{|y-x|}, |x-y|\big)| \to 0$$

as $y \to x$. So f is differentiable at x with $df(x) = \nabla f(x)$. $\square$

CHAPTER 7

The dual of L^p

7.1. The dual of L^p

Let $(X, \mathfrak{S}, \mu)$ be a measure space, and let $1 \le p \le \infty$. A **linear functional** on $L^p(\mu)$ is a linear mapping $\ell : L^p(\mu) \to \mathbb{C}$. A linear functional ℓ on $L^p(\mu)$ is **continuous** if

$$\|f_k - f\|_p \to 0 \quad \text{implies} \quad \ell(f_k) \to \ell(f),$$

or equivalently,

$$|\ell(f)| \le C\|f\|_p, \quad f \in L^p(\mu),$$

for some constant $C > 0$, i.e., ℓ is **bounded**. This equivalence holds on any normed space; see Lemma A.1. To see it directly, assume that $f_k \in L^p(\mu)$ so that

$$\Big|\ell\Big(\frac{f_k}{\|f_k\|_p}\Big)\Big| = \frac{|\ell(f_k)|}{\|f_k\|_p} \to \infty.$$

Then $g_k := f_k/\|f_k\|_p \in L^p(\mu)$ satisfies $\|g_k\|_p \le 1$ and

$$\Big\|\frac{g_k}{|\ell(g_k)|}\Big\|_p \to 0,$$

whereas

$$\Big|\ell\Big(\frac{g_k}{|\ell(g_k)|}\Big)\Big| = 1.$$

The **dual** of $L^p(\mu)$ is the set of all continuous linear functionals on $L^p(\mu)$; it is denoted by $L^p(\mu)^*$. The space $L^p(\mu)^*$ is a vector space and carries a natural norm, the **operator norm**,

$$\|\ell\| := \sup\{|\ell(f)| : \|f\|_p \le 1\} = \inf\{C : |\ell(f)| \le C\|f\|_p \text{ for all } f \in L^p(\mu)\}.$$

Let q be the conjugate exponent of p. Hölder's inequality 4.2 implies that a function $g \in L^q(\mu)$ defines a continuous linear functional ℓ_g on $L^p(\mu)$ via

$$\ell_g(f) := \int_X gf \, d\mu. \tag{7.1}$$

We shall see that every continuous linear functional on $L^p(\mu)$ has the form (7.1), if $1 < p < \infty$ and if $p = 1$ provided that μ is σ-finite. We will use the following result (compare with Proposition 4.4).

Proposition 7.1. *Let $1 \le p, q \le \infty$ be conjugate exponents. Suppose that $g : X \to \mathbb{C}$ is measurable and such that*

- *$fg \in L^1(\mu)$ for all $f \in S := \{\text{simple } f : \mu(\{x : f(x) \ne 0\}) < \infty\}$,*
- *the quantity $M_q(g) := \sup\{|\int fg\, d\mu| : f \in S, \|f\|_p = 1\}$ is finite,*
- *$\{x : g(x) \ne 0\}$ is σ-finite.*

Then $g \in L^q(\mu)$ and $M_q(g) = \|g\|_q$.

Proof. We claim that a bounded measurable function f with $\|f\|_p = 1$ that vanishes outside a set F of finite measure satisfies $|\int fg\, d\mu| \le M_q(g)$. By Corollary 3.7, there are simple functions s_i converging pointwise to f and satisfying $|s_i| \le |f|$.

Since $|s_i| \le \|f\|_\infty \chi_F$ and $\chi_F g \in L^1(\mu)$, we have $|\int fg\,d\mu| = \lim_{i\to\infty} |\int s_i g\,d\mu| \le M_q(g)$, by the dominated convergence theorem 3.22.

Suppose that $q < \infty$. By assumption, $E := \{x : g(x) \neq 0\} = \bigcup_{i=1}^\infty E_i$ where $E_i \subseteq E_{i+1}$ and $\mu(E_i) < \infty$. By Corollary 3.7, there are simple functions s_i converging pointwise to g and satisfying $|s_i| \le |g|$. If we set $g_i := s_i \chi_{E_i}$, then g_i converge pointwise to g, satisfy $|g_i| \le |g|$, and g_i vanishes outside of E_i. Define

$$f_i(x) := \begin{cases} \|g_i\|_q^{1-q} |g_i(x)|^{q-1} |g(x)|^{-1} \overline{g(x)} & g(x) \neq 0 \\ 0 & g(x) = 0 \end{cases}.$$

Then $\|f_i\|_p = 1$ and by Fatou's lemma 3.17,

$$\begin{aligned} \|g\|_q &\le \liminf_{i\to\infty} \|g_i\|_q = \liminf_{i\to\infty} \int |f_i g_i|\,d\mu \\ &\le \liminf_{i\to\infty} \int |f_i g|\,d\mu = \liminf_{i\to\infty} \int f_i g\,d\mu \le M_q(g), \end{aligned}$$

by the first paragraph. Thus, $M_q(g) = \|g\|_q$ by Hölder's inequality 4.2.

Assume that $q = \infty$. For $\epsilon > 0$ set $A := \{x : |g(x)| \ge M_\infty(g) + \epsilon\}$. If $\mu(A) > 0$ there is a subset $B \subseteq A$ with $0 < \mu(B) < \infty$, since $\{x : g(x) \neq 0\}$ is σ-finite. Set $f(x) := \mu(B)^{-1}\chi_B(x) g(x)/|g(x)|$ if $g(x) \neq 0$ and $f(x) := 0$ otherwise. Then $\|f\|_1 = 1$ and $\int fg\,d\mu = \mu(B)^{-1}\int_B |g|\,d\mu \ge M_\infty(g) + \epsilon$ which contradicts the first paragraph. Thus $\|g\|_\infty = M_\infty(g)$. □

Theorem 7.2 (Dual of L^p)**.** *Let $1 \le p, q \le \infty$ be conjugate exponents. For $1 < p < \infty$, the mapping $L^q(\mu) \ni g \mapsto \ell_g \in L^p(\mu)^*$, where*

$$\ell_g(f) = \int_X gf\,d\mu,$$

is an isometric isomorphism. The same is true for $p = 1$ provided that μ is σ-finite. For $p = \infty$ it is isometric but not surjective. So in all cases

$$\|\ell_g\| = \|g\|_q. \tag{7.2}$$

Proof. Hölder's inequality 4.2 implies that $\ell_g \in L^p(\mu)^*$ if $g \in L^q(\mu)$. That $\|\ell_g\| = \|g\|_q$ follows from Proposition 4.4.

Let us show surjectivity for $1 \le p < \infty$. Let $\ell \in L^p(\mu)^*$. Assume first that $\mu(X) < \infty$. Then, for each $E \in \mathfrak{S}$, $\chi_E \in L^p(\mu)$, and

$$\nu(E) := \ell(\chi_E), \quad E \in \mathfrak{S},$$

defines a complex measure. Indeed, if $E_i \in \mathfrak{S}$ are pairwise disjoint, then

$$\Big\| \sum_{i=1}^k \chi_{E_i} - \sum_{i=1}^\infty \chi_{E_i} \Big\|_p \to 0,$$

by the dominated convergence theorem 3.22, and hence, by continuity of ℓ,

$$\nu\Big(\bigcup_{i=1}^\infty E_i\Big) = \ell\Big(\sum_{i=1}^\infty \chi_{E_i}\Big) = \sum_{i=1}^\infty \ell(\chi_{E_i}) = \sum_{i=1}^\infty \nu(E_i).$$

If $\mu(E) = 0$, then $\chi_E = 0$ in $L^p(\mu)$, and thus $\nu(E) = 0$, i.e., $\nu \ll \mu$. By the Radon–Nikodym theorem 5.3, there exists $g \in L^1(\mu)$ so that

$$\ell(\chi_E) = \nu(E) = \int_E g\,d\mu = \int_X \chi_E g\,d\mu, \quad E \in \mathfrak{S}.$$

We may conclude that $\ell(f) = \int fg\,d\mu$ holds for each simple function f and that $|\int fg\,d\mu| \le \|\ell\| \|f\|_p$. Thus, $g \in L^q(\mu)$, by Proposition 7.1. Since ℓ and ℓ_g are continuous linear functionals on $L^p(\mu)$ that coincide on the set of simple functions, Proposition 4.13 implies that $\ell(f) = \ell_g(f)$ for all functions $f \in L^p(\mu)$.

If μ is σ-finite, there are sets $X_i \subseteq X_{i+1}$ so that $X = \bigcup_{i=1}^\infty X_i$ and $\mu(X_i) < \infty$. We may identify $L^p(X_i)$ with the subspace of $L^p(X)$ of functions that vanish on X_i^c. Then $\ell \in L^p(X_i)^*$ and so, by the preceding argument, there exists $g_i \in L^q(X_i)$ with $\|g_i\|_q = \|\ell|_{L^p(X_i)}\| \leq \|\ell\|$ and so that $\ell(f) = \ell_{g_i}(f)$ for all $f \in L^p(X_i)$. We have $g_i = g_j$ μ-a.e. on X_i if $i < j$. So we may define g on X by setting $g|_{X_i} = g_i$. By the monotone convergence theorem 3.14, $\|g\|_q = \lim_{i\to\infty} \|g_i\|_q \leq \|\ell\|$, thus $g \in L^q(\mu)$. And g satisfies $\ell(f) = \ell_g(f)$ for all $f \in L^p(\mu)$, since $f\chi_{X_i} \to f$ in $L^p(\mu)$ and therefore

$$\ell(f) = \lim_{i\to\infty} \ell(f\chi_{X_i}) = \lim_{i\to\infty} \ell_{g_i}(f\chi_{X_i}) = \lim_{i\to\infty} \int_{X_i} gf \, d\mu = \ell_g(f).$$

Finally, suppose that μ is arbitrary and that $p > 1$ (consequently $q < \infty$). By the previous paragraph, for each σ-finite subset $E \subseteq X$ there is a unique $g_E \in L^q(E)$ with $\ell(f) = \ell_{g_E}(f)$ for all $f \in L^p(E)$ and $\|g_E\|_q \leq \|\ell\|$. If F is σ-finite and $F \supseteq E$, then $g_F = g_E$ μ-a.e. on E and hence $\|\ell\| \geq \|g_F\|_q \geq \|g_E\|_q$. Then

$$M := \sup\big\{\|g_E\|_q : \sigma\text{-finite } E \subseteq X\big\} \leq \|\ell\|.$$

Let E_k be a sequence of σ-finite subsets in X such that $\|g_{E_k}\|_q \to M$, and set $F := \bigcup_{k=1}^\infty E_k$. Then F is σ-finite and $\|g_F\|_q = M$. If $G \supseteq F$ is σ-finite, then

$$\int |g_F|^q \, d\mu + \int |g_{G\setminus F}|^q \, d\mu = \int |g_G|^q \, d\mu \leq M^q = \int |g_F|^q \, d\mu,$$

whence $g_{G\setminus F} = 0$ and $g_G = g_F$ μ-a.e. In particular, if $f \in L^p(\mu)$ then the set $G := F \cup \{x : f(x) \neq 0\}$ is σ-finite (as $\{x : f(x) \neq 0\} = \bigcup_{i=1}^\infty \{x : |f(x)| > 1/i\}$), and thus $\ell(f) = \int fg_G \, d\mu = \int fg_F \, d\mu$. So we may take $g = g_F$. □

Corollary 7.3. *If $1 < p < \infty$ then $L^p(\mu)$ is reflexive.*

Proof. Let q be the conjugate exponent. By Theorem 7.2, we have an isometric isomorphism $L^p(\mu)^* \cong L^q(\mu)$. So if $h \in L^p(\mu)^{**} \cong L^q(\mu)^*$ then there exists $g \in L^p(\mu)$ such that

$$h(f) = \int_X gf \, d\mu, \quad f \in L^p(\mu)^* \cong L^q(\mu).$$

Consequently, h coincides with the evaluation mapping $\mathrm{ev}_g : f \mapsto f(g)$, hence $\mathrm{ev} : L^p(\mu) \to L^p(\mu)^{**}$ is surjective, i.e., $L^p(\mu)$ is reflexive. □

The dual space of $L^\infty(\mu)$ is much larger than $L^1(\mu)$, see the following example; its description will not be given here.

Example 7.4. Consider the interval $[0,1]$ with the Lebesgue measure λ. The mapping $\mathrm{ev}_0 : f \mapsto f(0)$ is a bounded linear functional on the subspace $C([0,1])$ of $L^\infty([0,1])$. By the Hahn–Banach theorem A.2, there exists $\ell \in L^\infty([0,1])^*$ such that $\ell(f) = f(0)$ for all $f \in C([0,1])$. Let $f_k \in C([0,1])$ be given by $f_k(x) := \max\{1 - kx, 0\}$. Then $\ell(f_k) = f_k(0) = 1$ for all k and $f_k(x) \to 0$ for all $x > 0$. So for any $g \in L^1([0,1])$ we have $\int_{[0,1]} f_k g \, d\lambda \to 0$, by the dominated convergence theorem 3.22. Thus ℓ cannot be of the form ℓ_g for any L^1-function g.

7.2. Weak convergence

Let $(X, \mathfrak{S}, \mu)$ be a measure space, and let $1 \leq p \leq \infty$. A sequence of functions $f_k \in L^p(\mu)$ is said to **converge weakly** to $f \in L^p(\mu)$, and we write $f_k \rightharpoonup f$, if

$$\ell(f_k) \to \ell(f) \quad \text{for all} \quad \ell \in L^p(\mu)^*.$$

Obviously, strong convergence implies weak convergence.

Proposition 7.5. *If $f \in L^p(\mu)$ and $\ell(f) = 0$ for all $\ell \in L^p(\mu)^*$, then $f = 0$ (where we assume that μ is σ-finite in the case $p = \infty$).*

Consequently, weak limits in $L^p(\mu)$ are unique.

Proof. This follows from (7.2), in fact, if q is conjugate to p, then

$$\|f\|_p = \|\ell_f\| = \sup_{\|g\|_q \le 1} \Big| \int fg\, d\mu \Big| = \sup_{\|g\|_q \le 1} |\ell_g(f)| = 0,$$

and thus $f = 0$. □

The following is a particular case of the **Banach–Alaoglu theorem**.

Theorem 7.6. *If $1 < p < \infty$ then a bounded sequence in $L^p(\mu)$ has a weakly convergent subsequence.*

Proof. This follows from a fundamental result of functional analysis which states that a Banach space is reflexive if and only if its closed unit ball is weakly sequentially compact, cf. [**2**].

We will give a direct proof in the case that X is an open subset of $\mathbb{R}^n$ and $\mu = \lambda$ is the Lebesgue measure. Let f_i be a bounded sequence in $L^p(X)$. By extending each f_i by 0 outside X we may assume that $f_i \in L^p(\mathbb{R}^n)$. By Theorem 7.2, we may identify $L^p(\mathbb{R}^n)^*$ with $L^q(\mathbb{R}^n)$, where q is conjugate to p. By Theorem 4.23, there is a dense sequence of functions $g_j \in L^q(\mathbb{R}^n)$.

Consider the sequence of numbers $C_{i1} := \int f_i g_1\, dx$ which is bounded, by Hölder's inequality 4.2. By passing to a subsequence denoted by f_i^1 we may assume that $C_{i1} \to C_1$. Repeating this argument with f_i^1, we can pass to a further subsequence f_i^2 so that $\int f_i^2 g_2\, dx \to C_2$, and inductively we obtain a countable family of subsequences such that for the kth subsequence (and all further subsequences) $\int f_i^k g_k\, dx \to C_k$ as $i \to \infty$. Then the sequence defined by $F_j := f_j^j$ satisfies $\int F_j g_k\, dx \to C_k$ as $j \to \infty$ for all k.

If $g \in L^q(\mathbb{R}^n)$ and $\epsilon > 0$, then $\|g - g_k\|_q \le \epsilon$ for some k. Thus

$$\begin{aligned} \Big| \int F_j g\, dx - \int F_i g\, dx \Big| &\le \int |F_j||g - g_k|\, dx + \int |F_i||g_k - g|\, dx \\ &\quad + \Big| \int F_j g_k\, dx - \int F_i g_k\, dx \Big| \\ &\le 2\epsilon \sup_j \|F_j\|_p + \epsilon, \end{aligned}$$

for sufficiently large i and j. Hence the limit $\lim_{j\to\infty} \int F_j g\, dx$ exists. Setting $\ell(g) := \lim_{j\to\infty} \int F_j g\, dx$ we obtain a bounded linear functional on $L^q(\mathbb{R}^n)$. By Theorem 7.2, there exists $f \in L^p(\mathbb{R}^n)$ such that $\ell(g) = \int fg\, dx$ for all $g \in L^q(\mathbb{R}^n)$. The proof is complete. □

7.3. Interpolation theorems

We have seen in Proposition 4.7 that $L^p(\mu) \cap L^r(\mu) \subseteq L^q(\mu) \subseteq L^p(\mu) + L^r(\mu)$ provided that $1 \le p < q < r \le \infty$, and the first inclusion is bounded. Now we investigate the question whether a linear operator which is bounded on $L^p(\mu)$ and $L^r(\mu)$ is also bounded on $L^q(\mu)$. We need a preliminary lemma from complex analysis.

Lemma 7.7 (Three lines lemma). *Let $S := \{z \in \mathbb{C} : 0 \leq \operatorname{Re} z \leq 1\}$ and let $f : S \to \mathbb{C}$ be bounded continuous and holomorphic in the interior of the strip S. If $|f(z)| \leq M_0$ for $\operatorname{Re} z = 0$ and $|f(z)| \leq M_1$ for $\operatorname{Re} z = 1$, then $|f(z)| \leq M_0^{1-t} M_1^t$ for $\operatorname{Re} z = t$ and $0 < t < 1$.*

PROOF. For $\epsilon > 0$ define $f_\epsilon(z) := f(z) M_0^{z-1} M_1^{-z} \exp(\epsilon z(z-1))$. Then f_ϵ satisfies the assumptions with M_0 and M_1 replaced by 1. Moreover, $|f_\epsilon(z)| \to 0$ as $|\operatorname{Im} z| \to \infty$ (uniformly for $0 \leq \operatorname{Re} z \leq 1$). So $|f_\epsilon(z)| \leq 1$ for z on the boundary of a rectangle $\{z : 0 \leq \operatorname{Re} z \leq 1,\ |\operatorname{Im} z| < A\}$. The maximum principle implies that $|f_\epsilon(z)| \leq 1$ for $z \in S$. Thus, for $\operatorname{Re} z = t$,

$$|f(z)| M_0^{t-1} M_1^{-t} = \lim_{\epsilon \to 0} |f_\epsilon(z)| \leq 1,$$

and the lemma is proved. □

We are ready to prove the **Riesz–Thorin interpolation theorem** which shows that the answer to the above question is yes.

Theorem 7.8 (Riesz–Thorin). *Let $(X, \mathfrak{S}, \mu)$ and $(Y, \mathfrak{T}, \nu)$ be measure spaces and let $p_0, p_1, q_0, q_1 \in [1, \infty]$. If $q_0 = q_1 = \infty$ we also assume that ν is σ-finite. Let p_t, q_t, $0 < t < 1$, be defined by*

$$\frac{1}{p_t} = \frac{1-t}{p_0} + \frac{t}{p_1}, \qquad \frac{1}{q_t} = \frac{1-t}{q_0} + \frac{t}{q_1}.$$

If $T : L^{p_0}(\mu) + L^{p_1}(\mu) \to L^{q_0}(\nu) + L^{q_1}(\nu)$ is a linear mapping such that

$$\|Tf\|_{q_0} \leq M_0 \|f\|_{p_0}, \qquad \textit{for all } f \in L^{p_0}(\mu),$$
$$\|Tf\|_{q_1} \leq M_1 \|f\|_{p_1}, \qquad \textit{for all } f \in L^{p_1}(\mu),$$

then for all $0 < t < 1$,

$$\|Tf\|_{q_t} \leq M_0^{1-t} M_1^t \|f\|_{p_t}, \qquad \textit{for all } f \in L^{p_t}(\mu). \tag{7.3}$$

PROOF. If $p_0 = p_1 = p$, then by Proposition 4.7,

$$\|Tf\|_{q_t} \leq \|Tf\|_{q_0}^{1-t} \|Tf\|_{q_1}^t \leq M_0^{1-t} M_1^t \|f\|_p$$

for all $f \in L^p(\mu)$, and we are done. So we may assume that $p_0 \neq p_1$, and thus $p_t < \infty$, for all $0 < t < 1$.

Let S_X be the class of simple functions s on X with $\mu(\{x : s(x) \neq 0\}) < \infty$, and S_Y the class of simple functions s on Y with $\nu(\{x : s(x) \neq 0\}) < \infty$. We shall show that (7.3) holds for all $f \in S_X$. Since S_X is dense in $L^p(\mu)$, by Proposition 4.13, we may conclude that $T|_{S_X}$ has a unique extension $\widetilde{T}$ to $L^{p_t}(\mu)$ satisfying the same estimate there. It remains to prove that $T = \widetilde{T}$ on $L^{p_t}(\mu)$. For $f \in L^{p_t}(\mu)$ choose a sequence $f_n \in S_X$ with $|f_n| \leq |f|$ and $f_n \to f$ pointwise; cf. Corollary 3.7. Set $E := \{x : |f(x)| > 1\}$, $g = \chi_E f$, and $g_n = \chi_E f_n$. If $p_0 < p_1$ (which we may assume without loss of generality), then $g \in L^{p_0}(\mu)$ and $f - g \in L^{p_1}(\mu)$ (cf. Proposition 4.7) and, by the dominated convergence theorem 3.22, $\|f_n - f\|_{p_t} \to 0$, $\|g_n - g\|_{p_0} \to 0$, and $\|(f_n - g_n) - (f - g)\|_{p_1} \to 0$. It follows that $\|Tg_n - Tg\|_{q_0} \to 0$ and $\|T(f_n - g_n) - T(f - g)\|_{q_1} \to 0$. By passing the a subsequence we get $Tg_n \to Tg$ ν-a.e. and $T(f_n - g_n) \to T(f - g)$ ν-a.e., by Corollary 4.11, and may conclude that $Tf_n \to Tf$ ν-a.e. By Fatou's lemma 3.17,

$$\|Tf\|_{q_t} \leq \liminf \|Tf_n\|_{q_t} \leq \liminf M_0^{1-t} M_1^t \|f_n\|_{p_t} = M_0^{1-t} M_1^t \|f\|_{p_t}$$

and (7.3) is proved.

Let us show that (7.3) holds for all $f \in S_X$. By Proposition 7.1,

$$\|Tf\|_{q_t} = \sup \Big\{ \Big| \int_Y (Tf) g \, d\nu \Big| : g \in S_Y,\ \|g\|_{q_t'} = 1 \Big\},$$

where q_t' is the conjugate exponent to q_t; the set $\{y : Tf(y) \neq 0\}$ is σ-finite either since $Tf \in L^{q_0}(\nu) \cap L^{q_1}(\nu)$ or, if $q_0 = q_1 = \infty$, by assumption. We may assume that $f \neq 0$ and that $\|f\|_{p_t} = 1$, by rescaling. Thus in order to show that (7.3) holds for all $f \in S_X$ it suffices to prove the following claim.

Claim: If $f \in S_X$, $\|f\|_{p_t} = 1$, then

$$\Big| \int_Y (Tf) g \, d\nu \Big| \leq M_0^{1-t} M_1^t, \qquad \text{for } g \in S_Y, \ \|g\|_{q_t'} = 1.$$

Let $f = \sum_{j=1}^m a_j \chi_{E_j}$ and $g = \sum_{k=1}^n b_k \chi_{F_k}$ be canonical representations, and write $a_j = |a_j| e^{i\varphi_j}$ and $b_k = |b_k| e^{i\psi_k}$. Define

$$\pi(z) := \frac{1-z}{p_0} + \frac{z}{p_1}, \quad \tau(z) = \frac{1-z}{q_0} + \frac{z}{q_1}, \quad z \in \mathbb{C},$$

so that $\pi(t) = 1/p_t$ and $\tau(t) = 1/q_t$ for $0 < t < 1$. Fix t and set

$$f_z := \sum_{j=1}^m |a_j|^{\frac{\pi(z)}{\pi(t)}} e^{i\varphi_j} \chi_{E_j};$$

note that $\pi(t) > 0$. If $\tau(t) < 1$ set

$$g_z := \sum_{k=1}^n |b_k|^{\frac{1-\tau(z)}{1-\tau(t)}} e^{i\psi_k} \chi_{F_k},$$

otherwise, if $\tau(t) = 1$, set $g_z = g$ for all z. Assume that $\tau(t) < 1$ (the case $\tau(t) = 1$ follows similarly). Consider the entire function

$$\Phi(z) := \int_Y (Tf_z) g_z \, d\nu = \sum_{j=1}^m \sum_{k=1}^n |a_j|^{\frac{\pi(z)}{\pi(t)}} |b_k|^{\frac{1-\tau(z)}{1-\tau(t)}} e^{i(\varphi_j + \psi_k)} \int_Y (T\chi_{E_j}) \chi_{F_k} \, d\nu$$

which is bounded on the strip $\{z \in \mathbb{C} : 0 \leq \operatorname{Re} z \leq 1\}$. By the three lines lemma 7.7, the claim follows if we show that $|\Phi(z)| \leq M_0$ for $\operatorname{Re} z = 0$ and $|\Phi(z)| \leq M_1$ for $\operatorname{Re} z = 1$. By Hölder's inequality 4.2, for $s \in \mathbb{R}$,

$$|\Phi(is)| \leq \|Tf_{is}\|_{q_0} \|g_{is}\|_{q_0'} \leq M_0 \|f_{is}\|_{p_0} \|g_{is}\|_{q_0'}.$$

Since $\pi(is) := 1/p_0 + is(1/p_1 - 1/p_0)$ and $1 - \tau(is) = (1 - 1/q_0) + is(1/q_1 - 1/q_0)$,

$$|f_{is}| = \sum_{j=1}^m |a_j|^{\frac{\operatorname{Re}(\pi(is))}{\pi(t)}} \chi_{E_j} = |f|^{\frac{\operatorname{Re}(\pi(is))}{\pi(t)}} = |f|^{\frac{p_t}{p_0}},$$

$$|g_{is}| = \sum_{k=1}^n |b_j|^{\frac{\operatorname{Re}(1-\tau(is))}{1-\tau(t)}} \chi_{F_k} = |g|^{\frac{\operatorname{Re}(1-\tau(is))}{1-\tau(t)}} = |f|^{\frac{q_t'}{q_0'}},$$

and hence $\|f_{is}\|_{p_0}^{p_0} = \|f\|_{p_t}^{p_t} = 1$ and $\|g_{is}\|_{q_0'}^{q_0'} = \|g\|_{q_t'}^{q_t'} = 1$. Thus, $|\Phi(z)| \leq M_0$ for $\operatorname{Re} z = 0$. A similar computation shows $|\Phi(z)| \leq M_1$ for $\operatorname{Re} z = 1$. The proof is complete. □

The second fundamental interpolation result is the **Marcinkiewicz interpolation theorem**. Let T be a mapping from some vector space $\mathcal{F}$ of measurable functions on $(X, \mathfrak{S}, \mu)$ to the space of measurable functions on $(Y, \mathfrak{T}, \nu)$. Then T is called **sublinear** if for all $f, g \in \mathcal{F}$ and $c > 0$, $|T(cf)| = c|Tf|$ and $|T(f+g)| \leq |Tf| + |Tg|$.

Theorem 7.9 (Marcinkiewicz)**.** *Let $(X, \mathfrak{S}, \mu)$ and $(Y, \mathfrak{T}, \nu)$ be measure spaces and let $p_0, p_1, q_0, q_1 \in [1, \infty]$ satisfy $p_0 \leq q_0$, $p_1 \leq q_1$, and $q_0 \neq q_1$. Let p_t, q_t, $0 < t < 1$, be defined by*

$$\frac{1}{p_t} = \frac{1-t}{p_0} + \frac{t}{p_1}, \quad \frac{1}{q_t} = \frac{1-t}{q_0} + \frac{t}{q_1}. \tag{7.4}$$

If T is a sublinear mapping on $L^{p_0}(\mu)+L^{p_1}(\mu)$ to the space of measurable functions on Y such that

$$\begin{aligned} \|Tf\|_{q_0,\infty} &\le M_0\|f\|_{p_0}, \quad \text{for all } f \in L^{p_0}(\mu), \\ \|Tf\|_{q_1,\infty} &\le M_1\|f\|_{p_1}, \quad \text{for all } f \in L^{p_1}(\mu), \end{aligned} \tag{7.5}$$

then for all $0 < t < 1$,

$$\|Tf\|_{q_t} \le M_t\|f\|_{p_t}, \quad \text{for all } f \in L^{p_t}(\mu), \tag{7.6}$$

where M_t depends only on M_i, p_i, q_i, t, for $i = 0, 1$.

In other words, if the sublinear mapping T is weak type (p_0, q_0) and (p_1, q_1) then T is **strong type** (p_t, q_t), i.e., T maps $L^{p_t}(\mu)$ to $L^{q_t}(\nu)$ and $\|Tf\|_{q_t} \le C\|f\|_{p_t}$ holds for all $f \in L^{p_t}(\mu)$.

In the proof we make use of the following simple lemma.

Lemma 7.10. *Let f be measurable and let $A > 0$. For $E_A = \{x \in X : |f(x)| > A\}$ set $h_A := f\chi_{E_A^c} + A(\operatorname{sgn} f)\chi_{E_A}$ and $g_A = f - h_A$. Then $d_{g_A}(\alpha) = d_f(\alpha + A)$ and $d_{h_A}(\alpha) = d_f(\alpha)$ if $\alpha < A$ and $d_{h_A}(\alpha) = 0$ if $\alpha \ge A$.*

Proof. Note that $g_A = (\operatorname{sgn} f)(|f| - A)\chi_{E_A}$ and thus $|g_A(x)| > \alpha$ if and only if $|f(x)| > \alpha + A$. This implies $d_{g_A}(\alpha) = d_f(\alpha + A)$. The second statement is obvious. ☐

Proof of Theorem 7.9. Assume that $p_0 = p_1 = p$ (and hence $p \ne \infty$) and (say) $q_0 < q_1 < \infty$. Then (7.5) implies

$$d_{Tf}(\beta) \le \Big(\frac{M_0\|f\|_p}{\beta}\Big)^{q_0}, \quad d_{Tf}(\beta) \le \Big(\frac{M_1\|f\|_p}{\beta}\Big)^{q_1}$$

and, by Proposition 4.29 (and Remark 4.30), with $A = \|f\|_p$ and $q = q_t$,

$$\begin{aligned} \|Tf\|_q^q &= q\int_0^\infty \beta^{q-1} d_{Tf}(\beta)\,d\beta \\ &= q\int_0^A \beta^{q-1} d_{Tf}(\beta)\,d\beta + q\int_A^\infty \beta^{q-1} d_{Tf}(\beta)\,d\beta \\ &\le qM_0^{q_0}\|f\|_p^{q_0}\int_0^A \beta^{q-q_0-1}\,d\beta + qM_1^{q_1}\|f\|_p^{q_1}\int_A^\infty \beta^{q-q_1-1}\,d\beta \\ &= \frac{qM_0^{q_0}}{q-q_0}\|f\|_p^q + \frac{qM_1^{q_1}}{q_1-q}\|f\|_p^q \end{aligned}$$

which implies the statement. If $q_1 = \infty$ then $\|Tf\|_\infty \le M_1\|f\|_p$ and thus $d_{Tf}(\beta) = 0$ if $\beta > M_1\|f\|_p$. So it suffices to repeat the computation with $A = M_1\|f\|_p$.

Let us now consider the case $p_0 < p_1$ and $q_0 < \infty$ and $q_1 < \infty$. Let $p = p_t$, $q = q_t$ and $f \in L^p(\mu)$. Then, with the notation of Lemma 7.10,

$$\begin{aligned} \int |g_A|^{p_0}\,d\mu &= p_0\int_0^\infty \alpha^{p_0-1} d_{g_A}(\alpha)\,d\alpha = p_0\int_0^\infty \alpha^{p_0-1} d_f(\alpha + A)\,d\alpha \\ &= p_0\int_A^\infty (\alpha - A)^{p_0-1} d_f(\alpha)\,d\alpha \le p_0\int_A^\infty \alpha^{p_0-1} d_f(\alpha)\,d\alpha, \\ \int |h_A|^{p_1}\,d\mu &= p_1\int_0^\infty \alpha^{p_1-1} d_{h_A}(\alpha)\,d\alpha = p_1\int_0^A \alpha^{p_1-1} d_f(\alpha)\,d\alpha, \end{aligned}$$

by Proposition 4.29 (and Remark 4.30). Moreover,

$$\int |Tf|^q\,d\nu = q\int_0^\infty \beta^{q-1} d_{Tf}(\beta)\,d\beta = 2^q q\int_0^\infty \beta^{q-1} d_{Tf}(2\beta)\,d\beta.$$

Since T is sublinear,

$$d_{Tf}(2\beta) \le d_{Tg_A}(\beta) + d_{Th_A}(\beta)$$

for all $\beta, A > 0$, by Lemma 4.28. Let us apply this for $A = \beta^r$, where

$$r := \frac{p_0(q_0 - q)}{q_0(p_0 - p)} = \frac{p_1(q_1 - q)}{q_1(p_1 - p)},$$

by (7.4). By assumption (7.5),

$$\begin{aligned}
\beta^{q_0} d_{Tg_A}(\beta) &\le (\|Tg_A\|_{q_0,\infty})^{q_0} \le (M_0\|g_A\|_{p_0})^{q_0}, \\
\beta^{q_1} d_{Th_A}(\beta) &\le (\|Th_A\|_{q_1,\infty})^{q_1} \le (M_1\|h_A\|_{p_1})^{q_1},
\end{aligned}$$

and thus

$$\begin{aligned}
\|Tf\|_q^q &\le 2^q q \int_0^\infty \beta^{q-1}\big(d_{Tg_A}(\beta) + d_{Th_A}(\beta)\big)\, d\beta \\
&\le 2^q q \int_0^\infty \beta^{q-1}\big((M_0\|g_A\|_{p_0}/\beta)^{q_0} + (M_1\|h_A\|_{p_1}/\beta)^{q_1}\big)\, d\beta \\
&\le 2^q q M_0^{q_0} p_0^{q_0/p_0} \int_0^\infty \beta^{q-q_0-1}\Big(\int_{\beta^r}^\infty \alpha^{p_0-1} d_f(\alpha)\, d\alpha\Big)^{q_0/p_0} d\beta \\
&\qquad + 2^q q M_1^{q_1} p_1^{q_1/p_1} \int_0^\infty \beta^{q-q_1-1}\Big(\int_0^{\beta^r} \alpha^{p_1-1} d_f(\alpha)\, d\alpha\Big)^{q_1/p_1} d\beta \\
&= \sum_{i=0}^1 2^q q M_i^{q_i} p_i^{q_i/p_i} \int_0^\infty \Big(\int_0^\infty \varphi_i(\alpha,\beta)\, d\alpha\Big)^{q_i/p_i} d\beta,
\end{aligned}$$

where

$$\begin{aligned}
\varphi_i(\alpha,\beta) &:= \chi_i(\alpha,\beta)\alpha^{p_i-1} d_f(\alpha)\beta^{(q-q_i-1)p_i/q_i}, \\
\chi_0 &:= \chi_{\{(\alpha,\beta):\alpha>\beta^r\}}, \quad \chi_1 := \chi_{\{(\alpha,\beta):\alpha<\beta^r\}}.
\end{aligned}$$

Since $q_i/p_i \ge 1$, Minkowski's integral inequality 4.5 gives

$$\int_0^\infty \Big(\int_0^\infty \varphi_i(\alpha,\beta)\, d\alpha\Big)^{q_i/p_i} d\beta \le \Big(\int_0^\infty \Big(\int_0^\infty \varphi_i(\alpha,\beta)^{q_i/p_i}\, d\beta\Big)^{p_i/q_i} d\alpha\Big)^{q_i/p_i}.$$

If $q_1 > q_0$, then $q - q_0 > 0$ and $r > 0$, and $\alpha > \beta^r$ if and only if $\alpha^{1/r} > \beta$, whence

$$\begin{aligned}
&\int_0^\infty \Big(\int_0^\infty \varphi_0(\alpha,\beta)^{q_0/p_0}\, d\beta\Big)^{p_0/q_0} d\alpha \\
&\quad = \int_0^\infty \Big(\int_0^{\alpha^{1/r}} \beta^{q-q_0-1}\, d\beta\Big)^{p_0/q_0} \alpha^{p_0-1} d_f(\alpha)\, d\alpha \\
&\quad = (q-q_0)^{-p_0/q_0} \int_0^\infty \alpha^{p_0-1+p_0(q-q_0)/(q_0 r)} d_f(\alpha)\, d\alpha \\
&\quad = (q-q_0)^{-p_0/q_0} \int_0^\infty \alpha^{p-1} d_f(\alpha)\, d\alpha \\
&\quad = |q-q_0|^{-p_0/q_0} p^{-1} \|f\|_p^p.
\end{aligned}$$

If $q_1 < q_0$, then $q - q_0 < 0$ and $r < 0$, and $\alpha > \beta^r$ if and only if $\alpha^{1/r} < \beta$, whence

$$\begin{aligned}
&\int_0^\infty \Big(\int_0^\infty \varphi_0(\alpha,\beta)^{q_0/p_0}\, d\beta\Big)^{p_0/q_0} d\alpha \\
&\quad = \int_0^\infty \Big(\int_{\alpha^{1/r}}^\infty \beta^{q-q_0-1}\, d\beta\Big)^{p_0/q_0} \alpha^{p_0-1} d_f(\alpha)\, d\alpha \\
&\quad = (q_0-q)^{-p_0/q_0} \int_0^\infty \alpha^{p_0-1+p_0(q-q_0)/(q_0 r)} d_f(\alpha)\, d\alpha
\end{aligned}$$

$$
\begin{aligned}
&= (q_0 - q)^{-p_0/q_0} \int_0^\infty \alpha^{p-1} d_f(\alpha)\, d\alpha \\
&= |q - q_0|^{-p_0/q_0} p^{-1} \|f\|_p^p.
\end{aligned}
$$

Similarly,

$$\int_0^\infty \Big(\int_0^\infty \varphi_1(\alpha,\beta)^{q_1/p_1}\, d\beta \Big)^{p_1/q_1} d\alpha \le |q - q_1|^{-p_1/q_1} p^{-1} \|f\|_p^p.$$

So for all $f \in L^p(\mu)$ with $\|f\|_p = 1$,

$$\|Tf\|_q \le 2q^{1/q} \Big(\sum_{i=0}^{1} \frac{M_i^{q_i} (p_i/p)^{q_i/p_i}}{|q - q_i|} \Big)^{1/q} =: M_p.$$

Since T is sublinear, in particular, $|T(cf)| = c|Tf|$ if $c > 0$, (7.6) follows.

In the remaining cases $q_0 = \infty$ or $q_1 = \infty$ we indicate how to modify the arguments.

If $p_1 = q_1 = \infty$ (hence $p_0 \le q_0 < \infty$), use $A = \beta/M_1$. Then $\|Th_A\|_\infty \le M_1 \|h_A\|_\infty \le \beta$ and thus $d_{Th_A}(\beta) = 0$.

If $p_0 < p_1 < \infty$ and $q_0 < q_1 = \infty$, use $A = (\beta/B)^r$ with $B = M_1 (p_1 \|f\|_p^p / p)^{1/p_1}$ and $r = p_1/(p_1 - p)$.

Similarly, if $p_0 < p_1 < \infty$ and $q_1 < q_0 = \infty$, use $A = (\beta/B)^r$ with B chosen such that $d_{Tg_A}(\beta) = 0$. $\square$

Let us apply the Marcinkiewicz interpolation theorem 7.9 to the Hardy–Littlewood maximal operator M defined by

$$Mf(x) = \sup_{r>0} \fint_{B_r(x)} |f(y)|\, dy, \quad f \in L^1_{\mathrm{loc}}(\mathbb{R}^n).$$

Corollary 7.11. *There is a constant $C > 0$ such that, for $1 < p < \infty$,*

$$\|Mf\|_p \le C \frac{p}{p-1} \|f\|_p, \quad f \in L^p(\mathbb{R}^n). \tag{7.7}$$

Proof. Clearly, $\|Mf\|_\infty \le \|f\|_\infty$ for $f \in L^\infty(\mathbb{R}^n)$, and by Theorem 6.3, $\|Mf\|_{1,\infty} \le C\|f\|_1$ for $f \in L^1(\mathbb{R}^n)$. Obviously, M is sublinear. Then (7.7) follows from the Marcinkiewicz interpolation theorem 7.9; the constant $Cp/(p-1)$ results from an inspection of the proof of Theorem 7.9. $\square$

CHAPTER 8

The Fourier transform

8.1. The Fourier transform on L^1

For a function $f \in L^1(\mathbb{R}^n)$ the **Fourier transform** $\widehat{f}$ is defined by

$$\widehat{f}(\xi) := \int_{\mathbb{R}^n} f(x) e^{-2\pi i \xi \cdot x}\, dx, \quad \xi \in \mathbb{R}^n, \tag{8.1}$$

where $\xi \cdot x := \xi_1 x_1 + \cdots + \xi_n x_n$; we shall also write $\mathscr{F} f = \widehat{f}$. It follows from Theorem 3.37 that $\widehat{f}$ is continuous on $\mathbb{R}^n$. Moreover, as

$$|\widehat{f}(\xi)| \le \int_{\mathbb{R}^n} |f(x)|\, dx = \|f\|_1,$$

$\widehat{f}$ is bounded and satisfies

$$\|\widehat{f}\|_\infty \le \|f\|_1. \tag{8.2}$$

Note that we have equality in (8.2) if $f \ge 0$:

$$|\widehat{f}(0)| = \int_{\mathbb{R}^n} f(x)\, dx = \|f\|_1 = \|\widehat{f}\|_\infty.$$

Next we collect elementary properties of the Fourier transform. For $y, \eta \in \mathbb{R}^n$ we consider the translation operator $T_y f(x) := f(x - y)$, $x \in \mathbb{R}^n$, cf. (4.6), and the **modulation** operator,

$$M_\eta f(x) := e^{2\pi i \eta \cdot x} f(x), \quad x \in \mathbb{R}^n. \tag{8.3}$$

We have the commutation relations

$$T_y M_\eta = e^{-2\pi i \eta \cdot y} M_\eta T_y.$$

Recall that $C_0(\mathbb{R}^n)$ denotes the space of all continuous functions $f : \mathbb{R}^n \to \mathbb{C}$ so that $|f(x)| \to 0$ as $|x| \to \infty$. Note that $C_0(\mathbb{R}^n)$ is the closure of $C_c(\mathbb{R}^n)$ with respect to $\|\ \|_\infty$. Indeed, if $f_i \in C_c(\mathbb{R}^n)$ converge uniformly to $f \in C(\mathbb{R}^n)$, then for each $\epsilon > 0$ there is $i \in \mathbb{N}$ such that $\|f_i - f\|_\infty < \epsilon$, and hence $|f(x)| < \epsilon$ if $x \notin \operatorname{supp} f_i$, i.e., $f \in C_0(\mathbb{R}^n)$. Conversely, for $f \in C_0(\mathbb{R}^n)$ and each positive integer consider the compact set $K_i := \{x : |f(x)| \ge 1/i\}$. Choose $g_i \in C_c(\mathbb{R}^n)$ so that $0 \le g_i \le 1$ and $g_i|_{K_i} = 1$. Then $f_i := f g_i \in C_c(\mathbb{R}^n)$ satisfies $\|f_i - f\|_\infty = \|f(g_i - 1)\|_\infty \le 1/i$.

Lemma 8.1. *Let $f, g \in L^1(\mathbb{R}^n)$, $y, \eta \in \mathbb{R}^n$, and $a > 0$. Then:*

(1) $(T_y f)\widehat{\ } = M_{-y} \widehat{f}$ *and* $(M_\eta f)\widehat{\ } = T_\eta \widehat{f}$.
(2) $(f(ax))\widehat{\ }(\xi) = a^{-n} \widehat{f}(a^{-1}\xi)$ *and* $(f(-x))\widehat{\ }(\xi) = \widehat{f}(-\xi)$.
(3) $(f * g)\widehat{\ } = \widehat{f}\widehat{g}$.
(4) *If $x \mapsto x^\alpha f(x)$ is in $L^1(\mathbb{R}^n)$ for all $|\alpha| \le k$, then $\widehat{f} \in C^k(\mathbb{R}^n)$ and*

$$\partial^\alpha \widehat{f} = ((-2\pi i x)^\alpha f(x))\widehat{\ }.$$

(5) *If $f \in C^k(\mathbb{R}^n)$, $\partial^\alpha f \in L^1(\mathbb{R}^n)$ for all $|\alpha| \le k$, and $\partial^\alpha f \in C_0(\mathbb{R}^n)$ for all $|\alpha| \le k - 1$, then*

$$(\partial^\alpha f)\widehat{\ }(\xi) = (2\pi i \xi)^\alpha \widehat{f}(\xi).$$

(6) $\int \widehat{f} g\,dx = \int f\widehat{g}\,dx$.

Proof. (1) We have

$$(T_y f)\widehat{\ }(\xi) = \int_{\mathbb{R}^n} f(x-y)e^{-2\pi i\xi\cdot x}\,dx = \int_{\mathbb{R}^n} f(x)e^{-2\pi i\xi\cdot(x+y)}\,dx = e^{-2\pi i\xi\cdot y}\widehat{f}(\xi)$$

and

$$(M_\eta f)\widehat{\ }(\xi) = \int_{\mathbb{R}^n} f(x)e^{-2\pi i(\xi-\eta)\cdot x}\,dx = \widehat{f}(\xi-\eta) = T_\eta\widehat{f}(\xi).$$

(2) Both assertions follow from

$$(f(ax))\widehat{\ }(\xi) = \int_{\mathbb{R}^n} f(ax)e^{-2\pi i\xi\cdot x}\,dx = |a|^{-n}\int_{\mathbb{R}^n} f(x)e^{-2\pi i a^{-1}\xi\cdot x}\,dx = |a|^{-n}\widehat{f}(a^{-1}\xi),$$

where either $a > 0$ or $a = -1$.

(3) By Young's inequality 4.15, $f * g \in L^1(\mathbb{R}^n)$ and so, by Fubini's theorem 3.27,

$$\begin{aligned}(f*g)\widehat{\ }(\xi) &= \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} f(x-y)g(y)e^{-2\pi i\xi\cdot x}\,dy\,dx\\ &= \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} f(x-y)e^{-2\pi i\xi\cdot(x-y)}g(y)e^{-2\pi i\xi\cdot y}\,dx\,dy\\ &= \widehat{f}(\xi)\int_{\mathbb{R}^n} g(y)e^{-2\pi i\xi\cdot y}\,dy = \widehat{f}(\xi)\widehat{g}(\xi).\end{aligned}$$

(4) By Theorem 3.38,

$$\partial^\alpha\widehat{f}(\xi) = \int_{\mathbb{R}^n}(-2\pi i x)^\alpha f(x)e^{-2\pi i\xi\cdot x}\,dx = ((-2\pi i x)^\alpha f(x))\widehat{\ }(\xi).$$

(5) By partial integration, cf. Corollary 6.14,

$$(\partial^\alpha f)\widehat{\ }(\xi) = \int_{\mathbb{R}^n}\partial^\alpha f(x)e^{-2\pi i\xi\cdot x}\,dx = (2\pi i\xi)^\alpha\int_{\mathbb{R}^n} f(x)e^{-2\pi i\xi\cdot x}\,dx = (2\pi i\xi)^\alpha\widehat{f}(\xi).$$

(6) Both integrals equal $\int\int f(x)g(\xi)e^{-2\pi i x\cdot\xi}\,dx\,d\xi$, by Fubini's theorem 3.27. The proof is complete. ☐

Let $\mathcal{S}(\mathbb{R}^n)$ denote the **Schwartz space** of rapidly decreasing functions:

$$\mathcal{S}(\mathbb{R}^n) := \{f \in C^\infty(\mathbb{R}^n) : \|f\|_{k,\alpha} < \infty \text{ for all } k \in \mathbb{N}, \alpha \in \mathbb{N}^n\},$$

where

$$\|f\|_{k,\alpha} := \sup_{x\in\mathbb{R}^n}(1+|x|)^k|\partial^\alpha f(x)|.$$

Lemma 8.2. *We have:*

(1) *If $f \in \mathcal{S}(\mathbb{R}^n)$ then $\partial^\alpha f \in L^p(\mathbb{R}^n)$ for all $\alpha \in \mathbb{N}^n$ and all $1 \le p \le \infty$.*
(2) *Let $f \in C^\infty(\mathbb{R}^n)$. Then $f \in \mathcal{S}(\mathbb{R}^n)$ if and only if $x^\beta\partial^\alpha f(x)$ is bounded for all α, β if and only if $\partial^\alpha(x^\beta f(x))$ is bounded for all α, β.*
(3) *$\mathcal{S}(\mathbb{R}^n)$ is a Fréchet space with the topology defined by the seminorms $\|\ \|_{k,\alpha}$.*

Proof. (1) If $f \in \mathcal{S}(\mathbb{R}^n)$ then $|\partial^\alpha f(x)| \le C(k)(1+|x|)^{-k}$ for all k, and $(1+|x|)^{-k} \in L^p(\mathbb{R}^n)$ if $k > n/p$, cf. (3.7).

(2) Clearly, $|x^\beta| \le (1+|x|)^k$ if $|\beta| \le k$. On the other hand, $\sum_{i=1}^n |x_i|^k$ is strictly positive on the unit sphere $|x| = 1$, thus it has a positive minimum m there. We may conclude that $\sum_{i=1}^n |x_i|^k \ge m|x|^k$, by homogeneity of both sides. Then

$$(1+|x|)^k \le 2^k\max\{1, |x|^k\} \le 2^k(1+|x|^k)$$

$$\leq 2^k\Big(1+m^{-1}\sum_{i=1}^n |x_i|^k\Big) \leq 2^k m^{-1}\sum_{|\beta|\leq k}|x^\beta|.$$

The first equivalence follows. The second equivalence is an easy consequence of the Leibniz formula.

(3) We must show completeness. Let f_m be a Cauchy sequence in $\mathcal{S}(\mathbb{R}^n)$, i.e., for all k,α, $\|f_m - f_\ell\|_{k,\alpha} \to 0$ as $m,\ell\to\infty$. Then for each α, the sequence $\partial^\alpha f_m$ converges uniformly to a continuous function f^α. Denoting e_j the standard unit vectors in $\mathbb{R}^n$, we have

$$f_m(x+te_j) - f_m(x) = \int_0^t \partial_j f_m(x+se_j)\,ds,$$

and letting $m\to\infty$ we obtain

$$f^0(x+te_j) - f^0(x) = \int_0^t f^{e_j}(x+se_j)\,ds,$$

and hence $f^{e_j} = \partial_j f^0$. By induction, we find that $f^\alpha = \partial^\alpha f^0$ for all α, thus $f := f^0 \in C^\infty(\mathbb{R}^n)$.

Let us show that $f\in\mathcal{S}(\mathbb{R}^n)$. Since f_m (being Cauchy) is bounded in $\mathcal{S}(\mathbb{R}^n)$, we have $\|f_m\|_{\alpha,k}\leq C_{\alpha,k}$ for all m, thus

$$|\partial^\alpha f_m(x)| \leq C_{\alpha,k}(1+|x|)^{-k}$$

for all x and all m. Letting $m\to\infty$ implies $|\partial^\alpha f(x)|\leq C_{\alpha,k}(1+|x|)^{-k}$ for all x, i.e., $\|f\|_{\alpha,k}\leq C_{\alpha,k}$.

Finally, we check that f_m converges to f in $\mathcal{S}(\mathbb{R}^n)$. For fixed α and k, set $g_m(x) := (1+|x|)^k\partial^\alpha f_m(x)$ and $g(x) := (1+|x|)^k\partial^\alpha f(x)$. Then g_m is a Cauchy sequence with respect to $\|\ \|_\infty$ which converges uniformly to g, since $g_m\to g$ pointwise and the limit is unique. That is $\|f_m - f\|_{\alpha,k} = \|g_m - g\|_\infty \to 0$ as required. □

Proposition 8.3. *The Fourier transform maps $\mathcal{S}(\mathbb{R}^n)$ continuously into itself.*

PROOF. If $f\in\mathcal{S}(\mathbb{R}^n)$ then $x^\alpha\partial^\beta f(x)\in L^1(\mathbb{R}^n)\cap C_0(\mathbb{R}^n)$ for all α,β, by Lemma 8.2. Thus, by Lemma 8.1, $\widehat{f}\in C^\infty(\mathbb{R}^n)$ and

$$\xi^\alpha\partial_\xi^\beta\widehat{f}(\xi) = (-1)^{|\beta|}(2\pi i)^{|\beta|}\xi^\alpha[x^\beta f(x)]\widehat{\ }(\xi) = (-1)^{|\beta|}(2\pi i)^{|\beta|-|\alpha|}[\partial_x^\alpha(x^\beta f(x))]\widehat{\ }(\xi).$$

Consequently,

$$\begin{aligned}|\xi^\alpha\partial_\xi^\beta\widehat{f}(\xi)| &\leq (2\pi)^{|\beta|-|\alpha|}\int_{\mathbb{R}^n}|\partial_x^\alpha(x^\beta f(x))|\,dx\\ &\leq (2\pi)^{|\beta|-|\alpha|}\int_{\mathbb{R}^n}(1+|x|)^{-n-1}\,dx\sup_{x\in\mathbb{R}^n}(1+|x|)^{n+1}|\partial_x^\alpha(x^\beta f(x))|\end{aligned}$$

which implies the statement in view of Lemma 8.2. □

Lemma 8.4 (Riemann–Lebesgue)**.** *$\mathscr{F}L^1(\mathbb{R}^n)\subseteq C_0(\mathbb{R}^n)$.*

PROOF. The Fourier transform maps functions in $C_c^\infty(\mathbb{R}^n)\subseteq\mathcal{S}(\mathbb{R}^n)$ to functions in $\mathcal{S}(\mathbb{R}^n)\subseteq C_0(\mathbb{R}^n)$. By Theorem 4.20, $C_c^\infty(\mathbb{R}^n)$ is dense in $L^1(\mathbb{R}^n)$, and if $\|f_k - f\|_1\to 0$ then $\|\widehat{f_k}-\widehat{f}\|_\infty\to 0$, by (8.2). This implies the statement, since $C_0(\mathbb{R}^n)$ is closed with respect to $\|\ \|_\infty$. □

At this point we compute the Fourier transform of a **Gaussian function**; this is a preparation for the Fourier inversion formula.

Lemma 8.5 (Fourier transform of the Gaussian). *For $f(x) = e^{-\pi a|x|^2}$, where $a > 0$, we have $\widehat{f}(\xi) = a^{-n/2}e^{-\pi|\xi|^2/a}$.*

PROOF. First suppose that $n = 1$. By Lemma 8.1,

$$(\widehat{f})'(\xi) = (-2\pi ixe^{-\pi ax^2})\widehat{\ }(\xi) = ia^{-1}(f')\widehat{\ }(\xi) = ia^{-1}2\pi i\xi\widehat{f}(\xi) = -2\pi a^{-1}\xi\widehat{f}(\xi),$$

hence $\partial_\xi(e^{\pi\xi^2/a}\widehat{f}(\xi)) = 0$, and so $e^{\pi\xi^2/a}\widehat{f}(\xi)$ is constant. Thus

$$e^{\pi\xi^2/a}\widehat{f}(\xi) = \widehat{f}(0) = \int_{\mathbb{R}} e^{-\pi ax^2}\,dx = a^{-1/2},$$

by Example 3.35. The case $n = 1$ and Fubini's theorem 3.27 imply the general case,

$$\widehat{f}(\xi) = \int_{\mathbb{R}^n} e^{-\pi a|x|^2}e^{-2\pi i\xi\cdot x}\,dx = \prod_{j=1}^{n}\int_{\mathbb{R}} e^{-\pi ax_j^2}e^{-2\pi i\xi_j\cdot x_j}\,dx_j = a^{-n/2}e^{-\pi|\xi|^2/a}. \qquad \square$$

Let us turn to inversion of the Fourier transform. For $f \in L^1(\mathbb{R}^n)$, we define

$$f^\vee(x) := \widehat{f}(-x) = \int_{\mathbb{R}^n} f(\xi)e^{2\pi i\xi\cdot x}\,d\xi, \quad x \in \mathbb{R}^n.$$

Theorem 8.6 (Fourier inversion theorem). *If $f \in L^1(\mathbb{R}^n)$ and $\widehat{f} \in L^1(\mathbb{R}^n)$, then f coincides a.e. with a continuous function f_0, and we have*

$$(\widehat{f})^\vee = (f^\vee)\widehat{\ } = f_0.$$

PROOF. For $t > 0$ and $x \in \mathbb{R}^n$, set

$$\psi(\xi) := e^{2\pi i\xi\cdot x - \pi t^2|\xi|^2} = M_x e^{-\pi t^2|\xi|^2}$$

By Lemmas 8.1 and 8.5,

$$\widehat{\psi}(y) = T_x(t^{-n}e^{-\pi|y|^2/t^2}) = t^{-n}e^{-\pi|x-y|^2/t^2} = \varphi_t(x-y),$$

for $\varphi(x) = e^{-\pi|x|^2}$, cf. (4.7). By Lemma 8.1,

$$\int_{\mathbb{R}^n} e^{-\pi t^2|\xi|^2}e^{2\pi i\xi\cdot x}\widehat{f}(\xi)\,d\xi = \int \widehat{f}(\xi)\psi(\xi)\,d\xi = \int f(y)\widehat{\psi}(y)\,dy = f * \varphi_t$$

which converges to f in $L^1(\mathbb{R}^n)$ as $t \to 0$, by Proposition 4.18. On the other hand, since $\widehat{f} \in L^1(\mathbb{R}^n)$,

$$\lim_{t\to 0}\int_{\mathbb{R}^n} e^{-\pi t^2|\xi|^2}e^{2\pi i\xi\cdot x}\widehat{f}(\xi)\,d\xi = \int_{\mathbb{R}^n} e^{2\pi i\xi\cdot x}\widehat{f}(\xi)\,d\xi = (\widehat{f})^\vee(x),$$

by the dominated convergence theorem 3.22. It follows that $f = (\widehat{f})^\vee$ a.e. and analogously $(f^\vee)\widehat{\ }$ a.e. Being Fourier transforms of L^1-functions, $(\widehat{f})^\vee$ and $(f^\vee)\widehat{\ }$ are continuous. $\square$

Corollary 8.7. *If $f \in L^1(\mathbb{R}^n)$ and $\widehat{f} = 0$, then $f = 0$ a.e.* $\square$

Corollary 8.8. *$\mathscr{F} : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$ is an isomorphism.*

PROOF. By Proposition 8.3, $\mathscr{F}$ maps $\mathcal{S}(\mathbb{R}^n)$ continuously into itself, and so does the mapping $f \mapsto f^\vee$, because $f^\vee(x) = \widehat{f}(-x)$. By Theorem 8.6, these mappings are inverse to each other. $\square$

The Fourier transform of an L^1-function need not be L^1 as illustrated by the following example.

Example 8.9 (The sinc function)**.** Clearly, the characteristic function of the interval $[-a, a]$ is in $L^1(\mathbb{R})$. Its Fourier transform

$$\widehat{\chi}_{[-a,a]}(\xi) = \int_{-a}^{a} e^{-2\pi i x\xi}\, dx = -\frac{e^{-2\pi i a\xi}}{2\pi i\xi} + \frac{e^{2\pi i a\xi}}{2\pi i\xi} = \frac{\sin(2\pi a\xi)}{\pi\xi}$$

however is not an element of $L^1(\mathbb{R})$. In particular, the Fourier transform of the rectangular function $\chi_{[-1/2,1/2]}$ is the (normalized) **sinc function** $\operatorname{sinc}(x) = \sin(\pi x)/(\pi x)$.

By the lemma of Riemann–Lebesgue 8.4, the Fourier transform is a bounded linear operator $\mathscr{F} : L^1(\mathbb{R}^n) \to C_0(\mathbb{R}^n)$. It is injective, but not surjective.

Proposition 8.10. *The bounded linear operator $\mathscr{F} : L^1(\mathbb{R}^n) \to C_0(\mathbb{R}^n)$ is injective, but not surjective.*

Proof. Assume $f, g \in L^1(\mathbb{R}^n)$ and $\widehat{f} = \widehat{g}$. Then $f - g \in L^1(\mathbb{R}^n)$ and $\widehat{f} - \widehat{g} = 0$. Thus Corollary 8.7 implies $f = g$ a.e.

Let us show that $\mathscr{F} : L^1(\mathbb{R}^n) \to C_0(\mathbb{R}^n)$ is not surjective. For simplicity let $n = 1$. It is more convenient to show that the inverse Fourier transform $(\ )^\vee : L^1(\mathbb{R}) \to C_0(\mathbb{R})$ is not surjective. The assertion is then an immediate consequence: if $g \in C_0(\mathbb{R}) \setminus (L^1(\mathbb{R}))^\vee$ then $g(-x) \in C_0(\mathbb{R}) \setminus \mathscr{F}L^1(\mathbb{R})$.

Assume that $(\ )^\vee : L^1(\mathbb{R}) \to C_0(\mathbb{R})$ is surjective. By the open mapping theorem A.3, there is a constant $C > 0$ such that

$$\|f\|_1 \le C\,\|f^\vee\|_\infty, \qquad \text{for all } f \in L^1(\mathbb{R}). \tag{8.4}$$

For $\epsilon > 0$ let $g_\epsilon(x) := \epsilon^{-1/2} e^{-\pi x^2/\epsilon}$ and $f_\epsilon := g_\epsilon * \chi_{[-1,1]}$. Then $f_\epsilon \in L^1(\mathbb{R})$, by (4.3), and $f_\epsilon \in C_0(\mathbb{R})$, by a simple computation. Thus by (8.4) and Example 3.35,

$$\|\widehat{f_\epsilon}\|_1 \le C\,\|f_\epsilon\|_\infty = C\,\|g_\epsilon * \chi_{[-1,1]}\|_\infty \le C\,\|g_\epsilon\|_1 = C.$$

By Lemmas 8.1 and 8.5,

$$\widehat{f_\epsilon}(\xi) = \widehat{g_\epsilon}(\xi)\,\widehat{\chi}_{[-1,1]}(\xi) = e^{-\pi\epsilon\xi^2}\,\widehat{\chi}_{[-1,1]}(\xi) \to \widehat{\chi}_{[-1,1]}(\xi)$$

pointwise as $\epsilon \to 0$. So, by Fatou's lemma 3.17,

$$\int_{\mathbb{R}} |\widehat{\chi}_{[-1,1]}|\, d\xi = \int_{\mathbb{R}} \lim_{k\to\infty} |\widehat{f}_{1/k}|\, d\xi \le \liminf_{k\to\infty} \int_{\mathbb{R}} |\widehat{f}_{1/k}|\, d\xi \le C,$$

a contradiction; see Example 8.9. □

8.2. The Fourier transform on L^2

In the previous section we have seen that the Fourier transform is a bounded linear operator (cf. (8.2) and Lemma 8.4)

$$\mathscr{F} : L^1(\mathbb{R}^n) \to C_0(\mathbb{R}^n).$$

If we abandon the requirement that $\mathscr{F}$ be defined pointwise by (8.1), it can be extended to other spaces.

Theorem 8.11 (Plancherel)**.** *If $f \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$, then $\widehat{f} \in L^2(\mathbb{R}^n)$, and $\mathscr{F}|_{L^1(\mathbb{R}^n)\cap L^2(\mathbb{R}^n)}$ extends uniquely to an isometric isomorphism on $L^2(\mathbb{R}^n)$.*

Proof. Let

$$F^1(\mathbb{R}^n) := \{f \in L^1(\mathbb{R}^n) : \widehat{f} \in L^1(\mathbb{R}^n)\}. \tag{8.5}$$

Then $F^1(\mathbb{R}^n) \subseteq L^2(\mathbb{R}^n)$, since $\widehat{f} \in L^1(\mathbb{R}^n)$ implies $f \in L^\infty(\mathbb{R}^n)$ (cf. (8.2)) and thus $f \in L^2(\mathbb{R}^n)$, by Proposition 4.7. Moreover, $F^1(\mathbb{R}^n)$ is dense in $L^2(\mathbb{R}^n)$, because

$\mathcal{S}(\mathbb{R}^n) \subseteq F^1(\mathbb{R}^n)$ and $\mathcal{S}(\mathbb{R}^n)$ is dense in $L^2(\mathbb{R}^n)$, by Theorem 4.20. Let $f, g \in F^1(\mathbb{R}^n)$, and set $h := \overline{\widehat{g}}$. By Theorem 8.6,

$$\widehat{h}(\xi) = \int_{\mathbb{R}^n} \overline{\widehat{g}(x)} e^{-2\pi i \xi \cdot x}\, dx = \int_{\mathbb{R}^n} \overline{\widehat{g}(x) e^{2\pi i \xi \cdot x}}\, dx = \overline{g(\xi)},$$

and hence, by Lemma 8.1,

$$\int_{\mathbb{R}^n} f(x)\overline{g(x)}\, dx = \int_{\mathbb{R}^n} f(x)\widehat{h}(x)\, dx = \int_{\mathbb{R}^n} \widehat{f}(x) h(x)\, dx = \int_{\mathbb{R}^n} \widehat{f}(x)\overline{\widehat{g}(x)}\, dx,$$

i.e., $\mathscr{F}|_{F^1(\mathbb{R}^n)}$ preserves the L^2-inner product. In particular,

$$\|\widehat{f}\|_2 = \|f\|_2. \tag{8.6}$$

Since $\mathscr{F}(F^1(\mathbb{R}^n)) = F^1(\mathbb{R}^n)$, by Theorem 8.6, $\mathscr{F}|_{F^1(\mathbb{R}^n)}$ extends by continuity to an isometric isomorphism $\widetilde{\mathscr{F}}$ on $L^2(\mathbb{R}^n)$.

It remains to check that $\widetilde{\mathscr{F}} = \mathscr{F}$ on $L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$. Let $f \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$ and $\varphi(x) := e^{-\pi|x|^2}$. Then $f * \varphi_\epsilon \in L^1(\mathbb{R}^n)$, by Young's inequality (4.15), and

$$(f * \varphi_\epsilon)\widehat{\ }(\xi) = \widehat{f}(\xi) e^{-\pi\epsilon^2|\xi|^2},$$

by Lemmas 8.1 and 8.5, and so $(f * \varphi_\epsilon)\widehat{\ } \in L^1(\mathbb{R}^n)$, since $\widehat{f}$ is bounded. That is $f * \varphi_\epsilon \in F^1(\mathbb{R}^n)$. By Proposition 4.18, $f * \varphi_\epsilon$ converges to f in $L^1(\mathbb{R}^n)$ and in $L^2(\mathbb{R}^n)$. We may infer $\|(f*\varphi_\epsilon)\widehat{\ } - \widehat{f}\|_\infty \to 0$, by (8.2), and $\|(f*\varphi_\epsilon)\widehat{\ } - \widetilde{\mathscr{F}} f\|_2 \to 0$, by (8.6). By Corollary 4.11, there is a subsequence $(f * \varphi_{\epsilon_k})\widehat{\ }$ that converges pointwise a.e. to $\widehat{f}$ as well as to $\widetilde{\mathscr{F}} f$. Therefore, $\widetilde{\mathscr{F}} f = \mathscr{F} f$ a.e. $\square$

We denote by $\widehat{f} = \mathscr{F} f$ also the Fourier transform of functions $f \in L^2(\mathbb{R}^n)$.

Corollary 8.12 (Parseval's theorem)**.** *If $f, g \in L^2(\mathbb{R}^n)$ then $\langle f, g\rangle = \langle \widehat{f}, \widehat{g}\rangle$, i.e., $\mathscr{F} : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$ is unitary.*

Proof. This follows from $\|\widehat{f}\|_2 = \|f\|_2$ by polarization,

$$2\langle f, g\rangle = \|f + g\|_2^2 - i\|f + ig\|_2^2 - (1-i)\|f\|_2^2 - (1-i)\|g\|_2^2. \qquad \square$$

The Fourier transform $\widehat{f}$ of a function $f \in L^2(\mathbb{R}^n)$ is not given by the formula (8.1); the integral in (8.1) may not exist. However, $\widehat{f}$ is the L^2-limit of the functions

$$(\chi_{B_r(0)} f)\widehat{\ }(\xi) = \int_{B_r(0)} f(x) e^{-2\pi i \xi \cdot x}\, dx$$

as $r \to \infty$. Here $\chi_{B_r(0)} f \in L^2(B_r(0)) \subseteq L^1(B_r(0))$, by Proposition 4.9, and so the integral exists. By the monotone convergence theorem 3.14, $\|\chi_{B_r(0)} f - f\|_2 \to 0$ as $r \to \infty$ and hence $\|(\chi_{B_r(0)} f)\widehat{\ } - \widehat{f}\|_2 \to 0$, by Theorem 8.11. By the same argument $\widehat{f}$ is the L^2-limit of the Fourier transform of *every* sequence of functions $f_m \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$ that converges to f in $L^2(\mathbb{R}^n)$. By Corollary 4.11 there is a subsequence that converges a.e., and so for $f \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$ the integral in (8.1) coincides a.e. with the extension provided by Theorem 8.11.

For instance, by Example 8.9,

$$\int_{[-r,r]} \widehat{\chi}_{[-a,a]}(\xi)\, e^{2\pi i \xi x}\, d\xi = \int_{[-r,r]} \frac{\sin(2\pi a \xi)}{\pi \xi}\, e^{2\pi i \xi x}\, d\xi$$

converges to $\chi_{[-a,a]}$ in $L^2(\mathbb{R})$ as $r \to \infty$.

Corollary 8.13. *The inversion formula $f = (\widehat{f})^\vee$ continuous to hold on $L^2(\mathbb{R}^n)$.*

Proof. By Theorem 8.11 the definition $f^\vee(x) := \widehat{f}(-x)$ makes sense for $f \in L^2(\mathbb{R}^n)$. Since $f = (\widehat{f})^\vee$ holds on $F^1(\mathbb{R}^n)$ (cf. (8.5)), by Theorem 8.6, and since $F^1(\mathbb{R}^n)$ is dense in $L^2(\mathbb{R}^n)$, we can conclude the assertion from Theorem 8.11 (which clearly holds also for $\widehat{f}$ replaced by $f^\vee$). $\square$

By Plancherel's theorem 8.11, the Fourier transform is a linear mapping $L^1(\mathbb{R}^n) + L^2(\mathbb{R}^n) \to L^\infty(\mathbb{R}^n) + L^2(\mathbb{R}^n)$ satisfying $\|\widehat{f}\|_\infty \leq \|f\|_1$ for $f \in L^1(\mathbb{R}^n)$ and $\|\widehat{f}\|_2 = \|f\|_2$ for $f \in L^2(\mathbb{R}^n)$. By the Riesz–Thorin interpolation theorem 7.8, we get the following result for immediate L^p-spaces.

Theorem 8.14 (Hausdorff–Young inequality)**.** *Let $1 \leq p \leq 2$ and let q be the conjugate exponent to p. If $f \in L^p(\mathbb{R}^n)$ then $\widehat{f} \in L^q(\mathbb{R}^n)$ and*

$$\|\widehat{f}\|_q \leq \|f\|_p.$$

Proof. Apply the Riesz–Thorin interpolation theorem 7.8. $\square$

In Lemma 8.5 we have seen by means of a Gaussian function that the Fourier transform maps an acute peak to a broadly spread peak. This is a general property of the Fourier transform that is called the **uncertainty principle**.

Theorem 8.15 (Heisenberg's uncertainty principle)**.** *If $f \in \mathcal{S}(\mathbb{R}^n)$, then*

$$\|f\|_2^2 \leq 4\pi \|(x_j - y_j)f(x)\|_2 \|(\xi_j - \eta_j)\widehat{f}(\xi)\|_2$$

for all $y, \eta \in \mathbb{R}^n$, $j = 1, \ldots, n$.

Thus f and $\widehat{f}$ cannot both be sharply localized about single points.

Proof. Replacing f by $M_{\eta_j e_j} T_{y_j e_j} f$, where e_j is the jth standard unit vector in $\mathbb{R}^n$, we may assume that $y = \eta = 0$, in view of Lemma 8.1. Integration by parts (cf. Corollary 6.14), Hölder's inequality 4.2, and (8.6) yield

$$\begin{aligned}
\|f\|_2^2 &= \int_{\mathbb{R}^n} f(x)\overline{f(x)} \partial_{x_j} x_j \, dx \\
&= -\int_{\mathbb{R}^n} (\partial_j f(x)\overline{f(x)} + f(x)\partial_j \overline{f(x)}) x_j \, dx \\
&\leq 2\|x_j f(x)\|_2 \|\partial_j f\|_2 \\
&= 4\pi \|x_j f(x)\|_2 \|\xi_j \widehat{f}(\xi)\|_2,
\end{aligned}$$

where in the last step we again used Lemma 8.1 and (8.6). $\square$

8.3. Paley–Wiener theorems

As seen in Lemma 8.1, the smoothness of a function is connected to the decay of its Fourier transform at infinity (and vice versa). We shall see below that in the extrem case, when f is compactly supported on $\mathbb{R}$, its Fourier transform $\widehat{f}$ extends to an entire function. Theorems that relate decay properties of a function (or distribution) at infinity with analyticity of its Fourier transform are called **Paley Wiener theorems**. We will investigate two such theorems.

The Fourier transform $\widehat{f}$ of a function f on $\mathbb{R}$ is by definition a function on $\mathbb{R}$. Often $\widehat{f}$ admits a holomorphic extension to some region in $\mathbb{C}$ which is not too surprising, since $e^{2\pi t z}$ is an entire function of z for every real t.

Let us formally consider the integral that defines the inverse Fourier transform

$$f(z) = \int_{-\infty}^{\infty} F(t) e^{2\pi i t z} \, dt \tag{8.7}$$

and allow z to be a complex number. In general, this integral may not be well-defined. We shall consider two situations which ensure the existence of this integral.

First we assume that F is supported on $\mathbb{R}_+ := \{x \in \mathbb{R} : x > 0\}$ and z lies in the upper half-plane $\mathbb{H} := \{z \in \mathbb{C} : \operatorname{Im} z > 0\}$. For $F \in L^2(\mathbb{R}_+)$ and $z \in \mathbb{H}$, the integral

$$f(z) = \int_0^\infty F(t) e^{2\pi i t z}\, dt, \quad z \in \mathbb{H}, \tag{8.8}$$

exists as Lebesgue integral, since $|e^{2\pi i t z}| = e^{-2\pi t \operatorname{Im} z}$ is in $L^2(\mathbb{R}_+)$ for each $z \in \mathbb{H}$.

Theorem 8.16 (Paley–Wiener I)**.** *If f is of the form* (8.8)*, then f is holomorphic in $\mathbb{H}$ and*

$$\sup_{y>0} \int_{-\infty}^\infty |f(x+iy)|^2\, dx = C < \infty. \tag{8.9}$$

Conversely, if f is holomorphic in $\mathbb{H}$ and satisfies (8.9)*, then there exists $F \in L^2(\mathbb{R}_+)$ such that f has the representation* (8.8) *and*

$$\int_0^\infty |F(t)|^2\, dt = C. \tag{8.10}$$

Proof. Assume that $F \in L^2(\mathbb{R}_+)$ and that f is given by (8.8). By Theorem 3.39 (applied to each half-plane $\{z : \operatorname{Im} z > \delta\}$, $\delta > 0$), f is holomorphic in $\mathbb{H}$. For fixed $y > 0$,

$$f(x+iy) = \int_0^\infty F(t) e^{-2\pi t y} e^{2\pi i t x}\, dt$$

and Plancherel's theorem 8.11, yields

$$\int_{-\infty}^\infty |f(x+iy)|^2\, dx = \int_0^\infty |F(t)|^2 e^{-4\pi t y}\, dt \le \int_0^\infty |F(t)|^2\, dt; \tag{8.11}$$

we may consider F as a function in $L^2(\mathbb{R})$ by extending it by 0 on $(-\infty, 0]$. This shows (8.9).

Now let f be holomorphic in $\mathbb{H}$ and satisfy (8.9). Fix $y > 0$, $\alpha > 0$, and let γ_α denote the rectangular path with vertices $\pm\alpha + i$ and $\pm\alpha + iy$. By Cauchy's theorem, for all $t \in \mathbb{R}$,

$$\int_{\gamma_\alpha} f(\zeta) e^{-2\pi i t \zeta}\, d\zeta = 0. \tag{8.12}$$

Let $\Phi(\beta)$, $\beta \in \mathbb{R}$, be the integral of $f(\zeta)e^{-2\pi i t\zeta}$ along the line segment from $\beta + i$ to $\beta + iy$. If I denotes the real interval with endpoints 1 and y, then by Hölder's inequality 4.2,

$$\begin{aligned} |\Phi(\beta)|^2 &= \left| \int_I f(\beta + is) e^{-2\pi i t(\beta + is)}\, ds \right|^2 \\ &\le \int_I |f(\beta+is)|^2\, ds \int_I e^{4\pi t s}\, ds =: \Psi(\beta) \int_I e^{4\pi t s}\, ds. \end{aligned} \tag{8.13}$$

By (8.9) and Fubini's theorem 3.27,

$$\int_{-\infty}^\infty \Psi(\beta)\, d\beta = \int_I \int_{-\infty}^\infty |f(\beta+is)|^2\, d\beta\, ds \le C\lambda(I) = C|1-y|.$$

It follows that there is a sequence $\alpha_k \to \infty$ such that $\Psi(\pm\alpha_k) \to 0$. Hence, in view of (8.13),

$$\Phi(\pm\alpha_k) \to 0, \tag{8.14}$$

for all t, and α_k is independent of t.

Let us consider

$$g_k(y,t) := \int_{-\alpha_k}^{\alpha_k} f(x+iy)e^{-2\pi itx}\,dx.$$

Then (8.12) and (8.14) imply

$$e^{2\pi ty}g_k(y,t) - e^{2\pi t}g_k(1,t) \to 0 \quad \text{as } k\to\infty. \tag{8.15}$$

If $f_y(x) := f(x+iy)$, then $f_y \in L^2(\mathbb{R})$ by (8.9). By Plancherel's theorem 8.11,

$$\|g_k(y,\cdot) - \widehat{f_y}\|_2 \to 0 \quad \text{as } k\to\infty.$$

By Corollary 4.11, there is a subsequence of $(g_k(y,t))_k$ which converges to $\widehat{f_y}(t)$ for a.e. t. Thus, if we define

$$F(t) := e^{2\pi t}\widehat{f_1}(t), \quad t\in\mathbb{R},$$

then (8.15) implies that, for each $y>0$, $F(t) = e^{2\pi ty}\widehat{f_y}(t)$ for a.e. $t\in\mathbb{R}$. Applying Plancherel's theorem 8.11 gives

$$\int_{-\infty}^{\infty} e^{-4\pi ty}|F(t)|^2\,dt = \int_{-\infty}^{\infty}|\widehat{f_y}(t)|^2\,dt \le C,$$

for all $y>0$, by (8.9). Letting $y\to\infty$ implies $F(t)=0$ for a.e. $t<0$, and letting $y\to 0$ gives

$$\int_0^\infty |F(t)|^2\,dt \le C. \tag{8.16}$$

This implies that $\widehat{f_y}(t) = e^{-2\pi ty}F(t)$ is in $L^1(\mathbb{R})$. Thus, by Corollary 8.13 (and the arguments preceding it),

$$f_y(x) = \int_{-\infty}^{\infty}\widehat{f_y}(t)e^{2\pi itx}\,dt,$$

that is

$$f(z) = \int_0^\infty F(t)e^{-2\pi ty}e^{2\pi itx}\,dt = \int_0^\infty F(t)e^{2\pi itz}\,dt, \quad z\in\mathbb{H}.$$

Finally, (8.10) follows from (8.16) and (8.11). □

Thanks to (8.9), the dominated convergence theorem 3.22 implies

$$\lim_{y\to 0^+}\int_{-\infty}^{\infty}|f(x+iy) - F^\vee(x)|^2\,dx = 0. \tag{8.17}$$

The theorem describes the structure of the **Hardy space** $H^2(\mathbb{H})$ of the upper half-plane, i.e.,

$$H^2(\mathbb{H}) := \{f : f \text{ holomorphic on } \mathbb{H},\ \|f\|_{H^2(\mathbb{H})} < \infty\},$$

which is a Hilbert space with norm given by

$$\|f\|_{H^2(\mathbb{H})} := \sup_{y>0}\Big(\int_{-\infty}^{\infty}|f(x+iy)|^2\,dx\Big)^{1/2}.$$

Indeed, the above theorem implies the following corollary.

Corollary 8.17. *The mapping $F \mapsto f(z) = \int_0^\infty F(t)e^{2\pi itz}\,dt$ yields an isomorphism between $L^2(\mathbb{R}_+)$ and $H^2(\mathbb{H})$.* □

Another way to make sense of the integral (8.7) is to require that F is compactly supported. If $0 < A < \infty$ and $F \in L^2([-A, A])$, then

$$f(z) = \int_{-A}^{A} F(t) e^{2\pi i t z}\, dt, \quad z \in \mathbb{C}, \tag{8.18}$$

clearly is well-defined.

Theorem 8.18 (Paley–Wiener II). *If f is of the form* (8.18)*, then f is entire and there exists $C > 0$ such that*

$$|f(z)| \leq C e^{2\pi A |z|}, \quad z \in \mathbb{C}, \tag{8.19}$$

and $f|_{\mathbb{R}} \in L^2(\mathbb{R})$. Conversely, if f is an entire function satisfying (8.19) *for some positive constants A and C, and $f|_{\mathbb{R}} \in L^2(\mathbb{R})$, then there exists $F \in L^2([-A, A])$ such that f has the representation* (8.18).

Entire functions f satisfying (8.19) are said to be of **exponential type**.

Proof. If f is of the form (8.18), then f is entire by Theorem 3.39, and

$$|f(z)| \leq \int_{-A}^{A} |F(t)| e^{-2\pi t \operatorname{Im} z}\, dt \leq \int_{-A}^{A} |F(t)|\, dt\, e^{2\pi A |\operatorname{Im} z|}$$

which implies (8.19). By Plancherel's theorem 8.11, $f|_{\mathbb{R}} \in L^2(\mathbb{R})$.

Assume that f is an entire function satisfying (8.19) for some positive constants A and C, and $f|_{\mathbb{R}} \in L^2(\mathbb{R})$. Define $f_\epsilon(x) := f(x) e^{-2\pi\epsilon |x|}$, for $\epsilon > 0$ and $x \in \mathbb{R}$. We claim that

$$\lim_{\epsilon \to 0} \int_{-\infty}^{\infty} f_\epsilon(x) e^{-2\pi i t x}\, dx = 0 \quad \text{ for } t \in \mathbb{R} \setminus [-A, A]. \tag{8.20}$$

This claim will imply the theorem as follows. By the dominated convergence theorem 3.22, $\|f_\epsilon - f|_{\mathbb{R}}\|_2 \to 0$ as $\epsilon \to 0$, and so by Plancherel's theorem 8.11, $\|\widehat{f_\epsilon} - \widehat{f|_{\mathbb{R}}}\|_2 \to 0$. Then, by (8.20) and Corollary 4.11, $F := \widehat{f|_{\mathbb{R}}}$ vanishes a.e. outside $[-A, A]$. By Corollary 8.13, the representation (8.18) holds for a.e. real z, and hence for all $z \in \mathbb{C}$, because both sides of (8.18) are entire functions.

Let us prove (8.20). For real α let γ_α be the ray defined by $\gamma_\alpha(s) := s e^{i\alpha}$, $s \in [0, \infty)$. Define

$$\Phi_\alpha(w) := \int_{\gamma_\alpha} f(z) e^{-2\pi w z} dz = e^{i\alpha} \int_0^\infty f(s e^{i\alpha}) e^{-2\pi w s e^{i\alpha}}\, ds,$$

for $w \in \Pi_\alpha := \{w \in \mathbb{C} : \operatorname{Re}(w e^{i\alpha}) > A\}$. By (8.19),

$$|f(s e^{i\alpha}) e^{-2\pi w s e^{i\alpha}}| \leq C e^{2\pi A s} e^{-2\pi s \operatorname{Re}(w e^{i\alpha})} = C e^{-2\pi s (\operatorname{Re}(w e^{i\alpha}) - A)},$$

and so, by Theorem 3.39, Φ_α is holomorphic on the half-plane Π_α. More is true for $\alpha = 0$ and $\alpha = \pi$. Since $f|_{\mathbb{R}} \in L^2(\mathbb{R})$,

$$\Phi_0(w) = \int_0^\infty f(s) e^{-2\pi w s}\, ds$$

is holomorphic in $\{w \in W : \operatorname{Re} w > 0\}$ and

$$\Phi_\pi(w) = -\int_0^\infty f(-s) e^{2\pi w s}\, ds = -\int_{-\infty}^0 f(s) e^{-2\pi w s}\, ds$$

is holomorphic in $\{w \in W : \operatorname{Re} w < 0\}$. Now, for $t \in \mathbb{R}$,

$$\int_{-\infty}^{\infty} f_\epsilon(x) e^{-2\pi i t x}\, dx = \Phi_0(\epsilon + it) - \Phi_\pi(-\epsilon + it). \tag{8.21}$$

We will show that any two of the functions Φ_α coincide on the intersection of their domains of definition (i.e., they are analytic continuations of each other). Then

$$\Phi_0(\epsilon+it)-\Phi_\pi(-\epsilon+it)=\begin{cases}\Phi_{\pi/2}(\epsilon+it)-\Phi_{\pi/2}(-\epsilon+it) & \text{if } t<-A,\\ \Phi_{-\pi/2}(\epsilon+it)-\Phi_{-\pi/2}(-\epsilon+it) & \text{if } t>A,\end{cases}$$

evidently tends to 0 as $\epsilon\to 0$, and (8.20) is proved.

Suppose that $0<\beta-\alpha<\pi$. If $w=|w|e^{-i(\alpha+\beta)/2}$, then

$$\operatorname{Re}(we^{i\alpha})=|w|\operatorname{Re}(e^{i(\alpha-\beta)/2})=|w|\cos\frac{\alpha-\beta}{2}=:|w|\eta>0,$$
$$\operatorname{Re}(we^{i\beta})=|w|\operatorname{Re}(e^{i(\beta-\alpha)/2})=|w|\cos\frac{\beta-\alpha}{2}=|w|\eta.$$

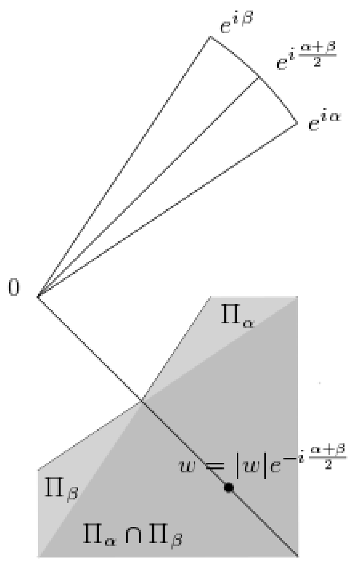
Thus, $w\in\Pi_\alpha\cap\Pi_\beta$ provided that $|w|>A/\eta$. Consider the path integral

$$\int_\gamma f(z)e^{-2\pi wz}dz,\quad \gamma(t)=re^{it},\ t\in[\alpha,\beta]. \tag{8.22}$$

Since $\operatorname{Re}(w\gamma(t))=|w|r\operatorname{Re}e^{i(t-(\alpha+\beta)/2)}\geq|w|r\eta$ and so, by (8.19),

$$|f(\gamma(t))e^{-2\pi w\gamma(t)}|\leq Ce^{2\pi r(A-|w|\eta)},$$

the path integral (8.22) tends to 0 as $r\to\infty$ if $|w|>A/\eta$. Thus, Cauchy's theorem implies that $\Phi_\alpha(w)=\Phi_\beta(w)$ if $w=|w|e^{-i(\alpha+\beta)/2}$ and $|w|>A/\eta$. By the identity theorem for holomorphic functions $\Phi_\alpha=\Phi_\beta$ on the intersection of their domains of definition. $\square$



APPENDIX A

Appendix

A.1. Basic set-theoretic operations

For an arbitrary index set A we have the distribution laws

$$E \cap \bigcup_{\alpha \in A} E_\alpha = \bigcup_{\alpha \in A} E \cap E_\alpha \quad \text{and} \quad E \cup \bigcap_{\alpha \in A} E_\alpha = \bigcap_{\alpha \in A} E \cup E_\alpha,$$

and de Morgan's laws

$$\Big(\bigcup_{\alpha \in A} E_\alpha \Big)^c = \bigcap_{\alpha \in A} (E_\alpha)^c \quad \text{and} \quad \Big(\bigcap_{\alpha \in A} E_\alpha \Big)^c = \bigcup_{\alpha \in A} (E_\alpha)^c,$$

$$E \setminus \bigcup_{\alpha \in A} E_\alpha = \bigcap_{\alpha \in A} E \setminus E_\alpha \quad \text{and} \quad E \setminus \bigcap_{\alpha \in A} E_\alpha = \bigcup_{\alpha \in A} E \setminus E_\alpha.$$

A map $f : X \to Y$ induces maps $f : \mathfrak{P}(X) \to \mathfrak{P}(Y)$ and $f^{-1} : \mathfrak{P}(Y) \to \mathfrak{P}(X)$ satisfying

$$f\Big(\bigcup_{\alpha \in A} E_\alpha \Big) = \bigcup_{\alpha \in A} f(E_\alpha) \quad \text{and} \quad f\Big(\bigcap_{\alpha \in A} E_\alpha \Big) \subseteq \bigcap_{\alpha \in A} f(E_\alpha),$$

$$f^{-1}\Big(\bigcup_{\alpha \in A} E_\alpha \Big) = \bigcup_{\alpha \in A} f^{-1}(E_\alpha) \quad \text{and} \quad f^{-1}\Big(\bigcap_{\alpha \in A} E_\alpha \Big) = \bigcap_{\alpha \in A} f^{-1}(E_\alpha),$$

$$f^{-1}(E^c) = (f^{-1}(E))^c,$$

$$E \subseteq F \Rightarrow f(E) \subseteq f(F) \quad \text{and} \quad E \subseteq F \Rightarrow f^{-1}(E) \subseteq f^{-1}(F),$$

$$E \subseteq f^{-1}(f(E)) \quad \text{and} \quad E \supseteq f(f^{-1}(E)).$$

A.2. Banach spaces

Let $\mathbb{K}$ denote either $\mathbb{R}$ or $\mathbb{C}$ and let X be a vector space over $\mathbb{K}$. A function $\| \ \| : X \to [0, \infty)$ is called a **seminorm** if

- $\|\lambda x\| = |\lambda| \|x\|$ for all $x \in X$ and $\lambda \in \mathbb{K}$,
- $\|x + y\| \le \|x\| + \|y\|$ for all $x, y \in X$,

and it is called a **norm** if additionally

- $\|x\| = 0$ if and only if $x = 0$.

A vector space equipped with a norm is called a **normed space**. The norm induces a metric $d(x, y) = \|x - y\|$ and hence a topology on X. A normed space that is complete with respect to the induced metric is called a **Banach space**.

A linear mapping $T : X \to Y$ between normed spaces is called **bounded** if it is bounded on bounded sets, i.e., there is a constant $C \ge 0$ such that

$$\|Tx\| \le C\|x\| \quad \text{for all } x \in X.$$

Lemma A.1. *For a linear mapping $T : X \to Y$ between normed spaces, the following are equivalent:*

(1) *T is bounded,*

(2) *T is continuous,*
(3) *T is continuous at 0.*

Proof. (1) $\Rightarrow$ (2) We have $\|Tx-Ty\| = \|T(x-y)\| \leq C\|x-y\| \leq \epsilon$ whenever $\|x-y\| \leq \epsilon/C$.

(2) $\Rightarrow$ (3) Obvious.

(3) $\Rightarrow$ (1) By assumption there is $\delta > 0$ so that $\|Tx\| \leq 1$ when $\|x\| \leq \delta$. Thus,

$$1 \geq \|T(\delta\|x\|^{-1}x)\| = \delta\|x\|^{-1}\|Tx\|,$$

and so T is bounded. □

The space $L(X,Y)$ of all bounded linear mappings between normed spaces X and Y is a normed space with respect to the **operator norm**

$$\|T\| := \sup_{\|x\|=1} \|Tx\| = \sup_{x\neq 0} \frac{\|Tx\|}{\|x\|} = \inf\big\{C : \|Tx\| \leq C\|x\| \text{ for all } x\big\}.$$

It is easy to see that $L(X,Y)$ is complete if so is Y. For $T \in L(X,Y)$ and $S \in L(Y,Z)$ we have $ST \in L(X,Z)$ with

$$\|ST\| \leq \|S\|\|T\|,$$

in particular, $L(X,X)$ is an algebra. If X is complete, $L(X,X)$ is a Banach algebra.

A bounded linear mapping $T \in L(X,Y)$ is an **isomorphism** if T is bijective and T^{-1} is bounded. We say that T is an **isometry** if $\|Tx\| = \|x\|$ for all $x \in X$. An isometry is an isomorphism onto its image.

The **dual space** X^* of a normed space X is the space of bounded linear functionals on X, i.e., $X^* = L(X,\mathbb{K})$. It is always a Banach space with respect to the operator norm. That there are plenty of bounded linear functionals on a normed space is a consequence of the Hahn–Banach theorem.

Theorem A.2 (Hahn–Banach theorem)**.**
Real version. *Let X be a real vector space, M a linear subspace of X, and ℓ a linear functional on M such that $\ell(x) \leq p(x)$ for $x \in M$, where $p : X \to \mathbb{R}$ satisfies $p(x+y) \leq p(x)+p(y)$ and $p(\lambda x) = \lambda p(x)$ for all $x,y \in X$ and $\lambda \geq 0$. Then there is a linear functional $\tilde{\ell}$ on X such that $\tilde{\ell}(x) \leq p(x)$ for all $x \in X$ and $\tilde{\ell}|_M = \ell$.*
Complex version. *Let X be a complex vector space, M a linear subspace of X, and ℓ a complex linear functional on M such that $|\ell(x)| \leq p(x)$ for $x \in M$, where p is a seminorm. Then there is a complex linear functional $\tilde{\ell}$ on X such that $|\tilde{\ell}(x)| \leq p(x)$ for all $x \in X$ and $\tilde{\ell}|_M = \ell$.*

Let M be a closed linear subspace of a normed space X and let $x \in X \setminus M$. Then there exists $\ell \in X^*$ such that $\ell(x) \neq 0$ and $\ell|_M = 0$. Indeed, if we let $\delta := \inf_{y\in M} \|x-y\|$ and define ℓ on $M + \mathbb{C}x$ by setting $\ell(y+\lambda x) := \lambda\delta$, then $|\ell(y+\lambda x)| = |\lambda|\delta \leq |\lambda|\|\lambda^{-1}y + x\| = \|y+\lambda x\|$ and the Hahn–Banach theorem implies the statement.

If we take $M = \{0\}$ and $x \neq 0$ we get $\ell \in X^*$ so that $\ell(x) \neq 0$. Thus, the bounded linear functionals on X separate points: if $x \neq y$ there is $\ell \in X^*$ with $\ell(x-y) \neq 0$, that is $\ell(x) \neq \ell(y)$.

For $x \in X$ we may consider the functional $\mathrm{ev}_x : X^* \to \mathbb{C}$ defined by $\mathrm{ev}_x(\ell) := \ell(x)$. Then the mapping $x \mapsto \mathrm{ev}_x$ is a linear isometry from X into X^{**}, in fact

$$|\mathrm{ev}_x(\ell)| = |\ell(x)| \leq \|\ell\|\|x\|$$

which implies $\|\mathrm{ev}_x\| \leq \|x\|$, on the other hand $\|x\| \leq \|\mathrm{ev}_x\|$, since by the previous paragraphs there is $\ell \in X^*$ such that $\|x\| = \ell(x) = \mathrm{ev}_x(\ell)$.

Since X^{**} is always complete, the closure $\mathrm{cl}(\hat{X})$ of $\hat{X} := \{\mathrm{ev}_x : x \in X\}$ in X^{**} is a Banach space; $\mathrm{cl}(\hat{X})$ is the **completion** of X, and $\mathrm{cl}(\hat{X}) = \hat{X}$ if X is complete. The mapping $x \mapsto \mathrm{ev}_x$ embeds X into $\mathrm{cl}(\hat{X})$ as a dense subspace.

If $\hat{X} = X^{**}$ then X is called **reflexive**. For instance, finite dimensional vector spaces are reflexive, since $\hat{X}$ and X^{**} have the same dimension.

Theorem A.3 (Open mapping theorem)**.** *Let X and Y be Banach spaces. Any surjective bounded linear mapping $T : X \to Y$ is **open**, i.e., T takes open sets to open sets.*

Consequently, a bijective bounded linear mapping between Banach spaces is an isomorphism.

Theorem A.4 (Closed graph theorem)**.** *Let X and Y be Banach spaces. Any **closed** linear mapping $T : X \to Y$, i.e., the graph $\Gamma(T) := \{(x, y) \in X \times Y : y = Tx\}$ is closed in $X \times Y$, is bounded.*

Theorem A.5 (Uniform boundedness principle or Banach–Steinhaus theorem)**.** *Let X be a Banach space, Y a normed space, and let A be a subset of $L(X, Y)$. If $\sup_{T\in A} \|Tx\| < \infty$ for all x in some nonmeager subset of X, then $\sup_{T\in A} \|T\| < \infty$.*

A.3. Hilbert spaces

Let H be a complex vector space. An **inner product** on H is a mapping $H \times H \to \mathbb{C} : (x, y) \mapsto \langle x, y\rangle$ such that

- $\langle ax + by, z\rangle = a\langle x, z\rangle + b\langle y, z\rangle$ for all $x, y \in H$ and $a, b \in \mathbb{C}$,
- $\langle x, y\rangle = \overline{\langle y, x\rangle}$ for all $x, y \in H$,
- $\langle x, x\rangle > 0$ for all $x \neq 0$.

A complex vector space equipped with an inner product is called a **pre-Hilbert space**. In a pre-Hilbert space we set $\|x\| := \sqrt{\langle x, x\rangle}$. Then we have the **Schwarz inequality**

$$|\langle x, y\rangle| \le \|x\|\|y\|, \quad \text{for all } x, y \in H,$$

with equality if and only if x and y are linearly dependent. Indeed, assume without loss of generality that $\|x\| = \|y\| = 1$. Then $\langle x, y\rangle \neq 0$ and $\langle x, y\rangle = a|\langle x, y\rangle|$ for some $a \in \mathbb{C}$ with $|a| = 1$. Now, for $t \in \mathbb{R}$,

$$0 \le \langle a^{-1}x - ty, a^{-1}x - ty\rangle = 1 - 2t\,\mathrm{Re}(a^{-1}\langle x, y\rangle) + t^2 = 1 - 2t|\langle x, y\rangle| + t^2.$$

The right-hand side is minimal for $t = |\langle x, y\rangle|$ and so $|\langle x, y\rangle| \le 1$ as required.

The Schwarz inequality implies that $\|\ \|$ is a norm on H,

$$\|x + y\|^2 = \|x\|^2 + \langle x, y\rangle + \langle y, x\rangle + \|y\|^2 \le (\|x\| + \|y\|)^2.$$

A pre-Hilbert space that is complete with respect to the norm $\|x\| = \sqrt{\langle x, x\rangle}$ is called a **Hilbert space**.

In any pre-Hilbert space we have the **parallelogram law**,

$$\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2), \quad \text{for all } x, y \in H.$$

Lemma A.6. *Any closed convex subset A of a Hilbert space H contains a unique element of smallest norm.*

Proof. Set $\delta := \inf_{x\in A} \|x\|$ and choose a sequence $x_n \in A$ such that $\|x_n\| \to \delta$. By the parallelogram law and convexity of A,

$$\|x_n - x_m\|^2 = 2(\|x_n\|^2 + \|x_m\|^2) - \|x_n + x_m\|^2 \le 2(\|x_n\|^2 + \|x_m\|^2) - 4\delta^2,$$

since $(x_n + x_m)/2 \in A$ and hence $\|x_n + x_m\| \geq 2\delta$. This implies that x_n is Cauchy and so $x_n \to x \in A$, since A is closed. As

$$|\|x_n\| - \|x\|| \leq \|x_n - x\| \to 0$$

we have $\|x\| = \delta$. If there is another $y \in A$ with $\|y\| = \delta$, then by the parallelogram law

$$\|x - y\|^2 = 2\delta^2 - 4\|(x + y)/2\|^2 \leq 0,$$

and hence $x = y$. □

Let H be a Hilbert space, and let A be a subset of H. We define the **orthogonal complement**

$$A^\perp := \{x \in H : \langle x, y\rangle = 0 \text{ for all } y \in A\}.$$

Then $A^\perp$ is a closed linear subspace of H; indeed, if $A^\perp \ni x_n \to x$ and $y \in A$, then

$$|\langle x, y\rangle| = |\langle x_n, y\rangle - \langle x, y\rangle| = |\langle x_n - x, y\rangle| \leq \|x_n - x\|\|y\| \to 0.$$

Proposition A.7. *If M is a closed linear subspace of H, then $H = M \oplus M^\perp$, i.e., each $x \in H$ is of the form $x = y + z$ for unique elements $y \in M$ and $z \in M^\perp$. Moreover, y and z are the unique elements in M and $M^\perp$ whose distance to x is minimal.*

Proof. By Lemma A.6, there is a unique element $y \in M$ such that $\|x - y\| \leq \|x - u\|$ for all $u \in M$. Set $z := x - y$. If $u \in M$, then after multiplication by a nonzero scalar we may assume that $\langle z, u\rangle \in \mathbb{R}$, and then

$$\mathbb{R} \ni t \mapsto \|z + tu\|^2 = \|z\|^2 + 2t\langle z, u\rangle + \|u\|^2$$

is real valued. Since $z + tu = x - (y - tu)$ and $y - tu \in M$, this function has a minimum at $t = 0$ and hence its first derivative vanishes at $t = 0$, that is $\langle z, u\rangle = 0$. It follows that $z \in M^\perp$.

If $z' \in M^\perp$ then $\|x - z'\|^2 = \|x - z\|^2 + \|z - z'\|^2 \geq \|x - z\|^2$, and thus z is the unique element in $M^\perp$ with minimal distance to x.

If $x = y' + z'$ with $y' \in M$ and $z' \in M^\perp$, then $y - y' = z - z' \in M \cap M^\perp$ and so $y - y' = z - z' = 0$. □

Theorem A.8 (Riesz)**.** *Let ℓ be a bounded linear functional on a Hilbert space H. Then there is a unique $y \in H$ such that $\ell(x) = \langle x, y\rangle$ for all $x \in H$.*

Proof. If $\ell = 0$ choose $y = 0$. Otherwise $M = \ker \ell$ is a proper closed subspace of H and there exists a unit vector $z \in M^\perp$, by Proposition A.7. Since $\ell(x)z - \ell(z)x \in M$, for each $x \in H$, we find

$$0 = \langle \ell(x)z - \ell(z)x, z\rangle = \ell(x) - \ell(z)\langle x, z\rangle,$$

i.e., $y := \overline{\ell(z)}z$ is as required.

If $u \in H$ so that $\ell(x) = \langle x, u\rangle$ for all $x \in H$, then $\langle x, v - u\rangle = 0$ for all x, and hence $u = v$. □

For $y \in H$, $\ell_y(x) := \langle x, y\rangle$ defines a bounded linear functional on H satisfying $\|\ell_y\| = \|y\|$, by the Schwarz inequality. So the mapping $y \mapsto \ell_y$ is a conjugate-linear isometry from H onto H^*, by Theorem A.8. It follows that a Hilbert space H is reflexive in a strong sense: H is naturally isomorphic to H^*, not only to H^{**}.

A subset $\{x_\alpha\}_{\alpha \in A}$ of a pre-Hilbert space H is called **orthonormal** if

$$\langle x_\alpha, x_\beta\rangle = \begin{cases} 1 & \alpha = \beta \\ 0 & \alpha \neq \beta \end{cases}.$$

An orthonormal set $\{x_\alpha\}_{\alpha\in A}$ and any element x satisfy **Bessel's inequality**:

$$\sum_{\alpha\in A} |\langle x, x_\alpha\rangle|^2 \le \|x\|^2,$$

where the sum is understood as $\sup\big\{\sum_{\alpha\in A_0} |\langle x, x_\alpha\rangle|^2 : A_0 \subseteq A \text{ finite}\big\}$. Indeed, let M be the linear subspace generated by $\{x_\alpha\}_{\alpha\in A_0}$. By Proposition A.7, $x = \sum_{\alpha\in A_0} c_\alpha x_\alpha + y$ for $y \in M^\perp$, and so $\langle x, x_\alpha\rangle = c_\alpha$ and $\|x\|^2 = \sum_{\alpha\in A_0} |c_\alpha|^2 + \|y\|^2$, by orthonormality. It follows that the sum in Bessel's inequality has only countably many nonzero terms.

An orthonormal set $\{x_\alpha\}_{\alpha\in A}$ in a Hilbert space H is called **complete** if its orthogonal complement is $\{0\}$. If $\{x_\alpha\}_{\alpha\in A}$ is a complete orthonormal set $\{x_\alpha\}_{\alpha\in A}$ in H, then each $x \in H$ can be written in the form

$$x = \sum_{\alpha\in A} \langle x, x_\alpha\rangle x_\alpha, \tag{A.1}$$

where the sum has only countably many nonzero terms and converges in the norm topology. To see this, let us enumerate by $\alpha_1, \alpha_2, \ldots$ the indices α for which $\langle x, x_\alpha\rangle \ne 0$. By Bessel's inequality, the series $\sum_{i=1}^\infty |\langle x, x_{\alpha_i}\rangle|^2$ converges, and hence $\|\sum_{i=m+1}^n \langle x, x_{\alpha_i}\rangle x_{\alpha_i}\|^2 = \sum_{i=m+1}^n |\langle x, x_{\alpha_i}\rangle|^2 \to 0$ as $m, n \to \infty$. So $\sum_{i=1}^\infty \langle x, x_{\alpha_i}\rangle x_{\alpha_i}$ converges, since H is complete. The difference $x - \sum_{i=1}^\infty \langle x, x_{\alpha_i}\rangle x_{\alpha_i}$ is zero, because $\{x_\alpha\}_{\alpha\in A}$ is complete, and (A.1) is shown. From this we obtain **Parseval's identity**

$$\|x\|^2 = \lim_{n\to\infty} \|\sum_{i=1}^n \langle x, x_{\alpha_i}\rangle x_{\alpha_i}\|^2 = \lim_{n\to\infty} \sum_{i=1}^n |\langle x, x_{\alpha_i}\rangle|^2 = \sum_{\alpha\in A} |\langle x, x_{\alpha_i}\rangle|^2,$$

which in turn implies that $\{x_\alpha\}_{\alpha\in A}$ is complete. For this reason a complete orthonormal set in a Hilbert space is called a **Hilbert basis**.

Every Hilbert space has a Hilbert basis. For by Zorn's lemma there exists a maximal orthonormal set and it is easy to see that it must be complete. A Hilbert space is separable if and only if it has a countable Hilbert basis (then every Hilbert basis is countable). This can be proved using the Gram–Schmidt orthonormalization process; in this case the existence of a Hilbert basis follows without using Zorn's lemma.

An invertible linear mapping $U : H_1 \to H_2$ between Hilbert spaces that preserves inner products, i.e.,

$$\langle Ux, Uy\rangle = \langle x, y\rangle, \qquad \text{for all } x, y \in H_1,$$

is called **unitary**. Unitary mappings are isometries, and conversely, surjective isometries between Hilbert spaces are unitary which follows from the **polarization identity**

$$4\langle x, y\rangle = \|x + y\|^2 + \|x - y\|^2 + i\|x + iy\|^2 - i\|x - iy\|^2.$$

Let H be a Hilbert space with Hilbert basis $\{x_\alpha\}_{\alpha\in A}$. For $x \in H$ consider the element $\hat{x}$ in the Hilbert space $\ell^2(A)$ defined by $\hat{x}(\alpha) := \langle x, x_\alpha\rangle$. The mapping $x \mapsto \hat{x}$ is an isometry from H to $\ell^2(A)$ by Parseval's identity. It is surjective and thus also unitary. For if $f \in \ell^2(A)$ then $\sum_{\alpha\in A} |f(\alpha)|^2 < 0$ and so the partial sums of the series $\sum f(\alpha)x_\alpha$ form a Cauchy sequence (by similar arguments as before). Then $x := \sum f(\alpha)x_\alpha$ exists in H and $\hat{x} = f$. This implies the following theorem.

Theorem A.9. *All separable infinite dimensional Hilbert spaces are isomorphic to* $\ell^2(\mathbb{N})$.

A.4. Fréchet spaces

A **topological vector space** is a vector space endowed with a topology in which addition and multiplication by scalars are continuous. A **locally convex space** is a topological vector space whose topology has a basis consisting of convex sets.

If X is a vector space and $\{p_\alpha\}_{\alpha\in A}$ is a family of seminorms on X, then the topology generated by the balls $B(x,\alpha,\epsilon) := \{y \in X : p_\alpha(x-y) < \epsilon\}$, for $x \in X$, $\alpha \in A$, and $\epsilon > 0$, turns X into a locally convex space. Actually, in every locally convex space the topology can be defined by means of a family of seminorms $\{p_\alpha\}_{\alpha\in A}$.

Let $T : X \to Y$ be a linear mapping between locally convex spaces X and Y with topologies defined by families $\{p_\alpha\}_{\alpha\in A}$ and $\{q_\beta\}_{\beta\in B}$ of seminorms, respectively. Then T is continuous if and only if for each $\beta \in B$ there are $\alpha_1, \dots, \alpha_n \in A$ and $C > 0$ such that $q_\beta(Tx) \le C\sum_{i=1}^n p_{\alpha_i}(x)$.

A locally convex space X with topologies definded by a family $\{p_\alpha\}_{\alpha\in A}$ of seminorms is Hausdorff if and only if for each $x \neq 0$ there exists $\alpha \in A$ so that $p_\alpha(x) \neq 0$. If X is Hausdorff and A is countable, then the topology of X is given by the translation invariant metric

$$d(x,y) := \sum_\alpha 2^{-\alpha}\frac{p_\alpha(x-y)}{1+p_\alpha(x-y)};$$

we say that X is **metrizable**. A complete Hausdorff locally convex space whose topology is defined by a countable family of seminorms is called a **Fréchet space**. The open mapping theorem and the closed graph theorem remain valid for Fréchet spaces.

Bibliography

1. L. Ambrosio, N. Fusco, and D. Pallara, *Functions of bounded variation and free discontinuity problems*, Oxford Mathematical Monographs, The Clarendon Press Oxford University Press, New York, 2000.
2. J. B. Conway, *A course in functional analysis*, Graduate Texts in Mathematics, vol. 96, Springer-Verlag, New York, 1985.
3. L. C. Evans and R. F. Gariepy, *Measure theory and fine properties of functions*, Studies in Advanced Mathematics, CRC Press, Boca Raton, FL, 1992. MR 1158660 (93f:28001)
4. H. Federer, *Geometric measure theory*, Die Grundlehren der mathematischen Wissenschaften, Band 153, Springer-Verlag New York Inc., New York, 1969. MR 0257325 (41 #1976)
5. G. B. Folland, *Real analysis*, second ed., Pure and Applied Mathematics (New York), John Wiley & Sons Inc., New York, 1999, Modern techniques and their applications, A Wiley-Interscience Publication. MR 1681462 (2000c:00001)
6. L. Grafakos, *Classical Fourier analysis*, second ed., Graduate Texts in Mathematics, vol. 249, Springer, New York, 2008. MR 2445437 (2011c:42001)
7. Y. Katznelson, *An introduction to harmonic analysis*, third ed., Cambridge Mathematical Library, Cambridge University Press, Cambridge, 2004. MR 2039503 (2005d:43001)
8. S. G. Krantz, *Explorations in harmonic analysis*, Applied and Numerical Harmonic Analysis, Birkhäuser Boston, Inc., Boston, MA, 2009, With applications to complex function theory and the Heisenberg group, With the assistance of Lina Lee. MR 2508404 (2011a:43017)
9. S. G. Krantz and H. R. Parks, *Geometric integration theory*, Cornerstones, Birkhäuser Boston, Inc., Boston, MA, 2008. MR 2427002
10. G. Leoni, *A first course in Sobolev spaces*, Graduate Studies in Mathematics, vol. 105, American Mathematical Society, Providence, RI, 2009. MR 2527916 (2010m:46049)
11. E. H. Lieb and M. Loss, *Analysis*, second ed., Graduate Studies in Mathematics, vol. 14, American Mathematical Society, Providence, RI, 2001. MR 1817225 (2001i:00001)
12. W. Rudin, *Real and complex analysis*, third ed., McGraw-Hill Book Co., New York, 1987.
13. E. M. Stein and R. Shakarchi, *Real analysis*, Princeton Lectures in Analysis, III, Princeton University Press, Princeton, NJ, 2005, Measure theory, integration, and Hilbert spaces. MR 2129625 (2005k:28024)
14. F. Trèves, *Topological vector spaces, distributions and kernels*, Academic Press, New York, 1967. MR 0225131 (37 #726)
15. S. Wagon, *Mathematica in action. Problem solving through visualization and computation. With CD-ROM. 3rd ed.*, 3rd ed. ed., New York, NY: Springer, 2010 (English).

Index