Analysis I

Piotr Hajłasz

1 Measure theory

1.1 σ-algebra.

DEFINITION. Let X be a set. A collection $\mathfrak{M}$ of subsets of X is *σ-algebra* if $\mathfrak{M}$ has the following properties

1. $X \in \mathfrak{M}$;
2. $A \in \mathfrak{M} \implies X \setminus A \in \mathfrak{M}$;
3. $A_1, A_2, A_3, \ldots \in \mathfrak{M} \implies \bigcup_{i=1}^{\infty} A_i \in \mathfrak{M}$.

The pair $(X, \mathfrak{M})$ is called *measurable space* and elements of $\mathfrak{M}$ are called *measurable sets.*

Exercise. *Let $(X, \mathfrak{M})$ be a measurable space. Show that*

1. $\emptyset \in \mathfrak{M}$*;*
2. $A, B \in \mathfrak{M} \implies A \setminus B \in \mathfrak{M}$*;*
3. $\mathfrak{M}$ *is closed under finite unions, finite intersections and countable intersections.*

EXAMPLES.

1. 2^X, the family of all subsets of X is a σ-algebra.
2. $\mathfrak{M} = \{\emptyset, X\}$ is a σ-algebra.
3. If $E \subset X$ is a fixed set, then $\mathfrak{M} = \{\emptyset, X, E, X \setminus E\}$ is a σ-algebra.

Proposition 1 *If $\{\mathfrak{M}_i\}_{i\in I}$ is a family of σ-algebras, then*

$$\mathfrak{M} = \bigcap_{i\in I} \mathfrak{M}_i$$

is a σ-algebra.

A simple proof is left to the reader.

Let $\mathcal{R}$ be a family of subsets of X. By $\sigma(\mathcal{R})$ we will denote the intersection of all σ-algebras that contain $\mathcal{R}$. Note that there is at least one σ-algebra that contains $\mathcal{R}$, namely 2^X.

1. $\sigma(\mathcal{R})$ is a σ-algebra that contains $\mathcal{R}$
2. $\sigma(\mathcal{R})$ is the smallest σ-algebra that contains $\mathcal{R}$ in the sense that if $\mathfrak{M}$ is a σ-algebra that contains $\mathcal{R}$, then $\sigma(\mathcal{R}) \subset \mathfrak{M}$.

We say that $\sigma(\mathcal{R})$ is a *σ-algebra generated by* $\mathcal{R}$.

EXAMPLE. If $\mathcal{R} = \{E\}$, then $\sigma(\mathcal{R}) = \{\emptyset, X, E, X \setminus E\}$.

In the sequel we will need the following result.

Proposition 2 *The family $\mathfrak{M}$ of subsets of a set X is a σ-algebra if and only if it satisfies the following three properties*

1. *$X \in \mathfrak{M}$;*
2. *If $A, B \in \mathfrak{M}$, then $A \setminus B \in \mathfrak{M}$.*
3. *If $A_1, A_2, A_3, \ldots \in \mathfrak{M}$ are pairwise disjoint, then $\bigcup_{i=1}^\infty A_i \in \mathfrak{M}$.*

We leave the proof as an exercise cf. the proof of Theorem 3(d). $\square$

1.2 Borel sets.

Let X be a metric space (or more generally a topological space). By $\mathfrak{B}(X)$ we will denote the σ-algebra generated by the family of all open sets in X. Elements of $\mathfrak{B}(X)$ are called *Borel sets* and $\mathfrak{B}(X)$ is called *σ-algebra of Borel sets*.

The following properties are obvious.

1. If $U_1, U_2, U_3, \ldots \subset X$ are open, then $\bigcap_{i=1}^\infty U_i$ is a Borel set;
2. All closed sets are Borel;
3. If $F_1, F_2, F_3, \ldots \subset X$ are closed, then $\bigcup_{i=1}^\infty F_i$ is a Borel set.

1.3 Measure.

Definition. Let $(X, \mathfrak{M})$ be a measurable space. A *measure* (called also *positive measure*) is a function

$$\mu : \mathfrak{M} \to [0, \infty]$$

such that

1. $\mu(\emptyset) = 0$;
2. μ is *countably additive* i.e. if $A_1, A_2, A_3, \ldots \in \mathfrak{M}$ are pairwise disjoint, then

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

The triple $(X, \mathfrak{M}, \mu)$ is called *measure space.*

If $\mu(X) < \infty$, then μ is called *finite measure.*

If $\mu(X) = 1$, then μ is called *probability* or *probability measure.*

If we can write $X = \bigcup_{i=1}^{\infty} A_i$, where $A_i \in \mathfrak{M}$ and $\mu(A_i) < \infty$ for all $i = 1, 2, 3, \ldots$, then we say that μ is σ*-finite.*

Example. If X is an arbitrary set, and $\mu : 2^X \to [0, \infty]$ is defined by $\mu(E) = m$ if E is finite and has m elements, $\mu(E) = \infty$ if E is infinite, then μ is a measure. It is called *counting measure.*

Theorem 3 (Elementary properties of measures) *Let $(X, \mathfrak{M}, \mu)$ be a measure space. Then*

(a) If the sets $A_1, A_2, \ldots, A_n \in \mathfrak{M}$ are pairwise disjoint then $\mu(A_1 \cup \ldots \cup A_n) = \mu(A_1) + \ldots + \mu(A_n)$.

(b) If $A, B \in \mathfrak{M}$, $A \subset B$ and $\mu(B) < \infty$, then $\mu(B \setminus A) = \mu(B) - \mu(A)$.

(c) If $A, B \in \mathfrak{M}, A \subset B$, then $\mu(A) \leq \mu(B)$.

(d) If $A_1, A_2, A_3, \ldots \in \mathfrak{M}$, then

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) \leq \sum_{i=1}^{\infty} \mu(A_i).$$

(e) If $A_1, A_2, A_3, \ldots \in \mathfrak{M}$, $\mu(A_i) = 0$, for $i = 1, 2, 3, \ldots$ then $\mu(\bigcup_{i=1}^{\infty} A_i) = 0$.

(f) If $A_1, A_2, A_3, \ldots \in \mathfrak{M}$, $A_1 \subset A_2 \subset A_3 \subset \ldots$ then

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \lim_{i\to\infty} \mu(A_i).$$

(g) If $A_1, A_2, A_3, \ldots \in \mathfrak{M}$, $A_1 \supset A_2 \supset A_3 \supset \ldots$ and $\mu(A_1) < \infty$, then

$$\mu\left(\bigcap_{i=1}^{\infty} A_i\right) = \lim_{i\to\infty} \mu(A_i).$$

Proof.

(a) The sets $A_1, A_2, \ldots, A_n, \emptyset, \emptyset, \emptyset, \ldots$ are pairwise disjoint. Hence

$$\begin{aligned}\mu(A_1 \cup \ldots \cup A_n) &= \mu(A_1 \cup \ldots \cup A_n \cup \emptyset \cup \emptyset \cup \emptyset \cup \ldots) \\ &= \mu(A_1) + \ldots + \mu(A_n) + \mu(\emptyset) + \mu(\emptyset) + \mu(\emptyset) + \ldots \\ &= \mu(A_1) + \ldots + \mu(A_n).\end{aligned}$$

(b) Since $B = A \cup (B \setminus A)$ and the sets A, $B \setminus A$ are disjoint we have

$$\mu(B) = \mu(A) + \mu(B \setminus A) \tag{1}$$

and the claim easily follows.

(c) The claim follows from (1).

(d) We have

$$\begin{aligned}A &= A_1 \cup A_2 \cup A_3 \cup \ldots \\ &= \underbrace{A_1}_{B_1} \cup \underbrace{(A_2 \setminus A_1)}_{B_2} \cup \underbrace{(A_3 \setminus (A_1 \cup A_2))}_{B_3} \cup \underbrace{(A_4 \setminus (A_1 \cup A_2 \cup A_3))}_{B_4} \cup \ldots\end{aligned}$$

The sets B_i are pairwise disjoint and $B_i \subset A_i$. Hence

$$\mu(A) = \sum_{i=1}^{\infty} \mu(B_i) \le \sum_{i=1}^{\infty} \mu(A_i).$$

(e) It easily follows from (d).

(f) Since $A_i \subset A_{i+1}$, we have

$$\begin{aligned}A &= A_1 \cup A_2 \cup A_3 \cup \ldots \\ &= \underbrace{A_1}_{B_1} \cup \underbrace{(A_2 \setminus A_1)}_{B_2} \cup \underbrace{(A_3 \setminus A_2)}_{B_3} \cup \underbrace{(A_4 \setminus A_3)}_{B_4} \cup \ldots\end{aligned}$$

The sets B_i are pairwise disjoint and hence

$$\mu(A) = \sum_{i=1}^{\infty} \mu(B_i) = \lim_{i\to\infty} (\mu(B_1) + \ldots + \mu(B_i)) = \lim_{i\to\infty} \mu(B_1 \cup \ldots \cup B_i) = \lim_{i\to\infty} \mu(A_i).$$

The limit exist because $\mu(A_i) \leq \mu(A_{i+1})$ (since $A_i \subset A_{i+1}$).

(g) It suffice to apply (f) to the sets $A_1 \setminus A_i$.

The proof is complete. □

Exercise. *Provide a detailed proof of (g). Where do we use the assumption that $\mu(A_1) < \infty$? Show an example that the conclusion of (g) need not be true without the assumption that $\mu(A_1) < \infty$.*

1.4 Outer measure and Carathéodory construction.

It is quite difficult to construct a measure with desired properties, but it is much easier to construct so called outer measure which has less restrictive properties. The Carathéodory theorem shows then how to extract a measure from an outer measure.

DEFINITION. Let X be a set. A function

$$\mu^* : 2^X \to [0, \infty]$$

is called *outer measure* if

1. $\mu^*(\emptyset) = 0$;
2. $A \subset B \implies \mu^*(A) \leq \mu^*(B)$;
3. for all sets $A_1, A_2, A_3, \ldots \subset X$

$$\mu^*\left(\bigcup_{i=1}^{\infty} A_i\right) \leq \sum_{i=1}^{\infty} \mu^*(A_i).$$

We call a set $E \subset X$ *μ^*-measurable* if

$$\forall A \subset X \quad \mu^*(A) = \mu^*(A \cap E) + \mu^*(A \setminus E). \tag{2}$$

Condition (2) is called *Carathéodory condition.* Since the inequality

$$\mu^*(A) \leq \mu^*(A \cap E) + \mu^*(A \setminus E)$$

is always satisfied, in order to verify measurability of a set E it suffices to show that

$$\mu^*(A) \geq \mu^*(A \cap E) + \mu^*(A \setminus E).$$

Proposition 4 *All sets with $\mu^*(E)=0$ are μ^* measurable.*

Proof. If $\mu^*(E)=0$, then for an arbitrary set $A\subset X$ we have

$$\mu^*(A)\geq\mu^*(A\setminus E)=\mu^*(A\setminus E)+\underbrace{\mu^*(A\cap E)}_{0},$$

because $A\cap E\subset E$. □

Let $\mathfrak{M}^*$ be the class of all μ^*-measurable sets.

Theorem 5 (Carathéodory) $\mathfrak{M}^*$ *is a σ-algebra and*

$$\mu^*:\mathfrak{M}^*\to[0,\infty]$$

is a measure.

Proof. We will split the proof into several steps.

Step I: $X\in\mathfrak{M}^*$

This is obvious.

Step II: $E,F\in\mathfrak{M}^* \implies E\cup F\in\mathfrak{M}^*$.

Let $A\subset X$. We have

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

Since $A\cap E=A\cap(E\cup F)\cap E$ and $(A\setminus E)\cap F=(A\cap(E\cup F))\setminus E$, we have

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

Step III: $E,F\in\mathfrak{M}^* \implies E\setminus F\in\mathfrak{M}^*$.

It follows from the symmetry of the condition (2) that $X\setminus E\in\mathfrak{M}^*$. Since $E\setminus F=X\setminus((X\setminus E)\cup F)$ the claim follows.

Step IV: If the sets $E_1,E_2,E_3,\ldots\in\mathfrak{M}^*$ are pairwise disjoint, then for every $A\subset X$

$$\mu^*(A\cap\bigcup_{i=1}^{\infty}E_i)=\sum_{i=1}^{\infty}\mu^*(A\cap E_i).$$

If $E,F\in\mathfrak{M}^*$ are pairwise disjoint, the for every $A\subset X$

$$\mu^*(A\cap(E\cup F))=\mu^*(A\cap(E\cup F)\cap E)+\mu^*(A\cap(E\cup F)\setminus E)=\mu^*(A\cap E)+\mu^*(A\cap F).$$

This, Step II and the induction argument implies that for every n

$$\mu^*(A \cap \bigcup_{i=1}^{n} E_i) = \sum_{i=1}^{n} \mu^*(A \cap E_i). \tag{3}$$

Hence

$$\mu^*(A \cap \bigcup_{i=1}^{\infty} E_i) \geq \sum_{i=1}^{n} \mu^*(A \cap E_i)$$

and thus in the limit

$$\mu^*(A \cap \bigcup_{i=1}^{\infty} E_i) \geq \sum_{i=1}^{\infty} \mu^*(A \cap E_i).$$

Since the opposite inequality is a property of an outer measure, the claim follows.

STEP V: If the sets $E_1, E_2, E_3, \ldots \in \mathfrak{M}^*$ are pairwise disjoint, then $\bigcup_{i=1}^{\infty} E_i \in \mathfrak{M}^*$.

Step II and the induction argument implies that for every n

$$\bigcup_{i=1}^{n} E_i \in \mathfrak{M}^*.$$

Hence (3) gives

$$\mu^*(A) = \mu^*\left(A \cap \bigcup_{i=1}^{n} E_i\right) + \mu^*\left(A \setminus \bigcup_{i=1}^{n} E_i\right) \geq \sum_{i=1}^{n} \mu^*(A \cap E_i) + \mu^*\left(A \setminus \bigcup_{i=1}^{\infty} E_i\right),$$

and in the limit

$$\mu^*(A) \geq \sum_{i=1}^{\infty} \mu^*(A \cap E_i) + \mu^*\left(A \setminus \bigcup_{i=1}^{\infty} E_i\right) = \mu^*\left(A \cap \bigcup_{i=1}^{\infty} E_i\right) + \mu^*\left(A \setminus \bigcup_{i=1}^{\infty} E_i\right).$$

Since the opposite inequality is a property of an outer measure the claim follows.

STEP VI: Final step. It follows from Steps I, III, V and Proposition 2 that $\mathfrak{M}^*$ is a σ-algebra and then Step IV with $A = X$ implies that μ^* restricted to $\mathfrak{M}^*$ is a measure. The proof is complete. □

We say that a measure $\mu : \mathfrak{M} \to [0, \infty]$ is *complete* if every subset of a set of measure zero is measurable (and hence has measure zero). Therefore it follows from Proposition 4 that the measure described by the Carathéodory theorem is complete.

Let (X, d) be a metric space. For $E, F \subset X$ we define

$$\operatorname{dist}(E, F) = \inf_{x \in E, y \in F} d(x, y), \qquad \operatorname{diam} E = \sup_{x, y \in E} d(x, y).$$

DEFINITION. An outer measure μ^* defined on subsets of a metric space is called *metric outer measure* if

$$\mu^*(E \cup F) = \mu^*(E) + \mu^*(F) \quad \text{whenever } \operatorname{dist}(E, F) > 0.$$

Theorem 6 *If μ^* is a metric outer measure, then all Borel sets are μ^*-measurable i.e. $\mathfrak{B}(X) \subset \mathfrak{M}^*$.*

Proof. Since $\mathfrak{M}^*$ is a σ-algebra, it suffices to show that all open sets belong to $\mathfrak{M}^*$. Let $G \subset X$ be open. It suffices to show that for every $A \subset X$

$$\mu^*(A) \geq \mu^*(A \cap G) + \mu^*(A \setminus G) \tag{4}$$

(because the opposite inequality is obvious). We can assume that $\mu^*(A) < \infty$ as otherwise (4) is obviously satisfied.

For each positive integer n define

$$G_n = \{x \in G : \operatorname{dist}(x, X \setminus G) > \frac{1}{n}\}.$$

Then

$$\operatorname{dist}(G_n, X \setminus G) \geq \frac{1}{n} > 0. \tag{5}$$

Let

$$D_n = G_{n+1} \setminus G_n = \{x \in G : \frac{1}{n+1} < \operatorname{dist}(x, X \setminus G) \leq \frac{1}{n}\}.$$

Clearly

$$G \setminus G_n = \bigcup_{i=n}^{\infty} D_i \tag{6}$$

and

$$\operatorname{dist}(D_i, D_j) \geq \frac{1}{i+1} - \frac{1}{j} > 0 \quad \text{provided } i + 2 \leq j.$$

Since the mutual distances between the sets $D_1, D_3, D_5, \ldots, D_{2n-1}$ are positive we have

$$\mu^*(A \cap D_1) + \mu^*(A \cap D_3) + \ldots \mu^*(A \cap D_{2n-1}) = \mu^*(A \cap (D_1 \cup D_3 \cup \ldots \cup D_{2n-1})) \leq \mu^*(A)$$

and similarly

$$\mu^*(A \cap D_2) + \mu^*(A \cap D_4) + \ldots + \mu^*(A \cap D_{2n}) \leq \mu^*(A).$$

Hence

$$\sum_{i=1}^{\infty} \mu^*(A \cap D_i) \leq 2\mu^*(A) < \infty.$$

Now (6) yields

$$\mu^*(A \cap (G \setminus G_n)) \leq \sum_{i=n}^{\infty} \mu^*(A \cap D_i)$$

and hence

$$\mu^*(A \cap (G \setminus G_n)) \to 0 \quad \text{as } n \to \infty.$$

Inequality (5) gives

$$\mu^*(A\cap G_n)+\mu^*(A\setminus G)=\mu^*((A\cap G_n)\cup(A\setminus G))\le\mu^*(A)$$

and thus

$$\begin{aligned}\mu^*(A\cap G)+\mu^*(A\setminus G) &\le \mu^*(A\cap G_n)+\mu^*(A\cap(G\setminus G_n))+\mu^*(A\setminus G)\\ &\le \mu^*(A)+\mu^*(A\cap(G\setminus G_n))\end{aligned}$$

and after passing to the limit

$$\mu^*(A\cap G)+\mu^*(A\setminus G)\le\mu^*(A).$$

The proof is complete. □

We will use outer measures to prove the following important result.

Theorem 7 *Let X be a metric space and μ a measure in $\mathfrak{B}(X)$. Suppose that X is a union of countably many open sets of finite measure. Then*

$$\mu(E)=\inf_{\substack{U\supset E\\ U-\text{open}}}\mu(U)=\sup_{\substack{C\subset E\\ C-\text{closed}}}\mu(C). \tag{7}$$

Before proving the theorem let's discuss two of its corollaries. The first one shows that in a situation described by the above theorem in order to prove that two measures are equal it suffices to compare them on the class of open sets.

Corollary 8 *If μ is as in Theorem 7 and ν is another measure on $\mathfrak{B}(X)$ that satisfies*

$$\nu(U)=\mu(U)\quad\textit{for all open sets } U\subset X$$

then

$$\nu(E)=\mu(E)\quad\textit{for all } E\subset\mathfrak{B}(X).$$

The second corollary describes an important class of measures satisfying assumptions of theorem 7.

Definition. Let X be a metric space and μ a measure defined on the σ-algebra of Borel sets. We say that μ is a *Radon measure* if $\mu(K)<\infty$ for all compact sets and

$$\mu(E)=\inf_{\substack{U\supset E\\ U-\text{open}}}\mu(U)=\sup_{\substack{K\subset E\\ K-\text{compact}}}\mu(K)\quad\text{for all } E\in\mathfrak{B}(X). \tag{8}$$

Corollary 9 *If X is a locally compact and separable metric space*[1] *and μ is a measure in $\mathfrak{B}(X)$ such that $\mu(K)<\infty$ for every compact set K, then X is a union of countably many open sets of finite measure and μ is a Radon measure.*

[1]Being *locally compact* means that ever point has a neighborhood whose closure is compact. $\mathbb{R}^n$ is locally compact, but also $\mathbb{R}^n\setminus\{0\}$ is locally compact.

Proof. It follows from locally compactness of X that X is a union of a family of open sets with compact closures. Since X is separable, every covering of X by open sets has countable subcovering and hence we can write

$$X = \bigcup_{i=1}^{\infty} U_i, \qquad \overline{U_i} \text{ — compact.}$$

This proves the first part of the corollary because $\mu(U_i) \leq \mu(\overline{U_i}) < \infty$. Now (8) follows from (7) and Theorem 3(f) because

$$C = \bigcup_{n=1}^{\infty} \underbrace{\left(C \cap \bigcup_{i=1}^{n} \overline{U_i} \right)}_{\text{compact}}.$$

and hence

$$\mu(C) = \sup_{\substack{K \subset C \\ K-\text{compact}}} \mu(K)$$

for every closed set C. The proof is complete □

Proof of Theorem 7. For $E \subset X$ we set

$$\mu^*(E) = \inf_{\substack{U \supset E \\ U-\text{open}}} \mu(U).$$

It is easy to see that μ^* is a metric outer measure. Hence μ^* restricted to the class of Borel sets is a measure. Clearly

$$\mu(U) = \mu^*(U) \quad \text{for all open sets } U \subset X, \tag{9}$$

and

$$\mu(E) \leq \mu^*(E) \quad \text{for all } E \in \mathfrak{B}(X). \tag{10}$$

We can assume that X is a union of an increasing sequence of open sets with finite measure. Indeed, if X is the union of open sets U_i of finite measure, then the sets $V_n = U_1 \cup U_2 \cup \ldots \cup U_n$ satisfy

$$X = \bigcup_{n=1}^{\infty} V_n, \qquad V_n \subset V_{n+1}, \qquad \mu(V_n) < \infty.$$

Now inequality (10) implies that for all $E \in \mathfrak{B}(X)$

$$\mu(V_n \setminus E) \leq \mu^*(V_n \setminus E) \quad \text{and} \quad \mu(V_n \cap E) \leq \mu^*(V_n \cap E), \tag{11}$$

because both $V_n \setminus E$ and $V_n \cap E$ are Borel. Actually we have equalities in both inequalities of (11) because otherwise we would have a sharp inequality

$$\mu(V_n) = \mu(V_n \setminus E) + \mu(V_n \cap E) < \mu^*(V_n \setminus E) + \mu^*(V_n \cap E) = \mu^*(V_n),$$

which contradicts (9). Hence in particular $\mu(V_n \cap E) = \mu^*(V_n \cap E)$. Since

$$\mu(V_n \cap E) \to \mu(E) \quad \text{and} \quad \mu^*(V_n \cap E) \to \mu^*(E)$$

as $n \to \infty$ by Theorem 3(f) we conclude that $\mu(E) = \mu^*(E)$ and thus

$$\mu(E) = \inf_{\substack{U \supset E \\ U-\text{open}}} \mu(U) \tag{12}$$

by the definition of μ^*. To prove that the measure of a set $E \in \mathfrak{B}(X)$ can be approximated by measures of closed subsets of E observe that $V_n \setminus E$ has finite measure and hence it follows from (12) that there is an open set G_n such that

$$V_n \setminus E \subset G_n, \qquad \mu(G_n \setminus (V_n \setminus E)) < \frac{\varepsilon}{2^n}.$$

The set $G = \bigcup_{n=1}^\infty G_n$ is open and $C = X \setminus G \subset E$ is closed. Now it suffices to observe that

$$E \setminus C = E \cap \bigcup_{n=1}^\infty G_n \subset \bigcup_{n=1}^\infty G_n \setminus (V_n \setminus E)$$

and hence $\mu(E \setminus C) < \varepsilon$. The proof is complete. □

1.5 Hausdorff measure.

Let $\omega_s = \pi^{s/2}/\Gamma(1 + \frac{s}{2})$, $s \geq 0$. If $s = n$ is a positive integer, then ω_n is volume of the unit ball in $\mathbb{R}^n$.

Let X be a metric space. For $\varepsilon > 0$ and $E \subset X$ we define

$$\mathcal{H}^s_\varepsilon(E) = \inf \frac{\omega_s}{2^s} \sum_{i=1}^\infty (\operatorname{diam} A_i)^s$$

where the infimum is taken over all possible coverings

$$E \subset \bigcup_{i=1}^\infty A_i \quad \text{with } \operatorname{diam} A_i < \varepsilon.$$

Since the function $\varepsilon \mapsto \mathcal{H}^s_\varepsilon(E)$ is nonincreasing, the limit

$$\mathcal{H}^s(E) = \lim_{\varepsilon \to 0} \mathcal{H}^s_\varepsilon(E)$$

exists. $\mathcal{H}^s$ is called *Hausdorff measure.*

It is easy to see that if $s = 0$, $\mathcal{H}^0$ is the counting measure.

Theorem 10 *$\mathcal{H}^s$ is a metric outer measure.*

Proof. Clearly

$$\mathcal{H}^s : 2^X \to [0,\infty],$$
$$\mathcal{H}^s(\emptyset) = 0,$$
$$A \subset B \implies \mathcal{H}^s(A) \le \mathcal{H}^s(B).$$

In order to show that $\mathcal{H}^s$ is an outer measure it remains to show that

$$\mathcal{H}^s\left(\bigcup_{n=1}^{\infty} E_n\right) \le \sum_{n=1}^{\infty} \mathcal{H}^s(E_n).$$

We can assume that the right hand side is finite. Hence

$$\sum_{n=1}^{\infty} \mathcal{H}^s_\varepsilon(E_n) < \infty \quad \text{for every } \varepsilon > 0.$$

Fix $\delta > 0$. We can find a covering of each set E_n

$$E_n \subset \bigcup_{i=1}^{\infty} A_{ni}, \quad \operatorname{diam} A_{ni} < \varepsilon$$

such that

$$\mathcal{H}^s_\varepsilon(E_n) \ge \frac{\omega_s}{2^s} \sum_{i=1}^{\infty} (\operatorname{diam} A_{ni})^s - \frac{\delta}{2^n}$$

Hence

$$\sum_{n=1}^{\infty} \mathcal{H}^s_\varepsilon(E_n) \ge \frac{\omega_s}{s^s} \sum_{i,n=1}^{\infty} (\operatorname{diam} A_{ni})^s - \delta \ge \mathcal{H}^s_\varepsilon\left(\bigcup_{n=1}^{\infty} E_n\right) - \delta,$$

because $\{A_{ni}\}_{i,n=1}^{\infty}$ forms a covering of $\bigcup_{n=1}^{\infty} E_n$. Passing to the limit with $\delta \to 0$ yields

$$\sum_{n=1}^{\infty} \mathcal{H}^s(E_n) \ge \sum_{n=1}^{\infty} \mathcal{H}^s_\varepsilon(E_n) \ge \mathcal{H}^s_\varepsilon\left(\bigcup_{n=1}^{\infty} E_n\right).$$

Now passing to the limit with $\varepsilon \to 0$ yields the result. We are left with the verification of the metric condition. Let $\operatorname{dist}(E,F) > 0$. It suffices to show that

$$\mathcal{H}^s_\varepsilon(E \cup F) = \mathcal{H}^s_\varepsilon(E) + \mathcal{H}^s_\varepsilon(F) \tag{13}$$

for all $\varepsilon < \operatorname{dist}(E,F)$. Let

$$E \cup F \subset \bigcup_{i=1}^{\infty} A_i, \quad \operatorname{diam} A_i < \varepsilon$$

be a covering of $E \cup F$. We can assume that

$$A_i \cap (E \cup F) \ne \emptyset \tag{14}$$

for all i, as otherwise we could remove A_i from the covering. Since $\operatorname{diam} A_i < \operatorname{dist}(E, F)$ it follows from (14) that A_i has a nonempty intersection with exactly one set E or F. Accordingly, the family $\{A_i\}_i$ splits into two *disjoint* subfamilies

$$\begin{aligned} \{B_j\}_j &= \text{all the sets } A_i \text{ such that } A_i \cap E \neq \emptyset; \\ \{C_j\}_j &= \text{all the sets } A_i \text{ such that } A_i \cap F \neq \emptyset. \end{aligned}$$

Hence

$$\begin{aligned} \frac{\omega_s}{2^s} \sum_{i=1}^{\infty} (\operatorname{diam} A_i)^s &= \frac{\omega_s}{2^s} \sum_j (\operatorname{diam} B_j)^s + \frac{\omega_s}{2^s} \sum_j (\operatorname{diam} C_j)^s \\ &\geq \mathcal{H}^s_\varepsilon(E) + \mathcal{H}^s_\varepsilon(F). \end{aligned}$$

Since $\{A_i\}_i$ was an arbitrary covering of $E \cup F$ such that $\operatorname{diam} A_i < \varepsilon$, we conclude upon taking the infimum that

$$\mathcal{H}^s_\varepsilon(E \cup F) \geq \mathcal{H}^s_\varepsilon(E) + \mathcal{H}^s_\varepsilon(F).$$

Since the opposite inequality is obvious (13) follows. □

Exercise. *Show that if*

- $\mathcal{H}^s(E) < \infty$, *then* $\mathcal{H}^t(E) = 0$ *for all* $t > s$;
- $\mathcal{H}^s(E) > 0$, *then* $\mathcal{H}^t(E) = \infty$ *for all* $0 < t < s$.

DEFINITION. The *Hausdorff dimension* is defined as follows. If $\mathcal{H}^s(E) > 0$ for all $s \geq 0$, then $\dim_H(E) = \infty$. Otherwise we define

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

It follows from the exercise that there is $s \in [0, \infty]$ such that $\mathcal{H}^t(E) = 0$ for $t > s$ and $\mathcal{H}^t(E) = \infty$ for $0 < t < s$. Hausdorff dimension of E equals s.

1.6 Lebesgue measure.

For a closed interval

$$P = [a_1, b_1] \times \ldots \times [a_n, b_n] \subset \mathbb{R}^n \tag{15}$$

we define its volume by

$$|P| = (b_1 - a_1) \cdot \ldots \cdot (b_n - a_n)$$

and for $A \subset \mathbb{R}^n$ we define

$$\mathcal{L}^*_n(A) = \inf \sum_{i=1}^{\infty} |P_i|$$

where the infimum is taken over all coverings

$$A \subset \bigcup_{i=1}^{\infty} P_i$$

by intervals as in (15).

Arguments similar to those in the proof that the Hausdorff measure is an outer metric measure give the following result

Theorem 11 *$\mathcal{L}_n^*$ is a metric outer measure.*

DEFINITION. $\mathcal{L}_n^*$ is called *outer Lebesgue measure.* $\mathcal{L}_n^*$-measurable sets are called Lebesgue measurable. $\mathcal{L}_n^*$ restricted to Lebesgue measurable sets is called *Lebesgue measure* and is denoted by $\mathcal{L}_n$.

Corollary 12 *All Borel sets are Lebesgue measurable.*

All sets with $\mathcal{L}_n^*(A) = 0$ are Lebesgue measurable. Then $\mathcal{L}_n(A) = 0$. A set has Lebesgue measure zero if and only it it can be covered by a sequence of intervals with whose sum of volumes can be arbitrarily small. Note that this is the same as the sets of measure zero in the setting of Riemann integral.

Theorem 13 *If P is a closed interval as in (15), then*

$$\mathcal{L}^n(P) = |P|.$$

In what follows we will often write $|A|$ to denote the Lebesgue measure of a Lebesgue measurable set A.

Proof. In order to prove the theorem we have to show that if $P \subset \bigcup_{i=1}^{\infty} P_i$, then $|P| \leq \sum_{i=1}^{\infty} |P_i|$.

We will need the following quite obvious special case of this fact.

Lemma 14 *If $P \subset \bigcup_{i=1}^{k} P_i$ is a finite covering of a closed interval P by closed intervals P_i, then $|P| \leq \sum_{i=1}^{k} |P_i|$.*

Proof. This lemma follows from the theory of Riemann integral.[2] Indeed, volume of the interval equals the integral of the characteristic function. Hence inequality $\chi_P \leq \sum_{i=1}^{k} \chi_{P_i}$ implies

$$|P| = \int_{\mathbb{R}^n} \chi_P \leq \sum_{i=1}^{k} \int_{\mathbb{R}^n} \chi_{P_i} = \sum_{i=1}^{k} |P_i|.$$

[2]This lemma can be proved in a more elementary way without using the Riemann integral, but the proof would be longer.

The proof is complete. $\square$

Now we can complete the proof of the theorem. Observe that each interval P_i is contained in a slightly bigger open interval P_i^ε such that

$$|\overline{P_i^\varepsilon}| = |P_i| + \frac{\varepsilon}{2^i},$$

where $\overline{P_i^\varepsilon}$ is the closure of P_i^ε. Now from the covering $P \subset \bigcup_{i=1}^\infty P_i^\varepsilon$ we can select a finite subcovering

$$P \subset \bigcup_{j=1}^k P_{i_j}^\varepsilon$$

(compactness) and hence the lemma yields

$$|P| \leq \sum_{j=1}^k |P_{i_j}^\varepsilon| \leq \sum_{i=1}^\infty |P_i^\varepsilon| = \sum_{i=1}^\infty |P_i| + \varepsilon.$$

Since $\varepsilon > 0$ was arbitrary the theorem follows. $\square$

Proposition 15

- *If* $P = (a_1, b_1) \times \ldots (a_n, b_n)$ *is a bounded interval, then*

$$\mathcal{L}^n(P) = (b_1 - a_1) \cdot \ldots \cdot (b_n - a_n).$$

- $\mathcal{L}^n(\partial P) = 0$.

- *If* P *is an unbounded interval in* $\mathbb{R}^n$ *and has nonempty interior (i.e. each side has positive length), then* $\mathcal{L}^n(P) = \infty$.

Proof. The first claim follows from the observation that P can be approximated from inside and from outside by closed intervals that differs arbitrarily small in volume. The second claim follows from the first one $\mathcal{L}^n(\partial P) = \mathcal{L}^n(\overline{P}) - \mathcal{L}^n(\operatorname{int} P) = 0$, and the lats claim follows from the fact that P contains closed intervals of arbitrarily large measure. $\square$

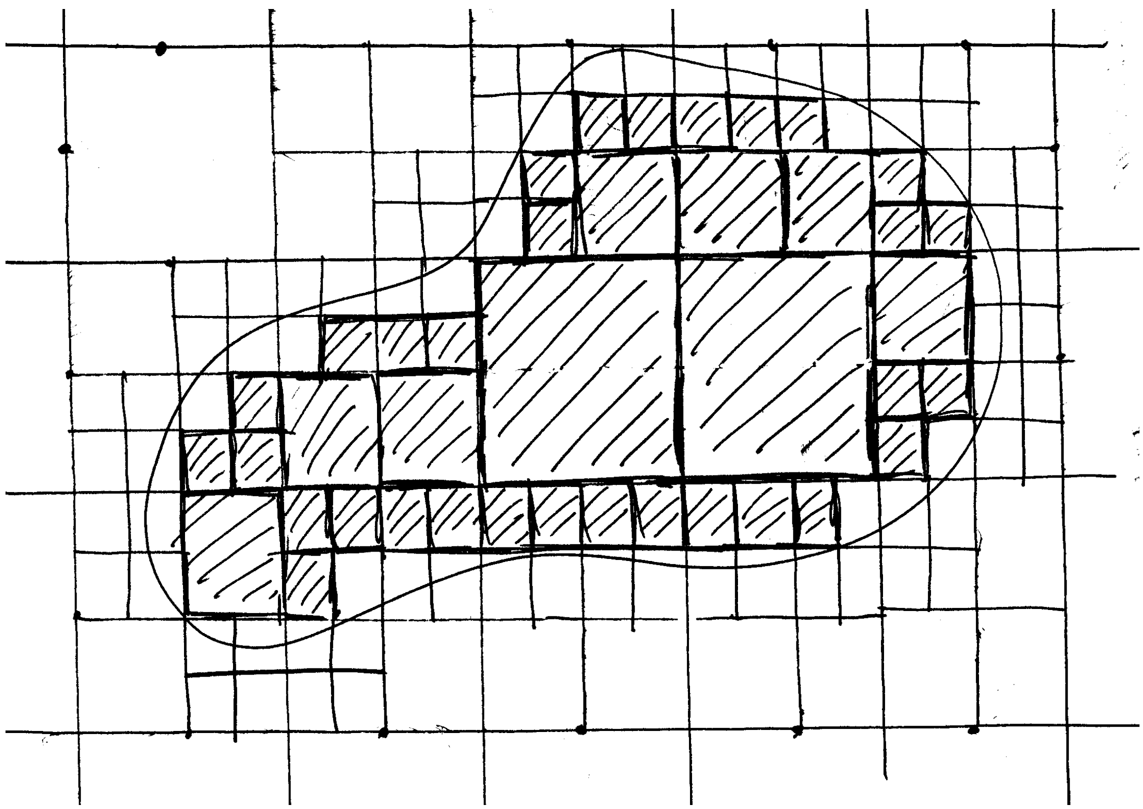
$\mathbb{R}^k \subset \mathbb{R}^n$, $k < n$ is a subset generated by the first k coordinates.

Corollary 16 *If* $k < n$, *then* $\mathcal{L}^n(\mathbb{R}^k) = 0$.

The next result shows that we can compute Lebesgue measure of an open set by representing it as a union of cubes. We call a cube dyadic if its sidelength equals 2^k for some $k \in \mathbb{Z}$.

Theorem 17 *An arbitrary open set in* $\mathbb{R}^n$ *is a union of closed dyadic cubes with pairwise disjoint interiors. Hence Lebesgue measure of the open set equals the sum of the measures of these cubes.*

Proof. Instead of a formal proof we will show a picture which explains how to represent any open set as a union of cubes with sidelength 2^{-k}, $k = 0, 1, 2, 3, \ldots$



The proof is complete □

Compare the following result with Theorem 7.

Theorem 18 *For an arbitrary set* $E \subset \mathbb{R}^n$

$$\mathcal{L}_n^*(E) = \inf \mathcal{L}^n(U),$$

where the infimum is taken over all open sets U *such that* $E \subset U$.

Proof. $\mathcal{L}_n^* = \inf \sum_i |P_i|$ and each interval P_i is contained in an open set (interval) of slightly bigger measure. □

Every Borel set B is Lebesgue measurable. Every set E with $\mathcal{L}^*(E) = 0$ is Lebesgue measurable. Hence sets of the form $B \cup E$ are Lebesgue measurable. These are all Lebesgue measurable sets. More precisely we have the following characterization.

Definition. Let X be a metric space. By a G_δ set we mean a set of the form $A = \bigcap_{i=1}^\infty G_i$, where the sets $G_i \subset X$ are open and by F_σ we mean a set of the form $B = \bigcup_{i=1}^\infty F_i$, where the sets $F_i \subset X$ are closed. Clearly all G_δ and F_σ sets are Borel.

Theorem 19 *Let $A \subset \mathbb{R}^n$. Then the following conditions are equivalent*

1. *A is Lebesgue measurable;*
2. *For every $\varepsilon > 0$ there is an open set G such that $A \subset G$ and $\mathcal{L}_n^*(G \setminus A) < \varepsilon$;*
3. *There is a G_δ set H such that $A \subset H$ and $\mathcal{L}_n^*(H \setminus A) = 0$;*
4. *For every $\varepsilon > 0$ there is a closed set F such that $F \subset A$ and $\mathcal{L}_n^*(A \setminus F) < \varepsilon$;*
5. *There is a F_σ set M such that $M \subset A$ and $\mathcal{L}_n^*(A \setminus M) = 0$;*
6. *For every $\varepsilon > 0$ there is an open set G and a closed set F such that $F \subset A \subset G$ and $\mathcal{L}_n(G \setminus F) < \varepsilon$.*

Proof. (1)⇒(2) Every measurable set can be represented as a union of sets with finite measure $A = \bigcup_{i=1}^\infty A_i$, $\mathcal{L}_n(A_i) < \infty$. It follows from Theorem 18 that for every i there is an open set G_i such that $A_i \subset G_i$ and $\mathcal{L}_n(G_i \setminus A_i) < \varepsilon/2^i$. Hence $A \subset G = \bigcup_{i=1}^\infty G_i$, $\mathcal{L}_n(G \setminus A) < \varepsilon$.

(2)⇒(3) We define $H = \bigcap_{i=1}^\infty U_i$, where U_i are open sets such that $A \subset U_i$, $\mathcal{L}_n^*(U_i \setminus A) < 1/i$.

(3)⇒(1) This implication is obvious.

(1)⇔(4)⇔(5) this equivalence follows from the equivalence of the conditions (1), (2) and (3) applied to the set $\mathbb{R}^n \setminus A$.

(1)⇒(6) If A is Lebesgue measurable then the existence of the sets F and G follows from the conditions (2) and (4).

(6)⇒(1) Take closed and open sets F_i, G_i such that $F_i \subset A \subset G_i$, $\mathcal{L}_n(G_i \setminus F_i) < 1/i$. Then the set $H = \bigcap_{i=1}^\infty G_i$ is G_δ and $\mathcal{L}_n^*(H \setminus A) = 0$. □

The Lebesgue measure has an important property of being invariant under translations i.e. if A is a translation of a Lebesgue measurable set B, then A is Lebesgue measurable and $\mathcal{L}^n(A) = \mathcal{L}^n(B)$. This property follows immediately from the definition of the Lebesgue measure. We will show not that this property implies that there is a set which is not Lebesgue measurable.

Theorem 20 (Vitali) *There is a set $E \subset \mathbb{R}^n$ which is not Lebesgue measurable.*[3]

[3]The reader will easily check that the same proof applies to any translation invariant measure μ such that the measure of the unit cube if positive and finite; see also Theorem 21.

Proof. We will prove the theorem for $n = 1$ only, but the same argument works for an arbitrary n. The idea of the proof is to find a countable family of pairwise disjoint and isometric (by translation)sets whose union E contains $(0,1)$ and is contained in $(-1,2)$. If the sets in the family are isometric to a set V, then V cannot be Lebesgue measurable, because measurability of V would imply that a number between 1 and 3 (measure of E) be equal to infinite sum of equal numbers (equal to measure of V) which is impossible. It should not be surprising that the construction of the set V has to involve axiom of choice.

For $x, y \in (0,1)$ we write $x \sim y$ if $x - y$ is a rational number. Clearly $\sim$ is an equivalence relation and hence $(0,1)$ is the union of a family $\mathcal{F}$ of pairwise disjoint sets

$$[x] = \{y \in (0,1) : x \sim y\}.$$

It follows from the axiom of choice that there is a set $V \subset (0,1)$ which contains exactly one element from each set in the family $\mathcal{F}$. Let

$$E = \bigcup_{a \in \mathbb{Q} \cap (-1,1)} V_a, \quad \text{where } V_a = V + a = \{x + a : x \in V\}.$$

It is easy to see that

- $V_a \cap V_b = \emptyset$ if $a, b \in \mathbb{Q}$, $a \neq b$;
- $(0,1) \subset E \subset (-1,2)$.

Suppose that V is measurable. Then each V_a is measurable, $\mathcal{L}^1(V) = \mathcal{L}^1(V_a)$ and E is measurable. If $\mathcal{L}^1(V) > 0$, then $\mathcal{L}^1(E) = \infty$ which is impossible and if $\mathcal{L}^1(V) = 0$, then $\mathcal{L}^1(E) = 0$ which is also impossible. That proves that the set V cannot be measurable.
□

The Lebesgue measure is translation invariant and $\mathcal{L}^n([0,1]^n) = 1$. It turns out that the above two properties uniquely determine Lebesgue measure. This proves that the Lebesgue measure is in a sense the only natural way of measuring length, area, volume of general sets in one, two, three,... dimensional Euclidean spaces.

Theorem 21 *If μ is a measure on $\mathfrak{B}(\mathbb{R}^n)$ such that $\mu(a + E) = \mu(E)$ for all $a \in \mathbb{R}^n$, $E \in \mathfrak{B}(\mathbb{R}^n)$ and $\mu([0,1]^n) = 1$, then $\mu(E) = \mathcal{L}^n(E)$ on $\mathfrak{B}(\mathbb{R}^n)$.*

Proof. According to Corollary 8 it suffices to prove that $\mu(U) = \mathcal{L}^n(U)$ for all open sets U.

Consider an open cube $Q_k = (0, 2^{-k})^n$, where k is a positive integer. There are 2^{kn} pairwise disjoint open cubes contained in the unit cube $[0,1]^n$, each being a translation of Q^k. Since the measure is invariant under translations we conclude that

$$\mu(Q_k) \leq 2^{-kn}.$$

It is easy to see that we can cover the boundary ∂Q by translations of Q_k in a way that the sum of μ-measures of the cubes covering ∂Q goes to zero as $k \to \infty$. This proves that $\mu(\partial Q) = 0$.

Now $[0,1]^n$ can be covered by 2^{kn} cubes each being a translation of the closed cube $\overline{Q}^k$. Since the measure is invariant under translations we conclude that

$$\mu(\overline{Q}_k) \geq 2^{-kn}$$

and hence

$$\mu(Q_k) = \mu(\overline{Q}_k) = 2^{-kn} = \mathcal{L}^n(Q_k). \tag{16}$$

As we know (Theorem 17) an arbitrary open set U is a union of cubes with pairwise disjoint interiors and sidelength 2^{-k} (k depends on the cube). Cubes are not disjoint, but meet along the boundary which has μ-measure zero. Hence the $\mu(U)$ equals the sum of the μ-measures of the cubes which according to (16) equals $\mathcal{L}^n(U)$. The proof is complete. □

Theorem 22 *$\mathcal{H}^n = \mathcal{L}_n$ on $\mathfrak{B}(\mathbb{R}^n)$.*

Sketch of the proof. Observe that the Hausdorff measure $\mathcal{H}^n$ in $\mathbb{R}^n$ has the property that

$$\mathcal{H}^n(E + a) = \mathcal{H}^n(E) \quad \text{for all } E \in \mathfrak{B}(\mathbb{R}^n).$$

Therefore in order to prove that $\mathcal{H}^n = \mathcal{L}_n$ on $\mathfrak{B}(\mathbb{R}^n)$ it suffices to prove that

$$\mathcal{H}^n([0,1]^n) = 1$$

This innocently looking result is however very difficult and we will not prove it here. □

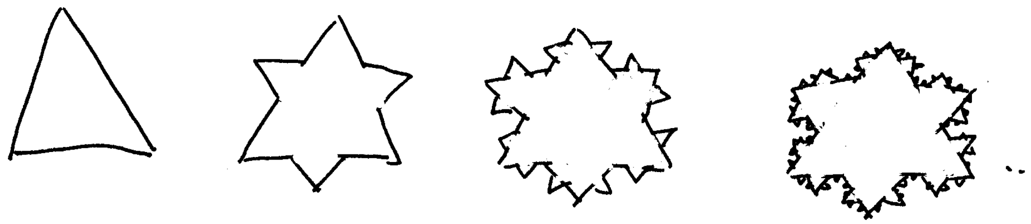
Exercise. *Using Theorem 22 prove that $\mathcal{H}^n = \mathcal{L}_n^*$ on the class of all subsets of $\mathbb{R}^n$.*

If $s < n$, then $\mathcal{H}^s$ is the s-dimensional measure in $\mathbb{R}^n$. Not that s is *not* required to be an integer. If $k < n$ is an integer and $M \subset \mathbb{R}^n$ is a k-dimensional manifold, then one can prove that $\mathcal{H}^k$ coincides on M with the natural k-dimensional Lebesgue measure on M.[4]

Examples. If $C \subset [0,1]$ is the standard ternary Cantor set, then one can prove that $\dim_H(C) = \log 2/\log 3 = 0.6309\ldots$ Moreover $\mathcal{H}^s(C) = 1$ for $s = \log 2/\log 3$.

The *Van Koch curve* is defined as the limiting curve for the following construction

[4] We have already seen how in the course of Advanced Calculus to measure subsets of M in the setting of the Riemann integral. This construction can easily be generalized to the Lebesgue measure on M. We will discuss it with more details later.



The Van Koch curve, denoted by K is homeomorphic to the circle S^1, but $\dim_H K = \log 4/\log 3 = 1,261\ldots$

Proposition 23 *A set $E \subset \mathbb{R}^n$ has Lebesgue measure zero if and only if for every $\varepsilon >$ there is a family of balls $B(x_i, r_i)\}_{i=1}^\infty$ such that*

$$E \subset \bigcup_{i=1}^\infty B(x_i, r_i), \qquad \sum_{i=1}^\infty r_i^n < \varepsilon.$$

Proof. $\Rightarrow$ Each bass $B(x_i, r_i)$ is contained in a cube with sidelength $2r_i$ and hence E can be covered by a sequence of cubes whose sum of volumes is $\sum_{i=1}^\infty (2r_i)^n < 2^n\varepsilon$.

$\Leftarrow$ If a set $E \subset \mathbb{R}^n$ has Lebesgue measure zero, then for every $\varepsilon > 0$ there is an open set U such that $E \subset U$ and $\mathcal{L}_n(U) < \varepsilon$. The set U can be represented as union of cubes with pairwise disjoint interiors (Theorem 17). each cube Q of sidelength ℓ is contained in a ball of radius $r = \sqrt{n}\ell$ and hence $r^n = n^{n/2}|Q|$. Therefore E can be covered by a sequence of balls $\{B(x_i, r_i)\}_{i=1}^\infty$ such that $\sum_{i=1}^\infty r_i^n < n^{n/2}\varepsilon$. □

DEFINITION. We say that a function between metric spaces $f : (X, d) \to (Y, \rho)$ is *Lipschitz continuous* if there is a constant $L > 0$ such that

$$\rho(f(x), f(y)) \leq Ld(x, y) \quad \text{for all } x, y \in X.$$

In this case we say that f is L-Lipschitz continuous and call L Lipschitz constant if f.

Proposition 24 *If $f : \mathbb{R}^n \to \mathbb{R}^n$ is Lipschitz continuous and $E \subset \mathbb{R}^n$ has Lebesgue measure zero, then $f(E)$ has Lebesgue measure zero too.*

Proof. It follows from Proposition 23 and the fact that the L-Lipschitz image of a ball of radius r is contained in a ball of radius Lr. □

Exercise. *Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be a homeomorphism. Prove that $f(A)$ is Borel if and only if A is Borel.*

Although the theory of Lebesgue measure seems to work very well something very important is missing: we haven't proved so far that if Q is a unit cube whose sides are *not* parallel to coordinate axes, then its measure equals 1. Indeed we considered only intervals with sides parallel to coordinate directions. Fortunately this important fact is contained in a result that we will prove now.

Theorem 25 *If $L : \mathbb{R}^n \to \mathbb{R}^n$ is a nondegenerate linear transformation with the matrix A ($\det A \neq 0$), then $L(E) \subset \mathbb{R}^n$ is a Borel (Lebesgue measurable) set if and only if $E \subset \mathbb{R}^n$ is Borel (Lebesgue measurable) set. Moreover*

$$\mathcal{L}_n(L(E)) = |\det A| \mathcal{L}_n(E). \tag{17}$$

Proof. Since L is a homeomorphism it preserves the class of Borel sets (see Exercise). Since both L and L^{-1} are Lipschitz continuous L preserves the class of sets of measure zero (Proposition 24). Therefore L preserves the class of Lebesgue measurable sets (because each Lebesgue measurable set is the union of a Borel set and a set of measure zero). It follows from Proposition 24 that the formula (17) is satisfied for sets of measure zero. Therefore it remains to prove (24) for Borel sets.

Define a new measure $\mu(E) = \mathcal{L}_n(L(E))$. Clearly μ is a measure on $\mathfrak{B}(\mathbb{R}^n)$ that is translation invariant. Let $\mu([0,1]^n) = a > 0$. Hence the measure $a^{-1}\mu$ satisfies all the assumptions of Theorem 21 and therefore $a^{-1}\mu(E) = \mathcal{L}_n(E)$, so we have

$$\mathcal{L}_n(L(E)) = a\mathcal{L}_n(E) \quad \text{for all } E \in \mathfrak{B}(\mathbb{R}^n).$$

It remains to prove that $a = |\det A|$. Clearly $a = a(A) : \mathrm{GL}\,(n) \to (0, \infty)$ is a function defined on the class of invertible matrices. We have

$$a(A_1 A_2) = a(A_1)a(A_2). \tag{18}$$

Indeed, if A_1, A_2 are matrices representing linear transformations L_1 and L_2, then

$$a(A_1A_2)\mathcal{L}_n(E) = \mathcal{L}_n(L_1(L_2(E)) = a(L_1)\mathcal{L}_n(L_2(E)) = a(L_1)a(L_2)\mathcal{L}_n(E).$$

Also

$$a(s^n I) = s^n, \quad s > 0. \tag{19}$$

Here I is the identity matrix. This property is quite obvious and we leave a formal proof to the reader. Now it remains to prove the following fact.

Lemma 26 *Let $a : \mathrm{GL}\,(n) \to (0, \infty)$ be a function with properties (18) and (19). Then $a(A) = |\det A|$.*

Proof. Let $A_i(s)$ be the diagonal matrix with $-s$ on the ith place and s on all other places on the diagonal. Since $A_i(s)^2 = s^2 I$ we have $a(A_i(s))^2 = a(A_i(s)^2) = s^{2n}$ and hance $a(A_i(s)) = |s^n| = |\det A_i(s)|$.

For $k \neq l$ let $B_{kl}(s) = [a_{ij}]$ be the matrix such that $a_{kl} = s$, $a_{ii} = 1$, $i = 1, 2, \ldots$ and all other entries equal zero. Multiplication by the matrix $B_{kl}(s)$ from the right (left) is equivalent to adding kth column (ℓth row) multiplied by s to ℓth column (kth row).

It is well known from linear algebra (and easy to prove) that applying such operations to any nonsingular matrix A it can be transformed to a matrix of the form tI or

$A_n(t)$. Since multiplication by $B_{kl}(s)$ does not change determinant, $t = |\det A|^{1/n}$. It remains to prove that $a(B_{kl}(s)) = 1$. Since $B_{kl}(-s) = A_k(1)B_{kl}(s)A_k(1)$ we have that $a(B_{kl}(-s)) = a(B_{kl}(s))$. On the other hand $B_{kl}(s)B_{kl}(-s) = I$ and hence $a(B_{kl}(s))^2 = 1$, so $a(B_{kl}(s)) = 1$. The proof of the lemma and hence the proof of the theorem is complete. □

2 Integration

2.1 Measurable functions.

DEFINITION. Let $(X, \mathfrak{M})$ be a measurable space and Y a metric space. We say that a function $f : X \to Y$ is *measurable* if

$$f^{-1}(U) \in \mathfrak{M} \quad \text{for every open set } U \subset Y.$$

If $E \in \mathfrak{M}$ is a measurable set, then we say that $f : E \to Y$ is measurable if $f^{-1}(U) \in \mathfrak{M}$ for every open $U \subset Y$.

Clearly, $f : E \to Y$ is measurable if it can be extended to a measurable function $\tilde{f} : X \to Y$, but there is also another point of view.

$$\mathfrak{M}_E = \{A \subset E : \ A \in \mathfrak{M}\} = \{B \cap E : \ B \in \mathfrak{M}\}$$

is a σ-algebra and measurability of $f : E \to Y$ is equivalent to measurability of f with respect to $\mathfrak{M}_E$.

Theorem 27 *Let $(X, \mathfrak{M})$ be a measurable space, $f : X \to Y$ a measurable function and $g : Y \to Z$ a continuous function. Then the function $g \circ f : X \to Z$ is measurable.*

Proof. For every open set $U \subset Z$ we have

$$(g \circ f)^{-1}(U) = f^{-1}(\underbrace{g^{-1}(U)}_{\text{open in } Y}) \in \mathfrak{M}.$$

The proof is complete. □

Theorem 28 *Let $(X, \mathfrak{M})$ be a measurable space and Y a metric space. If the functions $u, v : X \to \mathbb{R}$ are measurable and $\Phi : \mathbb{R}^2 \to Y$ is continuous, then the function*

$$h(x) = \Phi(u(x), v(x)) : X \to Y$$

is measurable.

Proof. Put $f(x) = (u(x), v(x))$, $f : X \to \mathbb{R}^2$. According to Theorem 27 it suffices to prove that f is measurable. Let $R = (a,b) \times (c,d)$ be an open rectangle. Then

$$f^{-1}(R) = f^{-1}((a,b) \times (c,d)) = u^{-1}((a,b)) \cap u^{-1}((c,d)) \in \mathfrak{M}.$$

Every open set $U \subset \mathbb{R}^2$ is a countable union of open rectangles

$$U = \bigcup_{i=1}^{\infty} R_i$$

and hence

$$f^{-1}(U) = \bigcup_{i=1}^{\infty} f^{-1}(R_i) \in \mathfrak{M}.$$

The proof is complete. □

This theorem has many applications.

(a) *If $f = u + iv : X \to \mathbb{C}$, where $u, v : X \to \mathbb{R}$ are measurable, then f is measurable.*

Indeed, we take $\Phi(u,v) = u + iv$.

(b) *If $f = u + iv : X \to \mathbb{C}$ is measurable, then the functions $u, v, |f| : X \to \mathbb{R}$ are measurable.*

Indeed, we take $\Phi_1(u + iv) = u$, $\Phi_2(u + iv) = v$, $\Phi_3(u + iv) = |u + iv|$.

(c) *If $f, g : X \to \mathbb{C}$ are measurable, then the functions $f+g, fg : X \to \mathbb{C}$ are measurable.*

Indeed, we take $\Phi_1 : \mathbb{C} \times \mathbb{C} \to \mathbb{C}$, $\Phi_1(z_1, z_2) = z_1 + z_2$ and $\Phi_2 : \mathbb{C} \times \mathbb{C} \to \mathbb{C}$, $\Phi_2(z_1, z_2) = z_1 z_2$.

(d) *A set $E \subset X$ is measurable if and only if the characteristic function if E*

$$\chi_E(x) = \begin{cases} 1 & \textit{if } x \in E, \\ 0 & \textit{if } x \notin E, \end{cases}$$

is measurable.

Indeed, this follows from the fact that for any open set $U \subset \mathbb{R}$

$$(\chi_E)^{-1}(U) = \begin{cases} \emptyset & \text{if } 0 \notin U,\, 1 \notin U, \\ E & \text{if } 0 \notin U,\, 1 \in U, \\ X \setminus E & \text{if } 0 \in U,\, 1 \notin U, \\ X & \text{if } 0 \in U,\, 1 \in U, \end{cases}$$

(e) *If* $f : X \to \mathbb{C}$ *is a measurable function, then there is a measurable function* $\alpha : X \to \mathbb{C}$ *such that* $|\alpha| = 1$ *and* $f = \alpha|f|$.

The set $E = \{x : f(x) = 0\}$ is measurable because

$$E = \bigcap_{i=1}^{\infty} f^{-1}\Big(\underbrace{B(0, \frac{1}{i})}_{\text{disc of radius } 1/i}\Big).$$

Hence the function $h(x) = f(x) + \chi_E(x)$ is measurable and $h(x) \neq 0$ for all x. Therefore h is a measurable function as a function $h : X \to \mathbb{C} \setminus \{0\}$. Since $\varphi : \mathbb{C} \setminus \{0\} \to \mathbb{C}$, $\varphi(z) = z/|z|$ is continuous we conclude that

$$\alpha(x) = \varphi(f(x) + \chi_E(x))$$

is measurable. If $x \in E$, then $\alpha(x) = 1$ and if $x \notin E$, then $\alpha(x) = f(x)/|f(x)|$. Therefore $f(x) = \alpha(x)|f(x)|$ for all $x \in X$.

Definition. Let X, Y be two metric spaces. We say that $f : X \to Y$ is a *Borel mapping* if it is measurable with respect to the σ-algebra of Borel sets in X i.e.

$$f^{-1}(U) \in \mathfrak{B}(X) \quad \text{for every open set } U \subset Y.$$

Clearly every continuous mapping is Borel.

Theorem 29 *Let* $(X, \mathfrak{M})$ *be a measurable space and* Y *a metric space. Let also* $f : X \to Y$ *be a measurable mapping.*

(a) If $E \subset Y$ *is a Borel set, then* $f^{-1}(E) \in \mathfrak{M}$.

(b) If Z *is another metric space and* $g : Y \to Z$ *is a Borel mapping, then* $g \circ f : X \to Z$ *is measurable.*

Proof. (a) Let $\mathcal{R} = \{A \subset Y | f^{-1}(A) \in \mathfrak{M}\}$. Clearly $\mathcal{R}$ contains all open sets. If we prove that $\mathcal{R}$ is a σ-algebra, we will have that $\mathfrak{B}(Y) \subset \mathcal{R}$ which is the claim.

- $Y \in \mathcal{R}$

Indeed, $f^{-1}(Y) = X \in \mathfrak{M}$.

- If $A \in \mathcal{R}$, then $Y \setminus A \in \mathcal{R}$.

Indeed, $f^{-1}(Y \setminus A) = X \setminus f^{-1}(A) \in \mathfrak{M}$.

- If $A_1, A_2, \ldots \in \mathcal{R}$, then $\bigcup_{i=1}^{\infty} A_i \in \mathcal{R}$.

Indeed, $f^{-1}(\bigcup_{i=1}^{\infty} A_i) = \bigcup_{i=1}^{\infty} f^{-1}(A_i) \in \mathcal{R}$. The above three properties imply that $\mathcal{R}$ is a σ-algebra.

(b) If $U \subset Z$ is open, then $g^{-1}(U) \subset Y$ is Borel and hence $(g \circ f)^{-1}(U) = f^{-1}(g^{-1}(U)) \in \mathfrak{M}$ by (a). $\square$

Definition. Let $(X, \mathfrak{M})$ be a measurable space. Let $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$. We say that a function $f : X \to \overline{\mathbb{R}}$ is measurable if $f^{-1}(U)$ is measurable for every open set $U \subset \mathbb{R}$ and if the sets $f^{-1}(+\infty)$, $f^{-1}(-\infty)$ are measurable. Similarly for $E \in \mathfrak{M}$ we define measurable functions $f : E \to \overline{\mathbb{R}}$.

Exercise. *Let $(X, \mathfrak{M})$ be a measurable space. Prove that a function $f : X \to \overline{\mathbb{R}}$ is measurable if and only if $f^{-1}((a, \infty]) \in \mathfrak{M}$ for all $a \in \mathbb{R}$.*

Let $\{a_n\}$ be a sequence in $\overline{\mathbb{R}}$. For $k = 1, 2, \ldots$ put

$$b_k = \sup\{a_k, a_{k+1}, a_{k+2}, \ldots\} \tag{20}$$

and

$$\beta = \inf\{b_1, b_2, \ldots\}\,. \tag{21}$$

We call β *upper limit* of $\{a_n\}$ and write

$$\beta = \limsup_{n \to \infty} a_n$$

Clearly $\{b_k\}$ is a decreasing sequence, so $\beta = \lim_{k \to \infty} b_k$. It is easy to see that there is a subsequence $\{a_{n_i}\}$ of $\{a_n\}$ such that $\lim_{i \to \infty} a_{n_i} = \beta$. Moreover β is the largest possible value for limits of all subsequences of $\{a_n\}$.

The *lower limit* is defined by

$$\liminf_{n \to \infty} a_n = -\limsup_{n \to \infty}(-a_n). \tag{22}$$

Equivalently the lower limit can be defined by interchanging sup and inf in (20) and (21). The lower limit equals to the lowest possible value for limits of subsequences of $\{a_n\}$.

For a sequence of functions $f_n : X \to \overline{\mathbb{R}}$ we define new functions $\sup_n f_n, \limsup_n f_n : X \to \overline{\mathbb{R}}$ by

$$\left(\sup_n f_n\right)(x) = \sup_n f_n(x) \quad \text{for every } x \in X,$$

$$\left(\limsup_n f_n\right)(x) = \limsup_n f_n(x) \quad \text{for every } x \in X.$$

Similarly we can define functions $\inf_n f_n$ and $\liminf_{n\to\infty} f_n$.

If $f(x) = \lim_{n\to\infty} f_n(x)$ for every $x \in X$, then we say that f is a *pointwise limit* of $\{f_n\}$.

Theorem 30 *Suppose that $f_n : X \to \overline{\mathbb{R}}$ is a sequence of measurable functions. Then the functions $\sup_n f_n$, $\inf_n f_n$, $\limsup_{n\to\infty} f_n$, $\liminf_{n\to\infty} f_n$ are measurable.*

Proof. Since

$$\left(\sup_n f_n\right)^{-1}((a,\infty]) = \bigcup_{n=1}^{\infty} f_n^{-1}((a,\infty]) \in \mathfrak{M},$$

measurability of $\sup_n f_n$ follows from the exercise. By a similar argument we prove that the function $\inf_n f_n$ is measurable. The two facts imply measurability of $\limsup f_n$ because

$$\limsup_{n\to\infty} f_n = \inf_{k\geq 1}\{\sup_{i\geq k} f_i\}.$$

Now measurability of $\liminf_n f_n$ is an obvious consequence of the equality

$$\liminf_{n\to\infty} f_n = -\limsup_{n\to\infty}(-f_n).$$

The proof is complete. □

Corollary 31 *The limit of every pointwise convergent sequence of functions $f_n : X \to \mathbb{C}$ (or $f_n : X \to \overline{\mathbb{R}}$) is measurable.*

Corollary 32 *If $f, g : X \to \overline{\mathbb{R}}$ are measurable functions, then the functions $\max\{f,g\}$, $\min\{f,g\}$ are measurable. In particular the fuctions*

$$f^+ = \max\{f, 0\}, \qquad f^- = -\min\{f, 0\}$$

are measurable.

The functions f^+ and f^- are called *positive part* and *negative part* of f. Note that $f = f^+ - f^-$, $|f| = f^+ + f^-$.

Corollary 31 has the following generalization to the case of metric space valued mappings.

Theorem 33 *Let $f_n : X \to Y$ be a sequence of measurable mappings with values in a separable metric space. If $f_n(x) \to f(x)$ for every $x \in X$, then $f : X \to Y$ is measurable.*

Proof. Let $\{y_i\}_{i=1}^\infty$ be a countable dense set. Then the functions $d_i : Y \to \mathbb{R}$, $d_i(y) = d(y_i, y)$ are continuous. Hence $d_i \circ f_n : X \to \mathbb{R}$ are measurable. Since $d_i \circ f_n \to d_i \circ f$, the functions $d_i \circ f : X \to \mathbb{R}$ are measurable too by Corollary 31. In particular the sets

$$\{x : d(y_i, f(x)) \geq d(y_i, F)\} = (d_i \circ f)^{-1}([d(y_i, F), \infty))$$

are measurable. To prove measurability of f it suffices to prove measurability of $f^{-1}(F)$ for every closed set F and it follows from the equality

$$f^{-1}(F) = \bigcap_{i=1}^{\infty} \{x : \ d(y_i, f(x)) \geq d(y_i, F)\} \, . \tag{23}$$

To prove this equality observe that if $x \in f^{-1}(F)$, then $f(x) \in F$ and hence $d(y_i, f(x)) \geq d(y_i, F)$ for every i, so x belongs to the right hand side of (23). On the other hand if $x \notin f^{-1}(F)$, then $f(x) \notin F$ and hence there is i such that $d(y_i, f(x)) < d(y_i, F)$ (because $\{y_i\}_{i=1}^\infty$ is a dense subset of Y and so we can find y_i arbitrarily close to $f(x)$), therefore x cannot belong to the right hand side of (23). □

Definition. Let $(X, \mathfrak{M})$ be a measurable space. A measurable function $s : X \to \mathbb{C}$ is called *simple function* if it has only a finite number of values. Simple functions can be represented in the form

$$s = \sum_{i=1}^{n} \alpha_i \chi_{A_i} \, ,$$

where $\alpha_i \neq \alpha_j$ for $i \neq j$ and A_i are disjoint measurable sets.

Theorem 34 *If $f : X \to [0, \infty]$ is measurable, then there is a sequence of simple functions s_n on X such that*

1. $0 \leq s_1 \leq s_2 \leq \ldots \leq f$;
2. $\lim_{n\to\infty} s_n(x) = f(x)$ *for every* $x \in X$.

If in addition f is bounded, then s_n converges to f uniformly.

Proof. It is easy to see that the sequence

$$s_n(x) = \begin{cases} n & \text{if } f(x) \geq n, \\ \frac{k}{2^k} & \text{if } \frac{k}{2^n} \leq f(x) < \frac{k+1}{2^n} \leq n, \end{cases}$$

has all properties that we need. □

2.2 Integral

In the integration theory we will assume that $0 \cdot \infty = 0$.

Definition. If $s : X \to [0, \infty)$ is a simple function of the form

$$s = \sum_{i=1}^{n} \alpha_i \chi_{A_i} \,,$$

where $\alpha_i \neq \alpha_j$ for $i \neq j$ and A_i are pairwise disjoint measurable sets, then for any $E \in \mathfrak{M}$ we define

$$\int_E s \, d\mu = \sum_{i=1}^{n} \alpha_i \mu(A_i \cap E) \,.$$

According to our assumption $0 \cdot \infty = 0$, so $0 \cdot \mu(A_i \cap E) = 0$, even if $\mu(A_i \cap E) = \infty$.

If $f : X \to [0, \infty]$ is measurable, and $E \in \mathfrak{M}$, we define the *Lebesgue integral* of f over E by

$$\int_E f \, d\mu = \sup \int_E s \, d\mu,$$

where the supremum is over all simple functions s such that $0 \leq s \leq f$.

(a) If $0 \leq f \leq g$, then $\int_E f \, d\mu \leq \int_E g \, d\mu$.

(b) If $A \subset B$ and $f \geq 0$, then $\int_A f \, d\mu \leq \int_B f \, d\mu$.

(c) If $f \geq 0$ and c is a constant $0 \leq c < \infty$, then $\int_E f \, d\mu = c \int_E f \, d\mu$.

(d) If $f(x) = 0$ for all $x \in E$, then $\int_E f \, d\mu = 0$, even if $\mu(E) = \infty$.

(e) If $\mu(E) = 0$, then $\int_E f \, d\mu = 0$, even if $f(x) = \infty$ for all $x \in E$.

(f) If $f \geq 0$, then $\int_E f \, d\mu = \int_X \chi_E f \, d\mu$.

In the proof of the next two theorems we will need the following lemma.

Lemma 35 *Let s and t be nonnegative measurable simple functions on X. Then the function*

$$\varphi(E) = \int_E s \, d\mu, \quad E \in \mathfrak{M}$$

defines a measure in $\mathfrak{M}$. Moreover

$$\int_X (s + t) \, d\mu = \int_X s \, d\mu + \int_X t \, d\mu \,. \tag{24}$$

Proof. Since $\varphi(\emptyset) = 0$, it remains to prove that φ is countably additive. If $E_1, E_2, \ldots$ are disjoint measurable sets, then

$$\begin{aligned}\varphi\left(\bigcup_{k=1}^{\infty} E_k\right) &= \sum_{i=1}^{n} \alpha_i \mu\left(\bigcup_{k=1}^{\infty} E_k \cap A_i\right) = \sum_{i=1}^{n} \alpha_i \sum_{k=1}^{\infty} \mu(E_k \cap A_i) \\ &= \sum_{k=1}^{\infty}\sum_{i=1}^{n} \alpha_i \mu(E_k \cap A_i) = \sum_{k=1}^{\infty} \varphi(E_k)\,.\end{aligned}$$

This completes the proof of the first part. Let now

$$s = \sum_{i=1}^{k} \alpha_i \chi_{A_i}, \quad t = \sum_{j=1}^{m} \beta_j \chi_{B_j}$$

be a representation of the simple functions as in the definition, then for $E_{ij} = A_i \cap B_j$ we have

$$\int_{E_{ij}} s + t\, d\mu = (\alpha_i + \beta_j)\mu(E_{ij}) = \int_{E_{ij}} s\, d\mu + \int_{E_{ij}} t\, d\mu\,.$$

Since X is a union of disjoint sets E_{ij} we conclude (24) from the first part of the lemma. □

The following theorem is one of the most important results of the theory of integration.

Theorem 36 (Lebesgue monotone convergence theorem) *Let $\{f_n\}$ be a sequence of measurable functions on X such that:*

1. $0 \le f_1 \le f_2 \le \ldots \le \infty$ *for every* $x \in X$*;*
2. $\lim_{n\to\infty} f_n(x) = f(x)$ *for every* $x \in X$*.*

Then f is measurable and

$$\lim_{n\to\infty} \int_X f_n\, d\mu = \int_X f\, d\mu$$

Proof. Measurability of f follows from Corollary 31. The sequence $\int_X f_n\, d\mu$ is increasing and hence it has a limit $\alpha = \lim_{n\to\infty} \int_X f_n\, d\mu$. Since $f_n \le f$, we have $\int_X f_n\, d\mu \le \int_X f\, d\mu$ and hence $\alpha \le \int_X f\, d\mu$. Let $0 \le s \le f$ be a simple function. Fix a constant $0 < c < 1$ and define

$$E_n = \{x :\, f_n(x) \ge cs(x)\}, \quad n = 1, 2, 3, \ldots$$

It is easy to see that $E_1 \subset E_2 \subset \ldots$ and $\bigcup_{n=1}^{\infty} E_n = X$ (because $f_n(x) \to f(x)$ for every $x \in X$.) Hence

$$\int_X f_n\, d\mu \ge \int_{E_n} f_n\, d\mu \ge c \int_{E_n} s\, d\mu\,. \tag{25}$$

Since $\varphi(E) = \int_E s\,d\mu$ is a measure and $E_1 \subset E_2 \subset \ldots$ we conclude (Theorem 3(f)) that

$$\int_{E_n} s\,d\mu = \varphi(E_n) \to \varphi\left(\bigcup_{n=1}^{\infty} E_n\right) = \varphi(X) = \int_X s\,d\mu\,.$$

Thus after passing to the limit in (25) we obtain $\alpha \geq c\int_E s\,d\mu$ and hence

$$\int_X f\,d\mu \geq \alpha \geq \int_X s\,d\mu\,,$$

because $0 < c < 1$ was arbitrary. Taking supremum over all simple functions $0 \leq s \leq f$ yields

$$\int_X f\,d\mu \geq \alpha \geq \int_X f\,d\mu\,,$$

i.e. $\int_X f\,d\mu = \alpha$ and the theorem follows. □

Theorem 37 *If $f_n : X \to [0,\infty]$ is measurable, for $n = 1, 2, 3, \ldots$ and*

$$f(x) = \sum_{n=1}^{\infty} f_n(x) \quad \textit{for all } x \in X,$$

then f is measurable and

$$\int_X f\,d\mu = \sum_{i=1}^{\infty} \int_X f_n\,d\mu\,.$$

Proof. The function f is a limit of measurable functions, so it is measurable. Let $\{s_i\}$ and $\{t_i\}$ be increasing sequences of simple nonnegative function such that $s_i \to f_1$, $t_i \to f_2$ (Theorem 34). Then $s_i + t_i \to f_1 + f_2$ and hence (24) combined with the monotone convergence theorem gives

$$\int_X f_1 + f_2\,d\mu = \int_X f_1\,d\mu + \int_X f_2\,d\mu\,.$$

Applying an induction argument we obtain

$$\int_X (f_1 + f_2 + \ldots + f_N)\,d\mu = \sum_{n=1}^{N} \int_X f_n\,d\mu\,.$$

Since $g_N = f_1 + \ldots + f_N$ is an increasing sequence of measurable functions and $g_N \to f$ the theorem follows from the monotone convergence theorem. □

Theorem 38 (Fatou's lemma) *If $f_n : X \to [0,\infty]$ is measurable, for $n = 1, 2, 3, \ldots$, then*

$$\int_X \left(\liminf_{n\to\infty} f_n\right)\,d\mu \leq \liminf_{n\to\infty} \int_X f_n\,d\mu\,.$$

Exercise. *Construct a sequence of measurable functions $f_n : \mathbb{R} \to [0, \infty]$ such that the inequality in Fatou's lemma is sharp.*

Proof of the theorem. Let $g_n = \inf\{f_n, f_{n+1}, f_{n+2}, \ldots\}$. Then $g_n \le f_n$ and hence

$$\int_X g_n \, d\mu \le \int_X f_n \, d\mu \, . \tag{26}$$

Since $0 \le g_1 \le g_2 \le \ldots$, all the functions g_n are measurable (Theorem 30) and $g_n \to \liminf_{k\to\infty} f_n$, inequality (26) combined with the monotone convergence theorem yields the result. □

Theorem 39 *Let $f : X \to [0, \infty]$ be measurable. Then*

$$\varphi(E) = \int_E f \, d\mu \quad \textit{for } E \in \mathfrak{M}$$

defines a measure in $\mathfrak{M}$. Moreover

$$\int_X g \, d\varphi = \int_X gf \, d\mu \tag{27}$$

for every measurable function $g : X \to [0, \infty]$.

Proof. Clearly $\varphi(\emptyset) = 0$. To prove that φ is a measure it remains to prove countable additivity. Let $E_1, E_2, \ldots \in \mathfrak{M}$ be disjoint sets and $E = \bigcup_{n=1}^{\infty} E_n$. Then obviously

$$\chi_E f = \sum_{n=1}^{\infty} \chi_{E_n} f$$

and

$$\begin{aligned} \varphi(E) &= \int_X \chi_E f \, d\mu = \int_X \sum_{n=1}^{\infty} \chi_{E_n} f \, d\mu \\ &= \sum_{n=1}^{\infty} \int_X \chi_{E_n} f \, d\mu = \sum_{n=1}^{\infty} \varphi(E_n) \, , \end{aligned}$$

by Theorem 37. To prove (27) observe that it is true if $g = \chi_E$, $E \in \mathfrak{M}$. Hence it is also true for simple functions and the general case follows from the monotone convergence theorem combined with Theorem 34. □

DEFINITION. Let $(X, \mathfrak{M}, \mu)$ be a measure space. The space $L^1(\mu)$ consists of all measurable functions $f : X \to \mathbb{C}$ for which $\|f\|_1 = \int_X |f| \, d\mu < \infty$. Elements of $L^1(\mu)$ are called *Lebesgue integrable functions* or *summable functions.*

If $f = u + iv$, then the functions u^+, u^-, v^+, v^- are measurable and Lebesgue integrable as bounded by $|f|$. For a set $E \in \mathfrak{M}$ we define

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

If $f : X \to \overline{\mathbb{R}}$ is measurable and each of the integrals $\int_E f^+\, d\mu$, $\int_E f^-\, d\mu$ is finite, then we also define

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

Theorem 40 *Suppose that $f, g \in L^1(\mu)$. If $\alpha, \beta \in \mathbb{C}$, then $\alpha f + \beta g \in L^1(\mu)$ and*

$$\int_X (\alpha f + \beta g)\, d\mu = \alpha \int_X f\, d\mu + \beta \int_X g\, d\mu\,.$$

We leave the proof as a simple exercise.

This theorem says that $L^1(\mu)$ is a linear space and $f \mapsto \int_X f\, d\mu$ is a linear functional in $L^1(\mu)$.

Similarly we define $L^1(\mu, \mathbb{R}^n)$ as the class of all measurable mappings $f = (f_1, f_2, \ldots, f_n) : X \to \mathbb{R}^n$ such that $\int_X |f|\, d\mu < \infty$ and then we set

$$\int_X f\, d\mu = \left(\int_X f_1\, d\mu, \ldots, \int_X f_n\, d\mu\right).$$

Theorem 41 *If $f \in L^1(\mu, \mathbb{R}^n)$, then*

$$\left|\int_X f\, d\mu\right| \leq \int_X |f|\, d\mu\,.$$

Proof. Let $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{R}^n$ be a unit vector such that $|\int_X f\, d\mu| = \langle \alpha, \int_X f\, d\mu\rangle$. Then

$$\begin{aligned}\left|\int_X f\, d\mu\right| &= \langle \alpha, \int_X f\, d\mu\rangle = \sum_{i=1}^n \alpha_i \int_X f_i\, d\mu \\ &= \int_X \sum_{i=1}^n \alpha_i f_i\, d\mu = \int_X \langle \alpha, f\rangle\, d\mu \leq \int_X |f|\, d\mu\end{aligned}$$

by Schwarz inequality. □

Theorem 42 (Lebesgue's dominated convergence theorem) *Suppose $\{f_n\}$ is a sequence of complex valued functions on X such that the limit $f(x) = \lim_{n\to\infty} f_n(x)$ exists for every $x \in X$. If there is a function $g \in L^1(\mu)$ such that*

$$|f_n(x)| \leq g(x) \quad \textit{for every } x \in X \textit{ and all } n = 1, 2, \ldots,$$

then $f \in L^1(\mu)$,

$$\lim_{n\to\infty}\int_X |f - f_n|\, d\mu = 0$$

and

$$\lim_{n\to\infty}\int_X f_n\, d\mu = \int_X f\, d\mu.$$

Proof. Clearly f is measurable and $|f| \le g$. Hence $f \in L^1(\mu)$. It is easy to see that $2g - |f - f_n| \ge 0$ and hence Fatou's lemma yields

$$\begin{aligned}
\int_X 2g\, d\mu &\le \liminf_{n\to\infty}\int_X (2g - |f - f_n|)\, d\mu \\
&= \int_X 2g + \liminf_{n\to\infty}\left(-\int_X |f - f_n|\, d\mu\right) \\
&= \int_X 2g - \limsup_{n\to\infty}\int_X |f - f_n|\, d\mu\,.
\end{aligned}$$

The integral $\int_X 2g\, d\mu$ is finite and hence

$$\limsup_{n\to\infty}\int_X |f_n - f|\, d\mu \le 0\,.$$

This immediately implies

$$\lim_{n\to\infty}\int_X |f_n - f|\, d\mu = 0\,.$$

Now

$$\left|\int_X f_n\, d\mu - \int_X f\, d\mu\right| \le \int_X |f_n - f|\, d\mu \to 0\,,$$

so

$$\lim_{n\to\infty}\int_X f_n\, d\mu = \int_X f\, d\mu\,.$$

The proof is complete. □

2.3 Almost everywhere

DEFINITION. If there is a set $N \in \mathfrak{M}$ of measure zero, $\mu(N) = 0$ such that a property $P(x)$ is satisfied for all $x \in X \setminus N$, then we say that the property $P(x)$ is satisfied *almost everywhere* (*a.e.*)

For example $f_n \to f$ a.e. means that there is a set $N \in \mathfrak{M}$, $\mu(N) = 0$ such that $f_n(x) \to f(x)$ for all $x \in X \setminus N$.

If $f = g$ a.e., then we write $f \sim g$. It is easy to see that $\sim$ is an equivalence relation ($f \sim g, g \sim h \Rightarrow f \sim h$ follows from the fact that the union of two sets of measure zero is a set of measure zero). Note that if $f \sim g$, then

$$\int_E f\, d\mu = \int_E g\, d\mu \quad \text{for every } X \in \mathfrak{M}.$$

Therefore from the point of view of the integration theory, functions which are equal a.e. cannot be distinguished. Moreover if $\int_E f\, d\mu = \int_E g\, d\mu$ for all $E \in \mathfrak{M}$, then $f = g$ a.e. This follows from part (b) of Theorem 43 below applied to $f - g$. Thus two functions f and g cannot be distinguished in the integration theory if and only if they are equal a.e. Therefore it is natural to *identify* such functions.

Theorem 43

(a) *Suppose $f : X \to [0, \infty]$ is measurable, $E \in \mathfrak{M}$ and $\int_E f\, d\mu = 0$, then $f = 0$ a.e. in E.*

(b) *Suppose $f \in L^1(\mu)$ and $\int_E f\, d\mu = 0$ for every $E \in \mathfrak{M}$. Then $f = 0$ a.e. in X.*

Proof. (a) Suppose $A_n = \{x \in E : \ f(x) \geq 1/n\}$. Then

$$\frac{1}{n}\mu(A_n) \leq \int_{A_n} f\, d\mu \leq \int_E f\, d\mu = 0.$$

Hence $\mu(A_n) = 0$ and therefore the set $\{x \in E : \ f(x) > 0\} = \bigcup_{n=1}^{\infty} A_n$ has measure zero.

(b) Let $= u + iv$. Define $E = \{x \in X : \ u(x) \geq 0\}$. Then

$$0 = \text{ real part of} \int_E f\, d\mu = \int_E u\, d\mu = \int_E u^+\, d\mu,$$

and hence $u^+ = 0$ a.e. by (a). Similarly $u^-, v^+, v^- = 0$ a.e. Accordingly, $f = 0$ a.e. □

Recall that a measure μ is complete if every subset of a set of measure zero is measurable (and hence has measure zero). All the measures obtained through the Carathéodory construction are complete (Proposition 4), e.g. the Hausforff measure and the Lebesgue measure are complete. However, we can assume that any measure is complete, because we can add all subsets of sets of measure zero to the σ-algebra. More precisely one can prove.

Theorem 44 *Let $(X, \mathfrak{M}, \mu)$ be a measure space, let $\mathfrak{M}^*$ be the collection of all $E \subset X$ for which there exist sets $A, B \in \mathfrak{M}$ such that $A \subset E \subset B$ and $\mu(B \setminus A) = 0$ and define $\mu(E) = \mu(A)$ in this situation. Then $\mathfrak{M}^*$ is a σ-algebra and μ is a complete measure on $\mathfrak{M}^*$.*

For a proof, see Rudin page 28.

Therefore, whenever needed, we can assume that the measure is complete.

Remark. If μ is a measure on $\mathfrak{B}(\mathbb{R}^n)$ and there is an uncountable set $E \in \mathfrak{B}(\mathbb{R}^n)$ such that $\mu(E) = 0$, then μ cannot be complete, because one can prove that the cardinality of all Borel sets in $\mathfrak{B}(\mathbb{R}^n)$ is the same as the cardinality of real numbers and the cardinality of all subsets of E is the same as the cardinality of all subsets of real numbers which is bigger than the cardinality of real numbers. Hence in order to make the measure μ complete we have to enlarge $\mathfrak{B}(\mathbb{R}^n)$ to a bigger σ-algebra by including also sets which are not Borel. This is the case of the Lebesgue measure. The σ-algebra of Lebesgue measurable sets is larger, than the σ-algebra of Borel sets.

DEFINITION. Suppose now that μ is a complete measure. If a function with values in a metric space Y is defined only at almost all points of X, i.e.

$$f : X \setminus N \to Y, \quad \mu(N) = 0\,,$$

then we say that f is measurable if there is $y_0 \in Y$ such that

$$\tilde{f}(x) = \begin{cases} f(x) & \text{if } x \in X \setminus N, \\ y_0 & \text{if } x \in N, \end{cases}$$

is measurable. In such a situation we say that $f : X \to Y$ is a *measurable function defined a.e.*

Note that it follows from the assumption that μ is complete that f can be defined in N in an *arbitrary* way and still the resulting function $\tilde{\tilde{f}}$ is measurable and $\tilde{\tilde{f}} \sim \tilde{f}$.

In particular functions in $L^1(\mu)$ need to be defined a.e. only. $L^1(\mu)$ is a linear space. Since we identify functions that are equal a.e. we identify $L^1(\mu)$ with the quotient space $L^1(\mu)/\sim$ (check that $L^1(\mu)/\sim$ is a linear space). Therefore $f \in L^1(\mu)$ satisfies $f = 0$ in $L^1(\mu)$ if and only if $f = 0$ a.e., if and only if $\|f\|_1 = \int_X |f|\, d\mu = 0$ (see Theorem 43(a)).

In all the previously discussed convergence theorems (Theorem 36, Theorem 37, Theorem 38, Theorem 42) we can replace the requirement of everywhere convergence by the a.e. convergence. For example the Lebesgue dominated convergence theorem can be formulated as follows.

Theorem 45 (Lebesgue dominated convergence theorem) *Suppose that $\{f_n\}$ is a sequence of complex valued functino on X defined a.e. If the limit $f(x) = \lim_{n\to\infty} f_n(x)$ exists a.e. and if there is a function $g \in L^1(\mu)$ such that*

$$|f_n(x)| \leq g(x) \text{ a.e.}, \quad \text{for all } n = 1, 2, 3, \ldots$$

then $f \in L^1(\mu)$,

$$\lim_{n\to\infty} \int_X |f - f_n|\, d\mu = 0$$

and

$$\lim_{n\to\infty}\int_X f_n\,d\mu = \int_X f\,d\mu\,.$$

Not surprisingly, this theorem is an easy consequence of the previous version of the Lebesgue dominated convergence theorem. We leave details to the reader.

Recall that if Y is a metric space, then a mapping $f:\mathbb{R}^n\to Y$ is Lebesgue measurable if $f^{-1}(U)$ is a Lebesgue measurable set for every open $U\subset Y$. Since the σ-algebra of Lebesgue measurable sets is larger, than that of Borel sets, the mapping f need not be Borel. I turns out, however, that f equals a.e. to a Borel mapping, at least when Y is separable.

Theorem 46 *If Y is a separable metric space and $f:\mathbb{R}^n\to Y$ is Lebesgue measurable, then there is a Borel mapping $g:\mathbb{R}^n\to Y$ such that $f=g$ a.e.*

Proof. Since Y is separable, there is a countable family of open sets $\{U_i\}_{i=1}^\infty$ in Y such that every open set in Y is a union of open sets from the family. Indeed, it suffices to take the family of balls with rational length radii and centers in a countable dense set in Y.

For $i=1,2,\ldots$ let $M_i\subset\mathbb{R}^n$ be a F_σ set such that

$$M_i\subset f^{-1}(U_i),\qquad \mathcal{L}_n(f^{-1}(U_i)\setminus M_i)=0\,,$$

see Theorem 19. Clearly the set $E=\bigcup_{i=1}^\infty(f^{-1}(U_i)\setminus M_i)$ has Lebesgue measure zero and hence there is a G_δ set H such that $E\subset H$, $\mathcal{L}_n(H)=0$ (Theorem 19). Fix $y_0\in Y$ and define

$$g(x)=\begin{cases} f(x) & \text{if } x\notin H,\\ y_0 & \text{if } x\in H.\end{cases}$$

Clearly $f=g$ a.e. and it remains to prove that the mapping g is Borel. It is easy to see that[5] if $y_0\notin U_i$, then $g^{-1}(U_i)=M_i\setminus H$ and if $y_0\in U_i$, then $g^{-1}(U_i)=M_i\cup H$. In each case $g^{-1}(U_i)$ is a Borel set. Since every open set $U\subset Y$ can be represented as $U=\bigcup_{j=1}^\infty U_{i_j}$, the set

$$g^{-1}(U)=\bigcup_{j=1}^\infty g^{-1}(U_{i_j})$$

is Borel. □

[5] A good picture is useful here.

2.4 Theorems of Lusin and Egorov

Theorem 47 (Lusin) *Let X be a metric space and μ a measure in $\mathfrak{B}(X)$ such that X is a union of countably many open sets of finite measure. If $f : X \to Y$ is a Borel mapping with values in a separable metric space, then for every $\varepsilon > 0$ there is a closed set $F \subset X$ such that $\mu(X \setminus F) < \varepsilon$ and $f|_F$ is continuous.*

Proof. Let $\{U_i\}_{i=1}^\infty$ be a countable family of open sets such that every open set in Y is a union of open sets from the family. It follows from Theorem 7 that for every i there is an open set $V_i \subset X$ such that[6]

$$f^{-1}(U_i) \subset V_i, \qquad \mu(V_i \setminus f^{-1}(U_i)) < \varepsilon 2^{-i-1}\,.$$

Let

$$E = \bigcup_{i=1}^\infty V_i \setminus f^{-1}(U_i).$$

Clearly $\mu(E) < \varepsilon/2$. We will prove that the mapping $g = f|_{X\setminus E}$ is continuous. Observe that[7]

$$g^{-1}(U_i) = V_i \cap (X \setminus E)\,. \tag{28}$$

Indeed, $g^{-1}(U_i) \subset V_i \cap (X \setminus E)$, but on the other hand

$$V_i \cap (X \setminus E) \subset V_i \cap (X \setminus (V_i \setminus f^{-1}(U_i))) = f^{-1}(U_i)\,,$$

so

$$V_i \cap (X \setminus E) \subset f^{-1}(U_i) \cap (X \setminus E) = g^{-1}(U_i)\,.$$

In order to prove that g is continuous it suffices to prove that for every open set $U \subset Y$ there is an open set $V \subset X$ such that[8] $g^{-1}(U) = V \cap (X \setminus E)$. Let $U = \bigcup_{j=1}^\infty U_{i_j}$. Then (28) gives $g^{-1}(U) = (\bigcup_{j=1}^\infty V_{i_j}) \cap (X \setminus E)$ which proves continuity of g in $X \setminus E$. The set E need not be closed, but it follows from Theorem 7 that there is an open set $G \subset X$ such that $E \subset G$ and $\mu(G \setminus E) < \varepsilon/2$. Now $f|_F$ is continuous, where $F = X \setminus G$, F is closed and $\mu(X \setminus F) = \mu(G) = \mu(E) + \mu(G \setminus E) < \varepsilon$. □

As a corollary we will prove the following characterization of Lebesgue measurable functions also known as the Lusin theorem.

Let $E \subset \mathbb{R}^n$ be a Lebesgue measurable set. Recall that $f : E \to \mathbb{R}$ is Lebesgue measurable if $f^{-1}(U)$ is Lebesgue measurable for every open $U \subset \mathbb{R}$ ($f^{-1}(U)$ is a subset of E).

[6]Actually to prove this fact we need first to split $f^{-1}(U_i)$ into sets of finite measure and then apply Theorem 7 to each such set.

[7]It is easy to see if you make an appropriate picture, but a formal proof is enclosed below.

[8]Because the class of open sets in the metric space $X \setminus E$ is exactly the class of sets of the form $V \cap (X \setminus E)$, where $V \subset X$ is open.

Theorem 48 (Lusin) *Let $E \subset \mathbb{R}^n$ be Lebesgue measurable. Then a function $f : E \to \mathbb{R}$ is Lebesgue measurable if and only if for every $\varepsilon > 0$ there is a closed set $F \subset E$ such that $f|_F$ is continuous and $\mathcal{L}_n(E \setminus F) < \varepsilon$.*

Proof. Since f equals a.e. to a Borel function necessity follows from Theorem 47. To prove sufficiency it suffices to show that for every $a \in \mathbb{R}$ the set $E_a = \{x : f(x) \geq a\}$ is Lebesgue measurable (this is a variant of the exercise from page 25.) Let F be a closed set such that $f|_F$ is continuous and $\mathcal{L}_n(E \setminus F) < \varepsilon$. Then the set $F' = \{x : (f|_F)(x) \geq a\} = E_a \cap F$ is closed and $\mathcal{L}_n^*(E_a \setminus F') \leq \mathcal{L}_n(E \setminus F) < \varepsilon$, hence measurability of E_a follows from Theorem 19. □

In the next theorem μ is a measure on an arbitrary set X (not necessarily a metric space), but it is crucial that $\mu(X) < \infty$.

Theorem 49 (Egorov) *Let μ be a finite measure on a set X. Let $f_n : X \to Y$ be a sequence of measurable functions with values in a separable metric space. If $f_n \to f$ a.e., then f is measurable and for every $\varepsilon > 0$ there is a measurable set $E \subset X$ such that $\mu(X \setminus E) < \varepsilon$ and $f_n \rightrightarrows f$ uniformly on E.*

Proof. Measurability of f follows from Theorem 33. Let d be the metric in Y. Define

$$C_{i,j} = \bigcup_{n=j}^{\infty} \{x : \; d(f_n(x), f(x)) \geq 2^{-i}\}\,.$$

We will prove that this set is measurable. First observe that the mapping $F = (f_n, f) : X \to Y \times Y$ is measurable. Indeed, if $\{U_i\}_{i=1}^{\infty}$ is a collection of open sets in Y such that every open set in Y is a countable union of open sets from the family, it easily follows that every open set in $Y \times Y$ is a countable union of open sets of the form $U_i \times U_j$. Therefore to prove measurability of F it suffices to observe that each set

$$F^{-1}(U_i \times U_j) = f_n^{-1}(U_i) \cap f^{-1}(U_j)$$

is measurable. Since the function $d : Y \times Y \to \mathbb{R}$ is continuous, the composed map $d \circ F : X \to \mathbb{R}$ is measurable and hence

$$\begin{aligned} \{x : \; d(f_n(x), f(x)) \geq 2^{-i}\} &= F^{-1}\{(y_1, y_2) \in Y \times Y : d(y_1, y_2) \geq 2^{-i}\} \\ &= F^{-1}(d^{-1}([2^{-i}, \infty)) = (d \circ F)^{-1}([2^{-i}, \infty)) \end{aligned}$$

is measurable. That proves measurability of the sets $C_{i,j}$. Now observe that $C_{i,j}$ is a decreasing sequence of sets with respect to j, $\mu(C_{i,1}) \leq \mu(X) < \infty$ and hence

$$\mu(C_{i,j}) \to \mu(\bigcap_{k=1}^{\infty} C_{i,k}) = 0 \quad \text{as } j \to \infty \tag{29}$$

by Theorem 3(g). Indeed, this intersection has measure zero because if $x \in \bigcap_{k=1}^{\infty} C_{i,k}$, then there is a subsequence $f_{n_j}(x)$ such that $d(f_{n_j}(x), f(x)) \geq 2^{-i}$ for all j, which can be true only for x in a set of measure zero. Now it follows from (29) that for every i there is an integer $J(i)$ such that $\mu(C_{i,J(i)}) < \varepsilon 2^{-i}$. It is easy to see that $\mu(\bigcup_{i=1}^{\infty} C_{i,J(i)}) < \varepsilon$ and that f_n converges uniformly to f on $E = X \setminus \bigcup_{i=1}^{\infty} C_{i,J(i)}$, because for every i $d(f_n(x), f(x)) < 2^{-i}$ for all $n \geq J(i)$ and all $x \in E$. □

Again if X is a metric space and μ is a finite measure of $\mathfrak{B}(X)$ we can assume that the set E in the above theorem is closed, see Theorem 7.

Exercise. *Show an example that if $\mu(X) = \infty$, then the claim of Egorov's theorem may be false.*

2.5 Convergence in measure

Definition. We say that a sequence of measurable functions $f_n : X \to \mathbb{R}$ *converges in measure* to a measurable function $f : X \to \mathbb{R}$ if for every $\varepsilon > 0$

$$\lim_{n\to\infty} \mu(\{x : \ |f_n(x) - f(x)| \geq \varepsilon\}) = 0.$$

We denote convergence in measure by $f_n \xrightarrow{\mu} f$. As usual we assume that the functions f_n and f are defined a.e.

Theorem 50 (Lebesgue) *If $\mu(X) < \infty$ and a sequence of measurable functions $f_n : X \to \mathbb{R}$ converges to $f : X \to \mathbb{R}$ a.e., then it converges in measure, $f_n \xrightarrow{\mu} f$.*

Proof. Given $\varepsilon > 0$, let

$$E_i = \{x : \ |f_i(x) - f(x)| \geq \varepsilon\} .$$

The sequence of sets $\bigcup_{i=n}^{\infty} E_i$ is a decreasing sequence of sets of finite measure and hence

$$\mu\left(\bigcup_{i=n}^{\infty} E_i\right) \to \mu\left(\bigcap_{n=1}^{\infty}\bigcup_{i=n}^{\infty} E_i\right) = 0.$$

The last set has measure zero, because if $x \in \bigcup_{i=n}^{\infty} E_i$ for every n, then x belongs to infinitely many E_i's so there is a subsequence f_{n_i} such that $|f_{n_i}(x) - f(x)| \geq \varepsilon$ which is true only on a set of measure zero. Hence

$$\mu(\{x : \ |f_n(x) - f(x)| \geq \varepsilon\}) = \mu(E_n) \leq \mu(\bigcup_{i=n}^{\infty} E_i) \to 0,$$

which proves convergence in measure. □

Theorem 51 (Riesz) *If $f_n \xrightarrow{\mu} f$, then there is a subsequence f_{n_i} such that $f_{n_i} \to f$ a.e.*

Remark. In this theorem measure of the space can be infinite, $\mu(X) = \infty$.

Proof. For every i there is n_i such that

$$\mu(\{x : \ |f_{n_i}(x) - f(x)| \geq \frac{1}{i}\}) \leq \frac{1}{2^i}\,.$$

We can assume that $n_1 < n_2 < n_3 \ldots$ Let

$$F_k = X \setminus \bigcup_{i=k}^{\infty} \{x : \ |f_{n_i}(x) - f(x)| \geq \frac{1}{i}\}\,.$$

Clearly $\mu(X \setminus F_k) \leq \sum_{i=k}^{\infty} 2^{-i} = 2^{-k+1}$ and hence $\mu(X \setminus \bigcup_{k=1}^{\infty} F_k) = 0$. We will prove that $f_{n_i} \to f$ pointwise on $\bigcup_{k=1}^{\infty} F_k$. To this end it suffices to show that f_{n_i} converges to f pointwise on every set F_k, but we will actually show a stronger fact that $f_{n_i} \rightrightarrows f$ uniformly on F_k. Indeed, for all $x \in F_k$, x does not belong to the set $\bigcup_{i=k}^{\infty} \{x : |f_{n_i}(x) - f(x)| \geq 1/i\}$, so for every $j \geq k$ and all $x \in F_k$

$$|f_{n_j}(x) - f(x)| \leq \frac{1}{j}$$

which proves uniform convergence of f_{n_j} to f on F_k □

Remark. We actually proved not only convergence a.e. but also uniform convergence on subsets of X whose complement has arbitrary small measure. Note that Lebesgue's theorem combined with this stronger conclusion of Riesz' theorem implies Egorov's theorem for sequences of real valued functions.

Exercise. *Generalize the Lebesgue and the Riesz theorems to the case of sequences of measurable mappings with values into separable metric spaces.*

2.6 Absolute continuity of the integral

The following result is known as the absolute continuity of the integral.

Theorem 52 *Let $f \in L^1(\mu)$. Then for every $\varepsilon > 0$ there is $\delta > 0$ such that $\int_E |f|\, d\mu < \varepsilon$ whenever $\mu(E) < \delta$.*

Proof. For if not, there would exist $\varepsilon_0 > 0$ and sets $E_n \in \mathfrak{M}$ such that $\mu(E_n) \leq 2^{-n}$ and $\int_{E_n} |f|\, d\mu \geq \varepsilon_0$. The sequence of sets $A_k = \bigcup_{n=k}^{\infty} E_n$ is decreasing and

$$\mu\left(\bigcap_{k=1}^{\infty} A_k\right) = 0\,, \tag{30}$$

because $\mu(\bigcap_{k=1}^{\infty} A_k) \leq \mu(A_n) \leq \sum_{i=n}^{\infty} \mu(E_i) \leq 2^{-n+1}$ for every n. Since $\varphi(E) = \int_E |f|\, d\mu$ is a measure on $\mathfrak{M}$ (Theorem 39) and $\varphi(A_1) = \int_{A_1} |f|\, d\mu \leq \int_X |f|\, d\mu < \infty$, Theorem 3(g) implies that

$$\varphi(A_k) \to \varphi\left(\bigcap_{k=1}^{\infty} A_k\right) = \int_{\bigcap_{k=1}^{\infty} A_k} |f|\, d\mu = 0$$

by (30). This is, however, a clear contradiction, because $\varphi(A_k) \geq \varphi(E_k) = \int_{E_k} |f|\, d\mu \geq \varepsilon_0$. □

2.7 Riesz representation theorem

Definition. If $f : X \to \mathbb{R}$ is a continuous function on a metric space X, then $\operatorname{supp} f = \overline{\{x \in X : f(x) \neq 0\}}$ is called the *support* of f and $C_c(X) = C_c(X, \mathbb{R})$ denotes the linear space of continuous functions on X with compact support. We say that a linear functional $I : C_c(X) \to \mathbb{R}$ is *positive* if

$$I(f) \geq 0 \quad \text{whenever } f \geq 0.$$

Let X be a locally compact metric space and μ a measure on $\mathfrak{B}(X)$ such that $\mu(X) < \infty$ for every compact set $K \subset X$, i.e. μ is a Radon measure, see Corollary 9. Then

$$I(f) = \int_X f\, d\mu$$

is a positive linear functional on $C_c(X)$. Surprisingly, every positive linear functional on $C_c(X)$ can be represented as an integral with respect to a Radon measure.

Theorem 53 (Riesz representation theorem) *Let X be a locally compact metric space*[9] *and let $I : C_c(X) \to \mathbb{R}$ be a positive linear functional. Then there is a unique Radon measure μ such that*

$$I(f) = \int_X f\, d\mu\,. \tag{31}$$

Proof. For an open set $G \subset X$ denote by $\Gamma(G)$ the class of continuous functions $0 \leq f \leq 1$ with compact support in G.

Lemma 54 *Let $G_1, G_2, \ldots, G_k \subset X$ be open sets. If $f \in \Gamma(G_1 \cup \ldots \cup G_k)$, then there exist functions $f_i \in \Gamma(G_i)$, $i = 1, 2, \ldots$ such that $f = f_1 + \ldots + f_k$.*

[9] This theorem is true for a more general class of locally compact Hausdorff topological spaces.

Proof. Since X is locally compact, there is a compact set E such that

$$\operatorname{supp} f \subset \operatorname{int} E, \qquad E \subset G_1 \cup \ldots \cup G_k\,.$$

The sets $H_i = G_i \cap E$ form an open covering of the compact space E. If we can find continuous nonnegative functions $\varphi_1, \ldots, \varphi_k$ in E such that

$$\varphi = \varphi_1 + \ldots + \varphi_k > 0 \text{ in } E, \quad \operatorname{supp} \varphi_i \subset H_i,$$

then the functions

$$f_i(x) = \begin{cases} f(x)\varphi_i(x)/\varphi(x) & \text{if } x \in E, \\ 0 & \text{if } x \notin E \end{cases}$$

will have desired properties. To construct functions φ_i observe that there is an open covering of E by sets D_i, $i = 1, 2, \ldots, k$ such that $\overline{D_i} \subset H_i$ and then we define $\varphi_i(x) = \operatorname{dist}(X, E \setminus D_i)$. □

Let now $I : C_c(X) \to \mathbb{R}$ be a positive linear functional. Obviously

$$I(0) = 0 \quad \text{and} \quad I(f) \leq I(g) \text{ for } f \leq g$$

(because $I(g - f) \geq 0$).

Exercise. *Prove that if $\psi : [0, \infty) \to [0, \infty)$ is an increasing function such that $\psi(s + t) = \psi(s) + \psi(t)$, then $\psi(t) = t\psi(1)$.*

It follows from the exercise that

$$I(tf) = tI(f) \quad \text{for all } t \geq 0 \text{ and } f \in C_c(X). \tag{32}$$

Now for $E \in \mathfrak{B}(X)$ we define

$$\mu(E) = \inf_{E \subset G-\text{open}} \lambda(G)\,, \tag{33}$$

where

$$\lambda(G) = \sup_{f \in \Gamma(G)} I(f)\,.$$

In order to prove that μ is a measure on $\mathfrak{B}(X)$ it suffices to prove that for any open sets $G_1, G_2, \ldots$

$$\lambda(\bigcup_{n=1}^{\infty} G_n) \leq \sum_{n=1}^{\infty} \lambda(G_n) \tag{34}$$

and

$$\lambda(G_1 \cup G_2) = \lambda(G_1) + \lambda(G_2) \quad \text{provided } G_1 \cap G_2 = \emptyset. \tag{35}$$

Indeed, the two properties easily imply that the function μ^* defined on the class of *all* subsets of X by (33) is a metric outer measure.

Let $\{G_n\}_{n=1}^\infty$ be a sequence of open sets and $L < \lambda(\bigcup_{n=1}^\infty G_n)$. Then there is a function $f \in \Gamma(\bigcup_{n=1}^\infty G_n)$ such that $I(f) > L$. Since $\operatorname{supp} f$ is compact, $f \in \Gamma(\bigcup_{n=1}^N G_n)$ for some finite N. According to the lemma $f = f_1 + \ldots + f_N$, $f_i \in \Gamma(G_i)$, so

$$L < I(f) = \sum_{n=1}^N I(f_n) \leq \sum_{n=1}^\infty \lambda(G_n)$$

and (34) easily follows.

If $f \in \Gamma(G_1)$, $g \in \Gamma(G_2)$, $G_1 \cap G_2 = \emptyset$, then $f + g \in \Gamma(G_1 \cup G_2)$. Indeed, the condition $G_1 \cap G_2 = \emptyset$ implies that $f + g \leq 1$. Hence

$$I(f) + I(g) = I(f + g) \leq \lambda(G_1 \cup G_2)$$

and (35) follows upon taking the supremum over f and g.

We proved that μ is a measure in $\mathfrak{B}(X)$. The measure μ is finite on compact sets. Indeed, if $K \subset G \subset \overline{G}$ and $\overline{G}$ is compact, then there is a function $h \in C_c(X)$ such that $h = 1$ on $\overline{G}$ and therefore

$$\mu(K) \leq \sup_{f \in \Gamma(G)} I(f) \leq I(h) < \infty\,.$$

Clearly

$$\mu(G) = \sup_{f \in \Gamma(G)} I(f) \quad \text{for every open } G \subset X$$

and

$$\mu(E) = \inf_{E \subset G-\text{open}} \mu(G) \quad \text{for every } E \in \mathfrak{B}(X).$$

Let $g \in C_c(X)$, $g \geq 0$. Since $I_g(f) = I(fg)$ is a positive functional it follows from what we have already proved that

$$\mu_g(G) = \sup_{f \in \Gamma(G)} I(fg) \quad \text{for every open } G \subset X$$

and

$$\mu_g(E) = \inf_{E \subset G-\text{open}} \mu_g(G) \quad \text{for every } E \in \mathfrak{B}(X)$$

is a measure in $\mathfrak{B}(X)$ that is finite on compact sets.

Lemma 55 *$I(g) = \mu_g(X)$.*

Proof. For $f \in \Gamma(X)$, $fg \leq g$ and hence $I(fg) \leq I(g)$. On the other hand if $f \in \Gamma(X)$, $f = 1$ on $\operatorname{supp} g$, then $fg = g$ and hence $I(fg) = I(g)$. Therefore

$$\mu_g(X) = \sup_{f \in \Gamma(X)} I(fg) = I(g)\,.$$

□

Lemma 56 $(\inf_E g)\mu(E) \le \mu_g(E) \le (\sup_E g)\mu(E)$ *for every* $E \in \mathfrak{B}(X)$.

Proof. If $\operatorname{supp} f \subset G$, then it follows from (32) that

$$(\inf_G g) I(f) \le I(fg) \le (\sup_G g) I(f)\,.$$

Taking the supremum over $f \in \Gamma(G)$ yields

$$(\inf_G g)\mu(G) \le \mu_g(G) \le (\sup_G g)\mu(G)$$

and then the lemma follows after taking the infimum over all open sets G such that $E \subset G$. This is because continuity of g implies

$$\inf_{E\subset G-\text{open}} (\inf_G g) = \inf_E g \qquad \text{and} \qquad \inf_{E\subset G-\text{open}} (\sup_G g) = \sup_E g\,.$$

□

Lemma 57 $\mu_g(E) = \int_E g\, d\mu$ *for all* $E \in \mathfrak{B}(X)$.

Proof. The measure μ is finite on compact sets and hence every Borel set is a union of sets of finite measure (Corollary 9). Therefore it suffices to prove the equality for sets E of finite measure $\mu(E) < \infty$. Given $\varepsilon > 0$ let $s = \sum_{k=1}^n a_k \chi_{E_k}$ be a simple function such that $s \le g \le s + \varepsilon$ (Theorem 34). Here the sets E_k are pairwise disjoint and[10] $\bigcup_{k=1}^n E_k = X$.

Then

$$\begin{aligned} \int_E s\, d\mu &= \sum_{k=1}^n a_k \mu(E_k \cap E) \le \sum_{k=1}^n (\inf_{E_k\cap E} g)\mu(E_k \cap E) \\ &\le \sum_{k=1}^n \mu_g(E_k \cap E) = \mu_g(E) \end{aligned}$$

and

$$\begin{aligned} \int_E (s+\varepsilon)\, d\mu &= \sum_{k=1}^n (a_k + \varepsilon)\mu(E_k \cap E) \ge \sum_{k=1}^n (\sup_{E_k\cap E} g)\mu(E_k \cap E) \\ &\ge \sum_{k=1}^n \mu_g(E_k \cap E) = \mu_g(E)\,. \end{aligned}$$

We proved that

$$\int_E (g-\varepsilon)\, d\mu \le \int_E s\, d\mu \le \mu_g(E) \le \int_E (s+\varepsilon)\, d\mu \le \int_E (g+\varepsilon)\, d\mu\,.$$

[10] i.e. we include the set where $a_k = 0$. For example if $s = \chi_E$, then we write $s = 1 \cdot \chi_E + 0 \cdot \chi_{X\setminus E}$.

Since $\varepsilon > 0$ was arbitrary and $\mu(E) < \infty$, the lemma easily follows. $\square$

Comparing Lemma 55 and Lemma 57 we obtain that $I(g) = \int_X g\,d\mu$ which proves (31) for $g \geq 0$. The general case follows from the fact that every function in $C_c(X)$ is a difference of two nonnegative functions.

We are left with the proof of uniqueness of the measure. Suppose that

$$I(f) = \int_X f\,d\mu = \int_X f\,d\nu \quad \text{for all } f \in C_c(X),$$

where μ and ν are measures on X that are finite on compact sets. According to Corollary 9 and Corollary 8 it suffices to prove that the measures μ and ν agree on the class of open sets.

Let G be an open set and $K \subset G$ a compact set. Let $f \in \Gamma(G)$ be such that $f = 1$ on K. Then

$$\mu(G) \geq \int_G f\,d\mu = I(f) = \int_G f\,d\nu \geq \nu(K)\,.$$

Taking supremum over compact sets $K \subset G$ and applying Corollary 9 we obtain that $\mu(G) \geq \nu(G)$. Similarly we prove the opposite inequality. The proof of the theorem is complete. $\square$

3 Integration on product spaces

This should be a word for word copy of Chapter 8 from Rudin's book. There is no need to rewrite the whole chapter, so I will not include anything here.

4 Vitali covering lemma, doubling and Hausdorff measures

4.1 Doubling measures

If $\sigma > 0$ and B is a ball in a metric space, then by σB we will denote a ball concentric with B and with the radius σ times that of B.

DEFINITION. Let X be a metric space and μ a measure on $\mathfrak{B}(X)$. We say that μ is a *doubling measure* if there is a constant $C_d > 0$ such that for every ball B, $0 < \mu(B) < \infty$ and $\mu(2B) \leq C_d\mu(B)$.

The Lebesgue measure is an example of a doubling measure with $C_d = 2^n$. It turns out that general doubling measures inherit many properties of the Lebesgue measure.

Definition. We say that a metric space X is a *doubling-metric* space if there is a constant $M > 0$ such that every ball B can be covered by no more than M balls of half the radius.

Examples. $\mathbb{R}^n$ is metric-doubling. If X is a metric-doubling space, then every subset of X is metric-doubling. In particular every subset of $\mathbb{R}^n$ is metric-doubling. If X is an infinite discrete space, then X is not metric-doubling.

Proposition 58 *If a metric space is complete and metric-doubling, then it is locally compact.*

Proof. It suffices to prove that any closed ball is compact. It easily follows from the metric-doubling condition that each closed ball is totally bounded. Since it is complete (as a closed subset of a complete metric space) it is compact. One can also prove this theorem more directly by mimicking the proof that the interval $[0, 1]$ is compact. $\square$

Proposition 59 *If μ is a doubling measure on a metric space X, then X is metric-doubling.*

Proof. We will prove that X satisfies the metric-doubling condition with $M = C_d^4$. Let $\{x_i\}_{i\in I} \subset B(x, r)$ be a maximal $r/2$-separated set, i.e.

$$\forall_{i,j\in I} \qquad d(x_i, x_j) \geq \frac{r}{2}$$

$$\forall_{y\in B(x,r)}\ \exists_{i\in I} \qquad d(x_i, y) < \frac{r}{2}\,.$$

From the second condition we have $B(x, r) \subset \bigcup_{i\in I} B(x_i, r/2)$ and it suffices to show that the number of elements in I is bounded by C_d^4, i.e. $\#I \leq C_d^4$. The balls $B(x_i, r/4)$ are pairwise disjoint and $B(x_i, r/4) \subset B(x, 2r) \subset B(x_i, 4r)$. We have

$$\mu(B(x, 2r)) \geq \sum_{i\in I} \mu(B(x_i, r/4)) \geq C_d^{-4} \sum_{i\in I} \mu(B(x_i, 4r)) \geq C_d^{-4} \mu(B(x, 2r))(\#I)$$

and hence $\#I \leq C_d^4$. $\square$

Theorem 60 (Volberg-Konyagin) *If X is a complete metric-doubling metric space, then there is doubling measure on X.*

This theorem is very difficult and we will not prove it. Note that according to the previous proposition metric-doubling condition is necessary for the existence of a doubling measure, so Volberg-Konyagin's theorem is very sharp.

The Euclidean space is metric doubling and hence any subset of the Euclidean space is metric-doubling. Therefore we have

Corollary 61 *If $X \subset \mathbb{R}^n$ is closed, then there is a doubling measure on X.*

Proposition 62 *Let μ be a doubling measure on a metric space X. Then there is a constant $C > 0$ such that*

$$\frac{\mu(B(x,r))}{\mu(B(x_0,r_0))} \geq C\left(\frac{r}{r_0}\right)^s$$

whenever $x \in B(x_0,r_0)$, $r \leq r_0$ and $s = \log C_d/\log 2$.

Remark. For the Lebesgue measure in $\mathbb{R}^n$, $C_d = 2^n$, $\log C_d/\log 2 = n$ and hence $\mathcal{L}_n(B(x,r))/\mathcal{L}(B(0,1)) \geq Cr^n$ which is the sharp lower bound estimate for the Lebesgue measure of the ball. That shows that the exponent s in the above lemma cannot be any smaller.

Proof. Let $x \in B(x_0,r_0)$, $r \leq r_0$. Let k be the largest integer such that $2^{k-1}r \leq 2r_0$. Hence $2^k r > 2r_0$ and thus $B(x_0,r_0) \subset B(x,2^k r)$. We have

$$\mu(B(x_0,r_0)) \leq \mu(B(x,2^k r)) \leq C_d^k \mu(B(x,r)),$$

$$\frac{\mu(B(x,r))}{\mu(B(x_0,r_0)} \geq C_d^{-k}\,. \tag{36}$$

Since $2^{k-1}r \leq 2r_0$, $r/r_0 \leq 4^{-1}2^{-k}$ and hence

$$\left(\frac{r}{r_0}\right)^s = \left(\frac{r}{r_0}\right)^{\log C_d/\log 2} \leq \underbrace{(4^{-1})^{\log C_d/\log 2}}_{A}(2^{-k})^{\log C_d/\log 2} = AC_d^{-k}.$$

This estimate together with (36) yields the claim. □

4.2 Covering lemmas

Theorem 63 (5r-covering lemma) *Let $\mathcal{B}$ be a family of balls in a metric space such that $\sup\{\operatorname{diam} B : B \in \mathcal{B}\} < \infty$. Then there is a subfamily of pairwise disjoint balls $\mathcal{B}' \subset \mathcal{B}$ such that*

$$\bigcup_{B\in\mathcal{B}} B \subset \bigcup_{B\in\mathcal{B}'} 5B\,.$$

If the metric space is separable, then the family $\mathcal{B}'$ is countable and we can arrange it as a sequence $\mathcal{B}' = \{B_i\}_{i=1}^\infty$, so

$$\bigcup_{B\in\mathcal{B}} B \subset \bigcup_{i=1}^\infty 5B_i\,.$$

Remark. Here $\mathcal{B}$ can be either a family of open balls or closed balls. In both cases proof is the same.

Proof. Let $\sup\{\operatorname{diam} B : \; B \in \mathcal{B}\} = R < \infty$. Divide the family $\mathcal{B}$ according to the diameter of the balls

$$\mathcal{F}_j = \{B \in \mathcal{B} : \frac{R}{2^j} < \operatorname{diam} B \leq \frac{R}{2^{j-1}}\} .$$

Clearly $\mathcal{B} = \bigcup_{j=1}^{\infty} \mathcal{F}_j$. Define $\mathcal{B}_1 \subset \mathcal{F}_1$ to be the maximal family of pairwise disjoint balls. Suppose the families $\mathcal{B}_1, \ldots, \mathcal{B}_{j-1}$ are already defined. Then we define $\mathcal{B}_j$ to be the maximal family of pairwise disjoint balls in

$$\mathcal{F}_j \cap \{B : \; B \cap B' = \emptyset \text{ for all } B' \in \bigcup_{i=1}^{j-1} \mathcal{B}_i\} .$$

Next we define $\mathcal{B}' = \bigcup_{j=1}^{\infty} \mathcal{B}_j$. Observe that every ball $B \in \mathcal{F}_j$ intersects with a ball in $\bigcup_{i=1}^{j} \mathcal{B}_j$. Suppose that $B \cap B_1 \neq \emptyset$, $B_1 \in \bigcup_{i=1}^{j} \mathcal{B}_i$. Then

$$\operatorname{diam} B \leq \frac{R}{2^{j-1}} = 2 \cdot \frac{R}{2^j} \leq 2 \operatorname{diam} B_1$$

and hence $B \subset 5B_1$. The proof is complete. □

Remark. We have actually proved more: every ball $B \in \mathcal{B}$ is contained in $5B_1$ for some ball $B_1 \in \mathcal{B}'$ such that $B_1 \cap B \neq \emptyset$. This fact will be used in the proof of the next result.

Definition. We say that a family $\mathcal{F}$ of closed balls covers a set A in the *Vitali sense* if for every $a \in A$ there is a ball $\overline{B} \in \mathcal{F}$ of arbitrarily small radius such that $a \in \overline{B}$. In other words

$$\forall_{a \in A} \quad \inf\{\operatorname{diam} \overline{B} : \; a \in \overline{B} \in \mathcal{F}\} = 0.$$

Theorem 64 (Vitali covering theorem) *Let μ be a doubling measure on X and $A \subset X$ a measurable set. Let $\mathcal{F}$ be a family of closed balls that covers A in the Vitali sense. Then there is a subfamily $\mathcal{G} \subset \mathcal{F}$ of pairwise disjoint balls such that*

$$\mu\big(A \setminus \bigcup_{\overline{B} \in \mathcal{G}} \overline{B}\big) = 0 . \tag{37}$$

Proof. Let $\mathcal{F}' \subset \mathcal{F}$ be a subfamily of all balls with diameters less than 1. Observe that the family $\mathcal{F}'$ also covers A in the Vitali sense. According to the $5r$-covering lemma there is a countable subfamily $\mathcal{G} \subset \mathcal{F}'$ of pairwise disjoint balls such that

$$A \subset \bigcup_{\overline{B} \in \mathcal{F}'} \overline{B} \subset \bigcup_{\overline{B} \in \mathcal{G}} 5\overline{B} .$$

We will prove that the family $\mathcal{G}$ satisfies (37). To this end it suffices to prove that for any ball Q of radius bigger than 1

$$\mu\big(Q \cap (A \setminus \bigcup_{\overline{B}\in\mathcal{G}} \overline{B})\big) = 0\,.$$

Observe that

$$Q \cap \big(A \setminus \bigcup_{\overline{B}\in\mathcal{G}} \overline{B}\big) = Q \cap \big(A \setminus \bigcup_{\substack{\overline{B}\in\mathcal{G}\\ \overline{B}\subset 10Q}} \overline{B}\big)\,.$$

Indeed, if $\overline{B}$ is not contained in $10Q$, then[11] $\overline{B} \cap (Q \cap A) = \emptyset$. The family of balls $\overline{B}$ such that $\overline{B} \in \mathcal{G}$ and $\overline{B} \subset 10Q$ is countable and hence we can name the balls by $\overline{B}_1, \overline{B}_2, \overline{B}_3, \ldots$ We have

$$\sum_{i=1}^{\infty} \mu(5\overline{B}_i) \leq C \sum_{i=1}^{\infty} \mu(\overline{B}_i) \leq C\mu(10Q) < \infty\,.$$

The first inequality follows from the doubling condition and the second inequality from the fact that the balls $\overline{B}_i$ are pairwise disjoint and contained in $10Q$. The above inequality implies that

$$\mu\Big(\bigcup_{i=N}^{\infty} 5\overline{B}_i\Big) \leq \sum_{i=N}^{\infty} \mu(5\overline{B}_i) \to 0 \qquad \text{as } N \to \infty$$

and it remains to prove that

$$Q \cap \Big(A \setminus \bigcup_{\substack{\overline{B}\in\mathcal{G}\\ \overline{B}\subset 10Q}} \overline{B}\Big) \subset Q \cap \big(A \setminus \bigcup_{i=1}^{N} \overline{B}_i\big) \subset \bigcup_{i=N}^{\infty} 5\overline{B}_i\,. \tag{38}$$

The first inclusion is obvious. To prove the second one let $a \in Q \cap (A \setminus \bigcup_{i=1}^{N} \overline{B}_i)$. The set $\bigcup_{i=1}^{N} \overline{B}_i$ is closed and $a \notin \bigcup_{i=1}^{N} \overline{B}_i$, so there is a ball $\overline{B}_a$ of sufficiently small radius $r < 1$ such that $a \in \overline{B}_a \in \mathcal{F}'$ and

$$\overline{B}_a \cap \bigcup_{i=1}^{N} \overline{B}_i = \emptyset\,. \tag{39}$$

On the other hand it follows from the remark stated after the proof of the $5r$-covering lemma that there is $\overline{B} \in \mathcal{G}$ such that $\overline{B}_a \cap \overline{B} \neq \emptyset$, $\overline{B}_a \subset 5\overline{B}$. Since $\operatorname{diam} \overline{B} < 1$ and $a \in Q$, it follows that $\overline{B} \subset 10Q$ and hence $\overline{B} = \overline{B}_j$ for some j. Now (39) implies that $j > N$ and hence

$$a \in 5\overline{B}_j \subset \bigcup_{i=N}^{\infty} 5\overline{B}_i$$

which proves (38). $\square$

[11]Because $\operatorname{diam} \overline{B} < 1$ and the radius of Q is bigger than 1.

Corollary 65 *For an open set $U \subset \mathbb{R}^n$ and every $\varepsilon > 0$ there is a sequence of pairwise disjoint closed balls $\overline{B}_1, \overline{B}_2, \overline{B_3}, \dots$ contained in U and of radii less than ε such that*

$$\mathcal{L}_n\big(U \setminus \bigcup_{i=1}^{\infty} \overline{B}_i\big) = 0\,.$$

Exercise. *Show a direct proof of the corollary without using the Vitali covering lemma.*

4.3 Doubling measures and the Hausdorff measure

We will snow now how to use the covering lemmas to prove important results about Hausdorff measures.

Theorem 66 *For any Borel set $A \subset \mathbb{R}^n$ and any $0 < \delta \leq \infty$*

$$\mathcal{L}_n(A) = \mathcal{H}^n_\delta(A) = \mathcal{H}^n(A)\,.$$

Proof. Step 1. If $\mathcal{L}_n(Z) = 0$, then $\mathcal{H}^n_\delta(Z) = 0$.

Indeed, for every $\varepsilon > 0$, Z can be covered by cubes Q_i, each of diameter less than δ such that $\sum_i |Q_i| < \varepsilon$. Since[12] $(\operatorname{diam} Q_i)^n = C(n)|Q_i|$ we have that

$$\mathcal{H}^n_\delta(Z) \leq \frac{\omega_n}{2^n} \sum_i (\operatorname{diam} Q_i)^n = \frac{\omega_n}{2^n} C(n) \sum_i |Q_i| < \varepsilon \frac{\omega_n}{2^n} C(n)$$

and hence $\mathcal{H}^n_\delta(Z) = 0$.

Step 2. $\mathcal{H}^n_\delta(A) \leq \mathcal{L}_n(A)$.

Let $A \subset \bigcup_{i=1}^\infty \overline{Q}_i$, where the cubes $\overline{Q}_i$ have pairwise disjoint interiors and

$$\sum_{i=1}^{\infty} |Q_i| \leq \mathcal{L}_n(A) + \varepsilon\,.$$

Fix i. It follows from Corollary 65 that there are pairwise disjoint balls $\overline{B}_j \subset Q_i$, $j = 1, 2, 3, \dots$ such that

$$\mathcal{L}_n(\overline{Q}_i \setminus \bigcup_{j=1}^{\infty} \overline{B}_j) = 0\,.$$

[12] By $C(n)$ we denote a constant which depends on n only.

Now applying Step 1 w have

$$\begin{aligned} \mathcal{H}^n_\delta(\overline{Q}_i) &= \mathcal{H}^n_\delta(\bigcup_{i=1}^\infty \overline{B}_i) \leq \sum_{i=1}^\infty \mathcal{H}^n_\delta(\overline{B}_i) \leq \sum_{i=1}^\infty \omega_n r_i^n \\ &= \sum_{i=1}^\infty \mathcal{L}_n(\overline{B}_i) = \mathcal{L}_n(\bigcup_{i=1}^\infty \overline{B}_i) = \mathcal{L}_n(\overline{Q}_i)\,. \end{aligned}$$

Therefore

$$\mathcal{H}^n_\delta(A) \leq \mathcal{H}^n_\delta(\bigcup_{i=1}^\infty \overline{Q}_i) \leq \sum_{i=1}^\infty \mathcal{H}^n_\delta(\overline{Q}_i) \leq \sum_{i=1}^\infty \mathcal{L}_n(\overline{Q}_i) \leq \mathcal{L}_n(A) + \varepsilon\,.$$

Since $\varepsilon > 0$ was arbitrary, the claim follows.

Step 3. $\mathcal{L}_n(A) \leq \mathcal{H}^n_\delta(A)$.

We will need the following fact.

Lemma 67 (Isodiametric inequality) *For any Lebesgue measurable set $A \subset \mathbb{R}^n$*

$$\mathcal{L}_n(A) \leq \omega_n \left(\frac{\operatorname{diam} A}{2}\right)^n$$

We will prove this lemma in Section 6. Observe that if A is a ball, then we have the equality. Therefore the lemma says that among all sets of fixed diameter, ball has the largest volume.

Now we can complete the proof of Step 3. Let $A \subset \bigcup_{i=1}^\infty C_i$, $\operatorname{diam} C_i < \delta$. Then

$$\mathcal{L}_n(A) \leq \sum_{i=1}^\infty \mathcal{L}_n(C_i) \leq \sum_{i=1}^\infty \omega_n \left(\frac{\operatorname{diam} C_i}{2}\right)^n = \frac{\omega_n}{2^n} \sum_{i=1}^\infty (\operatorname{diam} C_i)^n\,.$$

Taking the infimum over all such coverings we obtain

$$\mathcal{L}_n(A) \leq \mathcal{H}^n_\delta(A)\,.$$

Step 4. General case. We proved in steps 2 and 3 that $\mathcal{L}_n(A) = \mathcal{H}^n_\delta(A)$ for all δ. Hence

$$\mathcal{L}_n(A) = \lim_{\delta\to 0} \mathcal{H}^n_\delta(A) = \mathcal{H}^n(A)\,.$$

The proof is complete. □

Now we will compare doubling measures with Hausdorff measures in metric spaces. Recall that in a given metric space there is $0 \leq s \leq \infty$ such that

$$\mathcal{H}^t(X) = \infty \quad \text{for } t < s \qquad \text{and} \qquad \mathcal{H}^t(X) = 0 \quad \text{for } t > s.$$

This unique number s is called Hausdorff dimension of X and is denoted by $s = \dim_H X$.

Theorem 68 *Let μ be a doubling measure on a metric space X, $\mu(2B) \leq C_d \mu(B)$ for every ball B. Let $s = \log C_d / \log 2$. Then*

$$\dim_H X \leq s\,.$$

Proof. It suffices to prove that $\mathcal{H}^s(B) < \infty$ for every ball B. Indeed, then $\mathcal{H}^t(B) = 0$ for $t > s$ and hence $\mathcal{H}^t(X) = 0$ for $t > s$. This, however, implies that $\dim_H X \leq s$.

Let $B = B(x_0, R)$. According to Proposition 62 there is a constant $C > 0$ such that for every $x \in B$ and $r < R$, $\mu(B(x,r)) \geq Cr^s$.

Fix $0 < \delta < R$. Let $\{x_i\}_{i=1}^\infty \subset B$ be a maximal $\delta/2$-separated set, i.e. $d(x_i, x_j) \geq \delta/2$ for $i \neq j$ and for every $x \in B$ there is $i \in I$ such that $d(x, x_i) < \delta/2$. Therefore the balls $B(x_i, \delta/4)$ are disjoint, contained in $B(x, 2R)$ and $B \subset \bigcup_{i\in I} B(x_i, r/2)$. We estimate the number of elements in I.

$$\mu(B(x,2R)) \geq \sum_{i\in I} \mu(B(x_i, \delta/4)) \geq \sum_{i \in I} C(\delta/4)^s = (\# I) C (\delta/4)^s\,.$$

Hence

$$\# I \leq \underbrace{\mu(B(x,2R))C^{-1}4^s}_{\text{does not depend on } \delta}\, \delta^{-s} = C'\delta^{-s}\,.$$

Now $B \subset \bigcup_{i\in I} B(x_i, \delta/2)$ is a covering of the ball by sets of diameters less than or equal to δ, and hence

$$\mathcal{H}^s_\delta(B) \leq \frac{\omega_s}{2^s} \sum_{i\in I} (\operatorname{diam} B(x_i, \delta/2))^s \leq \frac{\omega_s}{2^s} \delta^s (\# I) \leq \frac{\omega_s}{2^s}\delta^s C' \delta^{-s} = \frac{\omega_s}{2^s} C'$$

Hence

$$\mathcal{H}^s(B) = \lim_{\delta \to 0} \mathcal{H}^s_\delta(B) \leq \frac{\omega_s}{2^s} C' < \infty$$

Definition. We say that a Borel measure μ on a metric space X is *s-regular*, if there is a constant $C \geq 1$ such that

$$C^{-1} r^s \leq \mu(B(x,r)) \leq Cr^s$$

for all $x \in X$ and all $0 < r < \operatorname{diam} X$.

Clearly, every s-regular measure is doubling.

Theorem 69 *Let μ be an s-regular measure on X. Then there are constants $C_1, C_2 > 0$ such that*

$$C_1 \mu(E) \leq \mathcal{H}^s(E) \leq C_2 \mu(E)$$

for every Borel set $E \subset X$. In particular $\mathcal{H}^s$ is s-regular.

Proof. First we will prove that there is a constant C_1 such that $C_1\mu(E) \leq \mathcal{H}^s(E)$ for every Borel set E. Obviously we can assume that $\mathcal{H}^s(E) < \infty$. Since every bounded set is contained in a ball of radius $\operatorname{diam} A$, then $\mu(A) \leq C(\operatorname{diam} A)^s$ by the s-regularity of the measure. For every $\varepsilon > 0$ there is a covering $E \subset \bigcup_{i=1}^\infty E_i$ such that

$$\frac{\omega_s}{2^s}\sum_{i=1}^\infty (\operatorname{diam} E_i)^s \leq \mathcal{H}^s(E) + \varepsilon\,.$$

We have

$$\begin{aligned} \mu(E) &\leq \mu\Big(\bigcup_{i=1}^\infty E_i\Big) \leq \sum_{i=1}^\infty \mu(E_i) \leq C\sum_{i=1}^\infty (\operatorname{diam} E_i)^s \\ &= C\frac{2^s}{\omega_s}\frac{\omega_s}{2^s}\sum_{i=1}^\infty (\operatorname{diam} E_i)^s \leq C\frac{2^s}{\omega_s}(\mathcal{H}^s(E)+\varepsilon)\,. \end{aligned}$$

Since $\varepsilon > 0$ was arbitrary we conclude that $C_1\mu(E) \leq \mathcal{H}^s(E)$, where $C_1 = C^{-1}\omega_s/2^s$.

The proof of the opposite inequality is more difficult.

Lemma 70 *If μ is an s-regular measure on a metric space X, then there is a constant $M > 0$ such that for every ball B of radius R and any $r < R$, the ball B can be covered by no more than $M(R/r)^s$ balls of radius r.*

Proof. Let $\{x_i\}_{i\in I} \subset B$ be a maximal r-separated set, i.e. $d(x_i, x_j) \geq r$ for $i \neq j$ and for every $x \in B$ there is $i \in I$ such that $x \in B(x_i, r)$. Hence $B \subset \bigcup_{i\in I} B(x_i, r)$ and $B(x_i, r/2) \cap B(x_j, r/2) = \emptyset$ for $i \neq j$. Since $B(x_i, r/2) \subset 2B$ and $\mu(B(x_i, r/2)) \geq C^{-1}(r/2)^s$ we have

$$C(2R)^s \geq \mu(2B) \geq (\#I)C^{-1}(r/2)^s\,.$$

Hence $\#I \leq C^2 4^s (R/r)^s$ and hence the lemma follows with $M = C^2 4^s$. □

The lemma implies that $\mathcal{H}^s_\delta(B) \leq C'\mu(B)$ where B is a ball of radius R. Indeed, we can cover B by $M(R/r)^s$ balls of radius r and we can take r so small that $2r < \delta$. Hence

$$\mathcal{H}^s_\delta(B) \leq \frac{\omega_s}{2^s}M\left(\frac{R}{r}\right)^s (2r)^s = \omega_s M R^s \leq C'\mu(B)\,.$$

Since the constant C' is independent of δ we conclude that $\mathcal{H}^s(B) \leq C'\mu(B)$. We will now prove a stronger inequality which says that for an arbitrary open set $U \subset X$

$$\mathcal{H}^s(U) \leq C''\mu(U)\,.$$

This will imply that $\mathcal{H}^s(E) \leq C''\mu(E)$ for any Borel set $E \subset X$. Indeed, both measures μ and $\mathcal{H}^s$ have the property that X is a union of a countably family of open sets of finite measure. Therefore both measures have the property that the measure of any Borel set equals infimum of measures of open sets that contain given Borel set (Theorem 7).

It follows from the proof of Lemma 70 that X is separable, so according to the $5r$-covering lemma we can find disjoint balls $B_i \subset U$ such that $U \subset \bigcup_{i=1}^{\infty} 5B_i$. Hence

$$\mathcal{H}^s(U) \le \sum_{i=1}^{\infty} \mathcal{H}^s(5B_i) \le C' \sum_{i=1}^{\infty} \mu(5B_i) \le C'' \sum_{i=1}^{\infty} \mu(B_i) \le C''\mu(U).$$

The proof is complete. □

5 Differentiation

5.1 Lebesgue differentiation theorem

DEFINITION. We say that $f \in L^1_{\rm loc}(\mu)$, if for every $x \in X$ there is $r > 0$ such that $\int_{B(x,r)} |f|\, d\mu < \infty$.

The integral average of f over a measurable set E of finite measure will be denote by

$$f_E = \mu(E)^{-1} \int_E f\, d\mu = ⨍_E f\, d\mu\, .$$

Theorem 71 (Lebesgue differentiation theorem) *If μ is a doubling measure on a metric space X and $f \in L^1_{\rm loc}(\mu)$, then*

$$\lim_{r\to 0} ⨍_{B(x,r)} f\, d\mu = f(x) \qquad \textit{for } \mu\textit{-a.e. } x \in X\, . \tag{40}$$

Proof. We can assume that $f \ge 0$. Since X is separable, it is a countable union of balls such that f is integrable on each such ball. Therefore, it suffices to prove that if $\int_B f\, d\mu < \infty$, then (40) is satisfied for μ-a.e. $x \in B$.

If for a measurable set $A \subset B$ we have

$$\liminf_{r\to 0} ⨍_{B(x,r)} f\, d\mu \le t \qquad \text{for all } x \in A\, ,$$

then

$$\int_A f\, d\mu \le t\mu(A)\, .$$

Indeed, given $\varepsilon > 0$, let U be an open set such that $A \subset U$, $\mu(U) \le \mu(A) + \varepsilon$. For every $x \in A$ there is arbitrarily small $r > 0$ such that $\overline{B}(x,r) \subset U$ and

$$⨍_{\overline{B}(x,r)} f\, d\mu \le t + \varepsilon\, .$$

Such closed balls form a covering of U in the Vitali sense and hence we can select pairwise disjoint balls $\overline{B}_1, \overline{B}_2, \overline{B}_3, \dots$ from the covering such that

$$\mu(U \setminus \bigcup_{i=1}^{\infty} \overline{B}_i) = 0, \qquad ⨍_{\overline{B}_i} f\, d\mu \le t + \varepsilon\,.$$

We have

$$\begin{aligned}\int_A f\, d\mu &\le \int_U f\, d\mu = \sum_{i=1}^{\infty} \int_{\overline{B}_i} f\, d\mu \le (t+\varepsilon) \sum_{i=1}^{\infty} \mu(\overline{B}_i) \\ &= (t+\varepsilon)\mu(U) \le (t+\varepsilon)(\mu(A)+\varepsilon)\,.\end{aligned}$$

Since $\varepsilon > 0$ was arbitrary, we conclude

$$\int_A f\, d\mu \le t\mu(A)$$

Similarly, if $A \subset B$ is a measurable set such that

$$\limsup_{r\to 0} ⨍_{B(x,r)} f\, d\mu \ge t \qquad \text{for all } x \in A\,,$$

then

$$\int_A f\, d\mu \ge t\mu(A)\,. \tag{41}$$

Let

$$A_{s,t} = \{x \in B :\ \liminf_{r\to 0} ⨍_{B(x,r)} f\, d\mu \le s < t \le \limsup_{r\to 0} ⨍_{B(x,r)} f\, d\mu\}\,.$$

Then

$$t\mu(A_{s,t}) \le \int_{A_{s,t}} f\, d\mu \le s\mu(A_{s,t})$$

and hence $\mu(A_{s,t}) = 0$. This easily implies that

$$\liminf_{r\to 0} ⨍_{B(x,r)} f\, d\mu = \limsup_{r\to 0} ⨍_{B(x,r)} f\, d\mu \qquad \text{for } \mu\text{-a.e. } x \in B\,,$$

i.e. the limit $\lim_{r\to 0} ⨍_{B(x,r)} f\, d\mu$ exists for μ-a.e. $x \in B$. Denote this limit by

$$g(x) = \lim_{r\to 0} ⨍_{B(x,r)} f\, d\mu$$

whenever the limit exists. We have to prove that $f(x) = g(x)$ for μ-a.e. $x \in B$.

Fix a measurable set $F \subset B$ and $\varepsilon > 0$. Let

$$A_n = \{x \in F :\ (1+\varepsilon)^n \le g(x) < (1+\varepsilon)^{n+1}\}, \quad n \in \mathbb{Z}\,.$$

Observe that

$$\limsup_{r\to 0} \fint_{B(x,r)} f\, d\mu = g(x) \ge (1+\varepsilon)^n \qquad \text{for all } x\in A_n\,,$$

and hence $\int_{A_n} f\, d\mu \ge (1+\varepsilon)^n \mu(A_n)$ by (41). We have

$$\begin{aligned} \int_F g\, d\mu &= \sum_{n=-\infty}^{\infty} \int_{A_n} g\, d\mu \le \sum_{n=-\infty}^{\infty} (1+\varepsilon)^{n+1}\mu(A_n) \\ &\le (1+\varepsilon) \sum_{n=-\infty}^{\infty} \int_{A_n} f\, d\mu = (1+\varepsilon)\int_F f\, d\mu\,. \end{aligned}$$

Similarly we prove that

$$\int_F g\, d\mu \ge (1+\varepsilon)^{-1} \int_F f\, d\mu\,.$$

Since $\varepsilon>0$ was arbitrary we conclude that for every measurable set $F\subset B$

$$\int_F f\, d\mu = \int_F g\, d\mu$$

and hence $f=g$ μ-a.e. □

Among all representatives of f in the class of functions that are equal f a.e. we select the following one

$$f(x) := \limsup_{r\to 0} \fint_{B(x,r)} f\, d\mu\,. \tag{42}$$

It follows from the Lebesgue differentiation theorem that the right hand side of (42) which is defined *everywhere* equals f a.e.

Definition. Let $f\in L^1_{\mathrm{loc}}(\mu)$. We say that $x\in X$ is a Lebesgue point of f if

$$\lim_{r\to 0} \fint_{B(x,r)} |f(y)-f(x)|\, d\mu(y) = 0\,,$$

where $f(x)$ is defined by (42).

Theorem 72 *Let μ be a doubling measure and $f\in L^1_{\mathrm{loc}}(\mu)$. Then μ-a.e. point $x\in X$ is a Lebesgue point of f.*

Proof. For every $c\in\mathbb{R}$

$$\lim_{r\to 0} \fint_{B(x,r)} |f(y)-c|\, d\mu(y) = |f(x)-c| \qquad \mu\text{-a.e.} \tag{43}$$

Let

$$E = \{x \in X : \text{ (43) holds for all } c \in \mathbb{Q}\}.$$

Since $\mathbb{Q}$ is countable, $\mu(X \setminus E) = 0$. Let $x \in E$, be such that $|f(x)| < \infty$, and let $\mathbb{Q} \ni c_n \to f(x)$. Then it follows from (43) that

$$\lim_{r \to 0} \fint_{B(x,r)} |f(y) - f(x)|\, d\mu = |f(x) - f(x)| = 0.$$

□

Definition. We say that a family $\mathcal{F}$ of measurable sets in X is *regular at* $x \in X$ if there is a constant $C > 0$ such that for every $S \in \mathcal{F}$ there is a ball $B(x, r_S)$ such that

$$S \subset B(x, r_S), \qquad \mu(B(x, r_S)) \le C\mu(S),$$

and for every $\varepsilon > 0$ there is a set $S \in \mathcal{F}$ with $\mu(S) < \varepsilon$.

Theorem 73 *Let μ be a doubling measure and $f \in L^1_{\rm loc}(\mu)$. If $x \in X$ is a Lebesgue point of f and $\mathcal{F}$ is a regular family at $x \in X$, then*

$$\lim_{\substack{S \in \mathcal{F} \\ \mu(S) \to 0}} \fint_S f\, d\mu = f(x).$$

Proof. For $S \in \mathcal{F}$ let r_s be defined as above. Observe that if $\mu(S) \to 0$, then $r_s \to 0$ by Proposition 62. We have

$$\begin{aligned} \left| \fint_S f\, d\mu - f(x) \right| &\le \fint_S |f(y) - f(x)|\, d\mu(y) \le \mu(S)^{-1} \int_{B(x,r_S)} |f(y) - f(x)|\, d\mu(y) \\ = \ & \frac{\mu(B(x,r_S))}{\mu(S)} \fint_{B(x,r_S)} |f(y) - f(x)|\, d\mu(y) \le C \fint_{B(x,r_S)} |f(y) - f(x)|\, d\mu(y) \to 0 \end{aligned}$$

as $\mu(S) \to 0$. □

Corollary 74 *If $f \in L^1_{\rm loc}(\mathbb{R}^n)$ and $x \in \mathbb{R}^n$ is a Lebesgue point of f, then for any sequence of cubes Q_i such that $x \in Q_i$,* $\operatorname{diam} Q_i \to 0$ *we have*

$$\lim_{i \to \infty} \fint_{Q_i} f\, d\mu = f(x).$$

Corollary 75 *Let $F(x) = \int_a^x f(t)\, dt$, where $f \in L^1(a,b)$. Then $F'(x) = f(x)$ for a.e. $x \in (a,b)$.*

Proof. We have

$$\frac{F(x+h) - F(x)}{h} = \fint_x^{x+h} f(t)\, dt \to f(x) \qquad \text{as } h \to 0,$$

whenever x is a Lebesgue point of f. □

Theorem 76 *If $A \subset \mathbb{R}^n$ is a measurable set, then for almost all $x \in A$*

$$\lim_{r \to 0} \frac{|B(x,r) \cap A|}{|B(x,r)|} = 1$$

and for almost all $x \in \mathbb{R}^n \setminus A$

$$\lim_{r \to 0} \frac{|B(x,r) \cap A|}{|B(x,r)|} = 0\,.$$

Proof. Applying the Lebesgue differentiation theorem to $f = \chi_A \in L^1_{\rm loc}(\mathbb{R}^n)$ we have that for a.e. $x \in A$

$$\frac{|B(x,r) \cap A|}{|B(x,r)|} = ⨍_{B(x,r)} \chi_A \, d\mathcal{L}_n \to \chi_A(x) = 1$$

and for a.e. $x \in \mathbb{R}^n \setminus A$

$$\frac{|B(x,r) \cap A|}{|B(x,r)|} \to \chi_A(x) = 0\,.$$

□

Definition. Let $A \subset \mathbb{R}^n$ be a measurable set. We say that $x \in \mathbb{R}^n$ is a *density point* of A if

$$\lim_{r \to 0} \frac{|A \cap B(x,r)|}{|B(x,r)|} = 1\,.$$

Thus the above theorem says that almost every point of a measurable set $A \subset \mathbb{R}^n$ is a density point.

6 Brunn-Minkowski inequality

In this section we will prove the Brunn-Minkowski inequality, isodiametric inequality and the isoperimetric inequality.

Theorem 77 (Brunn-Minkowski inequality) *If $A, B \subset \mathbb{R}^n$ are compact sets, then*

$$\mathcal{L}_n(A+B)^{1/n} \geq \mathcal{L}_n(A)^{1/n} + \mathcal{L}_n(B)^{1/n}\,.$$

where $A + B = \{a + b : \ a \in A, \ b \in B\}$.

Proof. Observe that if we translate the sets, then the set $A + B$ will also translate, so the inequality will remain the same. In particular, we can locate the origin of the coordinate system at any point.

We divide the proof into several steps.

STEP 1. A and B are rectangular boxes with sides parallel to the coordinate axes.

Denote the edges of A and B by $a_1,\ldots,a_n$ and $b_1,\ldots,b_n$ respectively. Then $A+B$ is also a rectangular box with edges $a_1+b_1,\ldots,a_n+b_n$. Hence the Brunn-Minkowski inequality takes the form

$$\prod_{i=1}^{n}(a_i+b_i)^{1/n} \geq \prod_{i=1}^{n} a_i^{1/n} + \prod_{i=1}^{n} b_i^{1/n}$$

which is equivalent to

7 L^p spaces

DEFINITION. Let (X,μ) be a measure space. For $0<p<\infty$, $\tilde{L}^p(\mu)$ denotes the class of all complex valued measurable functions such that

$$\|f\|_p = \left(\int_X |f|^p\, d\mu\right)^{1/p} < \infty$$

and we define $L^p(\mu) = \tilde{L}^p(\mu)/\sim$, where $f\sim g$ if $f=g$ a.e.

For $p=\infty$ we define $\tilde{L}^\infty$ to be the class of essentially bounded measurable functions, i.e. such that there is $M>0$ with

$$|f(x)| \leq M \quad \mu\text{-a.e.} \tag{44}$$

It is easy to see that there is a smallest number M satisfying (44). We denote this number by $\|f\|_\infty$. Therefore

$$|f(x)| \leq \|f\|_\infty \quad \mu\text{-a.e.}$$

and for any $\varepsilon>0$, $\mu(\{x:\,|f(x)|>\|f\|_\infty-\varepsilon\})>0$. Finally we set $L^\infty(\mu)=\tilde{L}(\mu)/\sim$.

EXAMPLE. Let μ be a counting measure on $\mathbb{N}=\{1,2,3,\ldots\}$, i.e. for $A\subset\mathbb{N}$, $\mu(A)=\#A$, the number of elements in the set A if A is finite and $\mu(A)=\infty$ is A is infinite. Clearly if $f:\mathbb{N}\to\mathbb{C}$, then

$$\int_{\mathbb{N}} f\, d\mu = \sum_{n=1}^{\infty} f(n).$$

The functions $f:\mathbb{N}\to\mathbb{C}$ can be identified with complex sequences and it is easy to see that

$$L^p(\mu) = \{x=(x_n)_{n=1}^\infty:\ \|x\|_p = \big(\sum_{n=1}^{\infty}|x_n|^p\big)^{1/p}<\infty\}\,,$$

and

$$L^\infty(\mu) = \{x=(x_n)_{n=1}^\infty:\ \|x\|_\infty = \sup_{n\in\mathbb{N}}|x_n|<\infty\}\,.$$

In this particular situation we denote $L^p(\mu)$ by ℓ^p for all $0<p\leq\infty$.

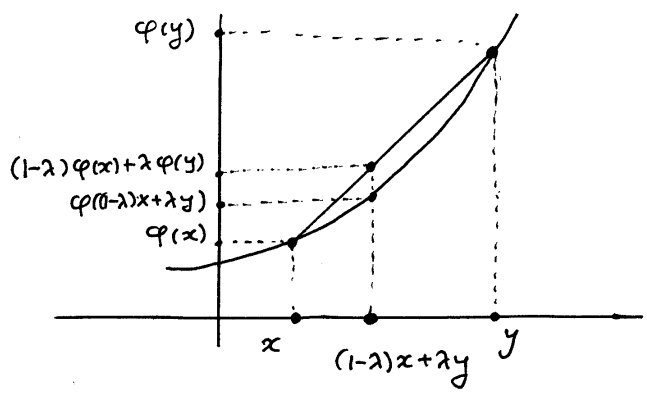
Theorem 78 (Jensen's inequality) *Let (X, μ) be a measure space such that $\mu(X) = 1$. If $f \in L^1(\mu)$ satisfies $a < f(x) < b$ for a.e. x and φ is a convex function on (a, b), then*

$$\varphi\left(\int_X f\,d\mu\right) \leq \int_X (\varphi \circ f)\,d\mu\,.$$

Remark. We *do not* exclude the cases $a = -\infty$ and $b = \infty$.

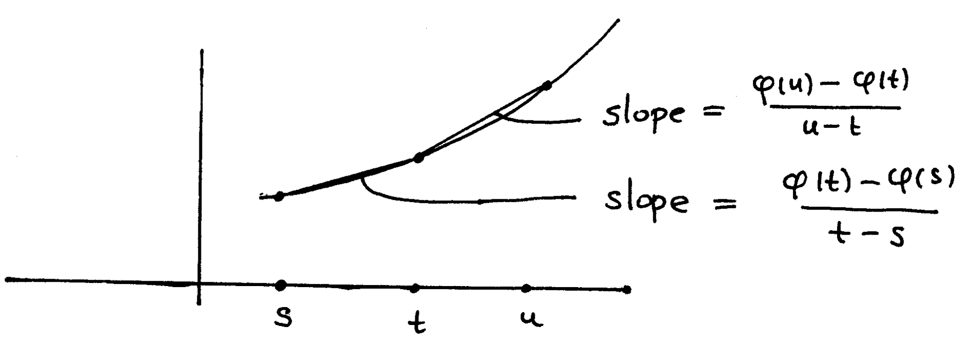
Recall that φ is convex on (a, b) if

$$\varphi((1-\lambda)x + \lambda y) \leq (1-\lambda)\varphi(x) + \lambda\varphi(y) \quad \text{for all } x, y \in (a, b) \text{ and } 0 \leq \lambda \leq 1.$$



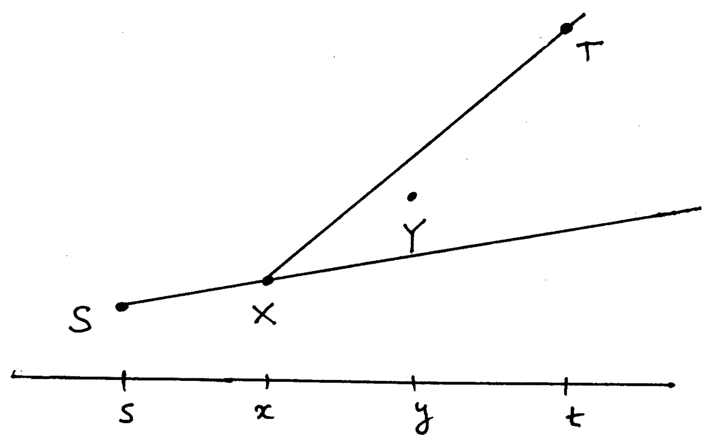
It is easy to see that this condition is equivalent to

$$\frac{\varphi(t) - \varphi(s)}{t - s} \leq \frac{\varphi(u) - \varphi(t)}{u - t} \quad \text{for all } s, t, u \text{ such that } a < s < t < u < b \tag{45}$$



Lemma 79 *Every convex function on (a, b) is continuous.*

Proof. Fix $a < s < t < b$ and take x, y such that $a < s < x < y < t < b$. Denote by S, X, Y, T the corresponding points on the graph of φ.



The point Y belongs to the angle with the vertex X formed by the half-lines $\overrightarrow{SX}$ and $\overrightarrow{XT}$. This easily implies that if y is decreasing to x, then $Y \to X$, which proves the right hand side continuity of φ. The proof of the left hand side continuity is similar. □

Proof of Jensen's inequality. For $t = \int_X f\, d\mu$ we have $a < t < b$. Fix this value of t and define

$$\beta = \sup_{a<s<t} \frac{\varphi(t) - \varphi(s)}{t - s}\, .$$

It follow from (45) that

$$\beta \leq \frac{\varphi(u) - \varphi(t)}{u - t} \quad \text{for all } t < u < b$$

and hence

$$\varphi(u) \geq \varphi(t) + \beta(u - t) \quad \text{for all } t < u < b. \tag{46}$$

On the other hand

$$\frac{\varphi(t) - \varphi(s)}{t - s} \leq \beta \quad \text{for all } a < s < t$$

and hence

$$\varphi(s) \geq \varphi(t) + \beta(s - t) \quad \text{for all } a < s < t \tag{47}$$

The inequalities (46) and (47) imply that

$$\varphi(s) \geq \varphi(t) + \beta(s - t) \quad \text{for all } a < s < b.$$

In particular

$$\varphi(f(x)) - \varphi(t) - \beta(f(x) - t) \geq 0 \quad \text{a.e.}$$

Integrating with respect to μ yields

$$\int_X \varphi(f(x)) - \underbrace{\varphi\Big(\int_X f\,d\mu\Big)}_{\text{constant}} - \beta\Big(f(x) - \underbrace{\int_X f\,d\mu}_{\text{constant}}\Big)\,d\mu \geq 0\,.$$

Since $\mu(X) = 1$ we conclude

$$\int_X (\varphi\circ f)\,d\mu - \varphi\Big(\int_X f\,d\mu\Big) - \beta\Big(\underbrace{\int_X f\,d\mu - \int_X f\,d\mu}_{0}\Big) \geq 0$$

and hence

$$\int_X (\varphi\circ f)\,d\mu \geq \varphi\Big(\int_X f\,d\mu\Big)\,.$$

□

The numbers $1 < p, q < \infty$ such that $1/p + 1/q = 1$ or $p = 1$ and $q = \infty$ are called *Hölder conjugate* or just *conjugate* exponents. Observe that if $1 < p < \infty$, then $q = p/(p-1)$.

Theorem 80 (Hölder's inequality) *Let (X, μ) be a measure space and $1 < p, q < \infty$ be conjugate exponents. Then for any two complex valued measurable functions f and g we have*

$$\int_X |fg|\,d\mu \leq \|f\|_p \|g\|_q\,.$$

Proof. We will need the following lemma.

Lemma 81 (Young's inequality) *If $a, b > 0$, then*

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}\,.$$

Proof. We can assume that $a, b > 0$, otherwise the inequality is obvious. Then $a = e^{s/p}$, $b = e^{t/q}$ for some $s, t \in \mathbb{R}$ and the inequality is equivalent to

$$e^{s/p + t/q} \leq e^s/p + e^t/q\,.$$

Now the lemma follows from the convexity of e^x. □

We can assume that $0 < \|f\|_p < \infty$ and $0 < \|g\|_q < \infty$, otherwise the Hölder inequality is obvious. The lemma yields

$$\frac{|f(x)g(x)|}{\|f\|_p\|g\|_q} \leq \frac{1}{p}\left(\frac{|f(x)|}{\|f\|_p}\right)^p + \frac{1}{q}\left(\frac{|g(x)|}{\|g\|_q}\right)^q\,.$$

Integrating this inequality yields

$$\int_X |fg|\,d\mu \le \|f\|_p\|g\|_q \underbrace{\Big(\frac{1}{p\|f\|_p^p}\int_X |f|^p\,d\mu + \frac{1}{q\|g\|_q^q}\int_X |g|^q\,d\mu\Big)}_{=1/p+1/q=1}$$

and the theorem follows. □

Theorem 82 (Minkowski's inequality) *Let (X,μ) be a measure space and $1 \le p \le \infty$. Then for any two complex valued measurable functions f and g we have*

$$\|f+g\|_p \le \|f\|_p + \|g\|_p\,.$$

Proof. If $p=1$, the inequality is obvious. If $p=\infty$, then

$$|f(x)+g(x)| \le |f(x)| + |g(x)| \le \|f\|_\infty + \|g\|_\infty \quad \text{a.e.}$$

Hence $\|f+g\|_\infty \le \|f\|_\infty + \|g\|_\infty$. If $1<p<\infty$, then we need to use the Hölder inequality.

$$\begin{aligned}
\int_X |f+g|^p\,d\mu &\le \int_X |f||f+g|^{p-1}\,d\mu + \int_X |g||f+g|^{p-1}\,d\mu \\
&\le \Big(\int_X |f|^p\,d\mu\Big)^{1/p}\Big(\int_X |f+g|^{(p-1)q}\,d\mu\Big)^{1/q} + \Big(\int_X |g|^p\,d\mu\Big)^{1/p}\Big(\int_X |f+g|^{(p-1)q}\,d\mu\Big)^{1/q} \\
&\le (\|f\|_p + \|g\|_p)\Big(\int_X |f+g|^p\,d\mu\Big)^{1/q}
\end{aligned}$$

and hence

$$\Big(\int_X |f+g|^p\,d\mu\Big)^{1-1/q} \le \|f\|_p + \|g\|_p.$$

Since $1-1/q = 1/p$ the theorem follows. □

Example. In the case of space ℓ^p, the Hölder and the Minkowski inequality read as follows

$$\sum_{n=1}^\infty |x_n y_n| \le \Big(\sum_{n=1}^\infty |x_n|^p\Big)^{1/p}\Big(\sum_{n=1}^\infty |y_n|^q\Big)^{1/q},$$

$$\Big(\sum_{n=1}^\infty |x_n+y_n|^p\Big)^{1/p} \le \Big(\sum_{n=1}^\infty |x_n|^p\Big)^{1/p} + \Big(\sum_{n=1}^\infty |y_n|^p\Big)^{1/p}.$$

Definition. Let K denote $\mathbb{R}$ or $\mathbb{C}$. The *normed* space is a pair $(X, \|\cdot\|)$, where X is a linear space over K and

$$\|\cdot\| : X \to [0,\infty)$$

is a function such that:

(a) $\|x+y\| \leq \|x\| + \|y\|$ for all $x, y \in X$;

(b) $\|\alpha x\| + |\alpha|\|x\|$ for all $x \in X$ and $\alpha \in K$;

(c) $\|x\| = 0$ iff $x = 0$.

The function $\|\cdot\|$ is called a *norm*.

Since $\|x-y\| = \|(x-z)+(z-y)\| \leq \|x-z\| + \|z-y\|$ we see that

$$d(x,y) = \|x-y\|$$

defines a metric in X. Normed spaces are always regarded as metric spaces with respect to this metric.

DEFINITION. We say that a normed space $(X, \|\cdot\|)$ is a *Banach space* if it is complete with respect to the metric $d(x,y) = \|x-y\|$.

It follows from the Minkowski inequality that $(L^p(\mu), \|\cdot\|_p)$ is a normed space for all $1 \leq p \leq \infty$.

Theorem 83 *$L^p(\mu)$ is a Banach space for all $1 \leq p \leq \infty$.*

Proof. We will prove this theorem for $1 \leq p < \infty$ and leave the case $p = \infty$ as an exercise. Let $\{f_n\}$ be a Cauchy sequence in $L^p(\mu)$. Let $\{f_{n_i}\}$ be a subsequence such that

$$\|f_{n_{i+1}} - f_{n_i}\|_p < 2^{-i}, \quad i = 1, 2, 3, \ldots$$

Let

$$g_k = \sum_{i=1}^{k} |f_{n_{i+1}} - f_{n_i}|, \qquad g = \sum_{i=1}^{\infty} |f_{n_{i+1}} - f_{n_i}|\,.$$

It follows from the Minkowski inequality that $\|g_k\|_p < 1$, and then Fatou's lemma applied to the sequence g_k^p yields

$$\int_X g^p\, d\mu = \int_X (\liminf_{k\to\infty} g_k^p)\, d\mu \leq \liminf_{k\to\infty} \int_X g_k^p\, d\mu \leq 1\,,$$

so $\|g\|_p \leq 1$. In particular $g(x) < \infty$ a.e. Thus the series

$$f_{n_1}(x) + \sum_{i=1}^{\infty} (f_{n_{i+1}}(x) - f_{n_i}(x))$$

converges absolutely for a.e. x. Since

$$f_{n_1}(x) + \sum_{i=1}^{k} (f_{n_{i+1}}(x) - f_{n_i}(x)) = f_{n_{k+1}}(x)$$

we conclude that the sequence $f_{n_k}(x)$ converges for a.e. x. Denote its limit by $f(x)$, i.e.

$$\lim_{k\to\infty} f_{n_k}(x) = f(x) \quad \text{a.e.}$$

It suffices to prove that $f \in L^p$ and $\|f - f_n\|_p \to 0$ as $n \to \infty$.

For every $\varepsilon > 0$ there is N such that $\|f_n - f_m\|_p < \varepsilon$ for all $n, m \geq N$. Applying Fatou's lemma to $|f_{n_i} - f_m|^p$ with $m \geq N$ fixed and $i = 1, 2, 3, \ldots$ we have

$$\int_X |f - f_m|^p \, d\mu \leq \liminf_{i\to\infty} \int_X |f_{n_i} - f_m|^p \, d\mu \leq \varepsilon^p \, .$$

Hence $f - f_m \in L^p$ and thus $f \in L^p$. Moreover $\|f - f_m\|_p \to 0$ as $m \to \infty$. □

Exercise. *Prove the above theorem for $p = \infty$.*

The following fact follows from the proof of the above theorem.

Corollary 84 *If $f_n \to f$ in $L^p(\mu)$, then there is a subsequence f_{n_i} which converges to f a.e.*

Lemma 85 *Let S be the class of complex, measurable, simple functions s on X such that*

$$\mu(\{x : \, s(x) \neq 0\}) < \infty \, .$$

If $1 \leq p < \infty$, then S is dense in $L^p(\mu)$.

Proof. Clearly $S \subset L^p(\mu)$. If $f \in L^p(\mu)$ and $f \geq 0$, let s_n be a sequence of simple functions such that $0 \leq s_n \leq f$, $s_n \to f$ pointwise. Since $s_n \in L^p$ it easily follows that $s_n \in S$. Now inequality $0 \leq |f - s_n|^p \leq f^p$ and the dominated convergence theorem implies that $s_n \to f$ in L^p. In the general case we write $f = (u^+ - u^-) + i(v^+ - v^-)$ and apply the above argument to each of the functions u^+, u^-, v^+, v^- separately. □

Theorem 86 *Let X be a locally compact metric space and μ a Radon measure on X. Then the class of compactly supported continuous functions $C_c(X)$ is dense in $L^p(\mu)$ for all $1 \leq p < \infty$.*

Proof. According to the lemma it suffices to prove that the characteristic function of a set of finite measure can be approximated in L^p by compactly supported continuous functions. Let E be a measurable set of finite measure. Given $\varepsilon > 0$ let $K \subset E$ be a compact set such that $\mu(E \setminus K) < (\varepsilon/2)^p$. Let U be an open set such that $K \subset U$, $\overline{U}$ is compact and $\mu(U \setminus K) < (\varepsilon/2)^p$. Finally let $\varphi \in C_c(X)$ be such that $\operatorname{supp} \varphi \subset U$, $0 \leq \varphi \leq 1$, $\varphi(x) = 1$ for $x \in K$. We have

$$\begin{aligned} \|\chi_E - \varphi\|_p &= \left(\int_{X\setminus K} |\chi_E - \varphi|^p \, d\mu \right)^{1/p} \leq \left(\int_{X\setminus K} |\chi_E|^p \, d\mu \right)^{1/p} + \left(\int_{X\setminus K} |\varphi|^p \, d\mu \right)^{1/p} \\ &\leq \mu(E \setminus K)^{1/p} + \mu(U \setminus K)^{1/p} < \varepsilon \, . \end{aligned}$$

□

7.1 Convolution

Let us start with an observation that since the Lebesgue measure in invariant under translations and under the mapping $x \mapsto -x$, for any $f \in L^1(\mathbb{R}^n)$ and any $y \in \mathbb{R}^n$ we have

$$\int_{\mathbb{R}^n} u(x)\, dx = \int_{\mathbb{R}^n} u(x+y)\, dx = \int_{\mathbb{R}^n} u(-x)\, dx\, .$$

Also since for every measurable set E and any $t > 0$ the set $tE = \{tx : x \in E\}$ has measure $\mathcal{L}_n(tE) = t^n \mathcal{L}_n(E)$ we easily conclude that the Lebesgue integral has the following scaling property

$$\int_{\mathbb{R}^n} u(x/t)\, dx = t^n \int_{\mathbb{R}^n} u(x)\, dx\, .$$

Observe that in the case of the Riemann integral the above equalities are direct consequences of the change of variables formula. We will prove a corresponding change of variables formula for the Lebesgue integral later, but as for the proof of the above equalities we do not have to refer to the general change of variables formula as they follow directly from the properties of the Lebesgue measure mentioned above.

DEFINITION. For measurable functions f and g on $\mathbb{R}^n$ we define the *convolution* by

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

and it is a question under what conditions the convolution is well defined.

If $f \in L^1_{\rm loc}(\mathbb{R}^n)$ and g is bounded, measurable and vanishes outside a bounded set, then the function $y \mapsto f(x-y)g(y)$ is integrable, so $(f*g)(x)$ is well defined and finite for every $x \in \mathbb{R}^n$. If $f, g \in L^1(\mathbb{R}^n)$, then it can happen that for a given x the function $y \mapsto f(x-y)g(y)$ is not integrable and hence $(f*g)(x)$ is not defined. However as a powerful application of the Fubini theorem we can prove the following surprising result.

Theorem 87 *If $f, g \in L^1(\mathbb{R}^n)$, then for a.e. $x \in \mathbb{R}^n$ the function $y \mapsto f(x-y)g(y)$ is integrable and hence $(f*g)(x)$ exists. Moreover $f * g \in L^1(\mathbb{R}^n)$ and*

$$\|f*g\|_1 \leq \|f\|_1 \|g\|_1\, .$$

Proof. The function $|f(x-y)g(y)|$ as a function of a variable $(x,y) \in \mathbb{R}^{2n}$ is measurable.[13] Hence Fubini's theorem yields

$$\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} |f(x-y)g(y)|\,dx\,dy = \int_{\mathbb{R}^n}\left(\int_{\mathbb{R}^n} |f(x-y)||g(y)|\,dx\right)dy$$
$$= \int_{\mathbb{R}^n}\left(\int_{\mathbb{R}^n} |f(x-y)|\,dx\right)|g(y)|\,dy = \|f\|_1\|g\|_1\,.$$

Therefore

$$\|f\|_1\|g\|_1 = \int_{\mathbb{R}^n}\left(\int_{\mathbb{R}^n} |f(x-y)g(y)|\,dy\right)dx \ge \int_{\mathbb{R}^n}\left|\int_{\mathbb{R}^n} f(x-y)g(y)\,dy\right|dx = \|f*g\|_1\,.$$

□

Theorem 88 *If $f, g, h \in L^1(\mathbb{R}^n)$ and $\alpha, \beta \in \mathbb{C}$, then*

1. $f * g = g * f$;
2. $f * (g * h) = (f * g) * h$;
3. $f * (\alpha g + \beta h) = \alpha f * g + \beta f * h$.

Remark. This theorem says that $(L^1, *)$ is a commutative algebra.

Proof. (1) We have

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

(2) We have

$$f*(g*h)(x) = \int_{\mathbb{R}^n} f(x-y)(g*h)(y)\,dy = \int_{\mathbb{R}^n} f(x-y)\left(\int_{\mathbb{R}^n} g(y-z)h(z)\,dz\right)dy$$
$$= \int_{\mathbb{R}^n} h(z)\left(\int_{\mathbb{R}^n} f(x-y)g(y-z)\,dy\right)dz = \int_{\mathbb{R}^n} h(z)\left(\int_{\mathbb{R}^n} f(-y)g(y-z+x)\,dy\right)dz$$
$$= \int_{\mathbb{R}^n} h(z)\left(\int_{\mathbb{R}^n} f(y)g((x-z)-y)\,dy\right)dz = \int_{\mathbb{R}^n} h(z)(f*g)(x-z)\,dz = (f*g)*h(x)\,.$$

(3) It is easy and left to the reader. □

Theorem 89 *If $f \in L^1_{\text{loc}}(\mathbb{R}^n)$ and g is continuous with compact support, then*

1. *$f * g$ is continuous on $\mathbb{R}^n$.*

[13] Why?

2. *If $f \in C^1(\mathbb{R}^n)$, then $f * g \in C^1(\mathbb{R}^n)$ and*

$$\frac{\partial}{\partial x_i}(f * g)(x) = \big(\frac{\partial f}{\partial x_i} * g\big)(x)\,.$$

3. *If $g \in C^1(\mathbb{R}^n)$, then $f * g \in C^1(\mathbb{R}^n)$ and*

$$\frac{\partial}{\partial x_i}(f * g)(x) = \big(f * \frac{\partial g}{\partial x_i}\big)(x)\,.$$

4. *If $f \in C^r(\mathbb{R}^n)$ or $g \in C^r(\mathbb{R}^n)$, then $f * g \in C^r(\mathbb{R}^n)$.*

Proof. (1) Since g is bounded and has compact support, $(f * g)(x)$ is well defined and finite for all $x \in \mathbb{R}^n$. Fix $x \in \mathbb{R}^n$. The function $y \mapsto f(y)g(x-y)$ vanishes outside a sufficiently large ball B (because g as compact support). Let $x_n \to x$. Then there is a ball B (perhaps larger than the one above), so that all the functions

$$y \mapsto f(y)g(x_n - y)$$

vanish outside B. Hence

$$|f(y)g(x_n - y)| \le \|g\|_\infty |f(y)| \chi_B(y) \in L^1(\mathbb{R}^n)$$

and the dominated convergence theorem yields

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

which proves continuity of $f * g$.

(2) The function $\partial f/\partial x_i$ is bounded on bounded subsets of $\mathbb{R}^n$ (because it is continuous). Fix $x \in \mathbb{R}^n$. For any $r > 0$ there is a constant $M > 0$ such that

$$\left|\frac{\partial f}{\partial x_i}\right| \le M \qquad \text{for all } z \in B(x, r+1).$$

Let $t_k \to 0$, $|t_k| < 1$. We have

$$\frac{f(x + t_k e_i - y) - f(x-y)}{t_k}\, g(y) \to \frac{\partial f}{\partial x_i}(x-y)g(y)\,.$$

The functions on the left hand side are bounded by $M\|g\|_\infty$ for all y with $|y| < r$ (by the mean value theorem). Take r so large that supp $g \subset B(0,r)$. Then

$$\begin{aligned}&\frac{(f * g)(x + t_k e_i) - (f * g)(x)}{t_k}\\ &= \int_{B(0,r)} \frac{f(x + t_k e_i - y) - f(x-y)}{t_k}\, g(y)\, dy \to \int_{B(0,r)} \frac{\partial f}{\partial x_i}(x-y)g(y)\, dy = \big(\frac{\partial f}{\partial x_i} * g\big)(x)\,.\end{aligned}$$

We could pass to the limit under the sign of the integral, because the functions were uniformly bounded.

(3) The proof is similar to that in the case (2); details are left to the reader.

(4) This result follows from (2) and (3) by induction. $\square$

Let $\varphi \in C_0^\infty(B(0,1))$ be such a function that $\varphi \geq 0$ and $\int_{\mathbb{R}^n} \varphi(x)\, dx = 1$. A function with the given properties can be constructed explicitly. Indeed,

$$\tilde{\varphi}(x) = \begin{cases} \exp\left(\frac{1}{|x|^2-1}\right) & \text{if } |x| < 1, \\ 0 & \text{if } |x| \geq 1. \end{cases}$$

belongs to $C_0^\infty(\mathbb{R}^n)$ and we define

$$\varphi(x) = \frac{\tilde{\varphi}(x)}{\int_{\mathbb{R}^n} \varphi(y)\, dy}.$$

For $\varepsilon > 0$ we set

$$\varphi_\varepsilon(x) = \varepsilon^{-n}\varphi(x/\varepsilon).$$

Then $\operatorname{supp} \varphi_\varepsilon \subset B(0,\varepsilon)$, $\varphi_\varepsilon \geq 0$ and $\int_{\mathbb{R}^n} \varphi_\varepsilon(x)\, dx = 1$. For $f \in L^1_{\text{loc}}(\mathbb{R}^n)$ we put

$$f_\varepsilon = f * \varphi_\varepsilon.$$

It follows from Theorem 89 that $f_\varepsilon \in C^\infty$.

Theorem 90 *Let f be a function defined on $\mathbb{R}^n$.*

1. *If f is continuous, then $f_\varepsilon \rightrightarrows f$ uniformly on compact sets as $\varepsilon \to 0$.*
2. *If $f \in C^r$, then $D^\alpha f_\varepsilon \rightrightarrows D^\alpha f$ uniformly on compact sets for all $|\alpha| \leq r$.*

Proof. (1) Uniform continuity of f on compact sets implies that for any compact set K

$$\sup_{|y|<\varepsilon} \sup_{x\in K} |f(x) - f(x-y)| \to 0 \qquad \text{as } \varepsilon \to 0.$$

Since $\int_{\mathbb{R}^n} \varphi_\varepsilon(y)\, dy = 1$, we have $\int_{\mathbb{R}^n} f(x)\varphi_\varepsilon(y)\, dy = f(x)$ and hence for any compact set K

$$\begin{aligned} \sup_{x\in K} |f(x) - f_\varepsilon(x)| &= \sup_{x\in K} \left| \int_{\mathbb{R}^n} (f(x) - f(x-y))\varphi_\varepsilon(y)\, dy \right| \\ &\leq \sup_{x\in K} \int_{B(0,\varepsilon)} |f(x) - f(x-y)|\varphi_\varepsilon(y)\, dy \to 0 \qquad \text{as } \varepsilon \to 0. \end{aligned}$$

(2) It follows from Theorem 89 and the induction argument that $D^\alpha(f*\varphi_\varepsilon) = (D^\alpha f)*\varphi_\varepsilon$ for all $|\alpha| \leq r$ and hence the theorem follows from the part (1). $\square$

Lemma 91 *If $f \in L^p$, $1 \le p < \infty$, then $f_\varepsilon \in L^p(\mathbb{R}^n)$ and $\|f_\varepsilon\|_p \le \|f\|_p$.*

Proof. If $p = 1$, then

$$\|f_\varepsilon\|_1 = \|f * \varphi_\varepsilon\|_1 \le \|f\|_1 \|\varphi_\varepsilon\|_1 = \|f\|_1$$

by Theorem 87. Let then $1 < p < \infty$. Hölder's inequality yields

$$\begin{aligned} |f_\varepsilon(x)|^p &= \left| \int_{\mathbb{R}^n} f(x-y)\varphi_\varepsilon(y)\, dy \right|^p \le \left(\int_{\mathbb{R}^n} |f(x-y)| \varphi_\varepsilon(y)^{1/p} \varphi_\varepsilon(y)^{1/q}\, dy \right)^p \\ &\le \left(\int_{\mathbb{R}^n} |f(x-y)|^p \varphi_\varepsilon(y)\, dy \right) \left(\int_{\mathbb{R}^n} \varphi_\varepsilon(y)\, dy \right)^{p/q} = |f|^p * \varphi_\varepsilon(x)\, . \end{aligned}$$

Hence

$$\|f_\varepsilon\|_p^p \le \int_{\mathbb{R}^n} |f|^p * \varphi_\varepsilon(x)\, dx = \|\, |f|^p * \varphi_\varepsilon\|_1 \le \|\, |f|^p\|_1 \|\varphi_\varepsilon\|_1 = \|f\|_p^p\, .$$

□

The following result proves density of the class of smooth functions in $L^p(\mathbb{R}^n)$.

Theorem 92 *If $f \in L^p(\mathbb{R}^n)$ and $1 \le p < \infty$, then $f_\varepsilon \in L^p(\mathbb{R}^n)$ and $\|f - f_\varepsilon\|_p \to 0$ as $\varepsilon \to 0$.*

Proof. We will need the following result.

Lemma 93 *If $f \in L^p(\mathbb{R}^n)$, $1 \le p < \infty$, then*

$$\lim_{y \to 0} \int_{\mathbb{R}^n} |f(x+y) - f(x)|^p\, dx = 0\, .$$

Proof. For $y \in \mathbb{R}^n$ let $\tau_y : L^p(\mathbb{R}^n) \to L^p(\mathbb{R}^n)$ be the translation operator defined by

$$(\tau_y f)(x) = f(x+y) \qquad \text{for } f \in L^p(\mathbb{R}^n).$$

The lemma says that τ_y is continuous, i.e. $\|\tau_y f - f\|_p \to 0$ as $y \to 0$. Given $\varepsilon > 0$ let g be a compactly supported continuous function such that $\|f - g\|_p < \varepsilon/3$, see Theorem 86. Then

$$\|\tau_y f - f\|_p \le \|\tau_y f - \tau_y g\|_p + \|\tau_y g - g\|_p + \|f - g\|_p = 2\|f - g\|_p + \|\tau_y g - g\|_p < 2\varepsilon/3 + \|\tau_y g - g\|_p\, .$$

Since $\tau_y g \rightrightarrows g$ converges uniformly, there is $\delta > 0$ such that $\|\tau_y g - g\|_p < \varepsilon/3$ for $|y| < \delta$ and hence

$$\|\tau_y f - f\|_p < \varepsilon \qquad \text{for } |y| < \delta,$$

which proves the lemma. □

Now we can complete the proof of the theorem. Assume first that $1 < p < \infty$. Hölder's inequality and Fubini's theorem yield

$$\begin{aligned}
\|f - f_\varepsilon\|_p &= \int_{\mathbb{R}^n} \left| \int_{\mathbb{R}^n} (f(x) - f(x-y))\varphi_\varepsilon(y)\, dy \right|^p dx \\
&\leq \int_{\mathbb{R}^n} \Big(\int_{\mathbb{R}^n} |f(x) - f(x-y)| \varphi_\varepsilon(y)^{1/p} \varphi_\varepsilon(y)^{1/q}\, dy \Big)^p dx \\
&\leq \int_{\mathbb{R}^n} \Big(\int_{\mathbb{R}^n} |f(x) - f(x-y)|^p \varphi_\varepsilon(y)\, dy \big(\underbrace{\int_{\mathbb{R}^n} \varphi_\varepsilon(y)\, dy}_{1} \big)^{p/q} \Big) dx \\
&= \int_{\mathbb{R}^n} \int_{B(0,\varepsilon)} |f(x) - f(x-y)|^p \varphi_\varepsilon(y)\, dy\, dx \\
&= \int_{B(0,\varepsilon)} \Big(\int_{\mathbb{R}^n} |f(x) - f(x-y)|^p\, dx \Big) \varphi_\varepsilon(y)\, dy \to 0 \qquad \text{as } \varepsilon \to 0.
\end{aligned}$$

by Lemma 93. If $p = 1$, then the above argument simplifies, because we do not have to use Hölder's inequality. □

Corollary 94 *C_0^∞ is dense in $L^p(\mathbb{R}^n)$ for all $1 \leq p < \infty$.*

Proof. Since continuous functions with compact support are dense in L^p it suffices to prove that every compactly supported continuous function can be approximated in L^p by C_0^∞. This, however, immediately follows from Theorem 92, because if f vanishes outside a bounded set, then f_ε has compact support. □

8 Functions of bounded variation

Definition. We say that a function $f : [a, b] \to \mathbb{R}$ has *bounded variation* if its *variation* defined by

$$\bigvee_a^b f = \sup \sum_{i=0}^{n-1} |f(x_{i+1}) - f(x_i)| < \infty$$

is finite, where the supremum is taken over all n and all partitions $a = x_0 < x_1 < \ldots < x_n = b$.

Clearly monotonic functions have bounded variation

$$\bigvee_a^b f = |f(b) - f(a)|\,.$$

If a function $f : [a,b] \to \mathbb{R}$ has bounded variation, then the function $x \mapsto \bigvee_a^x f$ is nondecreasing on $[a,b]$. Also the function $x \mapsto \bigvee_a^x f - f(x)$ is nondecreasing, because $f(y) - f(x) \leq \bigvee_x^y f$ for $y > x$. Therefore every function of bounded variation is a difference of two nondecreasing functions.

$$f(x) = \bigvee_a^x f - \Big(\bigvee_a^x f - f(x)\Big)\,. \tag{48}$$

We proved

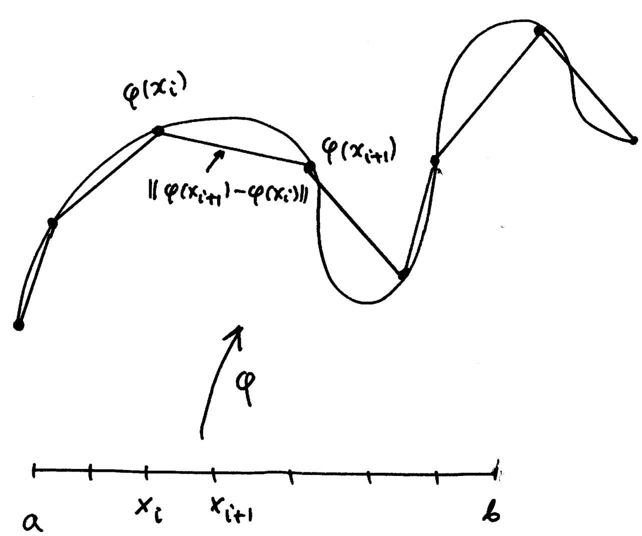
Theorem 95 (Jordan) *A function $f : [a,b] \to \mathbb{R}$ has bounded variation if and only if it is a difference of two nondecreasing functions.*

Exercise. *Prove that if $f : [a,b] \to \mathbb{R}$ is a continuous function of bounded variation, then the function $x \mapsto \bigvee_a^x f$ is continuous.*

The exercise and the formula (48) implies

Theorem 96 *A continuous function has bounded variation if and only if it is a difference of two continuous nondecreasing functions.*

Geometrically bounded variation is related to the length of a curve. Recall that if $\varphi = (\varphi_1, \ldots, \varphi_k) : [a,b] \to \mathbb{R}^k$ is a continuous curve, then its length is defined by



$$\begin{aligned}\ell(\varphi) &= \sup \sum_{i=0}^{n-1} \|\varphi(x_{i+1}) - \varphi(x_i)\| \\ &= \sup \sum_{i=1}^{n-1} \sqrt{(\varphi_1(x_{i+1}) - \varphi_1(x_i))^2 + \ldots + (\varphi_k(x_{i+1}) - \varphi_k(x_i))^2}\end{aligned}$$

where the supremum is taken over all n and all partitions $a = x_0 < x_1 < \ldots < x_n = b$.

DEFINITION. We say that a curve $\varphi : [a, b] \to \mathbb{R}^k$ is *rectifiable* if it has finite length $\ell(\varphi) < \infty$.

The following result easily follows from the definition of length and the definition of variation.

Theorem 97 *A continuous curve* $\varphi : [a, b] \to \mathbb{R}^k$ *is rectifiable if and only if the coordinate functions* $\varphi_i : [a, b] \to \mathbb{R}$ *are of bounded variation for* $i = 1, 2, \ldots, k$.

The following result is an easy exercise.

Theorem 98 *If functions* $f, g : [a, b] \to \mathbb{R}$ *have bounded variation, then the functions* $f \pm g$ *and* fg *have bounded variation. If in addition* $g \geq c > 0$*, then the function* f/g *has bounded variation.*

Exercise. *Prove the above theorem.*

Proposition 99 *Let* f *be a monotonic function on* (a, b)*. Then the set of points of* (a, b) *at which* f *is discontinuous is at most countable.*

Proof. Suppose f is nondecreasing. If f is discontinuous at $x \in (a, b)$, then

$$\lim_{t \to x^-} f(t) < \lim_{t \to x^+} f(t)$$

and hence there is a rational number $r(x)$ such that

$$\lim_{t \to x^-} f(t) < r(x) < \lim_{x \to t^+} f(t)\,.$$

It is easy to see that if f is discontinuous at x_1 and at x_2, $x_1 \neq x_2$, then $r(x_1) \neq r(x_2)$. Therefore there is a one-to-one correspondence between the set of points where f is discontinuous and the following subset of $\mathbb{Q}$

$$\{r(x) : \ f \text{ is discontinuous at } x\}$$

which clearly is at most countable. □

Corollary 100 *If* $f : [a, b] \to \mathbb{R}$ *has bounded variation, then the set of points at which* f *is discontinuous is at most countable.*

We will prove now a deep and important result.

Theorem 101 *Functions of bounded variation are differentiable a.e.*

Proof. It suffices to prove that nondecreasing functions are differentiable a.e.

Lemma 102 *If* $f : [a, b] \to \mathbb{R}$ *is nondecreasing, then*

$$\limsup_{h\to 0^+} \frac{f(x+h)-f(x)}{h} < \infty \qquad a.e. \tag{49}$$

Proof. Let E be the set of points where the upper limit (49) equals ∞. Fix an arbitrary $K > 0$. For every $x \in E$ there is an arbitrarily small interval $[x, x+h]$ such that

$$\frac{f(x+h)-f(x)}{h} > K\,,$$

i.e.

$$h < K^{-1}(f(x+h)-f(x))\,. \tag{50}$$

Such intervals cover the set E in the Vitali sense. Accordingly, we can select disjoint intervals $I_1, I_2, \ldots$ such that $|E \setminus \bigcup_{k=1}^\infty I_k| = 0$, so $|E| \leq \sum_{k=1}^\infty |I_k|$ and hence

$$|E| \leq \sum_{k=1}^\infty |I_k| < K^{-1}(f(b)-f(a))$$

by (50) and the fact that f is nondecreasing. Since the above inequality holds for any $K > 0$ it follows that $|E| = 0$. $\square$

DEFINITION. The *Dini derivatives* are defined as follows

$$D^+f(x) = \limsup_{h\to 0^+} \frac{f(x+h)-f(x)}{h}\,, \qquad D_+f(x) = \liminf_{h\to 0^+} \frac{f(x+h)-f(x)}{h}\,,$$

$$D^-f(x) = \limsup_{h\to 0^-} \frac{f(x+h)-f(x)}{h}\,, \qquad D_-f(x) = \liminf_{h\to 0^-} \frac{f(x+h)-f(x)}{h}\,.$$

The lemma says that $D^+f < \infty$ a.e. Similarly we prove that all other Dini derivatives D_+f, D^-f, D_-f are finite a.e.

We need to prove that

$$D^+f = D_+f = D^-f = D_-f \quad \text{a.e.} \tag{51}$$

To this end it suffices to prove that

$$D_-f \geq D^+f \quad \text{a.e.} \quad \text{and} \quad D_+f \geq D^-f \quad \text{a.e.}$$

because then (51) will follow from the inequality

$$D^+f \geq D_+f \geq D^-f \geq D_-f \geq D^+f \quad \text{a.e.}$$

We will prove only that $D_-f \geq D^+f$ a.e. as the prof of the inequality $D_+f \geq D^-f$ a.e. is very similar.

We need to prove that the set $\{D_-f < D^+f\}$ has measure zero. We have

$$\{D_-f < D^+f\} = \bigcup_{\substack{0<u<v \\ u,v\in\mathbb{Q}}} \underbrace{\{D_-f < u < v < D^+f\}}_{E(u,v)}\,.$$

This is a countable union of sets $E(u,v)$ and thus it suffices to prove that each set $E(u,v)$ has measure zero. Let U be an arbitrary open set such that $E(u,v) \subset U$. For every $x \in E(u,v)$ there is an arbitrarily small $h > 0$ such that $[x-h,x] \subset U$ and $(f(x-h) - f(x))/(-h) < u$, i.e.[14]

$$\bigvee_{x-h}^{x} f < hu\,. \tag{52}$$

Such intervals cover the set $E(u,v)$ in the Vitali sense, so we can select pairwise disjoint intervals $I_1, I_2, \ldots$ such that

$$|E(u,v) \setminus \bigcup_{k=1}^{\infty} I_k| = 0\,.$$

We have

$$\sum_{k=1}^{\infty} \bigvee_{I_k} f \leq u \sum_{k=1}^{\infty} |I_k| \leq u|U|\,.$$

The first inequality follows from (52) and the second inequality from the fact that the intervals $I_1, I_2, \ldots$ are disjoint and contained in U.

Denote[15] $G = E(u,v) \cap \bigcup_{k=1}^{\infty} \operatorname{int} I_k$. Clearly $|E(u,v) \setminus G| = 0$. For every $x \in G$ there is an arbitrarily short interval $[x, x+h] \subset \bigcup_k \operatorname{int} I_k$ such that $(f(x+h) - f(x))/h > v$ and hence

$$\bigvee_{x}^{x+h} f > hv\,.$$

[14] We take $[x-h,x]$ because $D_-f < u$ is a condition for the limit as $h \to 0^-$.

[15] $\operatorname{int} I_k$ denotes the interior of I_k.

These intervals cover the set G in the Vitali sense, so we can select pairwise disjoint intervals $J_1, J_2, \ldots \subset \bigcup_k I_k$ such that $|G \setminus \bigcup_k J_k| = 0$ and

$$\sum_{k=1}^{\infty} \bigvee_{J_k} f > v \sum_{k=1}^{\infty} |J_k| = v \left| \bigcup_{k=1}^{\infty} J_k \right| .$$

We have

$$v|E(u,v)| = v|G| \leq v \left| \bigcup_{k=1}^{\infty} J_k \right| < \sum_{k=1}^{\infty} \bigvee_{J_k} f \leq \sum_{k=1}^{\infty} \bigvee_{I_k} f \leq u|U| .$$

Since the measure of U can be arbitrarily close to the measure of $E(u,v)$ we obtain

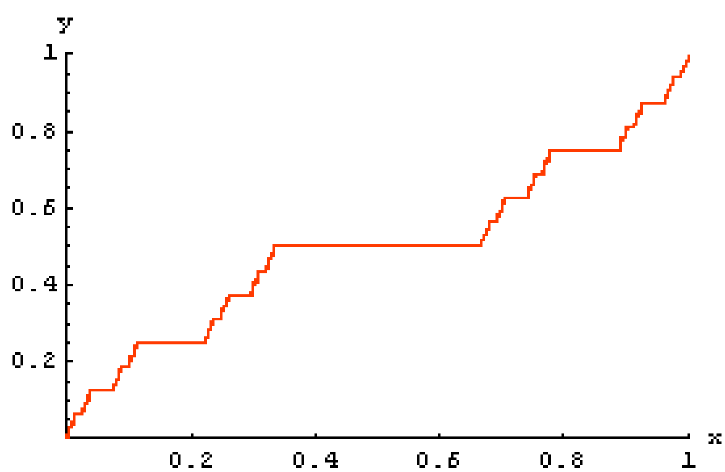
$$v|E(u,v)| \leq v|E(u,v)|$$

and hence $|E(u,v)| = 0$, because $v > u$. The proof is complete. $\square$

Theorem 103 *If a function $f : [a,b] \to \mathbb{R}$ has bounded variation, then $f' \in L^1(a,b)$. If in addition f is nondecreasing, then*

$$\int_a^b f'(x)\, dx \leq f(b) - f(a) .$$

Example. The Cantor Staircase Function. This is a continuous nondecreasing function $f : [0,1] \to [0,1]$ such that $f(0) = 0$, $f(1) = 1$ and f is constant in each of the intervals in the complement of the Cantor set. Since the Cantor set has measure zero we conclude that $f'(0) = 0$ a.e.



Therefore the function f has bounded variation (as a nondecreasing function) and

$$0 = \int_0^1 f'(x)\, dx < f(1) - f(0) = 1 .$$

Actually Sierpiński provided a more surprising example. He constructed a strictly increasing function with the derivative equal to zero a.e.

Proof of the theorem. We can assume that f is nondecreasing. Let $a_\varepsilon > a$ and $b_\varepsilon < b$ be points of differentiability of the function $F(x) = \int_a^x f(t)\,dt$ such that $|a - a_\varepsilon| < \varepsilon$, $|b - b_\varepsilon| < \varepsilon$ and $F'(a_\varepsilon) = f(a_\varepsilon)$, $F'(b_\varepsilon) = f(b_\varepsilon)$. Since

$$\lim_{h\to 0+} \frac{f(x+h) - f(x)}{h} = f'(x) \quad \text{a.e.}$$

Fatou's lemma yields

$$\begin{aligned} \int_{a_\varepsilon}^{b_\varepsilon} f'(x)\,dx &\le \liminf_{h\to 0+} \int_{a_\varepsilon}^{b_\varepsilon} \frac{f(x+h) - f(x)}{h}\,dx = \liminf_{h\to 0} \left(⨍_{b_\varepsilon}^{b_\varepsilon + h} f - ⨍_{a_\varepsilon}^{a_\varepsilon + h} f \right) \\ &= f(b_\varepsilon) - f(a_\varepsilon) \le f(b) - f(a)\,. \end{aligned}$$

Passing to the limit with $\varepsilon \to 0$ yields the result. □

8.1 Absolutely continuous functions

Definition. We say that a function $f : [a,b] \to \mathbb{R}$ is absolutely continuous if for every $\varepsilon > 0$ there is $\delta > 0$ such that if $(x_1, x_1 + h_1), \ldots, (x_k, x_k + h_k)$ are pairwise disjoint intervals in $[a,b]$ of total length less than δ, i.e. $\sum_{i=1}^k h_i < \delta$, then

$$\sum_{i=1}^{k} |f(x_i + h_i) - f(x_i)| < \varepsilon\,.$$

Theorem 104 *Absolutely continuous functions have bounded variation.*

We leave the proof as an exercise.

Theorem 105 *If $f, g : [a,b] \to \mathbb{R}$ are absolutely continuous, then also the functions $f \pm g$ and fg are absolutely continuous. If, in addition, $g \ge c > 0$ on $[a,b]$, then f/g is absolutely continuous.*

We leave the proof as an exercise.

The following result provides a complete characterization of absolutely continuous functions as the functions for which the fundamental theorem of calculus holds.

Theorem 106 *If $f \in L^1([a,b])$, then the function $F(x) = \int_a^x f(t)\,dt$ is absolutely continuous. On the other hand if F is absolutely continuous, then its derivative is integrable $F' \in L^1([a,b])$ and*

$$F(x) = F(a) + \int_a^x F'(t)\,dt \quad \textit{for all } x \in [a,b].$$

Proof. Absolute continuity of the function $F(x) = \int_a^x f(t)\,dt$ follows from the absolute continuity of the integral, Theorem 52.

If F is absolutely continuous, then it has bounded variation and hence $F' \in L^1([a,b])$ by Theorem 103. This implies that the function

$$x \mapsto \int_a^x F'(t)\,dt$$

is absolutely continuous. Hence the function

$$\varphi(x) = F(x) - F(a) - \int_a^x F'(t)\,dt$$

is absolutely continuous, $\varphi(a) = 0$, and[16] $\varphi'(x) = 0$ for a.e. x. It suffices to prove that $\varphi(x) = 0$ for all x. This will immediately follow from the next lemma.

Lemma 107 *If $\varphi : [a,b] \to \mathbb{R}$ is absolutely continuous and $\varphi'(x) = 0$ a.e., then φ is constant.*

Proof. Fix $\beta \in (a,b)$ and $\varepsilon > 0$. For almost every $x \in (a,\beta)$ there is an arbitrary small interval $[x, x+h] \subset (a,\beta)$ such that

$$\left| \frac{\varphi(x+h) - \varphi(x)}{h} \right| < \varepsilon\,.$$

Such intervals form a Vitali covering of the set of points in (a,β) where the derivative of φ equals 0 (almost all points in (a,β)). Hence we can select pairwise disjoint intervals from this covering $[x_1, x_1+h_1], [x_2, x_2+h_2], \ldots \subset (a,\beta)$ that cover almost all points of (a,β).

For a given $\varepsilon > 0$ we choose $\delta > 0$ as in the condition for the absolute continuity of φ. For N sufficiently large, the set $(a,\beta) \setminus \bigcup_{i=1}^N [x_i, x_i+h_i]$ has measure less than δ. This set is a union of $N+1$ open intervals. Denote these intervals by (ξ_j, ξ_j'), $j = 1, 2, \ldots, N+1$. We have

$$|\varphi(\beta) - \varphi(a)| \leq \sum_{i=1}^N |\varphi(x_i+h_i) - \varphi(x_i)| + \sum_{j=1}^{N+1} |\varphi(\xi_j') - \varphi(\xi_j)| \leq \varepsilon(\beta - a) + \varepsilon\,.$$

Since this inequality holds for any $\varepsilon > 0$ we conclude that $\varphi(\beta) = \varphi(a)$, so the function φ is constant. This proves the lemma and also completes the proof of the theorem. □

Theorem 108 (Integration by parts) *If the functions $f, g : [a,b] \to \mathbb{R}$ are absolutely continuous, then*

$$\int_a^b fg' = fg|_a^b - \int_a^b f'g\,.$$

[16]Because $\frac{d}{dt}\int_a^x F'(t)\,dt = F'(x)$ a.e. by the Lebesgue differentiation theorem.

Proof. Clearly $fg' = (fg)' - f'g$ a.e. and absolute continuity of fg implies

$$\int_a^b fg' = \int_a^b (fg)' - \int_a^b f'g = fg|_a^b - \int_a^b f'g\,.$$

□

Theorem 109 *Every function of bounded variation is a sum of an absolutely continuous function and a singular function, i.e. such a function that its derivative equals* 0 *a.e.*

Proof. The theorem follows immediately from the equality

$$f(x) = \Big(f(x) - \int_a^x f'(t)\,dt\Big) + \int_a^x f'(t)\,dt\,.$$

□

EXAMPLE. Every Lipschitz function $f : [a,b] \to \mathbb{R}$ is absolutely continuous and hence it is differentiable a.e. and satisfies the claim of the fundamental theorem of calculus.

DEFINITION. For a function $f : [a,b] \to \mathbb{R}$, the *Banach indicatrix* is defined by

$$N(f,y) = \# f^{-1}(y) \quad \text{for } y \in \mathbb{R},$$

i.e. $N(f,y)$ is the number of points in the preimage of y (it can be equal to 0 or to ∞ for some y).

Theorem 110 (Banach) *For a continuous function* $f : [a,b] \to \mathbb{R}$ *its indicatrix is measurable and*

$$\bigvee_a^b f = \int_{\mathbb{R}} N(f,y)\,dy\,. \tag{53}$$

Remark. We do not assume that the function f has bounded variation, so if the function f has unbounded variation, then the theorem says that the integral of the function on the right hand side of (53) equals ∞.

Proof. Under construction.

Theorem 111 *If* $f : [a,b] \to \mathbb{R}$ *is absolutely continuous, then*

$$\bigvee_a^b f = \int_a^b |f'(t)|\,dt\,.$$

Proof. Under construction.

The above two theorems immediately imply the following result.

Corollary 112 *If $f : [a, b] \to \mathbb{R}$ is absolutely continuous, then*

$$\int_a^b |f'(t)|\, dt = \int_{\mathbb{R}} N(f, y)\, dy\,.$$

Theorem 113 (Rademacher) *If $f : \Omega \to \mathbb{R}$ is Lipschitz continuous, where $\Omega \subset \mathbb{R}^n$ is open, then f is differentiable a.e., i.e.*

$$\nabla f(x) = \left\langle \frac{\partial f}{\partial x_1}(x), \ldots, \frac{\partial f}{\partial x_n}(x) \right\rangle$$

exists a.e. and

$$\lim_{y \to x} \frac{f(y) - f(x) - \nabla f(x) \cdot (y - x)}{\|y - x\|} = 0 \quad \textit{for a.e. } x \in \Omega\,.$$

Proof. Let $\nu \in S^{n-1}$ and

$$D_\nu f(x) = \frac{d}{dt} f(x + t\nu)|_{t=0}$$

be the directional derivative whenever it exists. Since $D_\nu f(x)$ exists precisely when the Borel-measurable functions

$$\limsup_{t \to 0} \frac{f(x + t\nu) - f(x)}{t} \quad \text{and} \quad \liminf_{t \to 0} \frac{f(x + t\nu) - f(x)}{t}$$

coincide, the set A_ν on which $D_\nu f$ fails to exist is Borel-measurable. The function $t \mapsto f(x+t\nu)$ is absolutely continuous and hence differentiable a.e. Thus the intersection of the set A_ν with any line parallel to ν has one dimensional measure zero and hence Fubini's theorem implies that A_ν has measure zero. Accordingly, for every $\nu \in S^{n-1}$

$$D_\nu f(x) \quad \text{exists for a.e. } x \in \Omega$$

Take any $\varphi \in C_0^\infty(\Omega)$ and note that for each sufficiently small $h > 0$

$$\int_\Omega \frac{f(x + h\nu) - f(x)}{h}\, \varphi(x)\, dx = -\int_\Omega \frac{\varphi(x - h\nu) - \varphi(x)}{-h}\, f(x)\, dx\,.$$

Indeed, invariance of the integral with respect to translations yields

$$\int_\Omega f(x + h\nu)\varphi(x)\, dx = \int_\Omega f(x)\varphi(x - h\nu)\, dx$$

and h has to be so small that $x+h\nu \in \Omega$ for all $x \in \operatorname{supp}\varphi$. The dominated convergence theorem yields

$$\int_\Omega D_\nu f(x)\varphi(x)\,dx = -\int_\Omega f(x) D_\nu\varphi(x)\,dx\,.$$

This is true for any $\nu \in S^{n-1}$ and in particular we have

$$\int_\Omega \frac{\partial f}{\partial x_i}(x)\,\varphi(x)\,dx = -\int_\Omega f(x)\,\frac{\partial\varphi}{\partial x_i}(x)\,dx \quad \text{for } i=1,2,\ldots,n.$$

Now

$$\begin{aligned}\int_\Omega D_\nu f(x)\varphi(x)\,dx &= -\int_\Omega f(x)D_\nu\varphi(x)\,dx = -\int_\Omega f(x)\,(\nabla\varphi(x)\cdot\nu)\,dx\\ &= -\sum_{i=1}^n\int_\Omega f(x)\,\frac{\partial\varphi}{\partial x_i}(x)\,\nu_i\,dx = \sum_{i=1}^n\int_\Omega \varphi(x)\,\frac{\partial f}{\partial x_i}(x)\,\nu_i\,dx\\ &= \int_\Omega \varphi(x)(\nabla f(x)\cdot\nu)\,dx\,.\end{aligned}$$

Lemma 114 *If $g \in L^1_{\mathrm{loc}}(\Omega)$ and $\int_\Omega g(x)\varphi(x)\,dx = 0$ for every $\varphi \in C_0^\infty(\Omega)$, then $g=0$ a.e.*

We leave the proof of this lemma as an exercise. The lemma implies that

$$D_\nu f(x) = \nabla f(x)\cdot\nu \quad \text{a.e.}$$

Now let $\nu_1,\nu_2,\ldots$ be a countable dense subset of S^{n-1} and let

$$A_k = \{x\in\Omega:\ \nabla f(x),\, D_{\nu_k}f(x) \text{ exists and } D_{\nu_k}f(x) = \nabla f(x)\cdot\nu_k\}\,.$$

Let $A = \bigcap_{k=1}^\infty A_k$. Clearly $|\Omega\setminus A| = 0$ and

$$D_{\nu_k}f(x) = \nabla f(x)\cdot\nu_k \quad \text{for all } x\in A \text{ and all } k=1,2,\ldots$$

We will prove that f is differentiable at each point of the set A. To this end, for any $x\in A$, $\nu\in S^{n-1}$ and $h>0$ define

$$Q(x,\nu,h) = \frac{f(x+h\nu)-f(x)}{h} - \nabla f(x)\cdot\nu$$

and it suffices to prove that if $x\in A$, then for every $\varepsilon>0$ there is $\delta>0$ such that

$$|Q(x,\nu,h)| < \varepsilon \quad \text{whenever } 0<h<\delta \text{ and } \nu\in S^{n-1}.$$

Let the function f be L-Lipschitz, i.e. $|f(x)-f(y)|\le L|x-y|$ for all $x,y\in\Omega$. This implies that $|\partial f/\partial x_i|\le L$ a.e. and thus $|\nabla f(x)|\le\sqrt{n}L$ a.e. Hence for any $x\in A$, $\nu,\nu'\in S^{n-1}$ and $h>0$

$$|Q(x,\nu,h)-Q(x,\nu',h)| \le (\sqrt{n}+1)L|\nu-\nu'|\,.$$

Given $\varepsilon > 0$ let p be so large that for each $\nu \in S^{n-1}$

$$|\nu - \nu_k| < \frac{\varepsilon}{2(\sqrt{n}+1)L} \quad \text{for some } k = 1, 2, \ldots, p.$$

Since

$$\lim_{h \to 0^+} Q(x, \nu_i, h) = 0 \quad \text{for all } x \in A \text{ and all } i = 1, 2, \ldots$$

we see that for a given $x \in A$ there is $\delta > 0$ such that

$$|Q(x, \nu_i, h)| < \varepsilon/2 \quad \text{whenever } 0 < h < \delta \text{ and } i = 1, 2, \ldots, p\,.$$

Now

$$|Q(x, \nu, h)| \leq |Q(x, \nu_k, h)| + |Q(x, \nu_k, h) - Q(x, \nu, h)| < \varepsilon/2 + (\sqrt{n}+1)L|\nu_k - \nu| < \varepsilon$$

whenever $\nu \in S^{n-1}$ and $0 < h < \delta$ and the theorem follows. □

Theorem 115 (McShane) *If $f : A \to \mathbb{R}$ is an L-Lipschitz function defined on a subset $A \subset X$ of a metric space X, then there is an L-Lipschitz function $\tilde{f} : X \to \mathbb{R}$ such that $\tilde{f}|_A = f$.*

Proof. For $x \in X$ we define

$$\tilde{f}(x) = \inf_{y \in A} (f(y) + Ld(x, y))$$

and it a a routine exercise to check that $\tilde{f}$ is L-Lipschitz and satisfies $\tilde{f}|_A = f$. □

Theorem 116 (Stepanov) *Let $\Omega \subset \mathbb{R}^n$ be open. Then a measurable function $f : \Omega \to \mathbb{R}$ is differentiable a.e. in Ω if and only if*

$$\limsup_{y \to x} \frac{|f(f) - f(x)|}{\|y - x\|} < \infty \quad \textit{for a.e. } x \in \Omega. \tag{54}$$

The implication $\Rightarrow$ is obvious. To prove the implication $\Leftarrow$ first we show that (54) implies that f is Lipschitz continuous on a big set A in the sense that the measure of $\Omega \setminus A$ is small. Then we extend $f|_A$ to a Lipschitz function on $\mathbb{R}^n$ using McShane's theorem. Now $\tilde{f}$ is differentiable a.e. by Rademacher theorem. In particular it is differentiable at a.e. point of A and it remains to prove that f is differentiable at every density point of A at which $\tilde{f}$ is differentiable. We leave details as an exercise. □

9 Signed measures

DEFINITION. Let $\mathfrak{M}$ be a σ-algebra. A function $\mu : \mathfrak{M} \to [-\infty, \infty]$ is called *signed measure* if

1. $\mu(\emptyset) = 0$;
2. μ attains at most one of the values $\pm\infty$;
3. μ is countably additive.

EXAMPLES.

1. $\mu = \mu_1 - \mu_2$, where μ_1 and μ_2 are (positive) measures and one of them is finite.
2. $$\mu(E) = \int_E f\, d\nu, \quad f \in L^1(\nu)\, .$$

DEFINITION. A measurable set $E \in \mathfrak{M}$ is called *positive* if $\mu(A) > 0$ for every measurable subset $A \subset E$.

EXAMPLE. If $E = E_1 \cup E_2$, $E_1 \cap E_2 = \emptyset$, $\mu(E_1) = 7$, $\mu(E_2) = -3$, then $\mu(E) = 4 > 0$, but E is not positive.

Lemma 117 *If μ is a signed measure, then every measurable set such that $0 < \mu(E) < \infty$ contains a positive set A of positive measure.*

Proof. If E is positive, then we take $A = E$. Otherwise there is $B \subset E$, $\mu(B) < 0$. Let n_1 be the smallest positive integer such that there is $B_1 \subset E$ with

$$\mu(B_1) \leq -\frac{1}{n_1}\, .$$

If $A_1 = E \setminus B_1$ is positive, then we take $A = A_1$ (observe that $\mu(A_1) > 0$ as otherwise $\mu(E) < 0$). If A_1 is not positive, then let n_2 be the smallest positive integer such that there is $B_2 \subset E \setminus B_1$ with

$$\mu(B_2) \leq -\frac{1}{n_2}\, .$$

If $A_2 = E \setminus (B_1 \cup B_2)$ is positive, then we take $A = A_2$ (observe that $\mu(A_2) > 0$). If A_2 is not positive, then we define A_3 as above and so on. If $A_m = E \setminus \bigcup_{j=1}^m B_j$ is positive for some m, then we take $A = A_m$ ($\mu(A_m) > 0$). Otherwise we obtain an infinite sequence B_j and we set

$$A = E \setminus \bigcup_{j=1}^{\infty} B_j\, .$$

We will prove that A is positive and $\mu(A) > 0$. We have

$$\mathbb{R} \ni \mu(A) = \mu(E) - \sum_{j=1}^{\infty} \mu(B_j)$$

and hence

$$0 < \mu(E) = \mu(A) + \sum_{j=1}^{\infty} \mu(B_j) \leq \mu(A) - \sum_{j=1}^{\infty} \frac{1}{n_j}$$

$$\sum_{j=1}^{\infty} \frac{1}{n_j} < \mu(A)\,. \tag{55}$$

Note that $\mu(A) < \infty$ as otherwise $\mu(E) = \mu(A) + \mu(E \setminus A) = \infty + \mu(E \setminus A) = \infty$. Hence (55) yields $\mu(A) > 0$ and $\sum_{j=1}^{\infty} 1/n_j < \infty$. In particular $n_j \to \infty$ ad $j \to \infty$. It remains to prove that A is positive. If $C \subset A = E \setminus \bigcup_{j=1}^{\infty} B_j$ is measurable, the for every m, $C \subset E \setminus \bigcup_{j=1}^{m} B_j$ and the definition of n_{m+1} yields

$$\mu(C) > -\frac{1}{n_{m+1} - 1} \to 0 \quad \text{as } m \to \infty$$

and hence $\mu(C) \geq 0$. This proves positivity of A. □

Theorem 118 (Hahn's decomposition theorem) *If μ is a signed measure, then there exist two disjoint measurable sets X^+ and X^- such that*

$$X = X^+ \cup X^-, \qquad X^+ \cap X^- = \emptyset\,,$$

μ is positive on X^+ and μ is negative on X^- (i.e. $-\mu$ is positive on X^-).

Therefore every signed measure is on the form

$$\mu = \mu^+ - \mu^-$$

where μ^+ and μ^- are positive measures defined by $\mu^+(E) = \mu(E \cap X^+)$, $\mu^-(E) = -\mu(E \cap X^-)$. Moreover the measures μ^+ and μ^- are concentrated on disjoint sets X^+ and X^-.

Exercise. *Prove that the decomposition $X = X^+ \cup X^-$ is unique up to sets of μ-measure zero, i.e. if $X = \tilde{X}^+ \cup \tilde{X}^-$ is another decomposition, then*

$$\mu(X^+ \setminus \tilde{X}^+) = \mu(\tilde{X}^+ \setminus X^+) = \mu(X^- \setminus \tilde{X}^-) = \mu(\tilde{X}^- \setminus X^-) = 0\,.$$

Example. Let $\mu(E) = \int_E f\, d\nu$, where $f \in L^1(\nu)$. Then μ is a signed measure and

$$\mu(E) = \int_E f^+\, d\nu - \int_E f^-\, d\nu\,.$$

Hence we can take

$$X^+ = \{x : \ f(x) \geq 0\}, \qquad X^- = \{x : \ f(x) < 0\}\,,$$

but we also can take

$$\tilde{X}^+ = \{x : \ f(x) > 0\}, \qquad \tilde{X}^- = \{x : \ f(x) \leq 0\}\,.$$

In general $X^+ \neq \tilde{X}^+$, $X^- \neq \tilde{X}^-$, however, the sets differ by the set where f equals 0 and this set has measure zero

$$\mu(\{x : \ f(x) = 0\}) = \int_{\{f=0\}} f\, d\nu = 0\,.$$

Proof of Hahn's theorem. Suppose μ does not attain value $+\infty$. Let

$$M = \sup\{\mu(A) : \ A \in \mathfrak{M},\ A \text{ is positive}\}\,.$$

Then there is a sequence of positive sets

$$A_1 \subset A_2 \subset A_3 \subset \ldots; \quad \mu(A_i) \to M$$

and if $A = \bigcup_{i=1}^\infty A_i$, then A is positive with $\mu(A) = M < \infty$. It remains to prove that $X \setminus A$ is negative (then we take $X^+ = A$, $X^- = X \setminus A$). For if not, there is $E \subset X \setminus A$ with $0 < \mu(E) < \infty$. Hence Lemma 117 implies that there is a positive subset $C \subset E$ of positive measure. Now $A \cup C$ is positive and

$$\mu(A \cup C) = \mu(A) + \mu(C) > M$$

which is an obvious contradiction. □

As an application we will prove the Radon-Nikodym-Lebesgue theorem and then we will classify all continuous linear functionals on $L^p(\mu)$ for $1 \leq p < \infty$.

Definition. Let μ be a (positive) measure in a σ-algebra $\mathfrak{M}$. If ν is another (positive) measure in $\mathfrak{M}$ such that

$$E \in \mathfrak{M},\ \mu(E) = 0 \Rightarrow \nu(E) = 0\,,$$

then we say that ν is *absolutely continuous with respect to* μ and we write

$$\nu \ll \mu\,.$$

Example. If $f \geq 0$ is measurable, and a measure ν is defined by

$$\nu(E) = \int_E f\, d\mu\,,$$

then $\nu \ll \mu$.

DEFINITION. Let λ be a measure in $\mathfrak{M}$. We say that λ is *concentrated* on $A \in \mathfrak{M}$ if

$$\lambda(E) = \lambda(A \cap E) \quad \text{for all } E \in \mathfrak{M}\,.$$

Equivalently λ is concentrated on $A \in \mathfrak{M}$ if $\lambda(X \setminus A) = 0$.

DEFINITION. Let λ_1, λ_2 be two measures in $\mathfrak{M}$. If there are sets

$$A, B \in \mathfrak{M}, \quad A \cap B = \emptyset$$

such that

$$\lambda_1 \text{ is concentrated on } A\,,$$
$$\lambda_2 \text{ is concentrated on } B\,,$$

then we say that the measures λ_1 and λ_2 are *mutually singular* and we write

$$\lambda_1 \perp \lambda_2\,.$$

Lemma 119 *Let λ_1, λ_2 be two measures in $\mathfrak{M}$. If there is a set $E \in \mathfrak{M}$ such that $\lambda_1(E) = 0$ and λ_2 is concentrated on E, then $\lambda_1 \perp \lambda_2$.*

Proof. It suffices to take $A = X \setminus E$ and $B = E$. □

EXAMPLES.

1. For $a \in \mathbb{R}^n$ the *Dirac mass* is a Borel measure defined by $\delta_a(E) = 1$ if $a \in E$ and $\delta_a(E) = 0$ if $a \notin E$. The lemma yields $\mathcal{L}_n \perp \delta_a$.
2. If M is a k-dimensional submanifold in $\mathbb{R}^n$, $k < n$, and $\lambda(E) = \mathcal{H}^k(E \cap M)$ for $E \subset \mathbb{R}^n$, then $\mathcal{L}_n \perp \lambda$.
3. if C is the standard ternary Cantor set and $\lambda(E) = \mathcal{H}^{\log 2/\log 3}(E \cap C)$ for $E \subset \mathbb{R}$, then $\lambda \perp \mathcal{L}_1$.

Theorem 120 (Radon-Nikodym-Lebesgue) *Let μ and ν be two (positive), σ-finite measures on $\mathfrak{M}$. Then*

(a) *There is a unique pair of measures λ_a, λ_s on $\mathfrak{M}$ such that*

$$\lambda = \lambda_a + \lambda_s, \qquad \lambda_a \ll \mu, \qquad \lambda_s \perp \mu\,.$$

(b) *There is a $\mathfrak{M}$-measurable function $h \geq 0$ on X such that*

$$\lambda_a(E) = \int_E h\, d\mu \qquad \text{for all } E \in \mathfrak{M}.$$

If in addition λ is finite (i.e. $\lambda(X) < \infty$), then the measures λ_a, λ_s are finite and $h \in L^1(\mu)$.

Remark. Part (a) is due to Lebesgue and part (b) due to Radon and Nikodym.

In particular if $\lambda \ll \mu$, then

$$\lambda(E) = \int_E h\, d\mu \tag{56}$$

for some $h \geq 0$ and all $E \in \mathfrak{M}$ which is a striking characterization of all absolutely continuous measures. The function h is called *Radon-Nikodym derivative* and is often denoted by $h = d\lambda/d\mu$. With this notation (56) can be rewritten as

$$\lambda(E) = \int_E \frac{d\lambda}{d\mu}\, d\mu \qquad \text{for all } E \in \mathfrak{M}\,.$$

Proposition 121 *If λ is a finite measure, absolutely continuous with respect to the Lebesgue measure, $\lambda \ll \mathcal{L}_n$, then*

$$\frac{d\lambda}{d\mathcal{L}_n} = \lim_{r\to 0^+} \frac{\lambda(B(x,r))}{|B(x,r)|} \quad \textit{for a.e. } x \in \mathbb{R}^n\,.$$

Proof. It follows from the Radon-Nikodym theorem that $d\lambda/d\mathcal{L}_n \in L^1(\mathbb{R}^n)$ and

$$\frac{\lambda(B(x,r))}{|B(x,r)|} = ⨍_{B(x,r)} \frac{d\lambda}{d\mathcal{L}_n}\, d\mathcal{L}_n \to \frac{d\lambda}{d\mathcal{L}_n}(x) \quad \text{as } r \to 0^+$$

for a.e. $x \in \mathbb{R}^n$ by the Lebesgue differentiation theorem. □

Proof of the Radon-Nikodym-Lebesgue theorem. First we will prove the second part of the theorem which is due to Radon and Nikodym. We can state it as follows.

Theorem 122 (Radon-Nikodym) *If λ and μ are σ-finite measures on a σ-algebra $\mathfrak{M}$ and $\lambda \ll \mu$, then there is a nonnegative measurable function $h : X \to [0,\infty]$ such that*

$$\lambda(E) = \int_E h\, d\mu \quad \textit{for all } E \in \mathfrak{M}\,.$$

If in addition λ is finite, then $h \in L^1(\mu)$.

Proof. We will prove the theorem under the additional assumption that $\mu(X) < \infty$ and $\lambda(X) < \infty$. The general case of σ-finite measures easily follows from this special case and the details are left to the reader. Let

$$\Phi = \{\varphi : X \to [0,\infty] :\ \varphi \text{ is measurable and } \int_E \varphi\, d\mu \leq \lambda(E) \text{ for all } E \in \mathfrak{M}\}\,.$$

Clearly $\Phi \neq \emptyset$, because $\varphi \equiv 0 \in \Phi$. If $\varphi_1, \varphi_2 \in \Phi$, then $\max\{\varphi_1, \varphi_2\} \in \Phi$. Indeed,

$$\begin{aligned}\int_E \max\{\varphi_1,\varphi_2\}\, d\mu &= \int_{E\cap\{\varphi_1\geq\varphi_2\}} \varphi_1\, d\mu + \int_{E\cap\{\varphi_1<\varphi_2\}} \varphi_2\, d\mu \\ &\leq \lambda(E\cap\{\varphi_1\geq\varphi_2\}) + \lambda(E\cap\{\varphi_1<\varphi_2\}) = \lambda(E)\,.\end{aligned}$$

Define

$$M = \sup_{\varphi \in \Phi} \int_X \varphi \, d\mu \le \lambda(X) < \infty \,.$$

There is a sequence $\varphi_n \in \Phi$ such that

$$\int_X \varphi_n \, d\mu \to M \quad \text{as } n \to \infty \,.$$

Let $h_n = \max\{\varphi_1, \varphi_2, \ldots, \varphi_n\} \in \Phi$. Then h_n is an increasing sequence of functions pointwise convergent to a function h. Since $\int_E h_n \, d\mu \le \lambda(E)$ for all $E \in \mathfrak{M}$ it follows from the monotone convergence theorem that $\int_E h \, d\mu \le \lambda(E)$ for all $E \in \mathfrak{M}$. Hence $h \in \Phi$. Moreover

$$\int_X h \, d\mu = \lim_{n\to\infty} \int_X h_n \, d\mu = M \,.$$

We will show that

$$\int_E h \, d\mu = \lambda(E) \quad \text{for all } E \in \mathfrak{M} \,.$$

Clearly

$$\eta(E) = \lambda(E) - \int_E h \, d\mu$$

defines a positive measure in $\mathfrak{M}$ and we need to show that $\eta \equiv 0$. For if not there would exist a set $A \in \mathfrak{M}$ such that $\eta(A) > 0$. This implies that $\lambda(A) > 0$ and hence $\mu(A) > 0$ (because $\lambda \ll \mu$). Thus there is $\varepsilon > 0$ such that $\eta(A) - \varepsilon\mu(A) > 0$. Define $\xi = \eta - \varepsilon\mu$. Then ξ is a signed measure and we can assume that A is positive with respect to ξ by Lemma 117. Hence if $E \in \mathfrak{M}$, $\xi(A \cap E) \ge 0$ and thus

$$0 \le \xi(A \cap E) = \eta(A \cap E) - \varepsilon\mu(A \cap E) = \lambda(A \cap E) - \int_{A\cap E} h \, d\mu - \varepsilon\mu(A \cap E) \,,$$

$$\int_{A\cap E} h \, d\mu + \varepsilon\mu(A \cap E) \le \lambda(A \cap E) \,,$$

but also

$$\int_{E\setminus A} h \, d\mu \le \lambda(E \setminus A)$$

because $h \in \Phi$. Adding up the two inequalities yields

$$\int_E (h + \varepsilon\chi_A) \, d\mu = \int_E h \, d\mu + \varepsilon\mu(A \cap E) \le \lambda(E) \,.$$

This implies that $h + \varepsilon\chi_A \in \Phi$. Since

$$\int_X (h + \varepsilon\chi_A) \, d\mu = M + \varepsilon\mu(A) > M \,,$$

we get a contradiction. The proof is complete. □

Now we can prove the first part of the theorem which is due to Lebesgue.

Theorem 123 (Lebesgue) *If λ and μ are σ-finite measures, then there exist unique measures λ_a and λ_s such that*

$$\lambda = \lambda_a + \lambda_s, \qquad \lambda_a \ll \mu, \qquad \lambda_s \perp \mu\,.$$

Proof. We leave the proof of the uniqueness as an exercise and so we are left with the proof of the existence of λ_a and λ_s. Let $\nu = \mu + \lambda$. Then $\mu \ll \nu$ and $\lambda \ll \nu$. By the Radon-Nikodym theorem

$$\mu(E) = \int_E f\, d\nu, \qquad \lambda(E) = \int_E g\, d\nu$$

for some measurable functions $f, g : X \to [0, \infty]$ and all $E \in \mathfrak{M}$. Define

$$\lambda_0(E) = \lambda(E \cap \{f = 0\}), \qquad \lambda_1(E) = \lambda(E \cap \{f > 0\})\,.$$

Since $\lambda = \lambda_0 + \lambda_1$ it suffices to prove that

$$\lambda_0 \perp \mu \quad \text{and} \quad \lambda_1 \ll \mu\,.$$

The measure λ_0 is concentrated on the set $\{f = 0\}$. Since $\mu(\{f = 0\}) = 0$, we conclude that $\lambda_0 \perp \mu$. To prove that $\lambda_1 \ll \mu$ let $\mu(E) = 0$. Then $f = 0$, ν-a.e. in the set E, i.e. $\nu(E \cap \{f > 0\}) = 0$ and thus

$$\lambda_1(E) = \int_{E \cap \{f>0\}} g\, d\nu = 0\,.$$

□

We will apply now the Radon-Nikodym theorem to classify all continuous linear functionals on $L^p(\mu)$ for $1 \le p < \infty$. Let us start first with some general facts about continuous linear mappings between normed spaces.

Theorem 124 *Let $L : X \to Y$ be a linear mapping between normed spaces. Then the following conditinos are equivalent*

(a) *L is continuous;*

(b) *L is continuous at 0;*

(c) *L is* bounded, *i.e. there is $C > 0$ such that*

$$\|Lx\| \le C\|x\| \quad \textit{for all } x \in X\,.$$

Remark. Formally we should use different symbols to denote norms in spaces X and Y. Since it will always be clear in which space we take the norm it is not dangerous to use the same symbol $\|\cdot\|$ to denote apparently different norms in X and Y.

Proof. The implication (a)⇒(b) is obvious.

(b)⇒(c). Suppose L is continuous at 0 but not bounded. Then there is a sequence $x_n \in X$ such that

$$\|Lx_n\| \geq n\|x_n\|\,.$$

Hence

$$0 \leftarrow \left\| L\frac{x_n}{n\|x_n\|} \right\| \geq 1$$

which is a contradiction[17].

(c)⇒(a). Let $x_n \to x$. Then

$$\|Lx - Lx_n\| = \|L(x - x_n)\| \leq C\|x - x_n\| \to 0$$

and hence $Lx_n \to Lx$ which proves continuity of L. □

In particular if X is a normed space over $\mathbb{R}$ (or $\mathbb{C}$), then a linear functional

$$\Lambda : X \to \mathbb{R} \quad (\text{or } \mathbb{C})$$

is continuous if and only if it is bounded, i.e. there is $C > 0$ such that

$$|\Lambda x| \leq C\|x\| \quad \text{for all } x \in X. \tag{57}$$

The number

$$\|\Lambda\| = \sup_{\|x\|\leq 1} |\Lambda x|$$

is called the *norm* of Λ.

Lemma 125 *The norm Λ is the smallest number C for which the inequality (57) is satisfied.*

Proof. Suppose $|\Lambda x| \leq C\|x\|$ for all $x \in X$. If $\|x\| \leq 1$, then $|\Lambda x| \leq C\|x\| \leq C$ and hence

$$\|\Lambda\| = \sup_{\|x\|\leq 1} |\Lambda x| \leq C\,.$$

Thus the number C is (57) cannot be smaller than $\|\Lambda\|$. It remains to prove that

$$|\Lambda x| \leq \|\Lambda\|\,\|x\| \quad \text{for all } x \in X\,.$$

[17]Convergence to 0 follows from the continuity of L at 0.

If $x = 0$, then the inequality is obvious. If $x \neq 0$, then

$$|\Lambda x| = \left|\Lambda \frac{x}{\|x\|}\right| \|x\| \leq \|\Lambda\|\|x\| .$$

The last inequality follows from the fact that $\|x/\|x\|\| = 1$. □

EXAMPLE. If $1/p + 1/q = 1$, $1 < p, q < \infty$ are Hölder conjugate, and $g \in L^q$, then

$$\Lambda_g f = \int_X fg \, d\mu$$

defines a bounded linear functional on $L^p(\mu)$ and $\|\Lambda_g\| \leq \|g\|_q$. Indeed, Hölder's inequality yields

$$|\Lambda_g f| = \left|\int_X fg \, d\mu\right| \leq \|g\|_q \|f\|_p .$$

Similarly if $g \in L^\infty(\mu)$, then

$$\Lambda_g f = \int_X fg \, d\mu$$

defines a bounded linear functional on $L^1(\mu)$ and $\|\Lambda_g\| \leq \|g\|_\infty$.

It turns out that under reasonable assumptions every functional on $L^p(\mu)$, $1 \leq p < \infty$ is on the form Λ_g for some $g \in L^q$ (or $g \in L^\infty$ if $p = 1$).

Theorem 126 *Suppose $1 \leq p < \infty$ and μ is a σ-finite measure. Assume that Λ is a bounded linear functional on $L^p(\mu)$. Then there is a unique $g \in L^q(\mu)$, $1/p + 1/q = 1$ such that*

$$\Lambda f = \int_X fg \, d\mu \quad \text{for all } f \in L^p(\mu) \ .$$

Moreover

$$\|\Lambda\| = \|g\|_q .$$

Proof. Uniqueness. Suppose

$$\Lambda f = \int_X f g_1 \, d\mu = \int_X f g_2 \, d\mu$$

for some $g_1, g_2 \in L^q(\mu)$ and all $f \in L^p(\mu)$. We need to show that $g_1 = g_2$ a.e. For $E \in \mathfrak{M}$ with $\mu(E) < \infty$ we have

$$\int_E g_1 \, d\mu = \Lambda(\chi_E) = \int_E g_2 \, d\mu .$$

Hence

$$\int_E (g_1 - g_2) \, d\mu = 0$$

for every $E \in \mathfrak{M}$ with $\mu(E) < \infty$. This easily implies that $g_1 - g_2 = 0$ a.e. and hence $g_1 = g_2$ μ-a.e.

We will prove the theorem under the additional assumption that $\mu(X) < \infty$. The general case of σ-finite measure can be reduced to the case of finite measure and we leave details to the reader. Define

$$\lambda(E) = \Lambda(\chi_E)\,.$$

It is obvious that λ is finitely additive

$$\lambda(A \cup B) = \lambda(A) + \lambda(B) \quad \text{for } A, B \in \mathfrak{M},\ A \cap B = \emptyset\ .$$

Indeed,

$$\lambda(A \cup B) = \Lambda(\chi_{A\cup B}) = \Lambda(\chi_A + \chi_B) = \Lambda(\chi_A) + \Lambda(\chi_B) = \lambda(A) + \lambda(B)\,.$$

We will use this fact to prove that actually λ is conuntably additive. Let $E = \bigcup_{i=1}^{\infty} E_i$, where $E_i \in \mathfrak{M}$ and $E_i \cap E_j = \emptyset$ for $i \neq j$. Let $A_k = E_1 \cup \ldots \cup E_k$. We have

$$\|\chi_E - \chi_{A_k}\|_p = \mu(E \setminus A_k)^{1/p} \to 0 \quad \text{as } k \to \infty\ ,$$

because the sets $E \setminus A_k$ form a decreasing sequence of sets with empty intersection and $\mu(E \setminus A_1) \leq \mu(X) < \infty$. Continuity of the functional λ implies

$$\Lambda\chi_{A_k} \to \Lambda\chi_E$$

and hence finite additivity of λ yields

$$\sum_{i=1}^{k} \lambda(E_i) = \lambda(A_k) \to \lambda(E) \quad \text{as } k \to \infty\,.$$

Accordingly

$$\sum_{i=1}^{\infty} \lambda(E_i) = \lambda(E)$$

which is countable additivity of λ. Note that λ is not necessarily positive, so

$$\lambda : \mathfrak{M} \to \mathbb{R}$$

is a signed measure. By the Hahn decomposition theorem $\lambda = \lambda^+ - \lambda^-$, where λ^+ and λ^- are positive finite measures concentrated on disjoint sets. Clearly $\mu(E) = 0$ implies that $\lambda(E) = 0$ and also $\lambda^+(E) = \lambda^-(E) = 0$. Hence $\lambda^+, \lambda^- \ll \mu$. The Radon-Nikodym theorem yields the existence of $0 \leq g^{\pm} \in L^1(\mu)$ such that

$$\lambda^{\pm}(E) = \int_E g^{\pm}\, d\mu\,.$$

Now

$$\lambda(E) = \int_E g\,d\mu \quad \text{where } g = g^+ - g^- \in L^1(\mu).$$

We can write the last statement as

$$\Lambda(\chi_E) = \int_X \chi_E g\,d\mu\,.$$

Since simple functions are linear combinations of characteristic functions

$$\Lambda f = \int_X fg\,d\mu \tag{58}$$

whenever f is a simple function. Now every $f \in L^\infty(\mu)$ is a uniform limit of a sequence of simple functions and hence (58) holds for all $f \in L^\infty(\mu)$. It remains to prove the following three statements

(a) $g \in L^q(\mu)$ (we know that $g \in L^1(\mu)$);

(b) $\|\Lambda\| = \|g\|_q$ (we know that $\|\Lambda\| \leq \|g\|_q$ by Hölder's inequality);

(c) $\Lambda f = \int_X fg\,d\mu$ for all $f \in L^p(\mu)$ (we know it for $f \in L^\infty(\mu)$) .

We will complete the proof in the case $1 < p < \infty$. The case $p = 1$ is easier and left to the reader. Let

$$\alpha(x) = \begin{cases} g(x)/|g(x)| & \text{if } g(x) \neq 0, \\ 0 & \text{if } g(x) = 0, \end{cases}$$

and

$$E_n = \{x:\ |g(x)| \leq n\}\,.$$

Define

$$f = |g|^{q/p}\alpha\chi_{E_n}\,.$$

Clearly $f \in L^\infty(\mu)$ (because $|f| \leq n^{q/p}$). Moreover

$$|f|^p \overset{\text{(a)}}{=} |g|^q\chi_{E_n} \overset{\text{(b)}}{=} fg\,.$$

Hence

$$\int_{E_n} |g|^q\,d\mu \overset{\text{by (b)}}{=} \int_X fg\,d\mu \overset{f\in L^\infty}{=} \Lambda f \leq \|\Lambda\|\|f\|_p \overset{\text{by (a)}}{=} \|\Lambda\|\left(\int_{E_n} |g|^q\,d\mu\right)^{1/p}.$$

This yields

$$\left(\int_{E_n} |g|^q\,d\mu\right)^{1-1/p} \leq \|\Lambda\|\,.$$

The left hand side converges to $\|g\|_q$ as $n \to \infty$ and hence $\|g\|_q \le \|\Lambda\|$. Since we already know the opposite inequality $\|\Lambda\| \le \|g\|_q$ we conclude

$$\|\Lambda\| = \|g\|_q\,.$$

This yields the claims (a) and (b). Now (c) follows from the continuity of the functional. Indeed, there is a sequence[18] $L^\infty \ni f_k \to f$ in L^p and hence

$$\begin{aligned}\left|\Lambda f - \int_X fg\,d\mu\right| &\le |\Lambda f - \Lambda f_k| + \left|\Lambda f_k - \int_X fg\,d\mu\right| = |\Lambda f - \Lambda f_k| + \left|\int_X (f_k - f)g\,d\mu\right| \\ &\le |\Lambda f - \Lambda f_k| + \|f_k - f\|_p \|g\|_q \to 0 \quad \text{as } k \to \infty\,.\end{aligned}$$

which proves (c). □

DEFINITION. Let X be a locally compact metric space. The class of *continuous functions vanishing at infinity* $C_0(X)$ consists of all continuous functions $u : X \to \mathbb{R}$ such that for every $\varepsilon > 0$ there is a compact set $K \subset X$ such that $|u(x)| < \varepsilon$ for all $x \in X \setminus K$.

Exercise. *Prove that* $u \in C_0(\mathbb{R}^n)$ *if and only if* u *is continuous on* $\mathbb{R}^n$ *and* $\lim_{|x|\to\infty} u(x) = 0$.

EXAMPLE. The set of positive integers $\mathbb{N}$ is a locally compact metric space with respect to the metric $|x - y|$. Functions on $\mathbb{N}$ can be identified with infinite sequences. Indeed, a sequence $(x_i)_{i=1}^\infty$ corresponds to a function $f(i) = x_i$. Every function on $\mathbb{N}$ is continuous and compact sets in $\mathbb{N}$ are exactly finite subsets. It easily follows that $C_0(\mathbb{N})$ consists of sequences such that $|x_i| \to 0$ as $i \to \infty$. This space is denoted by $c_0 = C_0(\mathbb{N})$. Thus

$$c_0 = \{(x_i)_{i=1}^\infty :\ x_i \in \mathbb{R},\, |x_i| \to 0 \text{ as } i \to \infty\}\,.$$

It is easy to see that c_0 is a Banach space with respect to the norm $\|x\|_\infty = \sup_{i=1,2,\ldots} |x_i|$, where $x = (x_i)_{i=1}^\infty$. Therefore c_0 is a closed subspace in ℓ^∞.

Recall that $C_c(X)$ consists of all continuous functions on X that have compact support. Clearly $C_c(X) \subset C_0(X)$.

Theorem 127 *Let X be a locally compact metric space. Then $C_0(X)$ is a Banach space with the norm*

$$\|u\|_\infty = \sup_{x\in X} |u(x)|\,.$$

Moreover $C_c(X)$ is a dense subset of $C_0(X)$.

We leave the proof as an exercise.

Observe that if μ is a measure on X, then $C_0(X)$ is a *closed* subspace of $L^\infty(\mu)$. Hence $C_0(X)$ is the closure of $C_c(X)$ in the metric of $L^\infty(\mu)$. If μ is a Radon measure on

[18] Because simple functions are dense in L^p.

a locally compact metric space, then the closure of $C_c(X)$ in the metric of $L^p(\mu)$ equals $L^p(\mu)$ for all $1 \leq p < \infty$ (Theorem 86). However the closure of $C_c(X)$ in the metric of $L^\infty(\mu)$ equals $C_0(X) \subsetneq L^\infty(\mu)$. We managed to characterize bounded linear functional on $L^p(\mu)$ for all $1 \leq p < \infty$ (Theorem 126). It is not so easy to characterize bounded linear functionals on $L^\infty(\mu)$. However Theorem 128 below provides a characterization of bounded linear functionals on $C_0(X)$.

Recall that if μ is a signed measure, then Hahn's decomposition theorem provides a unique representation $\mu = \mu^+ - \mu^-$, where μ^+ and μ^- are positive measures concentrated on disjoint sets. We define the measure $|\mu|$ by $|\mu| = \mu^+ + \mu^-$ and call the number $|\mu|(X)$ the *total variation of* μ.

If a signed measure μ is given by

$$\mu(E) = \int_E f\, d\lambda \quad \text{for some } f \in L^1(\lambda)\,,$$

then

$$|\mu|(E) = \int_E |f|\, d\lambda\,.$$

If μ is a signed measure and $f \in L^1(|\mu|)$, then we define

$$\int_X f d\mu = \int_X f\, d\mu^+ - \int_X f\, d\mu^-\,.$$

Theorem 128 (Riesz Representation Theorem) *Let X be a locally compact metric space and let Λ be a bounded linear functional on $C_0(X)$. Then there is a unique Borel signed measure μ of finite total variation such that*

$$\Lambda f = \int_X f\, d\mu \quad \textit{for all } f \in C_0(X)\,. \tag{59}$$

Moreover

$$\|\Lambda\| = |\mu|(X)\,.$$

In the proof we will need the following abstract and general result.

Lemma 129 *Let $X_0 \subset X$ be a dense linear subspace in a normed space X. If $\Lambda : X_0 \to \mathbb{R}$ is a bounded linear functional (with respect to the norm in X), then there is a unique functional $\tilde{\Lambda} : X \to \mathbb{R}$ (called extension of Λ) such that*

$$\tilde{\Lambda}x = \Lambda x \quad \textit{for all } x \in X_0 \quad \textit{and} \quad \|\tilde{\Lambda}\| = \|\Lambda\|\,.$$

Remarks. Here

$$\|\tilde{\Lambda}\| = \sup_{\substack{\|x\|\leq 1\\ x\in X}} |\tilde{\Lambda}x| \quad \text{and} \quad \|\Lambda\| = \sup_{\substack{\|x\|\leq 1\\ x\in X_0}} |\Lambda x|\,.$$

Usually the extension is denoted by the same symbol Λ rather than $\tilde{\Lambda}$.

EXAMPLE. Let X be a locally compact metric space and λ a Radon measure on X. If $\Lambda : C_c(X) \to \mathbb{R}$ is a linear functional such that

$$|\Lambda f| \leq M\|f\|_{L^1(\lambda)} \quad \text{for all } f \in C_c(X)\,,$$

then there is unique bounded linear functional $\tilde{\Lambda} : L^1(\lambda) \to \mathbb{R}$ such that $\tilde{\Lambda} f = \Lambda f$ for $f \in C_c(X)$ and

$$|\tilde{\Lambda} f| \leq M\|f\|_{L^1(\lambda)} \quad \text{for all } f \in L^1(\lambda)\,.$$

Indeed, $C_c(X)$ is a linear and dense subspace of $L^1(\lambda)$ by Theorem 86.

Proof of the lemma. Let $X_0 \ni x_n \to x \in X$. Then we define

$$\tilde{\Lambda} x = \lim_{n\to\infty} \Lambda x_n \tag{60}$$

and we have to prove that:

(a) the limit exists;

(b) the limit does not depend on the particular choice of a sequence $x_n \in X_0$ that converges to x;

(c) $\tilde{\Lambda} x = \Lambda x$ for $x \in X_0$;

(d) $\|\tilde{\Lambda}\| = \|\Lambda\|$.

If $x_n \to x$, $x_n \in X_0$, then

$$|\Lambda x_n - \Lambda x_m| = |\Lambda(x_n - x_m)| \leq \|\Lambda\|\|x_n - x_m\| \to 0 \quad \text{as } n, m \to \infty\,,$$

which shows that Λx_n is a Cauchy sequence of real numbers and hence it is convergent. This proves (a). If $x_n \to x$, $y_n \to x$, $x_n, y_n \in X_0$, then

$$|\Lambda x_n - \Lambda y_n| = |\Lambda(x_n - y_n)| \leq \|\Lambda\|\|x_n - y_n\| \to 0 \quad \text{as } n \to \infty$$

which proves (b). If $x \in X_0$, then taking $x_n = x$ we have $x_n \to x$ and hence $\tilde{\Lambda} x = \lim_{n\to\infty} \Lambda x = \Lambda x$ which proves (c). Clearly

$$\|\Lambda\| = \sup_{\substack{\|x\|\leq 1\\ x\in X_0}} |\Lambda x| \leq \sup_{\substack{\|x\|\leq 1\\ x\in X}} |\tilde{\Lambda} x| = \|\tilde{\Lambda}\|\,,$$

but we actually have equality because if $x \in X$, then there is a sequence $x_n \in X_0$, $x_n \to x$ and hence

$$|\tilde{\Lambda} x| = \lim_{n\to\infty} |\Lambda x_n| \leq \|\Lambda\| \lim_{n\to\infty} \|x_n\| = \|\Lambda\|\|x\|\,.$$

which proves that $\|\tilde{\Lambda}\| \leq \|\Lambda\|$ (see Lemma 125). The above two inequalities prove (d).
□

Remark. In the proof we used the fact that every Cauchy sequence in $\mathbb{R}$ is convergent, i.e. we used the fact that $\mathbb{R}$ is Banach space. Therefore the same proof gives a more general result about extensions of bounded linear mappings $L : X_0 \to Y$ defined on a dense linear subspace $X_0 \subset X$ of a noremd space X into a Banach space Y. The reader should find an appropriate statement.

Proof of Theorem 128. First we prove uniqueness of the measure μ. To this end we need to prove that if μ is a signed Borel measure of finite total variation such that $\int_X f\, d\mu = 0$ for all $f \in C_0(X)$, then $\mu \equiv 0$. Let $\mu = \mu^+ - \mu^-$ be the Hahn decomposition with measures μ^+ and μ^- concentrated on disjoint sets X^+ and X^- respectively. If

$$h(x) = \begin{cases} 1 & \text{if } x \in X^+, \\ -1 & \text{if } x \in X^-, \end{cases}$$

then $h^2 = 1$ and $d\mu = hd|\mu|$. Hence

$$|\mu|(X) = \int_X d|\mu| = \int_X h^2\, d|\mu|\,. \tag{61}$$

Let $f_n \in C_c(X)$ be a sequence that converges to h in $L^1(|\mu|)$ (Theorem 86). Note that

$$0 = \int_X f_n\, d\mu = \int_X f_n h\, d|\mu|\,. \tag{62}$$

Now (61) and (62) yield

$$|\mu|(X) = \int_X (h^2 - f_n h)\, d|\mu| = \int_X (h - f_n) h\, d|\mu| \leq \int_X |h - f_n|\, d|\mu| \to 0 \quad \text{as } n \to \infty\,.$$

Hence $|\mu|(X) = 0$ and thus $\mu = 0$.

Now let Λ be a bounded linear functional on $C_0(X)$. We can assume without lost of generality that[19] $\|\Lambda\| = 1$. We would like to apply the Riesz representation theorem for positive functionals on $C_c(X)$, but unfortunately the functional Λ is not necessarily positive. So we want to construct a positive functional Φ such that

$$|\Lambda f| \leq \Phi(|f|) \leq \|f\|_\infty \quad \text{for all } f \in C_c(X)\,. \tag{63}$$

Once the functional Φ is constructed we can finish the proof of the theorem as follows. By the Riesz representation theorem there is a (positive) Radon measure λ such that

$$\Phi(f) = \int_X f\, d\lambda \quad \text{for all } f \in C_c(X)\ .$$

[19] Otherwise we multiply Λ by $\|\Lambda\|^{-1}$.

Observe that

$$\lambda(X) = \sup\{\Phi(f) :\ 0 \le f \le 1,\ f \in C_c(X)\}\,. \tag{64}$$

Indeed, for $0 \le f \le 1$, $f \in C_c(X)$ we have $\int_X f\,d\lambda \le \lambda(X)$. On the other hand for every compact set $K \subset X$ there is $0 \le f \le 1$, $f \in C_c(X)$ such that $f = 1$ on K and hence $\int_X f\,d\lambda \ge \lambda(K)$. Now (64) follows from the fact that

$$\lambda(X) = \sup_{\substack{K \subset X \\ K-\text{compact}}} \lambda(K)\,.$$

Observe that (64) implies

$$\lambda(X) \le 1. \tag{65}$$

It follows from (63) that

$$|\Lambda(f)| \le \Phi(|f|) = \int_X |f|\,d\lambda = \|f\|_{L^1(\lambda)}\,.$$

Thus Λ is a bounded linear functional defined on a linear subspace[20] $C_c(X) \subset L^1(\lambda)$. Hence the functional can be uniquely extended to a bounded linear functional $\tilde{\Lambda}$ on $L^1(\lambda)$ such that $\tilde{\Lambda} f = \Lambda f$ for $f \in C_c(X)$ and

$$|\tilde{\Lambda} f| \le \|f\|_{L^1(\lambda)} \quad \text{for all } f \in L^1(\lambda)\,.$$

According to Theorem 126 there is a unique function $g \in L^\infty(\lambda)$, $\|g\|_\infty \le 1$ such that

$$\tilde{\Lambda} f = \int_X fg\,d\lambda \quad \text{for all } f \in L^1(\lambda).$$

If $f \in C_0(X)$ and $f_n \in C_c(X)$, $f_n \to f$ in the norm of $C_0(X)$, then

$$\tilde{\Lambda} f = \lim_{n\to\infty} \int_X f_n g\,d\lambda = \Lambda f_n \to \Lambda f$$

and hence $\tilde{\Lambda} f = \Lambda f$ for all $f \in C_0(X)$. Thus

$$\Lambda f = \int_X fg\,d\lambda \quad \text{for all } f \in C_0(X)\,.$$

Accordingly, we have the representation (59) with $d\mu = gd\lambda$. Since $\|\Lambda\| = 1$ and $d|\mu| = |g|\,d\lambda$ it follows that

$$|\mu|(X) = \int_X |g|\,d\lambda \ge \sup\{|\Lambda f| :\ f \in C_0(X), \|f\|_\infty \le 1\} = \|\Lambda\| = 1\,.$$

On the other hand $\|g\|_\infty \le 1$ and $\lambda(X) \le 1$ so

$$|\mu|(X) = \int_X |g|\,d\lambda \le 1\,.$$

[20]Bounded with respect to the norm of $L^1(\lambda)$.

Thus

$$|\mu|(X) = 1 = \|\Lambda\|$$

and the proof is complete. Therefore we are left with the construction of the positive functional Φ satisfying (63).

Denote by $C_c^+(X)$ the class of nonnegative functions in $C_c(X)$ and for $f \in C_c^+(X)$ define

$$\Phi(f) = \sup\{|\Lambda h| : \ h \in C_c(X), \ |h| \leq f\}.$$

If $f \in C_c(X)$, then $f = f^+ - f^-$, $f^+, f^- \in C_c^+(X)$ and we define

$$\Phi(f) = \Phi(f^+) - \Phi(f^-).$$

Clearly $\Phi(f) \geq 0$ for $f \in C_c^+(X)$, so Φ is positive and Φ satisfies (63). Therefore we only need to show that Φ is linear. Obviously $\Phi(cf) = c\Phi(f)$ and hence we are left with the proof that

$$\Phi(f+g) = \Phi(f) + \Phi(g) \quad \text{for all } f, g \in C_c(X). \tag{66}$$

Observe that it suffices to verify (66) for $f, g \in C_c^+(X)$. Indeed, for $f, g \in C_c(X)$ we will have then[21]

$$(f+g)^+ - (f+g)^- = f^+ - f^- + g^+ - g^-$$

and hence

$$(f+g)^+ + f^- + g^- = (f+g)^- + f^+ + g^+.$$

Now, the assumed linearity of Φ on $C_c(X)^+$ yields

$$\Phi((f+g)^+) + \Phi(f^-) + \Phi(g^-) = \Phi((f+g)^-) + \Phi(f^+) + \Phi(g^+)$$

$$\underbrace{\Phi((f+g)^+) - \Phi((f+g)^-)}_{\Phi(f+g)} = \underbrace{\Phi(f^+) - \Phi(f^-)}_{\Phi(f)} + \underbrace{\Phi(g^+) - \Phi(g^-)}_{\Phi(g)}.$$

Thus we are left with the proof of (66) for $f, g \in C_c(X)^+$. Fix $f, g \in C_c^+(X)$. Given $\varepsilon > 0$, there exist $h_1, h_2 \in C_c(X)$ such that $|h_1| \leq f$, $|h_2| \leq g$ and

$$\Phi(f) \leq |\Lambda h_1| + \varepsilon/2, \quad \Phi(g) \leq |\Lambda h_2| + \varepsilon/2.$$

Let $c_1 = \pm 1$, $c_2 = \pm 1$ be such that $|\Lambda h_1| = c_1 \Lambda h_1$, $|\Lambda h_2| = c_2 \Lambda h_2$. We have

$$\begin{aligned} \Phi(f) + \Phi(g) &\leq |\Lambda h_1| + |\Lambda h_2| + \varepsilon = \Lambda(c_1 h_1 + c_2 h_2) + \varepsilon \\ &\leq \Phi(|h_1| + |h_2|) + \varepsilon \leq \Phi(f+g) + \varepsilon \end{aligned}$$

and hence

$$\Phi(f) + \Phi(g) \leq \Phi(f+g).$$

[21] Because both sides are equal $f + g$.

Now we need to prove the opposite inequality. Let $h \in C_c(X)$ satisfy $|h| \le f+g$, let $V = \{x:\ f(x)+g(x) > 0\}$, and define

$$h_1(x) = \begin{cases} \frac{f(x)h(x)}{f(x)+g(x)} & \text{if } x \in V, \\ 0 & \text{if } x \notin V, \end{cases} \qquad h_2(x) = \begin{cases} \frac{g(x)h(x)}{f(x)+g(x)} & \text{if } x \in V, \\ 0 & \text{if } x \notin V. \end{cases}$$

It is easy to see that the functions h_1 and h_2 are continuous, $h = h_1 + h_2$, and $|h_1| \le f$, $|h_2| \le g$. We have

$$|\Lambda h| = |\Lambda h_1 + \Lambda h_2| \le |\Lambda h_1| + |\Lambda h_2| \le \Phi(f) + \Phi(g)$$

and after taking the supremum over all functions h we have

$$\Phi(f+g) \le \Phi(f) + \Phi(g)$$

which completes the proof. □

10 Maximal function

DEFINITION. For a function $u \in L^1_{\rm loc}(\mathbb{R}^n)$ we define the *Hardy-Littlewood maximal function* by

$$Mu(x) = \sup_{r>0} \fint_{B(x,r)} |u(y)|\, dy\, .$$

Clearly $Mu : \mathbb{R}^n \to [0,\infty]$ is a measurable function.

Theorem 130 (Mardy-Littlewood maximal theorem) *If $u \in L^p(\mathbb{R}^n)$, $1 \le p \le \infty$, then $Mu < \infty$ a.e. Moreover*

(a) If $u \in L^1(\mathbb{R}^n)$, then for every $t > 0$

$$|\{x:\ Mu(x) > t\}| \le \frac{C(n)}{t} \int_{\mathbb{R}^n} |u|\, . \tag{67}$$

(b) If $u \in L^p(\mathbb{R}^n)$, $1 < p \le \infty$, then $Mu \in L^p(\mathbb{R}^n)$ and

$$\|Mu\|_p \le C(n,p)\|u\|_p\, . \tag{68}$$

Remarks. Let $u = \chi_{[0,1]}$. Then for $x \ge 1$

$$Mu(x) \ge \frac{1}{2x}\int_0^{2x} \chi_{[0,1]}(y)\, dy = \frac{1}{2x} \notin L^1\, .$$

Hence the inequality (68) does not hold for $p = 1$. If $g \in L^1(\mathbb{R}^n)$, then

$$t|\{x:\ |g(x) > t\}| = \int_{\mathbb{R}^n}$$