Lectures on Real Analysis

Mikhail Tyaglov

14.06.2018

Contents

Chapter 1

Introduction to the set theory

This chapter is devoted to the set theory in Cantor's sense.

Notations

Let A and B be two sets.

- the set $A\bigcup B$ called the *union* of the sets A and B, is the set of all elements of A that belong to A or to B or to both of them.
- the set $A\bigcap B$ called the *intersection* of the sets A and B, is the set of all elements of A that belong to A and to B.
- the set $A\setminus B$ called the *difference* of the sets A and B, is the set of all elements of A that do not belong to B.
- the set $A \vartriangle B = (A\bigcup B)\setminus(A\bigcap B)$ is called the *symmetric difference* of the sets A and B.

1.1 Open, closed and compact sets

The **open ball** in $\mathbb{R}^n$ centered at x and of radius r is defined by

$$B_r(x) = \{y \in \mathbb{R}^n : |y-x| < r\}.$$

A subset $C \subset \mathbb{R}^n$ is **open** if for every $x \in C$ there exists $r > 0$ with $B_r(x) \subset C$. By definition, a set is **closed** if its complement is open.

We note that any (not necessarily countable) union of open sets is open, while in general the intersection of only finitely many open sets is open. A similar statement holds for the class of closed sets if one interchanges the roles of unions and intersections.

A set C is **bounded** if it is contained in some ball of finite radius. A bounded set is **compact** if it is also closed. Compact sets enjoy the Heine-Borel covering property:

- Assume C is compact, $C \subset \bigcup\limits_k A_k$, and each A_k is open. Then there are finitely many of the open sets, $A_{k_1}, A_{k_2}, \ldots, A_{k_m}$, such that $C \subset \bigcup\limits_{j=1}^{m} A_{k_j}$.

In words, *any* covering of a compact set by a collection of open sets contains a finite subcovering.

A point $x \in \mathbb{R}^n$ is a **limit point** of the set C if for every $r > 0$, the ball $B_r(x)$ contains points of C. This means that there are points in C which are arbitrarily close to x. An **isolated point** of C is a point $x \in C$ such that there exists an $r > 0$ where $B_r(x) \cap C$ is equal to $\{x\}$.

A point $x \in C$ is an **interior point** of C if there exists $r > 0$ such that $B_r(x) \subset C$. The set of all interior points of C is called the **interior** of C. Also, the **closure** $\overline{C}$ of the C consists of the union of C and all its limit points. The **boundary** of a set C, denoted by ∂C, is the set of points which are in the closure of C but not in the interior of C.

Note that the closure of a set is a closed set; every point in $\overline{C}$ is a limit point of C; and a set is closed if and only if it contains all its limit points. Finally, a closed set C is **perfect** if C does not have any isolated points.

The symbol $\varnothing$ denotes the empty set, the set with no elements at all. It is supposed to be open and closed simultaneously.

1.2 Comparison of sets

We start this section the following question. Given two sets, it is asked which one of them contains more elements than the other one. It is easy to answer this question if the sets contain only finitely (and rather few) many elements, or if one of the sets is finite, while the other one is infinite. However, it is not easy to answer the question if both sets are infinite, or contain quite a large number of elements that it is impossible to count them. For example, can we say what sets contains more elements: the set of natural numbers or even numbers; the set of rational numbers or integers; the set of irrational numbers or rational? As we show below, intuitive answers to these seemingly evident questions might be wrong. To find the correct answers to these questions we have to introduce the following definition.

Definition 1.2.1. Two sets A and B are called *equivalent*, and we write $A \sim B$, if there exists a one-to-one correspondence between their elements, that is, there exists a bijection $f : A \leftrightarrow B$ of the set A on the set B.

It is easy to see that this definition defines an equivalens relation, because $A \sim A$, $A \sim B$ implies $B \sim A$, and $A \sim B$, $B \sim C$ imply $A \sim C$.

Definition 1.2.2. If the sets sets A and B are equivalent, then they are said to be of equal *cardinalities* or *cardinal numbers*: $CardA = CardB$.

For finite sets, the equality of their cardinality numbers means the equality of the number of their elements. Analogously, we will say that if two infinite sets are equivalent, then they contain equal number of elements.

Example 1.2.3. Consider the set of all natural numbers $\mathbb{N}$ and the set of all even natural numbers $2\mathbb{N} = \{2n : n \in \mathbb{N}\}$. The mapping $f : \mathbb{N} \leftrightarrow 2\mathbb{N}$ that corresponds the number $2n$ to each natural number n is obviously a bijection. Therefore, the sets $\mathbb{N}$ and $2\mathbb{N}$ are equivalent.

Definition 1.2.4. The cardinal number of a set A is said to be less than or equal to the cardinal number of a set B:

$$CardA \leqslant CardB,$$

if the set A is equivalent to some subset B_0 of the set B (B_0 can coincide with B).

Definition 1.2.5. The cardinal number of a set A is said to be less than the cardinal number of a set B:

$$CardA < CardB,$$

if the set A is equivalent to some subset B_0 of the set B_0 but is not equivalent to B.

1.3 Countable sets

Definition 1.3.1. The cardinal number of the set of all natural numbers $\mathbb{N}$ is called the *countable cardinality* and will be denoted by the letter a, $Card\ \mathbb{N} = a$.

Definition 1.3.2. The sets equivalent to $\mathbb{N}$ are called *countable*.

Thus, if A is a countable set, $Card\ A = a$, then each element of A can be one-to-one corresponded to a natural number n. In other word, the elements of a countable set can be enumerated, that is, be represented as a sequence. So, if A is a countable set, then it can be written as follows

$$A = \{a_1, a_2, \ldots, a_n, \ldots\}.$$

The countable cardinality is the least cardinality among the ones of all infinite sets. This follows from the following theorem.

Theorem 1.3.3. *Any infinite set contains a countable subset.*

Proof. Let us choose two distinct elements of the set A and denote them a_1 and b_1. The set $A \setminus \{a_1, b_1\}$ is obviously infinite, so we can choose two distinct elements a_2 and b_2 of this set. We continue in this manner to get, at an n-th step the set

$$A \setminus \{\{a_1, a_2, \ldots, a_n\} \bigcup \{b_1, b_2, \ldots, b_n\}\}.$$

This set is infinite, so we can choose two distinct elements a_{n+1} and b_{n+1} of this set and so on. So we continue this procedure infinitely many times, and obtain two sequences

$$a_1, a_2, \ldots, a_n, \ldots; \quad b_1, b_2, \ldots, b_n, \ldots$$

of *distinct* elements of the set A. Let us introduce the notations

$$A_1 = \{a_1, a_2, \ldots, a_n, \ldots\}, \quad B_1 = \{b_1, b_2, \ldots, b_n, \ldots\}.$$

The set $A_1 \subset A$ is countable, as required. □

Remark 1.3.4. We proved even more than it was claimed in the assertion of the theorem. Namely, we established that we can always split an infinite set into two part, one of which is countable and the other one is infinite.

Indeed, the set $A \setminus A_1 = B_1 \bigcup (A \setminus (A_1 \bigcup B_1))$ contains the countable set B_1, so it is infinite.

We will use this Remark in what follows.

Theorem 1.3.5. *A union of at most countably many of at most countable sets is at most countable.*

Proof. The assertion of the theorem contains four statements which we proof separately.

1) *A union of finitely many finite sets is finite.*

   This statement is obvious.

2) *A union of countably many finite sets is at most countable set.*

   Let us have the sets

$$A_1 = \{a_{11}, a_{12}, \ldots, a_{1n_1}\},$$

$$A_2 = \{a_{21}, a_{22}, \ldots, a_{2n_2}\},$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$A_k = \{a_{k1}, a_{k2}, \ldots, a_{kn_k}\},$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

   Consider the set $A = \bigcup\limits_{k=1}^{\infty} A_k$, and show that it is at most countable. To do this it suffices to show that we can enumerate the elements of A or, equivalently, order them into a sequence.

Let us order the elements of A as follows: First, we write elements of the set A_1, then we write those elements of A_2 that do not belong to A_1, and so on. At the kth step, we write the elements of the set A_k omitting those ones of them that do belong to the previous sets, and so on. Thus, we finally obtain

$$A = \{a_{11}, a_{12}, \ldots, a_{1n_1}; a_{21}, a_{22}, \ldots, a_{2n_2}; \ldots; a_{k1}, a_{k2}, \ldots, a_{kn_k}; \ldots\}.$$

It is clear that in this manner, each element of each set A_k will be chosen, so we obtain the set A, indeed. It suffices now to enumerate the elements of the set A. If we use only finitely many numbers during enumeration, then the set A is finite (it might happen if, from some point, all elements of the rest sets are already written). If we use the whole set $\mathbb{N}$ during the enumeration, then A is countable.

3) *A union of finitely many countable sets is a countable set.*

Let us have the sets

$$A_1 = \{a_{11}, a_{12}, \ldots, a_{1n_1}, \ldots\},$$

$$A_2 = \{a_{21}, a_{22}, \ldots, a_{2n_2}, \ldots\},$$

....

$$A_k = \{a_{k1}, a_{k2}, \ldots, a_{kn_k}, \ldots\}.$$

Consider the set $A = \bigcup\limits_{j=1}^{k} A_j$, and show that it is countable. To do this, let us order the elements of A as follows: First, we write all the first elements of all the sets A_k. Then we write the second elements of A_k omitting those ones that already appeared in the list, and so on. Finally, we get

$$A = \{a_{11}, a_{21}, \ldots, a_{k1}; a_{12}, a_{22}, \ldots, a_{k2}; \ldots; a_{1n}, a_{2n}, \ldots, a_{kn}; \ldots\}.$$

It is clear that in such a manner, we mention all the elements of all sets. Moreover, the elements of A can be enumerated by natural numbers, so A is countable.

4) *A union of countably many countable sets is a countable set.*

Let us have the sets

$$A_1 = \{a_{11}, a_{12}, a_{13}, \ldots, a_{1n_1}, \ldots\},$$

$$A_2 = \{a_{21}, a_{22}, a_{23}, \ldots, a_{2n_2}, \ldots\},$$

$$A_3 = \{a_{31}, a_{32}, a_{33}, \ldots, a_{3n_2}, \ldots\},$$

....

$$A_k = \{a_{k1}, a_{k2}, a_{k3}, \ldots, a_{kn_k}, \ldots\}.$$

....

Consider the set $A = \bigcup\limits_{k=1}^{\infty} A_k$, and show that it is countable. In this case, we order the elements of A_k "along diagonals" omitting those ones that already appeared in the list:

$$A = \{a_{11}, a_{12}, a_{21}, a_{13}, a_{22}, a_{31}; \ldots; a_{1n}, a_{2,n-1}, a_{3,n-2}, \ldots, a_{n1}; \ldots\}.$$

Again, in such a manner every element of every sets will finally appear in the list. During enumeration of the elements of the set A, we will obviously use all the natural numbers, so A is countable.

Note that in the cases 3) and 4) it is not necessary to have *all* the sets A_k countable. It is enough if *some* of them are countable. $\square$

If we unite finitely many disjoint finite sets, their cardinalities are summed. Analogously, for infinite sets we can put, by Theorem 1.3.5,

$$n_1 + n_2 + \cdots + n_k + \cdots = a,$$

$$a + a + \cdots + a = a,$$

$$a + a + \cdots + a + \cdots = a.$$

Example 1.3.6. *Card* $\mathbb{Z} = a$, *Card* $\mathbb{Q} = a$, where $\mathbb{Z}$ is the set of all integers, and $\mathbb{Q}$ is the set of rational numbers.

Indeed, the set $\mathbb{Z}$ can be represented as

$$\mathbb{Z} = \mathbb{N} \bigcup \{0\} \bigcup (-\mathbb{N}),$$

where $-\mathbb{N} = \{-1, -2, -3, \ldots, -n, \ldots\}$ is a countable set. By Theorem 1.3.5, $\mathbb{Z}$ is a countable set.

Furthermore, the set of rational numbers $\mathbb{Q} = \{m/n : m \in \mathbb{Z}, n \in \mathbb{N}\}$ can be represented as

$$\mathbb{Q} = \bigcup_{n=1}^{\infty} \mathbb{Q}_n, \quad \text{where} \quad \mathbb{Q}_n = \{m/n : m \in \mathbb{Z}\},$$

so $\mathbb{Q}$ is countable as a countable union of countable sets.

Theorem 1.3.7. *It is possible to subtract a countable set from any infinite set such that the rest set is equivalent to the initial set.*

Proof. Let A be an infinite set. Select from A two countable sets A_1 and B_1 as it was done in the proof of Theorem 1.3.3 (see Remark 1.3.4). Then, we have

$$A = (A_1 \cup B_1) \cup (A \setminus (A_1 \cup B_1)),$$

$$A \setminus A_1 = B_1 \cup (A \setminus (A_1 \cup B_1)).$$

The sets $A_1 \cup B_1$ and B_1 are countable, so equivalent. Moreover, the set $A \setminus (A_1 \cup B_1)$ and $A_1 \cup B_1$ do not have common elements. Therefore, A and $A \setminus A_1$ are equivalent. □

Theorem 1.3.8. *The Cartesian product of finitely many countable sets is a countable set.*

Proof. Clearly, it is sufficient to prove the theorem for the Cartesian product of two countable sets.

Given two countable sets

$$A = \{a_1, a_2, \ldots, a_n, \ldots\}, \quad B = \{b_1, b_2, \ldots, b_n, \ldots\},$$

let us consider their Cartesian product

$$A \times B = \{(a_n, b_n) : \ a_n \in A, b_n \in B\}.$$

The sets $C_m = \{(a_n, b_m) : n \in \mathbb{N}\}$ are countable because there exists a bijection $f : (a_n, b_m) \leftrightarrow a_n$. Since

$$A \times B = \bigcup_{m=1}^{\infty} C_m,$$

the set $A \times B$ is countable by Theorem 1.3.5. □

Example 1.3.9. *Card* $\mathbb{Q}^m = a$. That is, the set of all m-dimensional vectors with rational coordinates is countable, since

$$\mathbb{Q}^m = \underbrace{\mathbb{Q} \times \mathbb{Q} \times \cdots \times \mathbb{Q}}_{m}.$$

Example 1.3.10. *The set of all polynomials with rational coefficients is countable.*

Let

$$\mathbb{Q}[x] = \{p_m(x) = r_0x^m + r_1x^{m-1} + \cdots + r_m : \ r_0, r_1, \ldots, r_m \in \mathbb{Q}, m \in \mathbb{N}\}$$

be the set of all polynomials with rational coefficients. Each polynomial $p_m(x) \in \mathbb{Q}[x]$ is uniquely defined by its coefficients $(r_0, r_1, \ldots, r_m)$, so if $\mathbb{Q}_m[x]$ is the set of all polynomials of degree m with rational coefficients, then $\mathbb{Q}_m[x] \sim \mathbb{Q}^{m+1}$, $\mathbb{Q}_m[x]$ is countable. Therefore, $\mathbb{Q}[x] = \bigcup\limits_{m=1}^{\infty} \mathbb{Q}_m[x]$ is countable as a countable union of countable sets.

Definition 1.3.11. A real number r is called *algebraic* if it is a root of a polynomial with rational coefficients.

Example 1.3.12. *The set of all algebraic numbers is countable.*

Indeed, any polynomial has finitely many roots. The set $\mathbb{Q}[x]$ is countable. Therefore, the set of all algebraic numbers is countable as a countable union of finite sets, according to Theorem 1.3.5.

Definition 1.3.13. A real number r which is not algebraic is called *transcendental*.

Do such numbers exist? To answer this question, we should answer the question about existence of infinite sets that are not countable, and define the cardinality of $\mathbb{R}$. We do this in the next Section.

1.4 Sets of cardinality continuum

To continue study of cardinalities of sets, let us remind the following classical fact from mathematical analysis.

Theorem 1.4.1 (Cantor)**.** *Any decreasing nested sequence of intervals has a unique common point.*

Recall that a sequence of closed intervals of the real line $\{[a_n, b_n]\}$, $n \in \mathbb{N}$, is called decreasing nested if

$$[a_1, b_1] \supset [a_2, b_2] \supset [a_3, b_3] \supset \ldots \supset [a_n, b_n] \supset \ldots,$$

and

$$\lim_{n\to\infty} (b_n - a_n) = 0.$$

We asked at the end Section 1.3 whether exist infinite sets that are not countable. The next theorem gives an affirmative answer to this question.

Theorem 1.4.2 (Cantor)**.** *The set of all points of the interval* $[0, 1]$ *is not countable.*

Proof. It is clear that $Card\ [0, 1] \geqslant a$, since the interval $[0, 1]$ contains the countable subset $\{1/n : \ n \in \mathbb{N}\}$ (in fact, we can also use Theorem 1.3.3). Suppose, by contradiction, that $Card\ [0, 1] = a$. Then we can enumerate all the points of the interval $[0, 1]$, that is, we can represent the interval $[0, 1]$ as follows $[0, 1] = \{x_1, x_2, \ldots, x_n, \ldots\}$. Let us now apply Cantor's procedure. To do this, let us denote the interval $[0, 1]$ as Δ_0 and split it into three equal (by length) subintervals: $[0, 1/3]$, $[1/3, 2/3]$, $[2/3, 1]$. Denote by Δ_1 one of them that does not contain the point x_1, that is, $x_1 \notin \Delta_1$ (If the point x_1 does not belong to two subintervals, we can choose any of them). Then, we split Δ_1 into three equal subintervals and denote by Δ_2 one of them that does not contain the point x_2, and so on. At the n-th step, we split the interval Δ_{n-1} into three equal subintervals and denote by Δ_n one of them that does not contain the point x_n, and so on.

Finally, we get a decreasing nested sequence of intervals

$$\Delta_0 \supset \Delta_1 \supset \Delta_2 \ldots \Delta_n \supset \ldots,$$

whose lengths tend to zero. By Cantor's theorem 1.4.1, there exists a unique point x_0 which belong to all the intervals Δ_n, $n = 1, 2, \ldots$. Since $x_0 \in [0, 1]$, there exists a number $m \in \mathbb{N}$ such that $x_0 = x_m$. Now, on the one hand, $x_m = x_0 \in \Delta_m$, while on the other hand, $x_m \notin \Delta_m$ by construction, a contradiction. So, the points of interval $[0, 1]$ cannot be enumerated by natural numbers, that is,

$$Card\ [0, 1] > a,$$

as required. □

Definition 1.4.3. The cardinal number of the set of the points of the interval $[0, 1]$ is called the *continuum cardinality* and will be denoted by the letter c, $Card\ [0, 1] = c$.

Theorem 1.4.4. *If $Card\ A \geqslant a$, $Card\ B \leqslant a$, then $Card\ (A \bigcup B) = Card\ A$.*

Proof. Without loss of generality, we can assume that $A \bigcap B = \varnothing$ (otherwise, we change B to $B \setminus A$, and this changing will not affect the set $A \bigcup B$).

Let us select from A a countable sets A_0 as it was done in the proof of Theorem 1.3.3. Then we have

$$A = A_0 \bigcup (A \setminus A_0),$$

$$A \bigcup B = (A_0 \bigcup B) \bigcup (A \setminus A_0).$$

The sets A_0 and $A_0 \bigcup B$ are both countable, and so equivalent. Since the first and the seconds summands in the unions above do not have common elements, the sets A and $A \bigcup B$ are equivalent, so $Card\ (A \bigcup B) = Card\ A$, as required. □

Theorem 1.4.5. *If $Card\ A > a$, $Card\ B \leqslant a$, then $Card\ (A \setminus B) = Card\ A$.*

Proof. The set $A \setminus B$ cannot be finite or countable, since, otherwise, the set $A \subset (A \setminus B) \bigcup B$ was finite or countable.

Obviously, $Card(A \setminus B) \leqslant CardA$, so by Theorem 1.4.4 we have

$$Card\ A \leqslant Card\ (A \setminus B) \bigcup B = Card\ (A \setminus B) \leqslant Card\ A,$$

therefore, $Card\ (A \setminus B) = Card\ A$, as required. □

Example 1.4.6. $Card\ (0, 1) = Card\ (0, 1] = Card\ [0, 1) = c$.

Example 1.4.7. $Card\ [a, b] = Card\ (a, b) = Card\ (a, b] = Card\ [a, b) = c\ (b > a)$.

The mapping $y = a + x(b - a)$ is a bijection between the intervals $[0, 1]$ and $[a, b]$, so $Card\ [a, b] = c$. Thus, any closed, half-closed, or open interval have cardinality continuum.

Example 1.4.8. $Card\ \mathbb{R} = c$, since the mapping $y = \arctan x$ set a bijection between $\mathbb{R}$ and the interval $(-\pi/2, \pi/2)$.

Corollary 1.4.9. *Transcendental numbers exist. The cardinality of the set of transcendental numbers is continuum.*

Proof. Let A be the set of all algebraic numbers, $CardA = a$, and T be the set of all transcendental numbers. Since $T = \mathbb{R} \setminus A$, we have $Card\ T = c$ by Theorem 1.4.5. □

It is clear that a finite union of sets X_k of cardinality continuum is of cardinality continuum whenever $X_k \bigcap X_j = \emptyset$ for $k \neq j$. Indeed, each X_k is equivalent to an interval of the form $[a, b)$ and the interval $[0, 1)$ can be represented as follows:

$$[0, 1) = \bigcup_{k=1}^{n} [c_{k-1}, c_k),$$

where $0 = c_0 < c_1 < c_2 < \ldots < c_{n-1} < c_n = 1$. Thus, $\bigcup\limits_{k=1}^{n} X_k \sim [0, 1)$.

Analogously, a countable union $\bigcup_{k=1}^{n} X_k$ of sets of cardinality continuum is of cardinality continuum if $X_k \bigcap X_j = \emptyset$ for $k \neq j$. In this case we represent the interval $[0, 1)$ as a countable union

$$[0,1) = \bigcup_{k=1}^{\infty} [c_{k-1}, c_k),$$

where $c_k = 1 - \dfrac{1}{k+1}$, $k = 0, 1, 2, \ldots$. Since $X_k \sim [c_{k-1}, c_k)$ for $k \in \mathbb{N}$, we obtain that $\bigcup_{k=1}^{\infty} X_k \sim [0, 1)$.

Thus,

$$c + c + \cdots + c = nc = c,$$

and

$$c + c + \cdots + c + \cdots = ac = c.$$

Theorem 1.4.10. *The cardinality of the set of all sequences whose elements are* 0*s and* 1*s, is continuum.*

Proof. Let A be the considered set:

$$A = \{(a_1, a_2, \ldots, a_n, \ldots) : \ a_n \in \{0; 1\}, \ n \in \mathbb{N}\}.$$

Together with the set A, let us consider the set of binary fractions:

$$F = \{0.\, a_1 a_2 \ldots a_n \ldots : \ a_n \in \{0; 1\}, \ n \in \mathbb{N}\}.$$

Each such a fraction defines a real number

$$x = 0.\, a_1 a_2 \ldots a_n \ldots = \sum_{k=1}^{+\infty} \frac{a_k}{2^k},$$

from the interval $[0, 1]$, since the minimal binary fraction from the set F is

$$0.\, 000 \ldots 0 \ldots = 0,$$

while the largest one is

$$0.\, 111 \ldots 1 \ldots = \sum_{k=1}^{+\infty} \frac{1}{2^k} = \frac{1}{2} \cdot \frac{1}{1 - \frac{1}{2}} = 1.$$

Conversely, every number $x \in [0, 1]$ can be expressed as

$$x = 0.\, a_1 a_2 \ldots a_n \ldots.$$

However, some numbers, namely, the numbers of the form $\frac{m}{2^k}$, and only they, have two representations by binary fractions. One of these representations have 1 in period, while the other one has 0 in period. For example,

$$\frac{3}{8} = 0.\, 011000 \ldots 00 \ldots = 0.\, 010111 \ldots 11 \ldots$$

Taking this fact into account, let us split the set F into two sets F_1 and F_2, so that F_1 consists only of binary fractions with 0 in period but the fraction $0.\, 000 \ldots 0 \ldots = 0$, and F_2 contains all other binary fractions. It is clear that $F_2 \sim [0, 1]$, so $Card\ F_2 = c$. The set F_1 is equivalent to the set of fractions of the form $\frac{m}{2^k}$ where $0 < m < 2^n$, $n \in \mathbb{N}$. This set is an infinite subset of all rational numbers, therefore $Card\ F_1 = a$. By Theorem 1.4.5, one has $Card\ F = Card\ (F_1 \bigcup F_2) = c$.

Since, obviously, $A \sim F$, we obtain $Card\ A = c$. $\square$

Example 1.4.11. *The cardinality of the of all (finite and infinite) sequences of natural numbers is continuum.*

Indeed, let L be the set of all binary fractions of the form

$$0.\, a_1 a_2 a_3 \dots a_n \dots, \quad a_k \in \{0, 1\}, \quad k = 1, 2, 3, \dots.$$

Since L is in one-to-one correspondence with the set of all sequences whose elements are 0s and 1s, we have $Card\ L = c$ by Theorem 1.4.10. Let us assign to every sequence $(n_1, n_2, \dots, n_m, \dots)$ of natural numbers a fraction $0.\, a_1 a_2 a_3 \dots a_n \dots$ according the following rule:

$$a_{n_1} = a_{n_1+n_2} = \dots = a_{n_1+\dots+n_m} = \dots = 1,$$

while all other a_k equal zero. Clearly, we construct a one-to-one correspondence between the set of all sequences (finite or infinite) of natural numbers and the set of the mentioned binary fractions.

Proposition 1.4.12. *The Cartesian product of finitely many sets of cardinality continuum is a set of cardinality continuum.*

Proof. Without loss of generality, it is enough to consider the Cartesian product of two sets. Let $CardA = CardB = c$. Consider $M = A \times B = \{(a, b) \ : \ a \in A, b \in B\}$. For every $m \in M$ we have $m = (a, b)$, and by Example 1.4.11, we have that there is a sequence $(p_1, p_2, \dots)$ of natural numbers corresponding to a, and there is a sequence $(q_1, q_2, \dots)$ of natural numbers corresponding to b. Thus, we can associate with the element m the following sequences of natural numbers:

$$(p_1, q_1, p_2, q_2, \dots)$$

Thus, we obtain a one-to-one correspondence between the set M and the set of all sequences of natural numbers which is of cardinality continuum by Example 1.4.11. □

Corollary 1.4.13. *For any $n \in \mathbb{N}$, $Card\ \mathbb{R}^n = c$.*

Corollary 1.4.14. *The continuum union of pair-wise non-intersecting sets of cardinality continuum is of cardinality continuum.*

Proof. Indeed, there exists a one-to-one corresponding between the considered union and the set of all straight lines in $\mathbb{R}^2$ parallel to the Ox axis. □

Theorem 1.4.15. *The Cartesian product of countably many sets of cardinality continuum is a set of cardinality continuum.*

Proof. Let $M = \bigotimes\limits_{n=1}^{\infty} A_k$, $CardA_k = c$, $k = 1, 2, \dots$. Every element m of the union M has the form $m = (a_1, a_2, a_3, \dots)$ where $a_k \in A_k$, $k \in \mathbb{N}$. By Example 1.4.11, with every a_k we can associate a sequence with natural elements. Thus we have

$$a_1 \sim (n_1, n_2, n_3, \dots),$$

$$a_2 \sim (m_1, m_2, m_3, \dots),$$

$$a_3 \sim (l_1, l_2, l_3, \dots),$$

$$\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots$$

so with the element m we associate the sequence

$$(n_1, n_2, m_1, n_3, m_2, l_1, \dots).$$

This is a necessary one-to-one correspondence between M and the set of all sequences with natural elements. □

The following example plays an important role for our further study of measures.

Example 1.4.16. *Cantor's set.*

Let $\mathcal{C}_0 = [0, 1]$. Split the interval $[0, 1]$ into three equal (by length) parts by the points $\frac{1}{3}$ and $\frac{2}{3}$, and delete the middle *open* interval $\left(\frac{1}{3}, \frac{2}{3}\right)$. The resulting set we denote as $\mathcal{C}_1$, so

$$\mathcal{C}_1 = \left[0, \frac{1}{3}\right] \bigcup \left[\frac{2}{3}, 1\right].$$

Let us now split each of the intervals $\left[0, \dfrac{1}{3}\right]$ and $\left[\dfrac{2}{3}, 1\right]$ split into three part by the numbers $\dfrac{1}{3^2}$, $\dfrac{2}{3^2}$, and $\dfrac{7}{3^2}$, $\dfrac{8}{3^2}$, respectively, and delete the middle intervals. The resulting set consisting of four intervals we denote as $\mathcal{C}_2$:

$$\mathcal{C}_2 = \left[0, \frac{1}{3^2}\right] \bigcup \left[\frac{2}{3^2}, \frac{1}{3}\right] \bigcup \left[\frac{2}{3}, \frac{7}{3^2}\right] \bigcup \left[\frac{8}{3^2}, 1\right].$$

Continuing the procedure, at the nth step we have the set $\mathcal{C}_{n-1}$ consisting of 2^{n-1} intervals. We split each of these intervals into three equal parts and delete the middle intervals. The resulting set consisting of 2^n intervals we denote as $\mathcal{C}_n$, and so on. Finally, we will have the sequence of sets $\mathcal{C}_n$: $\{\mathcal{C}_n\}_{n=0}^{\infty}$. Let us denote their intersection as $\mathcal{C}$:

$$\mathcal{C} = \bigcap_{n=0}^{\infty} \mathcal{C}_n.$$

The set $\mathcal{C}$ is called *Cantor's set.* Let us study some properties of this set.

1) *Card* $\mathcal{C} = c$.

   Indeed, let us represent the numbers of the interval $[0, 1]$ as ternary fractions

   $$x = 0.\, a_1 a_2 a_3 \ldots a_n \ldots,$$

   where a_k can assume only values 0, 1, or 2. Clearly, when we delete the middle interval at the first step of constructing Cantor's set, we exclude from $\mathcal{C}_0$ all the numbers whose first digit in the ternary fraction representation is 1. At the second step, we delete from $\mathcal{C}_1$ all the numbers whose second digit in the ternary fraction representation is 1, and so on. Therefore, Cantor's set consists of all the numbers of the interval $[0, 1]$ whose ternary fraction representations do not have 1, that is, all the digits in their representations are 0s and 2s. Thus, Cantor's set is equivalent to the set of all sequences whose elements are 0s and 1s, and *Card* $\mathcal{C} = c$ by Theorem 1.4.10.

2) *Cantor's set* $\mathcal{C}$ *is closed, so it is compact, since* $\mathcal{C} \subset [0, 1]$.

   Each set $\mathcal{C}_n$ is closed as a union finitely many closed intervals. But the intersection of any number of closed sets is a closed set, so $\mathcal{C}$ is closed.

3) *The "length" of Cantor's set* $\mathcal{C}$ *equals* 0.

   A strict definition of the "length" (measure) of a set is given below in Chapter 2. Now let us calculate the sum of the lengths of all intervals that was deleted during the process of constructing of Cantor's set.

   At the first step we deleted an interval of length $\frac{1}{3}$. At the second step we deleted two intervals of length $\dfrac{1}{3^2}$. At the third step, there were deleted four intervals of length $\frac{1}{3^3}$, and so on. Thus, the sum of the lengths of all deleted open intervals equals

   $$\frac{1}{3} + 2 \cdot \frac{1}{3^2} + 4 \cdot \frac{1}{3^3} + \cdots = \sum_{k=1}^{+\infty} \frac{2^{k-1}}{3^k} = \frac{1}{3} \cdot \frac{1}{1 - \frac{2}{3}} = 1.$$

   Since we deleted from $[0, 1]$ a union of intervals whose length equals 1, then the "length" of the rest set (Cantor's set) equals zero.

The "paradoxical" properties 1) and 3) of Cantor's set were ones of the reasons to develop the so-called measure theory, which we study below in Chapter 2.

We finish this section with the following theorem.

Theorem 1.4.17. *Every open set in $\mathbb{R}$ can represented as a countable union of disjoint open intervals.*

Proof. Let A be an open set in $\mathbb{R}$. For each $x \in A$, let I_x denote the largest open interval containing x and contained in A. More precisely, since A is open, x is contained in some small (non-trivial) interval, and therefore if

$$a_x = \inf\{a < x : (a, x) \subset A\} \qquad \text{and} \qquad b_x = \sup\{b > x : (x, b) \subset A\}$$

we must have $a_x < x < b_x$ (with possibly infinite values for a_x and b_x). If we now let $I_x = (a_x, b_x)$, then by construction we have $x \in I_x$ as well as $I_x \subset A$. Hence

$$A = \bigcup_{x \in A} I_x.$$

Now suppose that two intervals I_x and I_y intersect. Then their union (which is also an open interval) is contained in A and contains x. Since I_x is maximal, we must have $(I_x \cup I_y) \subset I_x$, and similarly $(I_x \cup I_y) \subset I_y$. This can happen only if $I_x = I_y$, therefore, any two distinct intervals in the collection $\mathcal{I} = \{I_x\}_{x \in A}$ must be disjoint. Moreover, the collection $\mathcal{I}$ of open intervals I_x is countable, since every open interval I_x contains a rational number. Since different intervals are disjoint, they must contain distinct rationals, and therefore $\mathcal{I}$ is countable, as required. $\square$

1.5 Sets of higher cardinalities

While the countable cardinality is the minimal among all the cardinal numbers, there is no maximal cardinal number, as the following theorem shows.

Theorem 1.5.1. *The cardinal numbers of the set of all subsets of a given set B is larger than the cardinal number of the set B.*

Proof. Let $B = \bigcup b$ be some set, and let $M = \bigcup m$ be the set of all subsets of B. Here we denote by b the elements of the set B, and by m the elements of the set M. We notice that the empty set $\varnothing$ and the set B itself are elements of the set M. There obviously exists a bijection between the elements b of the set B and the one-element sets $m = \{b\}$ of the set B, therefore, $Card\ M \geqslant Card\ B$. On the contrary, suppose that $Card\ M = Card\ B$. Then, there exists a bijection $f : M \leftrightarrow B$. This means that each element b of the set B can be represented as $b = f(m)$, where the element $m \in M$ is unique.

Let us split the set B into two classes. The first class contains all the elements b of the set B such that $b = f(m) \in m$. The second class contains all the elements b of the set B such that $b = f(m) \notin m$. Since f is a bijection, each element of B cannot belong to both these classes simultaneously, and must belong to one of them.

Consider the set m_0 of the elements of the second class. This class is a subset of B, so $m_0 \in M$, and the bijection f defines an element $b_0 \in B$ such that $b_0 = f(m_0)$. So, does the element b_0 belong to the first class or to the second one? If b_0 belongs to the first class, then $b_0 = f(m_0) \in m_0$, that is impossible, since m_0 is the set of elements of the second class, a contradiction. If b_0 belongs to the second class, then $b_0 = f(m_0) \notin m_0$. But m_0 contains *all* the elements of the second class, a contradiction again.

Thus, b_0 cannot belong neither to the first nor the second class. Therefore, no bijections $f : M \leftrightarrow B$ exist. $\square$

If the set B is finite, $Card\ B = n$, then the set M contains one empty set $(1 = \binom{n}{0})$, $n = \binom{n}{1}$ of one-element sets, $\binom{n}{2}$ of two-element sets,…, and the set B itself $(1 = \binom{n}{n})$. Since

$$1 + \binom{n}{1} + \binom{n}{2} + \cdots + \binom{n}{n-1} + 1 = (1+1)^n = 2^n,$$

then $Card\ M = 2^n$. By analogy, if B is an infinite set and $Card\ B = \alpha$, then it is assumed that

$$Card\ M = 2^{\alpha}.$$

Example 1.5.2. *The following formula is valid:*

$$2^a = c$$

Let T be the set of all subsets of the set $\mathbb{N}$, and let L be the set of all the binary fractions of the form

$$0.\,a_1a_2a_3\ldots a_n\ldots, \quad a_k \in \{0,1\}, \quad k = 1,2,3,\ldots.$$

It is clear that $Card\ T = 2^a$, since $Card\ \mathbb{N} = a$, while $Card\ L = c$ by Theorem 1.4.10.

To each element t of T (t is a finite or countable sequence of growing natural numbers) we can assign an element $x = 0.\,a_1a_2a_3\ldots a_n\ldots$ of L by the the following rule. If a natural number k belongs to the sequence t, then $a_k = 1$. If $k \notin t$, then $a_k = 0$. Clearly, in such a manner we obtain a one-to-one correspondence between T and L, so their cardinalities are equal.

It can be shown that any two cardinal numbers can be compared, that is, if A and B are two sets, then the following relations are exclusive (there exists only one of them):

$$Card\ A < Card\ B, \quad Card\ A = Card\ B, \quad Card\ A > Card\ B. \tag{1.5.1}$$

It can be proved by establishing first that if $A \supset A^* \supset A^{**}$ and $Card\ A = Card\ A^{**}$, then $Card\ A^* = Card\ A$. Then, the fact that the relations (1.5.1) are exclusive follows from the fact that if a set A is equivalent to a subset of a set B, while the set B is equivalent to a subset to the set A, then A and B are equivalent.

Remark 1.5.3. Despite of the well-developed theory of sets comparisons, there is no answer to the question known as ***continuum hypothesis***. The hypothesis (due to G. Cantor) claims that there is no sets whose cardinality is between a and c. However, it was proved in mid-XX century that it is *impossible* to prove or disprove the hypothesis in the system of axioms known as ZFC (Zermelo–Fraenkel set theory). In fact, we implicitly use this system of axioms in our life and, in particular, in these lecture notes. So we can only postulate here the affirm answer to the question as an additional axiom. For further details, see [2].

1.6 Problems

Problem 1.1. Prove explicitly (by finding a bijection) that

1) the sets $[0,1]$ and $(0,1)$ are equivalent;

2) the sets $[0,1]$ and $(a,b]$ are equivalent. Here $a,b \in \mathbb{R}$;

3) the sets $(0,1]$ and $[0,1)$ are equivalent;

4) the sets $[0,1]$ and (a,b) are equivalent, where $a,b \in \mathbb{R}$, $a < b$.

Hint: Find a bijection that transfer (one-to-one) a countable subset of the interval $[0,1]$ containing the point 1 to a countable subset of the the interval $[0,1)$ that does not contain the point 1. Use the composition of bijections.

Problem 1.2. Prove explicitly (by finding a bijection) that

1) the real axis $\mathbb{R}$ is equivalent to the closed interval $[-1,1]$;

2) the real axis $\mathbb{R}$ is equivalent to the open interval $(-a,a)$, $a > 0$;

3) the real axis $\mathbb{R}$ is equivalent to the interval $[-a,a)$, $a > 0$.

Hint: Find a bijection that transfer (one-to-one) a countable subset of the interval $[0,1]$ containing the point 1 to a countable subset of the the interval $[0,1)$ that does not contain the point 1. Use the composition of bijections.

Problem 1.3. Prove that the border of the square with vortices $A(-1,1)$, $B(1,1)$, $C(-1,1)$, and $D(-1,-1)$ is equivalent to the circumference $\{(x,y) : x^2+y^2=1\}$.

Problem 1.4. Prove that the circumference $\{(x,y) : x^2+y^2=1\}$ is equivalent to the ellipse $\{(x,y) : 2x^2+4y^2=16\}$.

Problem 1.5. Prove that the disc $\{(x,y) : x^2+y^2 \leqslant 9\}$ is equivalent to the set $\{(x,y) :\ |x| \leqslant 1,\ |y| \leqslant 1\}$.

Problem 1.6. Prove that the disc $\{(x,y) : x^2+y^2 \leqslant 25\}$ is equivalent to the set $\{(x,y) : 4x^2+8y^2 \leqslant 16\}$.

Problem 1.7. Prove that the set of points of discontinuity of a monotone function on an interval $[a,b]$, is at most countable.

Problem 1.8. Prove that the set of all intervals of $\mathbb{R}$, whose ends are rational, is countable.

Problem 1.9. Prove that the set of all triangles on the plane, whose vertices have rational coordinates, is countable.

Problem 1.10. Prove that the cardinal number of the set $\{(x,y) : x^2+y^2=1\}$ equals c (continuum).

Problem 1.11. Prove that any set of open disjoint discs on the plane is at most countable.

Problem 1.12. Let a set A be finite, and let a set B be countable. Prove that $A\bigcup B$ is equivalent to the set of all integers $\mathbb{Z}$.

Problem 1.13. Let a set A be finite, and let a set B be countable. Prove that $A\bigcup B$ is equivalent to the set of all numbers of the form 2^k, $k \in \mathbb{Z}$.

Problem 1.14. Let a set $A = B\bigcup C$ and $CardA = c$. Prove that at least one of the sets B and C is of cardinality c.

Problem 1.15. Prove that the cardinality the set of all (at most countable) sequences $\{u_k\}_{k=1}^{+\infty}$ with real elements ($u_k \in \mathbb{R}$) is c.
Hint: Use Theorem 1.4.15.

Problem 1.16. Prove that the cardinality of the set Φ of all *continuous* functions defined on $[0,1]$ is equal to c.
Hint: Prove that Card$\Phi \geqslant c$ by finding a subset of Φ of cardinality c. To prove Card$\Phi \leqslant c$, use the continuity of the functions in Φ, the fact that the set of rational numbers is countable, and the result of Problem 1.15.

Problem 1.17. Prove that the cardinality of the set $\mathcal{F}$ of all real functions defined on $[0,1]$ is greater than c. There is the notation Card $\mathcal{F} = f$.
Hint: Prove that $[0,1] \not\sim F$ and $\exists\, F^* \subset F$ such that $[0,1] \sim F^*$.

Problem 1.18. Prove that $2^c = f$.
Hint: First prove the fact that the cardinality of the strip $[0,1] \times \mathbb{R}$ is continuum.

Problem 1.19. Prove that the Cantor set $\mathcal{C}$ is totally disconnected and perfect. In other words, given two distinct points $x, y \in \mathcal{C}$, there is a point $z \notin \mathcal{C}$ that lies between x and y, and yet $\mathcal{C}$ has no isolated points.
Hint: If $x, y \in \mathcal{C}$ and $|x - y| > \frac{1}{3^k}$, then x and y belong to two different intervals in $\mathcal{C}_k$. Also, given any $x \in \mathcal{C}$ there is an end-point y_k of some interval in $\mathcal{C}_k$ that satisfies $x \neq y_k$ and $|x - y_k| \leqslant \frac{1}{3^k}$.

Problem 1.20. Prove that the interval $[0,1]$ cannot be represented as a countable union of disjoint *closed* intervals (where at least **two** of intervals are non-empty).
Hint: Use the Cantor's theorem about the sequence of nested intervals to get a contradiction.
Do not use Baire categories!!!

Problem 1.21. Prove that the interval $[0,1]$ cannot be represented as a countable union of disjoint *closed* sets (where at least **two** of sets are non-empty).

Hint: Use the fact that any open set on $[0,1]$ can be represented as a countable union of open sets.
Do not use Baire categories!!!

Chapter 2

Introduction to the measure theory

In this chapter, we give a short review of rings and algebras of sets and introduce the general measure theory.

2.1 Rings and algebras of sets

In what follows we consider collections of sets whose elements are also sets. We mainly assume that all the considered sets are subsets of some basic set (a *universe*).

Definition 2.1.1. A non-empty collection of sets $\mathcal{K}$ is called a *ring* if it possesses the following properties:

1) $A, B \in \mathcal{K} \implies A \bigcup B \in \mathcal{K}$;

2) $A, B \in \mathcal{K} \implies A \setminus B \in \mathcal{K}$.

Rings also possess the properties:

3) $\varnothing \in \mathcal{K}$;

4) $A, B \in \mathcal{K} \implies A \bigcap B \in \mathcal{K}$;

5) $A, B \in \mathcal{K} \implies A \vartriangle B \in \mathcal{K}$;

These properties follow from the properties 1) and 2) and from the equalities:

$$\varnothing = A \setminus A \quad (A \in \mathcal{K}),$$

$$A \bigcap B = A \setminus (A \setminus B),$$

$$A \vartriangle B = (A \setminus B) \bigcup (B \setminus A),$$

that can be easily checked.

Thus, a ring is a non-empty collection of sets containing the empty set and closed w.r.t. operations of union, intersection, substraction, and symmetric substraction.

Definition 2.1.2. A set $E \in \mathcal{K}$ is called the *unit* of the ring $\mathcal{K}$ if $A \bigcap E = A$ for any $A \in \mathcal{K}$.

Definition 2.1.3. A ring with unit is called *algebra*.

We denote algebras by the letter $\mathcal{A}$. Any algebra possesses one more property (additionally to the properties 1) – 5)).

6) $A \in \mathcal{A} \implies A^c := E \setminus A \in \mathcal{A}$.

The set A^c is called the *complement* of the set A in the algebra $\mathcal{A}$.

Example 2.1.4. Let X be a non-empty set. Then the collection $\mathcal{A} = \{\varnothing, X\}$ is an algebra with the unit X.

Example 2.1.5. Let X be a non-empty set. Then the collection of all its subsets $\mathcal{P}(X)$ is an algebra with the unit X.

Example 2.1.6. Let X be a non-empty set, and let $\mathcal{P}_f(X)$ be the collection of all finite subsets of X. Then $\mathcal{P}_f(X)$ is a ring. The ring $\mathcal{P}_f(X)$ is an algebra with the unit X if, and only if, X is a finite set. However, in this case, $\mathcal{P}_f(X)$ coincides with $\mathcal{P}(X)$.

Example 2.1.7. Let $\mathcal{P}_b(\mathbb{R})$ be the collection of all bounded subsets of $\mathbb{R}$. The set $\mathcal{P}_b(\mathbb{R})$ is a ring (without units).

Our next example is not as trivial as the previous ones, and requires some preparations. Consider the n-dimensional Euclidean space $\mathbb{R}^n$.

Definition 2.1.8. The set[1]

$$K = \{x = (x_i)_{i=1}^n : \ a_i < x_i < b_i \vee a_i \leqslant x_i < b_i \vee a_i < x_i \leqslant b_i \vee a_i \leqslant x_i \leqslant b_i, \ i = 1, 2, \ldots, n\}. \tag{2.1.1}$$

is called a *brick* or n-dimensional rectangle (parallelepiped).

In (2.1.1) we assume that $a_i \leqslant b_i$ for any $i = 1, 2, \ldots, n$, and that for different i there might hold different inequalities (one of the four mentioned inequalities). Thus, $\varnothing$ is a brick ($a_i = b_i$ for at least one index i, and $a_i < x_i < b_i$ for all i); the points are bricks ($a_i = b_i$ and $a_i \leqslant x_i \leqslant b_i$ for all i); finite intervals are bricks (for one fixed index i, $a_i < b_i$, and for all the rest indices i, $a_i = b_i$); parallelepipeds of any dimension k, $2 \leqslant k \leqslant n$, are bricks, etc.

Definition 2.1.9. A set $B \in \mathbb{R}^n$ is called *elementary* if it can be represented as a union of finitely many bricks:

$$B = \bigcup_{j=1}^{l} K_j. \tag{2.1.2}$$

The collection of all elementary sets of the space $\mathbb{R}^n$ is denoted $\mathcal{E}^n$.

Note that the representation of an elementary set in the form (2.1.2) is not unique, since the bricks K_j might have non-empty intersections. However, it is possible to find the representation of an elementary sets in the form of union of finitely many disjoint bricks. To do this, we can take any representation (2.1.2) of an elementary set and cut each brick in the representation by hyperplanes along all the sides of all the bricks in the representation. We finally obtain the representation

$$B = \bigsqcup_{i=1}^{m} K_i', \tag{2.1.3}$$

where the symbol $\bigsqcup$ stands for the union of pairwise disjoint sets (here $K_i' \bigcap K_j' = \varnothing$, $i \neq j$).

Theorem 2.1.10. *$\mathcal{E}^n$ is a ring.*

Proof. It is clear that the intersection of two bricks is a brick. So, if

$$B = \bigcup_{j=1}^{l} K_j, \quad C = \bigcup_{i=1}^{m} K_i'$$

[1] The symbol $\vee$ denotes the exclusive "or".

are two elementary sets, then the set

$$B\bigcap C=\left(\bigcup_{j=1}^{l}K_j\right)\bigcap\left(\bigcup_{i=1}^{m}K'_i\right)=\bigcup_{i=1}^{l}\bigcup_{j=1}^{m}\left(K_i\bigcap K'_j\right)$$

is elementary, as well.

The difference of two bricks $K\setminus K'$ is clearly an elementary set. Consequently, if K is a brick, and $B=\bigcup\limits_{i=1}^{l}K_i$ is an elementary set, then

$$K\setminus B=K\setminus\left(\bigcup_{i=1}^{l}K_i\right)=\bigcap_{i=1}^{l}(K\setminus K_i)$$

is an elementary set too.

Let now B and C be two elementary sets. They are bounded as finite unions of bounded sets (bricks), so there exists a brick K containing both sets. Therefore, the sets

$$B\bigcup C=K\setminus\left(K\setminus\left(B\bigcup C\right)\right)=K\setminus\left((K\setminus B)\bigcap(K\setminus C)\right)$$

$$B\setminus C=B\bigcap(K\setminus C).$$

are elementary. □

The ring $\mathcal{E}^n$ is not an algebra, since it contains no units. Nevertheless, the collection $\mathcal{E}^n(K)$ of all bricks containing in some brick K is an algebra whose unit is K.

Definition 2.1.11. An algebra $\mathcal{A}$ is called σ-algebra (sigma-algebra) if it is closed under countable union of its elements, that is, if it possesses the property

7) $(A_k)_{k=1}^{\infty}\subset\mathcal{A}\implies A=\bigcup\limits_{k=1}^{\infty}A_k\subset\mathcal{A}.$

If $\mathcal{A}$ is a σ-algebra, then it also possesses the following property

8) $(A_k)_{k=1}^{\infty}\subset\mathcal{A}\implies A=\bigcap\limits_{k=1}^{\infty}A_k\subset\mathcal{A}.$

Indeed, if $(A_k)_{k=1}^{\infty}\subset\mathcal{A}$, then by the laws of duality

$$A=\bigcap_{k=1}^{\infty}A_k=\left(\bigcup_{k=1}^{\infty}A_k^c\right)^c\in\mathcal{A}.$$

Here $A^c=E\setminus A$, where E is the unit of the algebra $\mathcal{A}$.

Thus, σ-algebra is an algebra closed under countable unions and intersections of its elements. Among the examples (2.1.4)–(2.1.7), the set $\mathcal{P}(X)$ is a σ-algebra only.

Example 2.1.12. Let $X=\mathbb{R}$, and let

$$\mathcal{A}=\{A\subset\mathcal{A}\ :\ A\text{ is countable or }A^c\text{ is countable}\}.$$

It is easy to check that $\mathcal{A}$ is a σ-algebra.

Example 2.1.13. Let $X=[0,1]$, and let $\mathcal{A}=\{\emptyset,X,[0,\frac{1}{2}],(\frac{1}{2},1]\}$. Then $\mathcal{A}$ is a σ-algebra.

Remark 2.1.14. For rings, there is a somewhat different situation. One has to differ σ-rings that are closed under countable unions of their elements, and δ-rings that are closed under countable intersections of their elements.

Let now X be a set, and let $(A_k)_{k=1}^{\infty}$ be a sequence of subsets of the set X.

Definition 2.1.15. The set $\overline{A}$ consisting of all the elements of X (and only them) that belong to infinitely many sets A_k of the sequence $(A_k)_{k=1}^{\infty}$, is called the limit superior of the sequence $(A_k)_{k=1}^{\infty}$. We will denote it as $\overline{A} = \limsup A_k = \overline{\lim}\, A_k$

Thus, a set $\overline{A}$ is the limit superior of the sequence $(A_k)_{k=1}^{\infty}$ if, and only if, for any $x \in \overline{A}$, there exists a sequence k_j, $j \in \mathbb{N}$, of indices such that $x \in \overline{A_{k_j}}$, $j = 1, 2, \ldots$.

Definition 2.1.16. The set $\underline{A}$ consisting of all the elements of X (and only them) that belong to all the sets A_k starting from a fixed index, is called the limit inferior of the sequence $(A_k)_{k=1}^{\infty}$. We will denote it as $\overline{A} = \liminf A_k = \underline{\lim}\, A_k$

Thus, a set $\underline{A}$ is the limit inferior of the sequence $(A_k)_{k=1}^{\infty}$ if, and only if, for any $x \in \underline{A}$, there exists an index $k_0 = k_0(x)$ such that $x \in \overline{A_k}$ for all $k \geqslant k_0$.

It obviously follows from the definitions 2.1.15–2.1.16 that any sequence has the limit superior and the limit inferior (probably empty sets), and that $\underline{A} \subset \overline{A}$. The embedding here can be strict as the following example shows.

Example 2.1.17. Let $A_k = [0, 1 + 1/k]$ for odd k, and $A_k = [1 - 1/k, 2]$ for even k. Then we have $\overline{\lim}\, A_k = [0, 2]$, $\underline{\lim}\, A_k = \{1\}$.

Definition 2.1.18. A sequence of sets $(A_k)_{k=1}^{\infty}$ is called *convergent* if

$$\underline{\lim}\, A_k = \overline{\lim}\, A_k = A.$$

In this case, A is called the *limit* of the sequence $(A_k)_{k=1}^{\infty}$.

Theorem 2.1.19. *The following relations hold:*

$$\overline{A} = \bigcap_{k=1}^{\infty} \left(\bigcup_{l=k}^{\infty} A_l \right), \qquad \underline{A} = \bigcup_{k=1}^{\infty} \left(\bigcap_{l=k}^{\infty} A_l \right). \tag{2.1.4}$$

Proof. Let us prove the first of the relations (2.1.4). Suppose that $x \in \overline{A}$. Then, by definition, there exists a sequence of indices k_j such that $x \in A_{k_j}$, $j = 1, 2, \ldots$. So, for any $k \in \mathbb{N}$, there exists a number $l = k_j \geqslant k$, so that $x \in \bigcup\limits_{l=k}^{\infty} A_l$ for any $k \in \mathbb{N}$. Consequently, $x \in \bigcap\limits_{k=1}^{\infty} \left(\bigcup\limits_{l=k}^{\infty} A_l \right)$.

Conversely, let $x \in \bigcap\limits_{k=1}^{\infty} \left(\bigcup\limits_{l=k}^{\infty} A_l \right)$. Then $x \in \bigcup\limits_{l=k}^{\infty} A_l$ for any $k \in \mathbb{N}$. This means that for any $k \in \mathbb{N}$, there exists a number $l \geqslant k$ such that $x \in A_l$, that is, x belongs to infinitely many sets of the sequence $(A_k)_{k=1}^{\infty}$.

Prove now the second relation (2.1.4). Let $x \in \underline{A}$. Then, by definition, there exists a index $k = k(x)$ such that $x \in A_l$ for all $l \geqslant k$. So $x \in \bigcap\limits_{l=k}^{\infty} A_l$, therefore, $x \in \bigcup\limits_{k=1}^{\infty} \left(\bigcap\limits_{l=k}^{\infty} A_l \right)$.

Conversely, let $x \in \bigcup\limits_{k=1}^{\infty} \left(\bigcap\limits_{l=k}^{\infty} A_l \right)$. This means that there exists a natural number k such that $x \in \bigcap\limits_{l=k}^{\infty} A_l$, that is, $x \in A_l$ for all $l \geqslant k$. $\square$

From this theorem we can obtain one more property of σ-algebras. Namely, if $\mathcal{A}$ is a σ-algebra, then it possesses the following property

9) $(A_k)_{k=1}^\infty \subset \mathcal{A} \implies \overline{\lim}\, A_k, \underline{\lim}\, A_k \in \mathcal{A}$. *Moreover, if the sequence* $(A_k)_{k=1}^\infty$ *converges, then its limit also belongs to* $\mathcal{A}$, $\lim A_k \in \mathcal{A}$.

This property is a consequence of the properties 7) and 8), and of the formulæ (2.1.4).

Definition 2.1.20. A sequence of sets $(A_k)_{k=1}^\infty$ is called monotone if it satisfies one the following relations

$$A_1 \subset A_2 \subset \cdots \subset A_k \subset \cdots \quad \text{(increasing sequence)},$$

or

$$A_1 \supset A_2 \supset \cdots \supset A_k \supset \cdots \quad \text{(decreasing sequence)}.$$

Lemma 2.1.21. *A monotone sequence of sets* $(A_k)_{k=1}^\infty$ *converges. Moreover, if the sequence is increasing, then* $\lim A_k = \bigcup\limits_{k=1}^\infty A_k$, *and if the sequence decreases, then* $\lim A_k = \bigcap\limits_{k=1}^\infty A_k$.

Proof. Let $(A_k)_{k=1}^\infty$ increases. Then $\bigcup\limits_{l=k}^\infty A_l = \bigcup\limits_{l=1}^\infty A_l$, therefore, $\bigcap\limits_{k=1}^\infty \left(\bigcup\limits_{l=k}^\infty A_l\right) = \bigcap\limits_{k=1}^\infty \left(\bigcup\limits_{l=1}^\infty A_l\right) = \bigcup\limits_{l=1}^\infty A_l$. Consequently, $\overline{A} = \bigcup\limits_{k=1}^\infty A_k$ (see (2.1.4)).

On the other hand, $\bigcap\limits_{l=k}^\infty A_l = A_k$, so $\underline{A} = \bigcup\limits_{k=1}^\infty \left(\bigcap\limits_{k=l}^\infty A_l\right) = \bigcup\limits_{k=1}^\infty A_k$.

Thus, $\underline{A} = \overline{A}$, which means that an increasing sequence of sets always converges, and $\lim A_k = \bigcup\limits_{k=1}^\infty A_k$.

The corresponding fact for decreasing sequences can be proved analogously. $\square$

2.2 General measure theory

The concept of the measure of a set is a natural extension of the such concepts as the length of an interval, the square of a flat body, the volume of a body, the increment $\varphi(b)-\varphi(a)$ of a non-decreasing function $\varphi(x)$ on the interval $[a,b]$, the integral of a nonnegative function over a certain domain, and many other notions.

Let $\mathcal{K}$ be a ring.

Definition 2.2.1. A function $\mu : \mathcal{K} \mapsto \mathbb{R}_+ := [0,\infty)$ is called a *measure* if it possesses the following property

1) $\mu\,(A_1 \dot\cup A_2) = \mu A_1 + \mu A_2$, *for any* $A_1, A_2 \in \mathcal{K}$.

The property 1) is called the *additivity property* of measures. Let us study some properties of measures.

2) $\mu\varnothing = 0$.

   Indeed, $\mu\varnothing = \mu(\varnothing \dot\cup \varnothing) = \mu\varnothing + \mu\varnothing = 2\mu\varnothing$, so $\mu\varnothing = 0$.

3) $A \subset B \implies \mu A \leqslant \mu B$ (*monotonicity*).

   Since $B = A \dot\cup (B \setminus A)$ and since $B \setminus A \in \mathcal{K}$, we have

$$\mu B = \mu A + \mu(B \setminus A) \geqslant \mu A, \tag{2.2.1}$$

   due to the additivity and nonnegativity of measures.

4) $A \subset B \implies \mu(B \setminus A) = \mu B - \mu A$.

   This property follows from (9.1.1).

5) $\mu(A\bigcup B) = \mu A + \mu B - \mu(A\bigcap B)$.

In fact, $A\bigcup B = A\dot{\bigcup}(B\setminus(A\bigcap B))$, so the property 5) follows now from the properties 1) and 4).

6) $\mu\left(\dot{\bigcup}_{k=1}^{l} A_k\right) = \sum_{k=1}^{l}\mu A_k$ (*finite additivity*).

This property follows from 1) by induction.

7) $\mu\left(\bigcup_{k=1}^{l} A_k\right) \leqslant \sum_{k=1}^{l}\mu A_k$ (*finite semi-additivity*).

Let us introduce the sets $A'_1 = A_1$, $A'_k = A_k \setminus \left(\bigcup_{j=1}^{k-1} A_j\right)$, $k = 2,3,\dots,l$. Since from each set we subtract all the elements that belong to the previous sets, we get $A'_k \subset A_k$ for all k, and $A'_k\bigcap A'_j = \varnothing$ whenever $k \neq j$. Obviously, $\bigcup_{k=1}^{l} A_k = \bigcup_{j=1}^{l} A'_j = \dot{\bigcup}_{j=1}^{l} A'_j$. Now from the properties 4) and 3) it follows

$$\mu\left(\bigcup_{k=1}^{l} A_k\right) = \mu\left(\dot{\bigcup}_{j=1}^{l} A'_j\right) = \sum_{j=1}^{l}\mu A'_j \leqslant \sum_{k=1}^{l}\mu A_k.$$

8) *If* $A \supset \dot{\bigcup}_{k=1}^{\infty} A_k$, *where* $A, A_k \in \mathcal{K}$, $k \in \mathbb{N}$, *then the series* $\sum_{k=1}^{\infty}\mu A_k$ *converges, and*

$$\mu A \geqslant \sum_{k=1}^{\infty}\mu A_k \tag{2.2.2}$$

Since $A \supset \dot{\bigcup}_{k=1}^{l} A_k$ for any $l \in \mathbb{N}$, so by the properties 3) and 6)

$$\mu A \geqslant \mu\left(\dot{\bigcup}_{k=1}^{l} A_k\right) = \sum_{k=1}^{l}\mu A_k, \qquad \forall l \in \mathbb{N}.$$

This estimate shows that the series $\sum_{k=1}^{\infty}\mu A_k$ converges, and by tending l to infinity we get (9.2.2).

The properties 1)–8) exhaust all the properties of additive measures on a ring $\mathcal{K}$. However, these properties turn to be not enough in a lot of cases, so it is useful to introduce a specific class of measures.

Definition 2.2.2. A measure μ defined on a ring $\mathcal{K}$ is called *countably additive*, or *σ-additive*, if it possesses the following property

9) *If* $A = \dot{\bigcup}_{k=1}^{\infty} A_k$, *where* $A, A_k \in \mathcal{K}$, $k \in \mathbb{N}$, *then*

$$\mu A = \sum_{k=1}^{\infty}\mu A_k. \tag{2.2.3}$$

Let us formulate two more definitions (properties).

Definition 2.2.3. A measure μ defined on a ring $\mathcal{K}$ is called *countably semi-additive*, or *σ-semi-additive*, if it possesses the following property

10) *If* $A \subset \bigcup\limits_{k=1}^{\infty} A_k$, *where* $A, A_k \in \mathcal{K}$, $k \in \mathbb{N}$, *then*

$$\mu A \leqslant \sum_{k=1}^{\infty} \mu A_k. \tag{2.2.4}$$

Definition 2.2.4. A measure μ defined on a ring $\mathcal{K}$ is called *continuous* if it possesses the property

11) *If a sequences* $(A_k)_{k=1}^{\infty} (\subset \mathcal{K})$ *is monotone, and* $\lim A_k \in \mathcal{K}$, *then*

$$\mu(\lim A_k) = \lim \mu A_k. \tag{2.2.5}$$

Theorem 2.2.5. *The properties of σ-additivity, σ-semi-additivity, and the continuity of the measures are equivalent.*

Proof. Theorem asserts that if a measure possesses one of the properties 9)–11), then it possesses the rest ones.

First, we prove that the property 9) is equivalent to the property 10).

Let the measure μ is σ-additive. Consider an arbitrary sequence of sets $(A_k)_{k=1}^{\infty} \subset \mathcal{K}$ and a set $A \in \mathcal{K}$ such that $A \subset \bigcup\limits_{k=1}^{\infty} A_k$. Introduce the following new sets

$$A_1' = A_1 \bigcap A, \quad A_k' = \left(A_k \bigcap A\right) \setminus \left(\bigcup_{j=1}^{k-1} A_j'\right).$$

It is easy to check that $A_k' \in \mathcal{K}$ $(k \in \mathbb{N})$, $A_k' \bigcap A_j' = \varnothing$ $(k \neq j)$, $A_k' \subset A_k$ $(k \in \mathbb{N})$, and $A = \dot{\bigcup\limits_{k=1}^{\infty}} A_k'$. Thus, by the property 3), one has $\mu A_k' \leqslant \mu A_k$ $(k \in \mathbb{N})$, and by the property 9)

$$\mu A = \sum_{k=1}^{\infty} \mu A_k' \leqslant \sum_{k=1}^{\infty} \mu A_k,$$

so 9) $\implies$ 10).

Conversely, let the measure μ is σ-semi-additive, and let $(A_k)_{k=1}^{\infty} \in \mathcal{K}$ be a sequence of *disjoint* sets, and $A = \dot{\bigcup\limits_{k=1}^{\infty}} A_k \in \mathcal{K}$. Then the inclusion $A \subset \bigcup\limits_{k=1}^{\infty} A_k$ holds, and due to the σ-semi-additivity of the measure μ we have

$$\mu A \leqslant \sum_{k=1}^{\infty} \mu A_k. \tag{2.2.6}$$

On the other hand, the opposite inclusion $A \supset \dot{\bigcup\limits_{k=1}^{\infty}} A_k$ holds, as well, so by the property 8)

$$\mu A \geqslant \sum_{k=1}^{\infty} \mu A_k. \tag{2.2.7}$$

From (9.4.3)–(2.2.7), it follows that μA is σ-additive whenever it is σ-semi-additive.

Let us prove now equivalence of the properties 9) and 11). Suppose that μ is σ-additive. Let a sequence $(A_k)_{k=1}^{\infty} (\subset \mathcal{K})$ be monotone (see Definition 2.1.20) and $A = \lim A_k \in \mathcal{K}$.

If the sequence $(A_k)_{k=1}^{\infty}$ increases, then according to Lemma 2.1.21,

$$A = \lim A_k = \bigcup_{k=1}^{\infty} A_k = \dot{\bigcup_{k=1}^{\infty}} (A_k \setminus A_{k-1}) \quad (A_0 = \varnothing),$$

so
$$\mu A = \sum_{k=1}^{\infty} \mu(A_k \setminus A_{k-1}) = \lim_{m\to\infty} \sum_{k=1}^{m} \mu(A_k \setminus A_{k-1}) = \lim_{m\to\infty} \sum_{k=1}^{m} (\mu A_k - \mu A_{k-1}) = \lim_{m\to\infty} \mu A_m.$$
by the properties 3), 4), and 9). Thus, 9) $\implies$ 11) whenever the sequence $(A_k)_{k=1}^{\infty}$ is increasing.

Suppose now that $(A_k)_{k=1}^{\infty}$ is decreasing. Then the sequence $(A_1 \setminus A_k)_{k=1}^{\infty}$ is increasing, and
$$\lim(A_1 \setminus A_k) = \bigcup_{k=1}^{\infty} (A_1 \setminus A_k) = A_1 \setminus \left(\bigcap_{k=1}^{\infty} A_k \right) = A_1 \setminus \lim A_k.$$
Since the measure μ is proved to be continuous for increasing sequences, we have
$$\mu(A_1 \setminus \lim A_k) = \lim \mu(A_1 \setminus A_k),$$
or, equivalently,
$$\mu A_1 - \mu(\lim A_k) = \lim(\mu A_1 - \mu A_k) = \mu A_1 - \lim \mu A_k,$$
so
$$\mu(\lim A_k) = \lim(\mu A_k).$$
Thus, we proved the implication 9) $\implies$ 11).

Conversely, let μ be a continuous measure, and let $(A_k)_{k=1}^{\infty} (\subset \mathcal{K})$ be a sequence of disjoint sets such that $A = \dot{\bigcup}_{k=1}^{\infty} A_k \in \mathcal{K}$. Introduce the sets $A'_k = \dot{\bigcup}_{j=1}^{k} A_j$, $k \in \mathbb{N}$. It is clear that the sequence $(A'_k)_{k=1}^{\infty}$ is increasing, and according to Lemma 2.1.21,
$$A = \dot{\bigcup_{k=1}^{\infty}} A_k = \bigcup_{k=1}^{\infty} A'_k = \lim A'_k.$$
Therefore, by the continuity and finite additivity of the measure μ, one obtains
$$\mu A = \lim \mu A'_k = \lim \mu \left(\dot{\bigcup_{j=1}^{k}} A_j \right) = \lim \sum_{j=1}^{k} \mu A_j = \sum_{j=1}^{\infty} \mu A_j.$$
Thus, 11) implies 9), as required. □

Theorem 2.2.6. *Let μ be a σ-additive measure defined on a σ-algebra $\mathcal{A}$, and let $(A_k)_{k=1}^{\infty}$ be an arbitrary sequence of elements of the algebra $\mathcal{A}$. Then the measure μ possesses the following properties:*

12) $\mu(\overline{\lim}\, A_k) \geqslant \overline{\lim}\, \mu A_k$*,*

13) $\mu(\underline{\lim}\, A_k) \leqslant \underline{\lim}\, \mu A_k$*,*

14) if the sequence $(A_k)_{k=1}^{\infty}$ converges, then
$$\mu(\lim A_k) = \lim \mu A_k.$$

Proof. We, first, prove the property 12). From (2.1.4) it follows that
$$\overline{A} = \overline{\lim}\, A_k = \bigcap_{k=1}^{\infty} \left(\bigcup_{l=k}^{\infty} A_l \right).$$
Consider the sequence of the sets $B_k = \bigcup_{l=k}^{\infty} A_l$, $k \in \mathbb{N}$ that belong to $\mathcal{A}$, since it is a σ-algebra by assumption. As k grows the union $\bigcup_{l=k}^{\infty} A_l$ decreases, so the sequence $(B_k)_{k=1}^{\infty}$ is decreasing, and $\overline{\lim}\, A_k = \bigcap_{k=1}^{\infty} B_k = \lim B_k$. Since the measure μ is continuous by Theorem 2.2.5, we have
$$\mu(\overline{\lim}\, A_k) = \lim \mu B_k. \tag{2.2.8}$$

Consider now the sequence $(\mu A_k)_{k=1}^{\infty}$. It is bounded from above by the number μE where E is the unit of the algebra $\mathcal{A}$, therefore, there exists finite $\overline{\lim}\,\mu A_k$. Consequently, in the sequence $(A_k)_{k=1}^{\infty}$ one can find a subsequence $(A_{k_j})_{j=1}^{\infty}$ converging to $\overline{\lim}\,\mu A_k$. Since obviously $B_{k_j} \supset A_{k_j}$, from (2.2.8) and from the monotonicity of the measure μ (the property 3)) we obtain

$$\mu(\overline{\lim}\, A_k) = \lim_k \mu B_k = \lim_j \mu B_{k_j} \geqslant \lim_j \mu A_{k_j} = \overline{\lim}\,\mu A_k,$$

so the property 12) is true. The property 13) can be proved analogously.

Suppose now that the sequence $(A_k)_{k=1}^{\infty}$ converges. Then

$$\underline{\lim}\, A_k = \overline{\lim}\, A_k = \lim A_k,$$

so by the properties 12) and 13)

$$\overline{\lim}\,\mu A_k \leqslant \mu(\overline{\lim} A_k) = \mu(\lim A_k) = \mu(\underline{\lim} A_k) \leqslant \underline{\lim}\,\mu A_k \leqslant \overline{\lim}\,\mu A_k,$$

therefore,

$$\mu(\lim A_k) = \lim \mu A_k,$$

as required. □

The difference between the property 11) and the property 14) of measures on σ-rings and on σ-algebras is that in the property 11) we have to consider only monotone sequences, while in the property 14) we avoid this restriction.

Definition 2.2.7. If there is defined a measure on a ring $\mathcal{K}$, then the sets of the ring $\mathcal{K}$ are called *measurable* w.r.t. the measure μ, or *μ-measurable.*

2.3 Problems

Problem 2.1. Prove that the definition of the ring given in the class is equivalent to the following one: *a non-empty system of sets $\mathcal{K}$ is called a ring if it possesses the following two properties*
1) $A, B \in \mathcal{K} \Rightarrow A \bigcap B \in \mathcal{K}$;
2) $A, B \in \mathcal{K} \Rightarrow A \vartriangle B \in \mathcal{K}$.

Problem 2.2. Let $X = \{a, b, c\}$ (a set consisting of three elements), and let $\mathcal{P}(X)$ is the set of all subsets of X.
a) Describe all rings that can be constructed with the elements of $\mathcal{P}(X)$.
b) Describe all algebras that can be constructed with the elements of $\mathcal{P}(X)$.

Problem 2.3. Prove that

a) intersection of rings is a ring,

b) intersection of algebras with common unit is an algebra,

c) intersection of σ-algebras with common unit is a σ-algebra.

Problem 2.4. Let $\mathcal{A}(X)$ be σ-algebra with the unit X, and let $X_0 \in \mathcal{A}(X)$. Show that the collection of sets

$$\mathcal{A}(X_0) = \{A \bigcap X_0 : A \in \mathcal{A}(X)\}$$

is a σ-algebra.

Problem 2.5. Let X be a set, and let $(E_n)_{n\in\mathbb{N}}$ be a sequence of sets such that $E_n \in X$ for any $n \in \mathbb{N}$. Prove that

$$X\backslash(\overline{\lim} E_n) = \underline{\lim}(X\backslash E_n); \qquad X\backslash(\underline{\lim} E_n) = \overline{\lim}(X\backslash E_n).$$

Problem 2.6. Let X be a set, and $\mathcal{M}(X)$ be a σ-algebra of subsets of X. Suppose a σ-additive measure μ is defined on $\mathcal{M}(X)$. Prove that the collection of all subsets of X of measure 0 is a σ-ring.

Problem 2.7. Let φ be an additive real function defined on a ring of sets $\mathcal{K}$. Prove that

1) if φ is finite on at least one element of $\mathcal{K}$, then $\varphi(\varnothing) = 0$,

2) for any $A, B \in \mathcal{K}$,
$$\varphi(A \cup B) + \varphi(A \cap B) = \varphi(A) + \varphi(B),$$

3) if $\mathcal{G}$ is the set of all elements of $\mathcal{K}$ such that $\varphi(A)$ is finite for any $A \in \mathcal{G}$, then $\mathcal{G}$ is a ring.

Problem 2.8. Let a measure μ (a nonnegative additive function) be defined on a ring $\mathcal{K}$. Prove that if

$$\mu(A \vartriangle B) = 0,$$

then $\mu(A) = \mu(B)$.

Problem 2.9. On the algebra of all subsets of rational numbers in $[0, 1]$, define a measure such that the measure of each rational number is positive and the measure of the whole set of all rational numbers on $[0, 1]$ equals 1.

Chapter 3

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

The Lebesgue measure is a natural extension of the concepts of the length of an interval, of the square of flat bodies, of the volume of bodies, and it is an extension of the Jordan measure (see homework for definition) to a more wide class of sets with the additional property of σ-additivity. Here we construct the Lebesgue measure in three steps. First, we define the measure on bricks, then we extend it to the ring $\mathcal{E}^n$ of all elementary sets, and then we, finally, extend the obtained measure to a more wide class of sets which we call the set of Lebesgue measurable sets.

3.1 The measure on bricks

Let K be a brick (see Definition 2.1.8). The following number

$$m'K = \prod_{i=1}^{n}(b_i - a_i) \tag{3.1.1}$$

is called the measure of the brick K.

From (3.1.1) it is easy to see that $m'K$ is the length of an interval for $n = 1$, the square of a rectangle for $n = 3$, and the volume of a parallelepiped for $n = 3$. It is also clear that if a brick is degenerate, that is, $b_i = a_i$ for at least one index i, then $m'K = 0$.

The measure of bricks introduced by (3.1.1) possesses the following properties:

1) $m'K \geqslant 0$,

2) $K = \dot{\bigcup\limits_{j=1}^{l}} K_j \implies m'K = \sum\limits_{j=1}^{l} m'K_j$.

The property 1) is obvious. For brevity, we prove the property 2) for the case $\mathbb{R}^2$.

Let, first, the brick K is split into the bricks K_j by vertical lines

$$x = x_s, \quad 0 \leqslant s \leqslant p,\ x_0 = a_1,\ x_p = b_1,$$

and by horizontal lines

$$y = y_t, \quad 0 \leqslant t \leqslant q,\ y_0 = a_2,\ y_q = b_2, \quad pq = l.$$

Then we have

$$\sum_{j=1}^{l} m'K_j = \sum_{s=1}^{p}\sum_{t=1}^{q}(x_s - x_{s-1})(y_t - y_{t-1}) = (b_1 - a_1)(b_2 - a_2) = m'K.$$

If now $K = \dot{\bigcup}_{j=1}^{l} K_j$ is an arbitrary representation of K as a union of disjoint bricks K_j, then we split the bricks K_j by vertical and horizontal lines drawn along the all sides of all the bricks K_j. The obtained new bricks we denote as K_i', $1 \leqslant i \leqslant r$. Thus,

$$K = \dot{\bigcup_{j=1}^{l}} K_j = \dot{\bigcup_{i=1}^{r}} K_i',$$

so

$$m'K = \sum_{i=1}^{r} m'K_i' = \sum_{j=1}^{l} \left(\sum_{i:K_i' \subset K_j} m'K_i' \right) = \sum_{j=1}^{l} m'K_j.$$

as we showed above.

Remark 3.1.1. Generally speaking, the measure m' on bricks is not a measure by Definition 2.2.1, since the collection of all bricks in $\mathbb{R}^n$ is not a ring. However, one can introduce the notion of semi-rings of sets (the collection of bricks is a semi-ring) and then to define measures on semi-rings. For details, see e.g. [6].

3.2 The measure on $\mathcal{E}^n$

Now we define a measure on the ring $\mathcal{E}^n$ of elementary sets in $\mathbb{R}^n$. Let B be an elementary set (see Definition 2.1.9). It can be represented (see (2.1.3)) as a union of *disjoint* bricks K_j'

$$B = \dot{\bigcup_{i=1}^{r}} K_i'.$$

We set

$$mB := \sum_{i=1}^{r} m'K_i'. \tag{3.2.1}$$

First of all, we must show that the measure (3.2.1) is defined correctly, that is, that $mK = m'K$ for any brick, and that the value mB does not depend on the representation of B as a union of bricks.

Let

$$B = \dot{\bigcup_{j=1}^{l}} K_j = \dot{\bigcup_{i=1}^{r}} K_i'.$$

We want to show that

$$\sum_{j=1}^{l} m'K_j = \sum_{i=1}^{r} m'K_i'. \tag{3.2.2}$$

To do this, let us introduce the following sets

$$\widetilde{K}_{j,i} = K_j \bigcap K_i', \quad j = 1, 2, \dots, l, \quad i = 1, 2, \dots, r.$$

The sets $\widetilde{K}_{j,i}$ are bricks as intersections of two bricks. Moreover, since the bricks K_j and K_i' are pairwise disjoint[1], we have

$$\widetilde{K}_{j_1,i_1} \bigcap \widetilde{K}_{j_2,i_2} = \varnothing, \qquad (j_1, i_1) \neq (j_2, i_2).$$

Obviously, $K_j = \dot{\bigcup}_{i=1}^{r} \widetilde{K}_{j,i}$ for any $j = 1, \dots, l$, and $K_i' = \dot{\bigcup}_{j=1}^{l} \widetilde{K}_{j,i}$ for any $i = 1, \dots, r$. Since the "measure" m' is additive (see the property 2) of m') on bricks, one gets

$$m'K_j = \sum_{i=1}^{r} m'\widetilde{K}_{j,i}, \quad j = 1, \dots, l, \qquad m'K_j' = \sum_{j=1}^{l} m'\widetilde{K}_{j,i}, \quad i = 1, \dots, r.$$

[1]That is, $K_{j_1} \bigcap K_{j_2} = \varnothing$ if $j_1 \neq j_2$, and, respectively, $K_{i_1}' \bigcap K_{i_2}' = \varnothing$ if $i_1 \neq i_2$.

Therefore,

$$\dot{\bigcup_{j=1}^{l}} K_j = \dot{\bigcup_{j=1}^{l}} \left(\dot{\bigcup_{i=1}^{r}} \widetilde{K}_{j,i} \right) = \dot{\bigcup_{j=1}^{l}} \dot{\bigcup_{i=1}^{r}} \widetilde{K}_{j,i} = \dot{\bigcup_{i=1}^{r}} \left(\dot{\bigcup_{j=1}^{l}} \widetilde{K}_{j,i} \right) = \dot{\bigcup_{i=1}^{r}} K'_i.$$

Now from additivity of m' on bricks, we obtain

$$\sum_{j=1}^{l} m'K_j = \sum_{j=1}^{l} \left(\sum_{i=1}^{r} m'\widetilde{K}_{j,i} \right) = \sum_{j=1}^{l} \sum_{i=1}^{r} m'\widetilde{K}_{j,i} = \sum_{i=1}^{r} \left(\sum_{j=1}^{l} m'\widetilde{K}_{j,i} \right) = \sum_{i=1}^{r} m'K'_i,$$

so the identity (3.2.2) is proved. In particular, from (3.2.2), it follows that

$$mK = m'K$$

for any brick, since the identity $K = K$ is one of the representations of the elementary set K as a finite union of bricks. Thus, the measure m is an extension of the "measure" m' to the ring $\mathcal{E}^n$ of all elementary sets in $\mathbb{R}^n$.

As well as m', the measure m possesses the following properties:

1) $mB \geqslant 0, \quad \forall B \in \mathcal{E}^n$,

2) $B = \dot{\bigcup}_{j=1}^{l} B_j \implies mB = \sum_{j=1}^{l} mB_j$.

The first property is evident. Moreover, it is sufficient to prove the property 2) for the case $l = 2$. So, let $B_1 = \dot{\bigcup}_{j=1}^{r_1} K_{1,j}$ and $B_2 = \dot{\bigcup}_{j=1}^{r_2} K_{2,j}$. Then we have

$$B = B_1 \dot{\bigcup} B_2 = \left(\dot{\bigcup_{j=1}^{r_1}} K_{1,j} \right) \dot{\bigcup} \left(\dot{\bigcup_{j=1}^{r_2}} K_{2,j} \right) = \dot{\bigcup_{i=1}^{2}} \dot{\bigcup_{j=1}^{r_i}} K_{i,j},$$

so

$$mB = \sum_{i=1}^{2} \sum_{j=1}^{r_i} m'K_{i,j} = \sum_{i=1}^{2} \left(\sum_{j=1}^{r_i} m'K_{i,j} \right) = mB_1 + mB_2.$$

Consequently, the function m defined on the ring $\mathcal{E}^n$ is a measure according to the Definition 2.2.1, therefore, it possesses the properties $2) - 8)$ of measures.

Theorem 3.2.1. *The measure m on $\mathcal{E}^n$ is σ-semi-additive.*

Proof. We need to prove that if B and B_k, $k \in \mathbb{N}$, are elementary sets, and if $B \subset \bigcup_{k=1}^{\infty} B_k$, then

$$mB \leqslant \sum_{k=1}^{\infty} mB_k. \tag{3.2.3}$$

If the series in the right hand side of (3.2.3) diverges, then we are done. Suppose now that this series converges, and fix some number $\varepsilon > 0$. For the set B, one can find a ***closed*** elementary set $\overline{B}_0 \subset B$ such that

$$m\overline{B}_0 > mB - \frac{\varepsilon}{2}. \tag{3.2.4}$$

It can be done as follows. Let $B = \dot{\bigcup}_{i=1}^{r} K_i$. In each brick K_i we choose a closed brick $\overline{K}_{0,i}$ so that

$$m'\overline{K}_{0,i} > m'K_i - \frac{\varepsilon}{2r},$$

and set

$$\overline{B}_0 = \dot{\bigcup_{i=1}^{r}} \overline{K}_{0,i}.$$

So, we obtain

$$m\overline{B}_0 = \sum_{i=1}^{r} m'\overline{K}_{0,i} > \sum_{i=1}^{r} \Big(m'K_i - \frac{\varepsilon}{2r} \Big) = \sum_{i=1}^{r} m'K_i - \frac{\varepsilon}{2} = mB - \frac{\varepsilon}{2}.$$

Furthermore, for each number $k \in \mathbb{N}$, let us choose an ***open*** elementary set $\widetilde{B}_k \supset B_k$ such that

$$m\widetilde{B}_k < mB_k + \frac{\varepsilon}{2^{k+1}}. \tag{3.2.5}$$

(It can be done in the same manner as we chose $\overline{B}_0$ above.)

Then we have

$$\overline{B}_0 \subset B \subset \bigcup_{k=1}^{\infty} B_k \subset \bigcup_{k=1}^{\infty} \widetilde{B}_k.$$

By the Heine-Borel theorem, from every cover of the closed set $\overline{B}_0$ by open sets, one can choose a finite subcover, that is, to choose sets $\widetilde{B}_{k_j}$, $j = 1, \ldots, l$, such that

$$\overline{B}_0 \subset \bigcup_{j=1}^{l} \widetilde{B}_{k_j}.$$

As we mentioned above, the measure m is finite semi-additive (satisfies the property 7) of measures). Therefore,

$$m\overline{B}_0 \leqslant \sum_{j=1}^{l} m\widetilde{B}_{k_j}.$$

Now from (3.2.4)–(3.2.5) we obtain

$$mB < m\overline{B}_0 + \frac{\varepsilon}{2} \leqslant \sum_{j=1}^{l} m\widetilde{B}_{k_j} + \frac{\varepsilon}{2} \leqslant \sum_{k=1}^{\infty} m\widetilde{B}_k + \frac{\varepsilon}{2} < \sum_{k=1}^{\infty} \Big(mB_k + \frac{\varepsilon}{2^{k+1}} \Big) + \frac{\varepsilon}{2} = \sum_{k=1}^{\infty} mB_k + \varepsilon.$$

Since ε is arbitrary, the inequality (3.2.3) holds. □

A simple consequence of Theorems 3.2.1 and 2.2.5 is the following fact.

Theorem 3.2.2. *The measure m on the ring $\mathcal{E}^n$ is σ-additive.*

Thus, by the formula (3.2.1), we define the σ-additive measure m on $\mathcal{E}^n$. Our next goal is to extend this measure onto a wider class of sets keeping its σ-additivity.

3.3 The outer measure

Let $A(\subset \mathbb{R}^n)$ be an arbitrary set.

Definition 3.3.1. The number

$$\mu^* A = \inf \left\{ \sum_j m'K_j : \quad A \subset \bigcup_j K_j \right\} \tag{3.3.1}$$

is called the *outer measure* of the set A.

We emphasize that the infimum in (3.3.1) is taken over *all* covers of the set A by finite or infinite collections of bricks. Also notice that μ^*A can be infinite (for example, in the case $A = \mathbb{R}^n$).

Let us study properties of the measure μ^*.

1) $\mu^*A \geqslant 0$, $A \subset \mathbb{R}^n$.

   This property is evident.

2) $A_1 \subset A_2 \implies \mu^*A_1 \leqslant \mu^*A_2$ (*The outer measure is monotone*).

   Since every cover of the set A_2 is also a cover of the set A_1, we have

$$\mu^*A_1 = \inf\left\{\sum_j m'K_j : \quad A_1 \subset \bigcup_j K_j\right\} \leqslant \inf\left\{\sum_j m'K_j : \quad A_2 \subset \bigcup_j K_j\right\} = \mu^*A_2.$$

3) $A \subset \bigcup_k A_k \implies \mu^*A \leqslant \sum_k \mu^*A_k$ (*The outer measure is σ-semi-additive*).

   Here $(A_k)_{k\geqslant 0}$ is a finite or countable collection of sets. As we mentioned above, the outer measure can be infinite. If one of the measures μ^*A_k is infinite or all of them are finite but the series $\sum_k \mu^*A_k$ is infinite, then the inequality $\mu^*A \leqslant \sum_k \mu^*A_k$ holds. Suppose now that all the measures μ^*A_k are finite and

$$\sum_k \mu^*A_k < +\infty.$$

   Let us fix a number $\varepsilon > 0$. By Definition 3.3.1, for every k there exists a collection of bricks $(K_{k,j})_{j=1}^{\infty}$ (finite or countable) such that $A_k \subset \bigcup_j K_{k,j}$ and

$$\mu^*A_k > \sum_j m'K_{k,j} - \frac{\varepsilon}{2^k}.$$

   Moreover, $A \subset \bigcup_k A_k \subset \bigcup_k \bigcup_j K_{k,j}$, so by Definition 3.3.1,

$$\mu^*A \leqslant \sum_k \sum_j m'K_{k,j} = \sum_k \left(\sum_j m'K_{k,j}\right) < \sum_k \left(m^*A_k + \frac{\varepsilon}{2^k}\right) \leqslant \sum_k \mu^*A_k + \varepsilon.$$

   Since ε is arbitrary, the inequality $\mu^*A \leqslant \sum_k \mu^*A_k$ holds.

4) $A \in \mathcal{E}^n \implies \mu^*A = mA \quad$ (*on $\mathcal{E}^n$ the outer measure coincides with m*).

   Let $A = \overset{r}{\underset{i=1}{\dot\bigcup}}\ K_i$ be an elementary set. Since $\overset{r}{\underset{i=1}{\dot\bigcup}}\ K_i$ is a cover of the set A and since μ^*A is the infimum over all covers, we have

$$\mu^*A \leqslant \sum_{i=1}^{r} m'K_i = mA. \tag{3.3.2}$$

   On the other hand, if $\bigcup_j K'_j$ is an arbitrary cover of the set A, then due to σ-additivity of the measure m (by Theorem 3.2.1), one obtains

$$mA \leqslant \sum_j mK'_j = \sum_j m'K'_j,$$

therefore,

$$mA \leqslant \inf \left\{ \sum_j m'K'_j : \quad A \subset \bigcup_j K_j \right\} = \mu^* A. \tag{3.3.3}$$

Now the property 4) follows from (3.3.2)–(3.3.3).

Example 3.3.2. *Let $A = \{a_i : \ a_i \in \mathbb{R}^n, i \in \mathbb{N}\}$ be a countable set. Then $\mu^* A = 0$.*

Indeed, let $\varepsilon > 0$ be arbitrary, and for every $i \in \mathbb{N}$, we take a brick K_i such that $a_i \in K_i$ and $m'K_i < \dfrac{\varepsilon}{2^i}$ (for example, K_i is the n-dimensional cube with edges of length $\sqrt[n]{\varepsilon/2^i}$). Thus, we have $A \subset \bigcup_i K_i$, and

$$\mu^* A \leqslant \sum_{i=1}^{\infty} m'K_i < \varepsilon \sum_{i=1}^{\infty} \frac{1}{2^i} = \varepsilon,$$

so $\mu^* A = 0$, since ε is arbitrary.

Let us introduce the class of sets in $\mathbb{R}^n$ with *finite* outer measure and denote it $\mathcal{M}^*(\mathbb{R}^n)$, that is,

$$\mathcal{M}^*(\mathbb{R}^n) = \{A \subset \mathbb{R}^n : \ \mu^* A < +\infty\}.$$

From the properties 2) and 3) of outer measure, it follows that $\mathcal{M}^*(\mathbb{R}^n)$ is a *ring*, since if $A_1, A_2 \in \mathcal{M}^*(\mathbb{R}^n)$, then $A_1 \bigcup A_2$ and $A_1 \setminus A_2 (\subset A_1)$ also have finite measures, so $A_1 \bigcup A_2 \in \mathcal{M}^*(\mathbb{R}^n)$ and $A_1 \setminus A_2 \in \mathcal{M}^*(\mathbb{R}^n)$.

3.4 The Lebesgue measure

Now we are in a position to introduce the Lebesgue measure and the Lebesgue measurable sets on $\mathbb{R}^n$.

Definition 3.4.1. A set $A \in \mathcal{M}^*(\mathbb{R}^n)$ is called Lebesgue measurable if for any $\varepsilon > 0$ there exists an elementary set B such that

$$\mu^*(A \triangle B) < \varepsilon. \tag{3.4.1}$$

The class of all Lebesgue measurable sets is denoted as $\mathcal{M}(\mathbb{R}^n)$.

Definition 3.4.2. The function μ^* restricted to $\mathcal{M}(\mathbb{R}^n)$ is called the Lebesgue measure, and is denoted μ.

Our next goal is to show that $\mathcal{M}(\mathbb{R}^n)$ is a ring, and that the function μ defined on this ring is a measure according to Definition 2.2.1. But before that we define the geometrical meaning of Lebesgue measurable sets. In fact, as it follows from (3.4.1), any Lebesgue measurable set can be approximated by elementary sets arbitrary accurately. That is, for any Lebesgue measurable set A, one can find an elementary set B such that A and B protrude from each other "not too far".

Before we start to solve the intended problem, let us notice the following properties of the function μ.

1) $\mu A \geqslant 0, \quad \forall A \in \mathcal{M}(\mathbb{R}^n)$.

   This property is obvious.

2) *$\mathcal{E}^n \subset \mathcal{M}(\mathbb{R}^n)$, and if $A \in \mathcal{E}^n$, then $\mu A = mA$.*

   Indeed, if $A \in \mathcal{E}^n$, then according to the property 4) of the outer measure, $A \in \mathcal{M}^*(\mathbb{R}^n)$, since $\mu^* A = mA < +\infty$. Let us set $B := A$, then for any $\varepsilon > 0$, we have

   $$\mu^*(A \triangle B) = \mu^* \varnothing = 0 < \varepsilon,$$

   therefore, $A \in \mathcal{M}(\mathbb{R}^n)$, and by the property 4) of the outer measure $\mu A = \mu^* A = mA$.

Lemma 3.4.3. *The following relations hold:*

a) $(A_1 \bigcup A_2) \vartriangle (B_1 \bigcup B_2) \subset (A_1 \vartriangle B_1) \bigcup (A_2 \vartriangle B_2)$,

b) $(A_1 \bigcap A_2) \vartriangle (B_1 \bigcap B_2) \subset (A_1 \vartriangle B_1) \bigcup (A_2 \vartriangle B_2)$,

c) $(A_1 \setminus A_2) \vartriangle (B_1 \setminus B_2) \subset (A_1 \vartriangle B_1) \bigcup (A_2 \vartriangle B_2)$,

d) $(A_1 \vartriangle A_2) \vartriangle (B_1 \vartriangle B_2) \subset (A_1 \vartriangle B_1) \bigcup (A_2 \vartriangle B_2)$.

Proof. We prove here the relations $a)$ and $c)$. The relations $b)$ and $d)$ can be proved analogously.

$a)$ Let $x \in (A_1 \bigcup A_2) \vartriangle (B_1 \bigcup B_2)$. Then either $x \in A_1 \bigcup A_2$ but $x \notin B_1 \bigcup B_2$, or $x \in B_1 \bigcup B_2$ but $x \notin A_1 \bigcup A_2$. If $x \in A_1 \bigcup A_2$ but $x \notin B_1 \bigcup B_2$, then $x \in A_1$ or $x \in A_2$ but $x \notin B_1$ *and* $x \notin B_2$. Therefore, $x \in A_1 \vartriangle B_1$ or $x \in A_2 \vartriangle B_2$. If $x \in B_1 \bigcup B_2$ but $x \notin A_1 \bigcup A_2$, then the inclusion $a)$ can be proved analogously.

$c)$ Let $x \in (A_1 \setminus A_2) \vartriangle (B_1 \setminus B_2)$. Then either $x \in A_1 \setminus A_2$ but $x \notin B_1 \setminus B_2$, or $x \notin A_1 \setminus A_2$ but $x \in B_1 \setminus B_2$. If $x \in A_1 \setminus A_2$ but $x \notin B_1 \setminus B_2$, then $x \in A_1$ but $x \notin A_2$, and either $x \in B_1$ and $x \in B_2$, or $x \notin B_1$. This means that either $x \in A_1 \vartriangle B_1$, or $x \in A_2 \vartriangle B_2$. If $x \notin A_1 \setminus A_2$ but $x \in B_1 \setminus B_2$, then the inclusion $c)$ can be prove analogously. □

Lemma 3.4.4. *For any $A_1, A_2 \in \mathcal{M}^*(\mathbb{R}^n)$, the following inequality holds*

$$|\mu^* A_1 - \mu^* A_2| \leqslant \mu^*(A_1 \vartriangle A_2) \tag{3.4.2}$$

Proof. It is easy to see that $A_1 \subset A_2 \bigcup (A_1 \vartriangle A_2)$. By additivity of the outer measure μ^*, one has

$$\mu^* A_1 \leqslant \mu^* A_2 + \mu^*(A_1 \vartriangle A_2),$$

that is,

$$\mu^* A_1 - \mu^* A_2 \leqslant \mu^*(A_1 \vartriangle A_2).$$

By interchanging A_1 and A_2, we get

$$\mu^* A_2 - \mu^* A_1 \leqslant \mu^*(A_1 \vartriangle A_2).$$

The last two inequalities give us (3.4.2), as required. □

Theorem 3.4.5. *The set $\mathcal{M}(\mathbb{R}^n)$ is a ring.*

Proof. Due to definition of rings, it is sufficient to prove that if $A_1, A_2 \in \mathcal{M}(\mathbb{R}^n)$, then $A_1 \bigcup A_2 \in \mathcal{M}(\mathbb{R}^n)$ and $A_1 \setminus A_2 \in \mathcal{M}(\mathbb{R}^n)$.

So, let $A_1, A_2 \in \mathcal{M}(\mathbb{R}^n)$. Then for an arbitrary $\varepsilon > 0$, one can find elementary sets B_1 and B_2 such that

$$\mu^*(A_1 \vartriangle B_1) < \frac{\varepsilon}{2} \quad \text{and} \quad \mu^*(A_2 \vartriangle B_2) < \frac{\varepsilon}{2}.$$

Let us denote $A := A_1 \bigcup A_2$ and $B := B_1 \bigcup B_2$. it is clear that $A_1, A_2 \in \mathcal{M}^*(\mathbb{R}^n)$, so $A \in \mathcal{M}^*(\mathbb{R}^n)$. The set B is an elementary set as the union of two elementary sets. By Lemma 3.4.4,

$$A \vartriangle B \subset (A_1 \vartriangle B_1) \bigcup (A_2 \vartriangle B_2).$$

Then by the property of semi-additivity of the outer measure, we obtain

$$\mu^*(A \vartriangle B) \leqslant \mu^*(A_1 \vartriangle B_1) + \mu^*(A_2 \vartriangle B_2) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

so $A = A_1 \bigcup A_2 \in \mathcal{M}(\mathbb{R}^n)$.

The proof of the fact that $A_1 \setminus A_2 \in \mathcal{M}(\mathbb{R}^n)$, can be proved in the same manner. □

Remark 3.4.6. In fact, using Lemma 3.4.4, it is possible to prove that if $A_1, A_2 \in \mathcal{M}(\mathbb{R}^n)$, then $A_1 \bigcap A_2 \in \mathcal{M}(\mathbb{R}^n)$ and $A_1 \triangle A_2 \in \mathcal{M}(\mathbb{R}^n)$. But this also follows from Theorem 3.4.5 and the properties 4) and 5) of rings.

Corollary 3.4.7. *The set $\mathcal{M}([0,1])$ of all Lebesgue measurable subsets of the interval $[0,1]$ is an algebra.*

Proof. Since $\mathcal{M}([0,1]) \subset \mathcal{M}(\mathbb{R})$ and $[0,1] \in \mathcal{M}([0,1])$, the set $\mathcal{M}([0,1])$ is a ring with the unit $[0,1]$. □

Theorem 3.4.8. *The function μ is additive on $\mathcal{M}(\mathbb{R}^n)$.*

Proof. We have to prove that if $A = A_1 \dot\cup A_2$, then

$$\mu A = \mu A_1 + \mu A_2. \tag{3.4.3}$$

So, let $A_1, A_2 \in \mathcal{M}(\mathbb{R}^n)$ and $A = A_1 \dot\cup A_2$. By Theorem 3.4.5 $A \in \mathcal{M}(\mathbb{R}^n)$. Since the outer measure is semi-additive by the property 3) of the outer measure, and since μ^* coincides with μ on $\mathcal{M}(\mathbb{R}^n)$, we have

$$\mu A \leqslant \mu A_1 + \mu A_2 \tag{3.4.4}$$

To prove the opposite inequality, let us fix a number $\varepsilon > 0$ and take elementary sets B_1 and B_2 such that

$$\mu^*(A_1 \triangle B_1) < \frac{\varepsilon}{6} \quad \text{and} \quad \mu^*(A_2 \triangle B_2) < \frac{\varepsilon}{6}. \tag{3.4.5}$$

Such sets B_1 and B_2 exist, since A_1 and A_2 are Lebesgue measurable by assumption.

Let us set $B := B_1 \bigcup B_2$. By assumption $A_1 \bigcap A_2 = \varnothing$, but B_1 and B_2 can have common elements. However, from Lemma 3.4.3 it follows that

$$B_1 \bigcap B_2 = (A_1 \bigcap A_2) \triangle (B_1 \bigcap B_2) \subset (A_1 \triangle B_1) \bigcup (A_2 \triangle B_2),$$

so

$$m(B_1 \bigcap B_2) = \mu^*(B_1 \bigcap B_2) \leqslant \mu^*(A_1 \triangle B_1) + \mu^*(A_2 \triangle B_2) < \frac{\varepsilon}{6} + \frac{\varepsilon}{6} = \frac{\varepsilon}{3}.$$

This inequality together with the property 5) of measures on rings imply

$$mB = mB_1 + mB_2 - m\left(B_1 \bigcap B_2\right) > mB_1 + mB_2 - \frac{\varepsilon}{3}. \tag{3.4.6}$$

Now, according to Lemma 3.4.4, we obtain

$$mB_1 \geqslant \mu A_1 - \mu^*(A_1 \triangle B_1) > \mu A_1 - \frac{\varepsilon}{6}, \tag{3.4.7}$$

and, analogously,

$$mB_2 > \mu A_2 - \frac{\varepsilon}{6}. \tag{3.4.8}$$

Recall the relation $A \triangle B \subset (A_1 \triangle B_2) \bigcup (A_1 \triangle B_2)$ (Lemma 3.4.3). By semi-additivity of the outer measure, one has

$$\mu^*(A \triangle B) \leqslant \mu^*(A_1 \triangle B_2) + \mu^*(A_1 \triangle B_2) < \frac{\varepsilon}{3},$$

so from Lemma 3.4.4 it follows that

$$\mu A \geqslant mB - \mu^*(A \triangle B) > mB - \frac{\varepsilon}{3}. \tag{3.4.9}$$

Now the estimates (3.4.6)–(3.4.9) imply

$$\mu A > mB - \frac{\varepsilon}{3} > mB_1 + mB_2 - \frac{2\varepsilon}{3} > \mu A_1 + \mu A_2 - \varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, we have

$$\mu A \geqslant \mu A_1 + \mu A_2.$$

This inequality together with (3.4.4) implies (3.4.3). □

Theorem 3.4.9. *The measure μ defined on the ring $\mathcal{M}(\mathbb{R}^n)$ is σ-additive.*

Proof. The measure μ is the restriction of the outer measure μ^* to the class $\mathcal{M}(\mathbb{R}^n)$. Since the function μ^* is σ-semi-additive, so is the measure μ on the ring $\mathcal{M}(\mathbb{R}^n)$. Now by Theorem 2.2.5, the measure μ is σ-additive. □

Theorem 3.4.10. *Let $A \in \mathcal{M}^*(\mathbb{R}^n)$ and $A = \bigcup\limits_{k=1}^{\infty} A_k$, where $A_k \in \mathcal{M}(\mathbb{R}^n)$, $k \in \mathbb{N}$. Then $A \in \mathcal{M}(\mathbb{R}^n)$.*

In other word, if the outer measure of a set is finite, and the set can be represented as a countable union of Lebesgue measurable sets, then it is Lebesgue measurable.

Proof. Let us set $A_0 = \varnothing$, and $A'_k = A_k \setminus \bigcup\limits_{j=1}^{k-1} A_j$, $k \in \mathbb{N}$. It is easy to see that by construction $A'_k \bigcap A'_l = \varnothing$ whenever $k \neq l$. At the same time, every point of A belongs to (at least) one of the sets A'_k. Therefore, $A = \dot{\bigcup\limits_{k=1}^{\infty}} A'_k$, and $A'_k \in \mathcal{M}(\mathbb{R}^n)$, since $\mathcal{M}(\mathbb{R}^n)$ is a ring by Theorem 3.4.5.

Now, since $\dot{\bigcup\limits_{k=1}^{l}} A'_k \subset A$ for any $l \in \mathbb{N}$, by the monotonicity of the outer measure (the property 2), and by additivity of the Lebesgue measure, we obtain

$$\sum_{k=1}^{l} \mu A'_k = \mu\left(\dot{\bigcup_{k=1}^{l}} A'_k\right) = \mu^*\left(\dot{\bigcup_{k=1}^{l}} A'_k\right) \leqslant \mu^* A < +\infty.$$

Thus, the partial sums of the series $\sum\limits_{k=1}^{\infty} \mu A'_k$ with nonnegative terms are bounded from above, so the series $\sum\limits_{k=1}^{\infty} \mu A'_k$ converges.

Let $\varepsilon > 0$ be an arbitrary positive number. Then there exists a number $m_0 \in \mathbb{N}$ such that

$$\sum_{k=m_0+1}^{\infty} \mu A'_k < \frac{\varepsilon}{2}.$$

The set $\dot{\bigcup\limits_{k=1}^{m_0}} A'_k$ is Lebesgue measurable as a finite union of Lebesgue measurable sets. Therefore, there exists an elementary set B such that

$$\mu^*\left(\left(\dot{\bigcup_{k=1}^{m_0}} A'_k\right) \vartriangle B\right) < \frac{\varepsilon}{2}$$

It is easy to check that

$$A \vartriangle B \subset \left(\left(\dot{\bigcup_{k=1}^{m_0}} A'_k\right) \vartriangle B\right) \bigcup \left(\dot{\bigcup_{k=m_0+1}^{\infty}} A'_k\right).$$

Thus, from the property of σ-additivity of the outer measure, we obtain

$$\mu^*(A \vartriangle B) \leqslant \mu^*\left(\left(\dot{\bigcup_{k=1}^{m_0}} A'_k\right) \vartriangle B\right) + \sum_{k=m_0+1}^{\infty} \mu A'_k < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} < \varepsilon,$$

as required. □

Corollary 3.4.11. *The union of countably many null sets (sets whose Lebesgue measure is 0) has Lebesgue measure 0, that is, it is a null set.*

Proof. Let $A = \bigcup\limits_{k=1}^{\infty} A_k$, where $A_k \in \mathcal{M}(\mathbb{R}^n)$ and $\mu(A_k) = 0$, $k \in \mathbb{N}$. Then due to σ-semi-additivity of the outer measure, we have

$$\mu^*(A) \leqslant \sum_{k=1}^{\infty} \mu(A_k) = 0.$$

Now by Theorem 3.4.10, the set A is Lebesgue measurable, so its Lebesgue measure coincides with its outer measure and equals 0. □

Definition 3.4.12. A measure μ defined on a ring $\mathcal{K}$ is called *complete* if every subset of a null set is measurable w.r.t. the measure μ.

Thus, if a measure μ defined on a ring $\mathcal{K}$ is complete, then by the monotonicity of measures, every subset of a null set is a null set.

Theorem 3.4.13. *The Lebesgue measure is complete.*

Proof. Let $\mu A = 0$, $A \in \mathcal{M}(\mathbb{R})$, and let $A_0 \subset A$. Then for any $\varepsilon > 0$ and for $B = \varnothing \in \mathcal{E}^n$, we have

$$\mu^*(A_0 \vartriangle B) = \mu^*(A_0 \vartriangle \varnothing) = \mu^* A_0 \leqslant \mu^* A = \mu A = 0 < \varepsilon.$$

Thus, by Definition 3.4.1, the set A_0 is Lebesgue measurable. □

3.5 Extension of the notion of measurability. The class of measurable sets.

The Lebesgue measure defined above possesses a substantial defect. Namely, only sets with finite (outer) measure can be Lebesgue measurable. So, many "good" infinite sets left unmeasurable w.r.t. our Lebesgue measure (e.g. $\mathbb{R}^n$, quadrants, strips, interiors of parabolas, etc). Let us eliminate this defect.

Let $\widehat{Q}_l$ be an n-dimensional cube with the centre at the origin and with edge of length $2l$, that is,

$$\widehat{Q}_l = \{x = (x_i)_{i=1}^n : \ |x_i| \leqslant l, \ i = 1, 2, \ldots, n\}, \quad l \in \mathbb{R}. \tag{3.5.1}$$

Definition 3.5.1. A set $A \subset \mathbb{R}^n$ is called Lebesgue measurable in extended sense, or σ-measurable, if for any $l \in \mathbb{N}$ the set $A \bigcap \widehat{Q}_l$ is Lebesgue measurable.

If a set A is σ-measurable, then we put

$$\mu A := \lim_{l \to \infty} \mu\left(A \bigcap \widehat{Q}_l\right). \tag{3.5.2}$$

The class of all σ-measurable sets is denoted as $\mathcal{M}_\sigma(\mathbb{R}^n)$.

Since the sequences $A \bigcap \widehat{Q}_l$ increases as l grows, then the sequences $\mu\left(A \bigcap \widehat{Q}_l\right)$ also increases. Therefore, the limit in (3.5.2) exists (but can be infinite).

Theorem 3.5.2. $\mathcal{M}(\mathbb{R}^n) \subset \mathcal{M}_\sigma(\mathbb{R}^n)$.

Proof. Let $A \in \mathcal{M}(\mathbb{R}^n)$. Then for any $l \in \mathbb{N}$, the set $A \bigcap \widehat{Q}_l$ is a Lebesgue measurable set as an intersection of two Lebesgue measurable sets. Then by Definition 3.5.1, the set A is σ-measurable. □

Theorem 3.5.3. *If for a set $A \in \mathcal{M}_\sigma(\mathbb{R}^n)$ the limit* (3.5.2) *is finite, then $A \in \mathcal{M}(\mathbb{R}^n)$.*

Proof. Let $A_l := A \bigcap \widehat{Q}_l$, $l \in \mathbb{N}$. The sequence A_l is increasing, and

$$A = \bigcup_{l=1}^{\infty} A_l = \dot{\bigcup_{l=1}^{\infty}} (A_l \setminus A_{l-1}), \qquad (A_0 = \varnothing).$$

Since $A \in \mathcal{M}_\sigma(\mathbb{R}^n)$, the sets A_l are Lebesgue measurable, and since $\mathcal{M}(\mathbb{R}^n)$ is a ring, $A_l \setminus A_{l-1} \in \mathcal{M}(\mathbb{R}^n)$ for any $l \in \mathbb{N}$.

The existence of finite limit (3.5.2), that is, of $\lim\limits_{l\to\infty} \mu A_l$, means that the following series converges

$$\sum_{l=1}^{\infty} \mu(A_l \setminus A_{l-1}),$$

since

$$\sum_{l=1}^{\infty} \mu(A_l \setminus A_{l-1}) = \lim_{l\to\infty} \sum_{k=1}^{l} \mu(A_k \setminus A_{k-1}) = \lim_{l\to\infty} \sum_{k=1}^{l} (\mu A_k - \mu A_{k-1}) = \lim_{l\to\infty} \mu A_l.$$

Now from σ-additivity of the outer measure, we have

$$\mu^* A \leqslant \sum_{l=1}^{\infty} \mu^*(A_l \setminus A_{l-1}) = \sum_{l=1}^{\infty} \mu(A_l \setminus A_{l-1}) < +\infty.$$

Therefore, $A \in \mathcal{M}^*(\mathbb{R}^n)$, so $A \in \mathcal{M}(\mathbb{R}^n)$ by Theorem 3.4.10. □

Theorem 3.5.3 shows that extension of $\mathcal{M}(\mathbb{R}^n)$ to $\mathcal{M}_\sigma(\mathbb{R}^n)$ is made by adding the sets of infinite measure to $\mathcal{M}(\mathbb{R}^n)$.

Theorem 3.5.4. *The set $\mathcal{M}_\sigma(\mathbb{R}^n)$ is a σ-algebra with the unit $\mathbb{R}^n$.*

Proof. First, we prove that $\mathcal{M}_\sigma(\mathbb{R}^n)$ is a ring. Let $A, B \in \mathcal{M}_\sigma(\mathbb{R}^n)$, then for any $l \in \mathbb{N}$, the sets $A \bigcap \widehat{Q}_l$ and $B \bigcap \widehat{Q}_l$ are Lebesgue measurable. But

$$(A \bigcup B) \bigcap \widehat{Q}_l = (A \bigcap \widehat{Q}_l) \bigcup (B \bigcap \widehat{Q}_l) \in \mathcal{M}(\mathbb{R}^n),$$

and

$$(A \setminus B) \bigcap \widehat{Q}_l = (A \bigcap \widehat{Q}_l) \setminus (B \bigcap \widehat{Q}_l) \in \mathcal{M}(\mathbb{R}^n).$$

So $A \bigcup B, A \setminus B \in \mathcal{M}_\sigma(\mathbb{R}^n)$. Obviously, $\mathbb{R}^n$ is the unit of $\mathcal{M}_\sigma(\mathbb{R}^n)$, therefore, $\mathcal{M}_\sigma(\mathbb{R}^n)$ is an algebra. It is left to show that it is a σ-algebra.

Consider $A_k \in \mathcal{M}_\sigma(\mathbb{R}^n)$, $k \in \mathbb{N}$, and put $A := \bigcup\limits_{k=1}^{\infty} A_k$. Then for any $l \in \mathbb{N}$, one has $A_k \bigcap \widehat{Q}_l \in \mathcal{M}(\mathbb{R}^n)$, and

$$A \bigcap \widehat{Q}_l = \left(\bigcup_{k=1}^{\infty} A_k \right) \bigcap \widehat{Q}_l = \bigcup_{k=1}^{\infty} (A_k \bigcap \widehat{Q}_l) \subset \widehat{Q}_l.$$

Since $\mu^*(\widehat{Q}_l) < +\infty$, by monotonicity of the outer measure we obtain $A \bigcap \widehat{Q}_l \in \mathcal{M}^*(\mathbb{R}^n)$. Now from Theorem 3.4.10 it follows that $A \bigcap \widehat{Q}_l \in \mathcal{M}(\mathbb{R}^n)$, so $\mathcal{M}_\sigma(\mathbb{R}^n)$ is a σ-algebra. □

Corollary 3.5.5. *The set $\mathcal{M}([0,1])$ of all Lebesgue measurable subsets of the interval $[0,1]$ is a σ-algebra.*

Proof. By Corollary 3.4.7, $\mathcal{M}([0,1])$ is an algebra. If $A_k \in \mathcal{M}([0,1])$, $k \in \mathbb{N}$, then

$$A = \bigcup_{k=1}^{\infty} A_k \subset [0,1],$$

so, as in the proof of Theorem 3.5.4, $\mu^*(A) \leqslant \mu^*([0,1]) = \mu([0,1]) = 1$. Thus, $A \in \mathcal{M}^*([0,1])$, and according to Theorem 3.4.10, one has $A \in \mathcal{M}([0,1])$, as required. □

Remark 3.5.6. The corollary is true for $\mathcal{M}(X)$ where X is a compact set in $\mathbb{R}^n$.

Definition 3.5.7. The triple $(X, \mathcal{M}, \widetilde{\mu})$ is called a measure space. If $\widetilde{\mu}(X) < +\infty$, then the measure $\widetilde{\mu}$ is called *finite*. If additionally $\mu(X) = 1$, it is called a *probability measure*. If $\mu(X) = +\infty$, but there exist sets X_i such that $\widetilde{\mu}(X_i) < +\infty$, $i \in \mathbb{N}$, and $X = \bigcup_{i=1}^{\infty} X_i$, then the measure $\widetilde{\mu}$ is called *σ-finite*.

The Lebesgue measure on $\mathcal{M}_\sigma(\mathbb{R}^n)$ is σ-finite.

Now let us discuss how large the class $\mathcal{M}_\sigma(\mathbb{R}^n)$ is.

Definition 3.5.8. Let K_1 and K_2 be two bricks, and let K_1° and K_2° denote their interiors, respectively. We say that bricks K_1 and K_2 are *almost disjoint* if $K_1^\circ \bigcap K_2^\circ = \varnothing$, meaning that they intersect at most along their boundaries.

Theorem 3.5.9. *Every open set in $\mathbb{R}^n$, $n \geqslant 1$, can represented as a countable union of almost disjoint closed bricks.*

Proof. Let $A \subset \mathbb{R}^n$ be open. We construct a family of closed cubes (bricks of equal sides) as follows. First, we bisect $\mathbb{R}^n$ into almost disjoint closed cubes $\{Q_i : \ i \in \mathbb{N}\}$ of side one with integer coordinates. If $Q_i \subset A$, we include Q_i in the family, and if Q_i is disjoint from A, we exclude it. Otherwise, we bisect the sides of Q_i to obtain 2^n almost disjoint closed cubes of side one-half and repeat the procedure. Iterating this process arbitrarily many times, we obtain a countable family of almost disjoint closed cubes.

The union of the cubes in this family is contained in A, since we only include cubes that are contained in A. Conversely, if $x \in A$, then since A is open some sufficiently small cube in the bisection procedure that contains x is entirely contained in A, and the largest such cube is included in the family. Hence the union of the family contains A, and is therefore equal to A. □

From this theorem and from Theorem 3.5.4, we now obtain the following fact.

Corollary 3.5.10. *Any open set in $\mathbb{R}^n$ is σ-measurable.*

Moreover, it is clear now that closed sets (as complements of open sets), countable and finite unions and intersections of open and closed sets in $\mathbb{R}^n$ are σ-measurable.

Recall that for $\mathbb{R}$ Theorem 3.5.9 has a more improved version, see Theorem 1.4.17.

3.6 Borel σ-algebra

Definition 3.6.1. The Borel σ-algebra $\mathcal{B}(\mathbb{R}^n)$ on $\mathbb{R}^n$ is the σ-algebra generated by the open sets, that is,

$$\mathcal{B}(\mathbb{R}^n) := \sigma(\mathcal{T}(\mathbb{R}^n)) = \bigcap \{\mathcal{A} \subset \mathcal{P}(\mathbb{R}^n) : \ \mathcal{A} \supset \mathcal{T}(\mathbb{R}^n) \ \text{ and } \ \mathcal{A} \ \text{ is a } \sigma\text{-algebra}\},$$

where $\mathcal{T}(\mathbb{R}^n)$ is the set of all open sets in $\mathbb{R}^n$.

A set that belongs to the Borel σ-algebra is called a *Borel set*.

Thus, the Borel σ-algebra is the *minimal* σ-algebra on $\mathbb{R}^n$ that contains all opens sets.

Since σ-algebras are closed under complementation, the Borel σ-algebra is also generated by the closed sets in $\mathbb{R}^n$. Moreover, since $\mathbb{R}^n$ is σ-compact (i.e. it is a countable union of compact sets) its Borel σ-algebra is generated by the compact sets.

So, all the finite and countable union and intersections of opens and closed sets on $\mathbb{R}^n$ (taken in arbitrary order) are Borel sets.

Proposition 3.6.2. *The Borel algebra $\mathcal{B}(\mathbb{R}^n)$ is generated by the collection of closed bricks $\mathcal{K}(\mathbb{R}^n)$:*

$$\mathcal{B}(\mathbb{R}^n) = \sigma(\mathcal{K}(\mathbb{R}^n)) := \bigcap \{\mathcal{A} \subset \mathcal{P}(\mathbb{R}^n) : \ \mathcal{A} \supset \mathcal{K}(\mathbb{R}^n) \ \textit{ and } \ \mathcal{A} \ \textit{ is a } \sigma\textit{-algebra}\}.$$

Every Borel set is Lebesgue measurable.

Proof. Since $\mathcal{K}(\mathbb{R}^n)$ is a subset of the set of closed sets, we have $\sigma(\mathcal{K}(\mathbb{R}^n)) \subset \mathcal{B}(\mathbb{R}^n)$. Conversely, by Theorem 3.5.9, $\sigma(\mathcal{K}(\mathbb{R}^n)) \supset \mathcal{T}(\mathbb{R}^n)$, so $\sigma(\mathcal{K}(\mathbb{R}^n)) \supset \sigma(\mathcal{T}(\mathbb{R}^n)) = \mathcal{B}(\mathbb{R}^n)$, and therefore $\mathcal{B}(\mathbb{R}^n) = \sigma(\mathcal{K}(\mathbb{R}^n))$. From the property 2) of Lebesgue measures and from Theorem 3.5.2, we obtain $\mathcal{K}(\mathbb{R}^n) \subset \mathcal{M}_\sigma(\mathbb{R}^n)$. Since $\mathcal{M}_\sigma(\mathbb{R}^n)$ is a σ-algebra, it follows that $\sigma(\mathcal{K}(\mathbb{R}^n)) \subset \mathcal{M}_\sigma(\mathbb{R}^n)$, so $\mathcal{B}(\mathbb{R}^n) \subset \mathcal{M}_\sigma(\mathbb{R}^n)$. □

Note that if

$$A = \bigcup_{j=1}^{\infty} K_j$$

is a decomposition of an open set A into a union of almost disjoint closed bricks, then

$$A \supset \dot{\bigcup_{j=1}^{\infty}} K_j^\circ$$

is a disjoint union, and therefore

$$\sum_{j=1}^{\infty} \mu(K_j^\circ) \leqslant \mu(A) \leqslant \sum_{j=1}^{\infty} \mu(K_j)$$

Since $\mu(K_k^\circ) = \mu(K_j)$, it follows that

$$\mu(A) = \sum_{j=1}^{\infty} \mu(K_j) \tag{3.6.1}$$

for any such decomposition and that the sum is independent of the way in which A is decomposed into almost disjoint rectangles.

The Borel σ-algebra $\mathcal{B}(\mathbb{R}^n)$ is not complete and is strictly smaller than the Lebesgue σ-algebra $\mathcal{M}_\sigma(\mathbb{R}^n)$. In fact, one can show that the cardinality of $\mathcal{B}(\mathbb{R}^n)$ is equal to the cardinality c of the real numbers, whereas the cardinality of $\mathcal{M}_\sigma(\mathbb{R}^n)$ is equal to 2^c. For example, the Cantor set is a set of measure zero with the same cardinality as $\mathbb{R}$ and every subset of the Cantor set is Lebesgue measurable (see Homework). We can obtain examples of sets that are Lebesgue measurable but not Borel measurable by considering subsets of sets of measure zero.

Examples of Lebesgue measurable sets that are not Borel sets may also arise from the theory of product measures in $\mathbb{R}^n$ for $n \geqslant 2$. For example, let $N = E \times \{0\} \subset \mathbb{R}^2$ where $E \subset \mathbb{R}$ is a non-Lebesgue measurable set in $\mathbb{R}$. Then N is a subset of the x-axis, which has two-dimensional Lebesgue measure zero, so N belongs to $\mathcal{M}_\sigma(\mathbb{R}^2)$ since Lebesgue measure is complete. One can show, however, that if a set belongs to $\mathcal{B}(\mathbb{R}^2)$ then every section with fixed x or y coordinate belongs to $\mathcal{B}(\mathbb{R})$; thus, N cannot belong to $\mathcal{B}(\mathbb{R}^2)$ since the $y = 0$ section E is not Borel (because it is not Lebesgue measurable).

3.7 Borel regularity

Regularity properties of measures refer to the possibility of approximating in measure one class of sets (for example, non-measurable sets) by another class of sets (for example, measurable sets). Lebesgue measure is Borel regular in the sense that Lebesgue measurable sets can be approximated in measure from the outside by open sets and from the inside by closed sets, and they can be approximated by Borel sets up to sets of measure zero. Moreover, there is a simple criterion for Lebesgue measurability in terms of open and closed sets. The following theorem expresses a fundamental approximation property of Lebesgue measurable sets by open and compact sets. Equations (3.7.1) and (3.7.2) are called outer and inner regularity, respectively.

Theorem 3.7.1. *If $A \subset \mathbb{R}^n$, then*

$$\mu^*(A) = \inf\{\mu(G) : \ A \subset G, \ G \ \textit{open}\} \tag{3.7.1}$$

and if $A \in \mathcal{M}_\sigma(\mathbb{R}^n)$, then

$$\mu(A) = \sup\{\mu(F) : \ F \subset A, \ F \ \textit{compact}\} \tag{3.7.2}$$

Proof. First, we prove (3.7.1). If $\mu^*(A) = +\infty$, we are done. Suppose now that $\mu^*(A)$ is finite. If $A \subset G$, then $\mu^*(A) \leqslant \mu^*(G)$ by the property 2) of the outer measure, so

$$\mu^*(A) \leqslant \inf\{\mu(G) : \ A \subset G, \ G \text{ open}\}$$

It is left to prove the reverse inequality

$$\mu^*(A) \geqslant \inf\{\mu(G) : \ A \subset G, \ G \text{ open}\} \tag{3.7.3}$$

Let $\varepsilon > 0$. By Definition 3.3.1 of the outer measure, there exists a cover of A by bricks K_i, $i \in \mathbb{N}$, such that

$$\sum_{i=1}^{\infty} \mu(K_i) \leqslant \mu^*(A) + \frac{\varepsilon}{2}$$

Let $\widetilde{K}_i$ be an *open* brick such that $K_i \subset \widetilde{K}_i$ and

$$\mu(\widetilde{K_i}) \leqslant \mu(K_i) + \frac{\varepsilon}{2^{i+1}}. \tag{3.7.4}$$

Then the collection of open bricks $\widetilde{K}_i$, $i \in \mathbb{N}$, covers A and

$$G := \bigcup_{i=1}^{\infty} \widetilde{K}_i$$

is an open set that contains A. Moreover, by (3.7.4) and by σ-semi-additivity of the Lebesgue measure, one has

$$\mu(G) \leqslant \sum_{i=1}^{\infty} \mu(\widetilde{K}_i) \leqslant \sum_{i=1}^{\infty} \mu(K_i) + \frac{\varepsilon}{2},$$

and therefore

$$\mu(G) \leqslant \mu^*(A) + \varepsilon, \tag{3.7.5}$$

which proves (3.7.3) since $\varepsilon > 0$ is arbitrary.

Next, we prove (3.7.2). If $F \subset A$, then by property 2) of the outer measure $\mu(F) \leqslant \mu(A)$, so

$$\sup\{\mu(F) : \ F \subset A, \ F \text{ compact}\} \leqslant \mu(A).$$

Therefore, we just need to prove the reverse inequality

$$\mu(A) \leqslant \sup\{\mu(F) : \ F \subset A, \ F \text{ compact}\}. \tag{3.7.6}$$

To do this we apply the previous result to the complement A^c and use the measurability of A.

First, suppose that A is a bounded measurable set, so $\mu(A) < +\infty$ since there exists a brick such that $A \subset K$, and $\mu(A) \leqslant \mu(K) < +\infty$. Let $H \subset \mathbb{R}^n$ be a compact set that contains A. By the preceding result, for any $\varepsilon > 0$, there is an open set $G \supset H \setminus A$ such that

$$\mu(G) \leqslant \mu(H \setminus A) + \varepsilon \tag{3.7.7}$$

Then $F = H \setminus G$ is a compact set such that $F \subset A$. Moreover, $H \subset F \bigcup G$ and $H = A \,\dot\cup\, (H \setminus A)$, so

$$\mu(H) \leqslant \mu(F) + \mu(G), \qquad \mu(H) = \mu(A) + \mu(H \setminus A). \tag{3.7.8}$$

It follows from (3.7.7)–(3.7.8) that

$$\mu(A) = \mu(H) - \mu(H \setminus A) \leqslant \mu(H) - \mu(G) + \varepsilon \leqslant \mu(F) + \varepsilon,$$

which implies (3.7.6) and proves the result for bounded measurable sets.

Now suppose that $A \in \mathcal{M}_\sigma(\mathbb{R}^n)$ is unbounded, and consider $A_l := A \bigcap \widehat{Q}_l$, $l \in \mathbb{R}^n$, where $\widehat{Q}_l$ is defined in (3.5.1). By Definition 3.5.1 (see (3.5.2)), we have

$$\mu(A) = \lim_{l \to +\infty} \mu(A_l). \tag{3.7.9}$$

If $\mu(A) = +\infty$, then $\mu(A_l) \to +\infty$ as $l \to +\infty$. Since A_l is bounded and measurable, by previous result, we can find a compact set $F_l \subset A_l \subset A$ such that

$$\mu(F_l) + 1 \geqslant \mu(A_l),$$

so that $\mu(F_k) \to +\infty$. Therefore,

$$\sup\{\mu(F) : \ F \subset A, \ F \text{ compact}\} = +\infty,$$

which proves the result in this case.

Finally, suppose that A is unbounded and $A \in \mathcal{M}(\mathbb{R}^n)$, so $\mu(A) < +\infty$. From (3.7.9), for any $\varepsilon > 0$ one can choose $l \in \mathbb{N}$ such that ($A_l \subset A_{l+1} \subset A$ for any l)

$$\mu(A) \leqslant \mu(A_l) + \frac{\varepsilon}{2}.$$

Moreover, since A_l is bounded, there is a compact set $F \subset A_l$ such that

$$\mu(A_l) \leqslant \mu(F) + \frac{\varepsilon}{2}.$$

Therefore, for any $\varepsilon > 0$, there exists a compact set $F \subset A$ such that

$$\mu(A) \leqslant \mu(F) + \varepsilon,$$

which gives (3.7.6) and completes the proof. □

It follows that we may determine the Lebesgue measure of a measurable set in terms of the Lebesgue measure of open or compact sets by approximating the set from the outside by open sets or from the inside by compact sets. The outer approximation in (3.7.1) does not require that A is measurable. Thus, for any set $A \subset \mathbb{R}^n$, given $\varepsilon > 0$, we can find an open set $G \supset A$ such that $\mu(G) - \mu^*(A) < \varepsilon$. If A is measurable, we can strengthen this condition to get that $\mu^*(G \setminus A) < \varepsilon$; in fact, this gives a necessary and sufficient condition for measurability.

Theorem 3.7.2. *$A \in \mathcal{M}_\sigma(\mathbb{R}^n)$ if and only if for every $\varepsilon > 0$ there is an open set $G \supset A$ such that*

$$\mu^*(G \setminus A) < \varepsilon \tag{3.7.10}$$

Proof. First we assume that A is σ-measurable and show that it satisfies the condition given in the theorem.

Suppose that $\mu(A) < +\infty$ and let $\varepsilon > 0$. From (3.7.5) there is an open set $G \supset A$ such that $\mu(G) \leqslant \mu^*(A) + \varepsilon$. Then, since A is measurable, $G \setminus A$ is measurable and

$$\mu^*(G \setminus A) = \mu(G \setminus A) = \mu(G) - \mu\left(G \bigcap A\right) = \mu(G) - \mu\left(A\right) = \mu(G) - \mu^*\left(A\right) < \varepsilon,$$

which proves the result when A has finite measure.

If $\mu(A) = +\infty$, define $A_l = A \bigcap \widehat{Q}_l \subset A$ with $\widehat{Q}_l$ given by (3.5.1), and let $\varepsilon > 0$. Since A_l is measurable with finite measure, the argument above shows that for each $l \in \mathbb{N}$, there is an open set $G_l \supset A_l$ such that

$$\mu(G_l \setminus A_l) < \frac{\varepsilon}{2^l}.$$

Then $G = \bigcup\limits_{l=1}^{\infty} G_l$ is an open set that contains A, and

$$\mu^*(G \setminus A) = \mu^*\left(\bigcup_{l=1}^{\infty}(G_l \setminus A)\right) \leqslant \sum_{l=1}^{\infty} \mu^*\left(G_l \setminus A\right) \leqslant \sum_{l=1}^{\infty} \mu^*\left(G_l \setminus A_l\right) < \varepsilon.$$

Conversely, suppose that $A \subset \mathbb{R}^n$ satisfies the condition of the theorem. Let $\varepsilon > 0$, and choose an open set G such that $\mu^*(G \setminus A) < \frac{\varepsilon}{2}$.

If $\mu^*(A) < +\infty$, then due to semi-additivity of the outer measure

$$\mu(G) = \mu^*(G) = \mu^*\left(A \dot{\bigcup} (G \setminus A)\right) \leqslant \mu^*(A) + \mu^*(G \setminus A) < \mu^*(A) + \frac{\varepsilon}{2},$$

So $\mu(G) < +\infty$. By Theorem 3.5.9, one can represent $G = \bigcup\limits_{i=1}^{\infty} K_i$, where K_i are almost disjoint closed bricks, and

$$\mu(G) = \sum_{i=1}^{\infty} m(K_i) < +\infty.$$

Consequently, there exists a number $i_0 \in \mathbb{N}$ such that

$$\sum_{i=i_0+1}^{\infty} m(K_i) < \frac{\varepsilon}{2}$$

Let $B := \bigcup\limits_{i=1}^{i_0} K_i \in \mathcal{E}^n$. Then

$$\mu^*(A \vartriangle B) = \mu^*\Big((A \setminus B) \bigcup (B \setminus A)\Big) \leqslant \mu^*(A \setminus B) + \mu^*(B \setminus A). \tag{3.7.11}$$

On the other hand, since $A \subset G$, we have

$$A \setminus B = A \setminus \left(\bigcup_{i=1}^{i_0} K_i\right) \subset \bigcup_{i=i_0+1}^{\infty} K_i,$$

and

$$B \setminus A = \left(\bigcup_{i=1}^{i_0} K_i\right) \setminus A \subset G \setminus A.$$

Now from (3.7.11) we obtain

$$\mu^*(A \vartriangle B) \leqslant \mu^*(A \setminus B) + \mu^*(B \setminus A) \leqslant \mu^*\left(\bigcup_{i=i_0+1}^{\infty} K_i\right) + \mu^*(G \setminus A) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Thus, $A \in \mathcal{M}_\sigma(\mathbb{R}^n)$.

Suppose now that $\mu^*(A) = +\infty$. We have to prove that $A_l = A \bigcap \widehat{Q}_l \in \mathcal{M}(\mathbb{R}^n)$, $l \in \mathbb{N}$, where $\widehat{Q}_l$ is defined in (3.5.1). Any cube $\widehat{Q}_l$ is closed by definition. Let us denote by $\widehat{Q}_l^\circ$ its interior. Then the set $G_l := G \bigcap \widehat{Q}_l^\circ$ is open, and we have

$$G_l \setminus A_l = \left(G \bigcap \widehat{Q}_l^\circ\right) \setminus \left(A \bigcap \widehat{Q}_l\right) \subset \left(G \bigcap \widehat{Q}_l\right) \setminus \left(A \bigcap \widehat{Q}_l\right) = (G \setminus A) \bigcap \widehat{Q}_l \subset G \setminus A.$$

From semi-additivity of the outer measure we obtain

$$\mu^*(G_l \setminus A_l) \leqslant \mu^*(G \setminus A) < \frac{\varepsilon}{2}, \qquad l \in \mathbb{N}.$$

Therefore, A_l is a bounded set with finite outer measure for which there exists an open set G_l such that $\mu^*(G_l \setminus A_l) < \frac{\varepsilon}{2}$. By the previous result, $A_l \in \mathcal{M}(\mathbb{R}^n)$, so $A \in \mathcal{M}_\sigma(\mathbb{R}^n)$. □

The following theorem gives another characterization of Lebesgue measurable sets, as ones that can be "squeezed" between open and closed sets.

Theorem 3.7.3. *$A \in \mathcal{M}_\sigma(\mathbb{R}^n)$ if, and only if, for every $\varepsilon > 0$ there is an open set G and a closed set F such that $F \subset A \subset G$, and*

$$\mu(G \setminus F) < \varepsilon \tag{3.7.12}$$

If $\mu(A) < +\infty$, then F may be chosen to be compact.

Proof. If for a given $A \subset \mathbb{R}^n$, for any $\varepsilon > 0$ there exist an open set $G \supset A$ and a closed set $F \subset A$ such that (3.7.12) satisfied, then by monotonicity of the outer measure we have $\mu^*(G \setminus A) \leqslant \mu^*(G \setminus F) < \varepsilon$, so $A \in \mathcal{M}_\sigma(\mathbb{R}^n)$ according to Theorem 3.7.2.

Conversely, let $A \in \mathcal{M}_\sigma(\mathbb{R}^n)$. Then $A^c \in \mathcal{M}_\sigma(\mathbb{R}^n)$, and by Theorem 3.7.2 given $\varepsilon > 0$, there exist open sets $G \supset A$ and $H \supset A^c$ such that

$$\mu^*(G \setminus A) < \frac{\varepsilon}{2}, \qquad \mu^*(H \setminus A^c) < \frac{\varepsilon}{2}.$$

Then, defining the closed set $F := H^c$, we have $G \supset A \supset F$ and

$$\mu(G \setminus F) \leqslant \mu^*(G \setminus A) + \mu^*(A \setminus F) = \mu^*(G \setminus A) + \mu^*(H \setminus A^c) < \varepsilon.$$

Finally, suppose that $\mu(A) < +\infty$ and let $\varepsilon > 0$. According to Theorem 3.7.1, there exists a compact set $F \subset A$ such that $\mu(A) < \mu(F) + \dfrac{\varepsilon}{2}$, and

$$\mu(A \setminus F) = \mu(A) - \mu(F) < \frac{\varepsilon}{2}.$$

As before, from Theorem 3.7.2, there is an open set $G \supset A$ such that

$$\mu(G) < \mu(A) + \frac{\varepsilon}{2}.$$

It follows that $G \supset A \supset F$, and

$$\mu(G \setminus F) = \mu(G \setminus A) + \mu(A \setminus F) < \varepsilon,$$

which shows that we may take F compact when $A \in \mathcal{M}(\mathbb{R}^n)$. □

From the previous results, we can approximate measurable sets by open or closed sets, up to sets of arbitrarily small but, in general, nonzero measure. By taking countable intersections of open sets or countable unions of closed sets, we can approximate measurable sets by Borel sets, up to sets of measure zero.

Definition 3.7.4. The collection of sets in $\mathbb{R}^n$ that are countable intersections of open sets is denoted by $G_\delta(\mathbb{R}^n)$, and the collection of sets in $\mathbb{R}^n$ that are countable unions of closed sets is denoted by $F_\sigma(\mathbb{R}^n)$.

$G_\delta(\mathbb{R}^n)$ and $F_\sigma(\mathbb{R}^n)$ sets are Borel. Thus, it follows from the next result that every Lebesgue measurable set can be approximated up to a set of measure zero by a Borel set. This is the Borel regularity of Lebesgue measure.

Theorem 3.7.5. *Suppose that $A \in \mathcal{M}_\sigma(\mathbb{R}^n)$. Then there exist sets $G \in G_\delta(\mathbb{R}^n)$ and $F \in F_\sigma(\mathbb{R}^n)$ such that*

$$G \supset A \supset F, \qquad \mu(G \setminus A) = \mu(A \setminus F) = 0.$$

Proof. For each $k \in \mathbb{N}$, choose an open set G_k and a closed set F_k such that $G_k \supset A \supset F_k$ and

$$\mu(G_k \setminus F_k) < \frac{1}{k}.$$

Then

$$G := \bigcap_{k=1}^{\infty} G_k, \qquad F := \bigcup_{k=1}^{\infty} F_k$$

are G_δ and F_σ sets with the required properties. □

As a corollary of this result, we get that the Lebesgue σ-algebra is the completion of the Borel σ-algebra w.r.t. Lebesgue measure in the sense that $\mathcal{M}_\sigma(\mathbb{R}^n)$ is the space of σ-Lebesgue measurable sets that differs from $\mathcal{B}(\mathbb{R}^n)$ only by subsets of Borel null sets, and $\mathcal{M}_\sigma(\mathbb{R}^n)$ is complete w.r.t. Lebesgue measure.

Theorem 3.7.6. *The Lebesgue σ-algebra $\mathcal{M}_\sigma(\mathbb{R}^n)$ is the completion of the Borel σ-algebra $\mathcal{B}_\sigma(\mathbb{R}^n)$.*

Proof. Lebesgue measure is complete from Theorem 3.4.13. By the previous theorem, if $A \in \mathcal{M}_\sigma(\mathbb{R}^n)$, then there is a F_σ set $F \subset A$ such that $M = A \setminus F$ has Lebesgue measure zero. It follows by the same theorem that there is a Borel set $N \in G_\delta$ with $\mu(N) = 0$ and $M \subset N$. Thus, $A = F \bigcup M$ where $F \in \mathcal{B}(\mathbb{R}^n)$ and $M \subset N \in \mathcal{B}(\mathbb{R}^n)$ with $\mu(N) = 0$, which proves that $\mathcal{M}_\sigma(\mathbb{R}^n)$ is the completion of $\mathcal{B}(\mathbb{R}^n)$. $\square$

3.8 Invariance properties of Lebesgue measure and non-Lebesgue measurable sets

A crucial property of Lebesgue measure in $\mathbb{R}^n$ is its translation-invariance, which can be stated as follows: if A is a measurable set and $h \in \mathbb{R}^n$, then the set $A + h = \{x + h : \ x \in A\}$ is also measurable, and $\mu(A + h) = \mu(A)$. The invariance of outer measure μ^* is an immediate consequence of Definition 3.3.1, since $\{K_i + h : \ i \in \mathbb{N}\}$ is a cover of $A + h$ if and only if $\{K_i : \ i \in \mathbb{N}\}$ is a cover of A, and $\mu(K + h) = \mu(K)$ for every brick K. Clearly, $\mu(B + h) = \mu B$ for any elementary set B. To prove the measurability of $A + h$ under the assumption that A is measurable, we note that for any $\varepsilon > 0$, one can find an elementary set B such that $\mu^*(A \bigtriangleup B) < \varepsilon$. And since μ^* is translation-invariant (as we proved above), one has $\mu^*((A + h) \bigtriangleup (B + h)) < \varepsilon$.

In the same way one can prove the relative dilation-invariance of Lebesgue measure. Suppose $\delta > 0$, and denote by δA the set $\{\delta x : \ x \in A\}$. We can then assert that δA is measurable whenever A is, and $\mu(\delta A) = \delta^n \mu(A)$. One can also easily see that Lebesgue measure is reflection-invariant. That is, whenever A is measurable, so is $-A = \{-x : x \in A\}$ and $\mu(-A) = \mu(A)$.

It also can be shown that the Lebesgue measure is rotation-invariant, see [4, Section 2.8].

Above we mentioned non-Lebesgue measurable sets. But do such sets exists? The following example answer this question affirmative.

Example 3.8.1 (Non-Lebesgue measurable set)**.** Let us consider the interval $[0, 2\pi)$ as the unit circumference $\Gamma = \{z \in \mathbb{C} : |z| = 1\}$ on the complex plane (that is, there is one-to-one correspondece between Γ and $[0, 2\pi)$). Then every point φ of the interval $[0, 2\pi)$ becomes the complex number $z = e^{i\varphi}$.

Let α be an irrational number. Define an equivalence relation $\sim$ on Γ by $z_2 \sim z_1$ if $z_2 = z_1 e^{\pi k \alpha i}$, $k \in \mathbb{Z}$. It is easy to see that the relation $\sim$ is reflexive ($z \sim z$), symmetric ($z_2 \sim z_1 \implies z_1 \sim z_2$) and transitive ($z_2 \sim z_1, z_3 \sim z_2 \implies z_3 \sim z_1$), so circumference Γ is split by this equivalence relation into equivalence classes (each class is evidently a countable set). Let the set Φ_0 contain exactly one element from each equivalence class, and let

$$\Phi_m = \{z e^{\pi m \alpha i} : \ z \in \Phi_0\}, \quad m \in \mathbb{Z}.$$

We now show that

$$\Gamma = \bigcup_{m \in \mathbb{Z}}^{\cdot} \Phi_m. \tag{3.8.1}$$

Indeed, $\Phi_m \bigcap \Phi_l = \varnothing$, $m \neq l$, since if, on the contrary, $z \in \Phi_m$ and $z \in \Phi_l$, then $z = z_0 e^{\pi \alpha m i}$ and $z = z_0' e^{\pi \alpha l i}$, where $z_0, z_0' \in \Phi_0$. So we obtain $z_0 e^{\pi \alpha m i} = z_0' e^{\pi \alpha l i}$, or $z_0' = z e^{\pi \alpha (m-l) i}$, that is, $z_0' \sim z_0$. But Φ_0 contains exactly one element from each equivalence class, so there are no equivalent elements in Φ_0. Therefore, $z_0' = z_0$, so $e^{\pi \alpha l i} = e^{\pi \alpha m i}$, or $e^{\pi \alpha (m-l) i} = 1$, or $\pi \alpha (m - l) i = 2\pi k i$, $k \in \mathbb{Z}$, or $\alpha = \frac{2k}{m-l} \in \mathbb{Q}$, a contradiction, since α is irrational by assumption. Thus, $\Phi_m \bigcap \Phi_l = \varnothing$ whenever $m \neq l$.

On the other hands, every point $z \in \Gamma$ belongs to one of the equivalence class by construction. Therefore, $z \sim z_0 \in \Phi_0$, so $z = z_0 e^{\pi \alpha m i}$, that is, $z \in \Phi_m$. Thus, the identity (3.8.1) is proved.

Suppose now that the set Φ_0 is Lebesgue measurable. Then each set Φ_m is also measurable, and $\mu\Phi_m = \mu\Phi_0$, $m \in \mathbb{Z}$, since Φ_m is obtained from Φ_0 by a rotation on Γ (a shift on $\mathbb{R}$), but the Lebesgue measure is invariant with respect to rotations and shifts. The Lebesgue measure is σ-additive, so we have

$$\mu\Gamma = 2\pi = \sum_{m=-\infty}^{+\infty} \mu\Phi_m.$$

However, this equality is impossible, since the sum of the series here is either equals zero (if $\mu\Phi_0 = 0$), or equals infinity (if $\mu\Phi_0 = a > 0$).

Thus, the set Φ_0 is non-Lebesgue measurable.

3.9 Other examples of measures

3.9.1 Lebesgue–Stieltjes measure

Here we shortly describe a generalization of one-dimensional Lebesgue measure, called Lebesgue-Stieltjes measure on $\mathbb{R}$. This measure is obtained from a non-decreasing and left continuous function $F : \mathbb{R} \mapsto \mathbb{R}$ by assigning to intervals the following measure

$$m'_F(a, b) = F(b) - F(a + 0),$$

$$m'_F[a, b] = F(b + 0) - F(a),$$

$$m'_F(a, b] = F(b + 0) - F(a + 0),$$

$$m'_F[a, b) = F(b) - F(a).$$

It is easy to see that the defined measure m' is nonnegative and additive (in fact, m' is not a measure by Definition 2.2.1, since the system of intervals on $\mathbb{R}$ is not a ring, see Remark 3.1.1). Note that the system of bricks on $\mathbb{R}$ is the system of intervals. As well, the measure m' defined on bricks (see (3.1.1)) is a particular case of m'_F as $n = 1$ (in this case, $F(t) = t, \quad t \in \mathbb{R}$).

Applying the extension process described in Sections 3.2–3.5 to the function m'_F, we define on $\mathbb{R}$ the ring $\mathcal{M}_F$ of sets measurable w.r.t. σ-additive measure μ_F (extension of m'_F) which is called the *Lebesgue–Stieltjes measure*. The ring $\mathcal{M}_F$ depends on F but it always contains all Borel sets (see [4, Section 2.9], [3, Section 1.5] for more details). If the function F has a finite variation on $\mathbb{R}$, that is, if $F(+\infty) - F(-\infty) < +\infty$, then, obviously, $\mathcal{M}_F$ is a σ-algebra with the unit $\mathbb{R}$, and $\mu_F(\mathbb{R}) = F(+\infty) - F(-\infty)$. Moreover, if $F(+\infty) - F(-\infty) = 1$, then the measure μ_F is called *normed* but if additionally $F(-\infty) = 0$ and $F(+\infty) = 1$, then μ_F is called a *probability measure.*

We emphasize once again that if $F(t) = t$, then $\mu_F = \mu$ is the ordinary Lebesgue measure.

3.9.2 Discrete measure

Let $X = \{x_1, x_2, \ldots, x_n, \ldots\}$ be an arbitrary countable set, and let $(p_n)_{n=1}^{\infty}$ be a sequence of *positive* numbers satisfying the condition

$$\sum_{n=1}^{\infty} p_n < +\infty.$$

For any subset A of the set X we define

$$mA = \sum_{k:\ x_k \in A} p_k. \tag{3.9.1}$$

The formula (3.9.1) defines a σ-additive measure m on the σ-algebra $\mathcal{P}(X)$ of all subsets of the set X.

If $X = \{x_1, x_2, \ldots, x_n, \ldots\} \subset \mathbb{R}$, then we connect this set with the function $F : \mathbb{R} \mapsto \mathbb{R}_+$ which defined as follows

$$F(x) = \sum_{k:\ x_k < x} p_k \tag{3.9.2}$$

This function is called a *saltus function*, because

$$F(x_k + 0) - F(x_k) = p_k,$$

and on the intervals free of points x_k, the function F is constant.

By the function (3.9.2), one can defined a Lebesgue-Stieltjes measure μ_F. Such measure is called *discrete measure*.

3.10 Problems

Problem 3.1. Let $\{A_k\}_{k=1}^{+\infty}$ be *decreasing* sequence of sets from $\mathcal{M}_\sigma(\mathbb{R}^n)$, and $\mu A_k = +\infty$ $(k \in \mathbb{N})$. Can the set $A = \bigcap\limits_{k=1}^{\infty} A_k$ have

a) infinite measure;

b) finite positive measure;

c) measure 0?

Problem 3.2. Given two Lebesgue measurable sets A_1 and A_2 on the interval $[0,1]$ such that $\mu A_1 + \mu A_2 > 1$. Prove that $\mu\left(A_1 \bigcap A_2\right) > 0$.

Problem 3.3. Given n Lebesgue measurable sets $A_1, A_2, \ldots, A_n$ on the interval $[0,1]$ such that $\sum\limits_{k=1}^{n} \mu A_k > n-1$. Prove that $\mu\left(\bigcap\limits_{k=1}^{n} A_k\right) > 0$.

Problem 3.4. Given n, construct Lebesgue measurable sets $A_1, A_2, \ldots, A_n$ on the interval $[0,1]$ such that $\sum\limits_{k=1}^{n} \mu A_k = n-1$ and

$$\mu\left(\bigcap_{k=1}^{n} A_k\right) = 0.$$

Definition. A set $A \in \mathbb{R}^n$ is called *measurable in Jordan sense* (or simply *Jordan measurable*) if for any $\varepsilon > 0$ there exist elementary sets $B_1, B_2 \in \mathcal{E}^n$ such that $B_1 \subset A \subset B_2$, and

$$m(B_2 \setminus B_1) < \varepsilon.$$

In this case, the (Jordan) measure μ^J of the set A is defined as follows:

$$\mu^J = \inf\{mB : \ A \subset B, B \in \mathcal{E}^n\}.$$

Problem 3.5. Prove that if $A \in \mathbb{R}^n$ is Jordan measurable, then it is Lebesgue measurable, and its Lebesgue measure equals its Jordan measure.

Hint: Prove first the semi-additivity of the Jordan measure.

Problem 3.6. Prove that the set $A := \mathbb{Q} \bigcap [0,1]$ is Lebesgue measurable but non-Jordan measurable. Find the Lebesgue measure of A.

Problem 3.7. Find the Lebesgue measure of the Cantor set and prove that any subset of the Cantor set is Lebesgue measurable.

Problem 3.8. Find the Lebesgue measure of the subset of the interval $[0,1]$ consisting of the numbers (of $[0,1]$) whose decimal form does not contain a digit n, $n = 0, 1, 2 \ldots, 9$.

Problem 3.9. Find the Lebesgue measure of a subset of the interval $[0,1]$ consisting of the numbers (of $[0,1]$) in whose decimal form the digit 2 always stays earlier than the digit 3.

Problem 3.10 (The Borel-Cantelli lemma)**.** Let $\{A_k\}_{k=1}^{+\infty}$ be a sequence of Lebesgue measurable sets such that

$$\sum_{k=1}^{+\infty} \mu A_k < +\infty.$$

Prove that $\overline{A} = \overline{\lim} A_k$ is Lebesgue measurable, and $\mu \overline{A} = 0$.

Problem 3.11. Find the Lebesgue measure of the subset of the plane square $\{(x, y) : 0 \leqslant x \leqslant 1, 0 \leqslant x \leqslant 1\}$ consisting of points (x, y) such that $0 \leqslant \sin x \leqslant \frac{1}{2}$ and $\cos(x + y)$ is irrational.

Problem 3.12. Find the Lebesgue measure of the subset of the plane square $\{(x, y) : 0 \leqslant x \leqslant 1, 0 \leqslant x \leqslant 1\}$ consisting of points (x, y) whose Descartes and polar coordinates are irrational.

Problem 3.13. Prove that any set of $\mathbb{R}$ with positive Lebesgue measure is of cardinality continuum.

Problem 3.14. Prove that the cardinality of $\mathcal{M}_\sigma(\mathbb{R}^n)$ is 2^c.

Problem 3.15. Let a set $A \subset [0, 1]$ is Lebesgue measurable. Prove that the function $f(x) = \mu\left(A \bigcap [0, x]\right)$ is continuous on $[0, 1]$.

Problem 3.16. Prove that the set $B = \bigcup\limits_{n=2}^{\infty} {}^{\!\!\cdot}\left[\frac{1}{n}, \frac{1}{n} + \frac{\alpha}{n(n-1)}\right] \subset [0, 1]$ is Lebesgue measurable and

$$\lim_{x \to +0} \frac{\mu\left(B \bigcap [0, x]\right)}{x} = \alpha,$$

where $\alpha \in (0, 1)$.

Problem 3.17. Let $A \subset [0, 1]$ be a measurable set w.r.t. the Lebesgue measure μ on $[0, 1]$, and let $\mu(A \cap (a, b)) \leqslant \alpha(b - a)$, $0 < \alpha < 1$, for any interval $(a, b) \subset [0, 1]$. Prove that $\mu A = 0$.

Problem 3.18. Let $A \subset [0, 1]$ and $\mu(A) > 0$. Prove that there exists a pair of points $x, y \in A$ such that the distance $|x - y|$ is irrational. Here μ is the Lebesgue measure.

Problem 3.19. Let $A \subset [0, 1]$ and $\mu(A) > 0$. Prove that there exist a pair of points $x, y \in A$ such that the distance $|x - y|$ is rational. Here μ is the Lebesgue measure.

Problem 3.20. Let a set $A (\in \mathbb{R})$ be non-Lebesgue measurable, and a set $A_0 (\in \mathbb{R})$ be of Lebesgue measure 0. Prove that the set $A \bigcap (A_0^c)$ is non-Lebesgue measurable. Is the statement true for the case when $A, A_0 \in \mathbb{R}^n$?

Problem 3.21. Construct a Lebesgue measurable set $A \subset [0, 1] \times [0, 1]$ such that both its projections on the coordinate axes Ox and Oy are non-Lebesgue measurable.

Problem 3.22. Let μ be the Lebesgue measure on $\mathbb{R}^n$, $A \in \mathcal{M}$, and $\mu(A) > 0$. Prove that there exists a Lebesgue non-measurable set $B \subset A$.

Problem 3.23. Construct a σ-additive measure on $\mathbb{R}^n$ such that any subset of $\mathbb{R}^n$ is measurable.

Problem 3.24. Construct a set $A \subset [0, 1] \times [0, 1]$ measurable w.r.t. the Lebesgue measure on $[0, 1] \times [0, 1]$ whose projections on the axes OX and OY are *not* measurable w.r.t. the Lebesgue measure on $[0, 1]$.

Problem 3.25. Construct a set $A \subset [0, 1] \times [0, 1]$ which is *not* measurable w.r.t. the Lebesgue measure on $[0, 1] \times [0, 1]$ and whose projections on the axes OX and OY are measurable w.r.t. the Lebesgue measure on $[0, 1]$.

Problem 3.26. Find the (two-dimensional) Lebesgue measure of the set

$$A = \left\{(x, y) \in [0, 1] \times [0, 1] \ : \ x \in [0, 1] \setminus \mathbb{Q}_{[0,1]} \quad \text{and} \quad \sin y < \frac{1}{2}\right\},$$

where $\mathbb{Q}_{[0,1]}$ is the set of all rational numbers of the interval $[0, 1]$.

Problem 3.27. Prove that the outer measure is not additive.

Hint: Use Lebesgue non-measurable sets.

Problem 3.28. Prove that $\mathbb{Q} \in F_\sigma(\mathbb{R})$ and $\mathbb{Q} \notin G_\delta(\mathbb{R})$.

Hint: Using the fact that a set A is dense in $\mathbb{R}$ if it has a common point with any interval on $\mathbb{R}$, prove that a countable intersection of sets in $G_\delta(\mathbb{R})$ dense in $\mathbb{R}$ is a set in $G_\delta(\mathbb{R})$ which is also dense in $\mathbb{R}$. Supposing that $\mathbb{Q}$ belongs to $G_\delta(\mathbb{R})$ find a sequence of G_δ-sets dense in $\mathbb{R}$ whose intersection is not dense in $\mathbb{R}$ that will give a contradiction.

Problem 3.29. Let

$$A = \{(x, y) : a \leqslant x \leqslant b \quad \text{and} \quad 0 \leqslant y \leqslant f(x)\},$$

where $f(x)$ is a *positive* continuous function on $[a, b]$. Prove that A is measurable w.r.t. the Lebesgue measure on $\mathbb{R}^2$ and its Lebesgue measure is

$$\mu A = \int_a^b f(x)dx,$$

where the integral is the Riemann integral.

Chapter 4

Measurable functions

In this chapter, we introduce measurable functions and study their main properties as a first step of the construction of the Lebesgue integral.

4.1 Definition of measurable functions

Let X be a set, $\mathcal{M} = \mathcal{M}(X)$ be a σ-algebra of subsets of the set X, which is the unit of this σ-algebra, and let μ be a **complete** σ-additive measure defined on $\mathcal{M}$. In what follows, we call the three $(X, \mathcal{M}, \mu)$ a *measure space* (or a μ-measure space).

Definition 4.1.1. A function $f : X \mapsto \mathbb{R}$ is called *measurable* on the set X w.r.t. the measure μ (or *μ-measurable*) if for any real number c the set $X(f > c) := \{x \in X : \ f(x) > c\}$ is μ-measurable.

We denote the set of all μ-measurable on X functions as $S(X, \mathcal{M}, \mu)$. If it is clear from the context what measure is used, then we omit the mention of the measure and denote the set of all measurable functions as $S(X)$.

Example 4.1.2. Let $\mu X = 0$. Then any function $f : X \mapsto \mathbb{R}$ is measurable on X, since from the completeness of the considered measure, we obtain for any $c \in \mathbb{R}$:

$$X(f > c) \in \mathcal{M}(X).$$

Lemma 4.1.3. *The function $f \in S(X)$ if, and only if, one of the following conditions holds:*

$$X(f \geqslant c) \in \mathcal{M}(X), \tag{4.1.1}$$

$$X(f < c) \in \mathcal{M}(X), \tag{4.1.2}$$

$$X(f \leqslant c) \in \mathcal{M}(X). \tag{4.1.3}$$

Proof. Let $S(x)$, and let $c \in \mathbb{R}$. First we show that

$$X(f \geqslant c) = \bigcap_{k=1}^{\infty} X\left(f > c - \frac{1}{k}\right). \tag{4.1.4}$$

Indeed,

$$x \in X(f \geqslant c) \Longleftrightarrow f(x) \geqslant c \ \Longleftrightarrow \forall k \in \mathbb{N} \ \ f(x) > c - \frac{1}{k} \ \Longleftrightarrow$$

$$\Longleftrightarrow \forall k \in \mathbb{N} \ \ x \in X\left(f(x) > c - \frac{1}{k}\right) \Longleftrightarrow \ x \in \bigcap_{k=1}^{\infty} X\left(f(x) > c - \frac{1}{k}\right).$$

If $f \in S(X)$, then any set in the right-hand side of (4.1.4) is measurable. Therefore, $X(f \geqslant c)$ is also measurable as a countable intersection of measurable sets, since $\mathcal{M}(x)$ is a σ-algebra by assumption. Thus, we proved that if $f \in S(X)$, then $X(f \geqslant c) \in \mathcal{M}(X)$.

Conversely, let the set $X(f \geqslant c)$ be measurable. Then from the identity (which can be easily checked)

$$X(f > c) = \bigcup_{k=1}^{\infty} X\left(f \geqslant c + \frac{1}{k}\right).$$

it follows that $f \in S(X)$.

The equivalence of the condition (4.1.2) to the inclusion $f \in S(X)$ follows from the following obvious identities:

$$X(f \leqslant c) = X \setminus X(f > c), \quad X(f > c) = X \setminus X(f \leqslant c).$$

The equivalence of the condition (4.1.3) to the inclusion $f \in S(X)$ follows from the obvious identities:

$$X(f < c) = X \setminus X(f \geqslant c), \quad X(f \geqslant c) = X \setminus X(f < c).$$

□

Lemma 4.1.3 claims that anyone of the conditions (4.1.1)–(4.1.3) can be used in the definition of measurable functions.

Corollary 4.1.4. *If $f \in S(X)$, then for any $a, b \in \mathbb{R}$, $a \leqslant b$, the following sets are measurable:*

$$X(a \leqslant f \leqslant b), \quad X(a < f \leqslant b), \quad X(a \leqslant f < b), \quad X(a < f < b), \quad X(f = a).$$

Proof. Indeed,

$$X(a \leqslant f \leqslant b) = X(f \leqslant b) \bigcap X(f \geqslant a) \in \mathcal{M}(X).$$

The rest assertions of the theorem can be proved analogously. □

Example 4.1.5. Let $X = \dot{\bigcup}_{k=1}^{m} X_k$. Define the function $h : X \mapsto \mathbb{R}$ putting $h(x) = c_k$ whenever $x \in X_k$, $k = 1, \dots, m$, where $c_k \in \mathbb{R}$ are arbitrary real numbers (that may coincide for different indices). Such function is called a *simple function*.

If all the sets X_k are measurable, then for any $c \in \mathbb{R}$, the set

$$X(h > c) = \dot{\bigcup}_{k:\ c_k > c} X_k$$

is measurable as a union of finitely many measurable sets (if $c \geqslant c_k$ for all $k = 1, \dots, m$, then the set $X(h > c) = \varnothing$ is measurable), therefore, the function h is measurable.

If all the numbers c_k are distinct, then from $h \in S(x)$ it follows that $X_k \in \mathcal{M}(X)$ for all $k = 1, \dots, m$.

Theorem 4.1.6. *Let X be $\mathbb{R}$ or an interval in $\mathbb{R}$. Any continuous function $f : X \mapsto \mathbb{R}$ is σ-Lebesgue measurable (that is, measurable w.r.t. the Lebesgue measure).*

Proof. Let X_0 be the set of all interior points of the set X (if X is an interval, then X_0 is the same interval but without the end points). For an arbitrary $c \in \mathbb{R}$, consider the set $X_0(f > c)$. If this set is empty, then it is measurable. Suppose now that $X_0(f > c) \neq \varnothing$. Let $x_0 \in X_0(f > c)$, and let $\varepsilon := \frac{f(x_0) - c}{2}$. Since f is continuous, there exists $\delta > 0$ such that for all $x \in X_0 \bigcap (x_0 - \delta, x_0 + \delta)$, we have $|f(x) - f(x_0)| < \varepsilon$. So,

$$f(x) > f(x_0) - \varepsilon = f(x_0) - \frac{f(x_0) - c}{2} = \frac{c + f(x_0)}{2} > c.$$

Clearly, we can choose $\delta > 0$ so small that $(x_0 - \delta, x_0 + \delta) \subset X_0$, and $(x_0 - \delta, x_0 + \delta) \subset X_0(f > c)$ as we proved above. Therefore, the set $X_0(f > c)$ is open, so $X_0(f > c)$ is σ-Lebesgue measurable by Corollary 3.5.10. The set $X(f > c)$ differs from $X_0(f > c)$ by one or two points (the ends of the interval). But one-point sets (degenerated bricks) are σ-Lebesgue measurable, so the set $X(f > c)$ is σ-Lebesgue measurable, thus $f \in S(X)$. □

4.2 Properties of measurable functions

Let us now study some properties of measurable functions.

1) $f \in S(X),\ l \in \mathbb{R} \implies f + l \in S(X)$.

   Obviously, $X(f + l > c) = X(f > c - l) \in \mathcal{M}(X)$ for any real c, so $f + l \in S(X)$.

2) $f \in S(X),\ k \in \mathbb{R} \implies kf \in S(X)$.

   This property follows from the identity

$$X(kf > c) = \begin{cases} X(f > c/k), & k > 0, \\ X(f < c/k), & k < 0, \\ X, & k = 0,\ c < 0, \\ \varnothing, & k = 0,\ c \geqslant 0, \end{cases}$$

   since all the sets in the right-hand side of this identity are measurable.

To establish the next property, we need the following lemma.

Lemma 4.2.1. *If $f, g \in S(X)$, then $X(f > g) \in \mathcal{M}(X)$.*

Proof. Let us enumerate all the rational numbers $\mathbb{Q} = \{r_k :\ k \in \mathbb{N}\}$. This is possible, since $Card\ \mathbb{Q} = a$.

Let $x \in X(f > g)$, that is, $f(x) > g(x)$. Then there exists a rational number r_k such that $f(x) > r_k > g(x)$, therefore, $x \in X(f > r_k) \bigcap X(g < r_k)$, so

$$X(f > g) \subset \bigcup_{k=1}^{\infty} \left(X(f > r_k) \bigcap X(g < r_k) \right).$$

The converse inclusion is obvious, so we have

$$X(f > g) = \bigcup_{k=1}^{\infty} \left(X(f > r_k) \bigcap X(g < r_k) \right).$$

Since $\mathcal{M}(X)$ is a σ-algebra by assumption, the right-hand side of the last identity is measurable, so is the set $X(f > g)$. □

Now we are in a position to establish the next properties of measurable functions.

3) $f, g \in S(X) \implies f + g \in S(X)$.

   Clearly,

$$X(f + g > c) = X(f > -g + c),$$

   so $f + g$ is measurable by Lemma 4.2.1 and by the properties 1) and 2) of measurable functions.

4) $f \in S(X) \implies f^2 \in S(X)$.

   Indeed, the set

$$X(f^2 > c) = \begin{cases} X(f > \sqrt{c}) \bigcup X(f < -\sqrt{c}), & c \geqslant 0, \\ X, & c < 0. \end{cases}$$

   is measurable for any real c.

Note that converse is not true. That is, if f^2 is measurable, then this does not mean that the function f is measurable, generally speaking.

Example 4.2.2. Let $X = [0, 1]$, and X_0 be a non-Lebesgue measurable subset of X. Suppose that

$$f(x) = \begin{cases} 1, & x \in X_0, \\ -1, & x \in X \setminus X_0. \end{cases}$$

The function f is non-Lebesgue measurable, since the set $X(f = 1)$ (which is exactly X_0) is non-Lebesgue measurable. At the same time, the function $f^2 \equiv 1$ is Lebesgue measurable on X.

5) $f, g \in S(X) \implies f \cdot g \in S(X)$.

This property follows from the identity

$$f \cdot g = \frac{(f+g)^2 - f^2 - g^2}{2}$$

and from the properties 2), 3), and 4).

6) $f, g \in S(X),\ g(x) \neq 0$ for $x \in X \implies \dfrac{f}{g} \in S(X)$.

It suffices to prove that the function $\dfrac{1}{g}$ is measurable and then to apply the property 5). So, from the identity

$$X\left(\frac{1}{g} > c\right) = \begin{cases} X(0 < g < 1/c), & c > 0, \\ X(g > 0), & c = 0, \\ X(g > 0) \bigcup X(g < 1/c), & c < 0, \end{cases}$$

we have $\dfrac{1}{g} \in S(X)$.

Definition 4.2.3. The function

$$f^+(x) = \begin{cases} f(x), & f(x) \geqslant 0, \\ 0, & f(x) < 0, \end{cases}$$

is called the *positive part of* f, and the function

$$f^-(x) = \begin{cases} 0, & f(x) > 0, \\ -f(x), & f(x) \leqslant 0, \end{cases}$$

is called the *negative part of* f.

7) $f \in S(X) \implies |f|, f^+, f^- \in S(X)$.

This property follows from the identities

$$X(|f| > c) = \begin{cases} X(f < -c) \bigcup X(f > c), & c \geqslant 0, \\ X, & c < 0, \end{cases}$$

and

$$f^+ = \frac{|f| + f}{2}, \quad f^- = \frac{|f| - f}{2}.$$

Example 4.2.2 shows that $|f| \in S(X) \not\Longrightarrow f \in S(X)$.

8) $f \in S(X),\ X_0 \in \mathcal{M}(X) \implies f \in S(X_0)$.

The proof of this property is based on the fact that *if $\mathcal{A}$ is a σ-algebra with the unit X, and if $X_0 \in \mathcal{A}$, then $\mathcal{A}(X_0) := \{A \bigcap X_0 :\ A \in \mathcal{A}(X)\}$ is a σ-algebra with the unit X_0* (see Homework no.4).

Now for any $c \in \mathbb{R}$, one has

$$X_0(f > c) = X_0 \bigcap X(f > c) \in \mathcal{M}(X_0).$$

9) *If* $f : X \mapsto \mathbb{R}$, $X = \bigcup\limits_{k=1}^{\omega} X_k$, *where* $X_k \in \mathcal{M}(X)$, *and*[1] $1 \leqslant \omega \leqslant +\infty$, *and if* $f \in S(X_k)$, $1 \leqslant k \leqslant \omega$, *then* $f \in S(X)$.

This property follows from the identity

$$X(f > c) = \bigcup_k X_k(f > c),$$

which can be easily checked.

10) *Let* $f \in S(X)$ *and* $\mu X(f = \pm\infty) = 0$, *then for any* $\varepsilon > 0$ *there exists* $g \in S(X)$ *such that* g *is bounded on* X *and* $\mu X(f \neq g) < \varepsilon$.

Consider the sets

$$A_k = X(|f| > k), \qquad A = X(f = \pm\infty) = X(|f| = +\infty).$$

By assumption, $\mu A = 0$. Moreover, $A = \bigcap\limits_{k=1}^{\infty} A_k$, and

$$A_1 \supset A_2 \supset A_3 \supset \dots$$

According to Lemma 2.1.21, $\lim\limits_{k\to\infty} A_k = A$, and due to σ-additivity of the measure μ, we obtain

$$\mu A = \mu\left(\lim_{k\to\infty} A_k\right) = \lim_{k\to\infty} \mu A_k = 0$$

by Theorem 2.2.5. Therefore, for any $\varepsilon > 0$ there exists a number $N \in \mathbb{N}$ such that $\mu A_N < \varepsilon$ for any $k \geqslant N$. Introduce the following function on X

$$g(x) = \begin{cases} f(x), & x \in X \setminus A_N, \\ 0, & x \in A_N. \end{cases}$$

It is clear that $g \in S(x)$. Moreover, it is bounded, since $|g(x)| \leqslant N$ on X. At the same time, $X(f \neq g) = A_N$, so $\mu X(f \neq g) < \varepsilon$.

Definition 4.2.4. Some property is said to hold on the set X "*almost everywhere*" (a.e.) if the set of the points of X for which this property does not hold, is a null set (a set of measure zero).

For example, the statement "*the function* f *equals zero almost everywhere on the set* X" means that $\mu X(f \neq 0) = 0$.

Thus, the property 10) shows that any measurable function which is *almost everywhere* finite on X becomes bounded if we neglect a subset of X of arbitrary small measure.

4.3 Sequences of measurable functions

In what follows, we consider operations on measurable functions related to passage to the limit. Since during this process we can obtain infinite values, we extend the set of functions under consideration. Namely, let us consider functions $f : X \mapsto \overline{\mathbb{R}} = [-\infty, +\infty]$. Since we need to operate with infinite values, let us define the following:

1) $a + (\pm\infty) = \pm\infty$, 2) $a - (\pm\infty) = \mp\infty$, $a \in \mathbb{R}$,
3) $(+\infty) + (+\infty) = +\infty$, 4) $-\infty + (-\infty) = -\infty$,
5) $a \cdot (\pm\infty) = \pm\infty$, $a > 0$, 6) $a \cdot (\pm\infty) = \mp\infty$, $a < 0$, 7) $0 \cdot (\pm\infty) = 0$,
8) $(\pm\infty) \cdot (\pm\infty) = +\infty$, 9) $(\pm\infty) \cdot (\mp\infty) = -\infty$,

[1] That is, the union can be finite or countable.

10) $|\pm\infty| = +\infty$, 11) $\dfrac{a}{\pm\infty} = 0$, $a \in \mathbb{R}$.

The symbols $(\pm\infty) - (\pm\infty)$, $\dfrac{\pm\infty}{\pm\infty}$, $\dfrac{\pm\infty}{\mp\infty}$, $\dfrac{a}{0}$ are considered to be meaningless.

The definition of measurability for the functions $f : X \mapsto \overline{\mathbb{R}}$ is left the same as Definition 4.1.1. Therefore, all the properties proved for measurable functions are valid, provided the corresponding operations are acceptable.

Theorem 4.3.1. *Let $(f_k)_{k=1}^{\infty}$ be a sequence of functions measurable on X. Then the functions* $\sup\{f_k\}$, $\inf\{f_k\}$, $\overline{\lim}\{f_k\}$, $\underline{\lim}\{f_k\}$, $\lim f_k$ *(if any) are measurable on X.*

All the operations mentioned in the theorem are pointwise. For instance, the function $f^* = \sup\{f_k\}$ is defined as follows

$$f^*(x) = \sup\{f_k(x) : \ k \in \mathbb{N}\}, \quad x \in X. \tag{4.3.1}$$

Proof of Theorem 4.3.1. First we prove that $f^* = \sup\{f_k\}$ is measurable. To do this, let us establish the following identity

$$X(f^* \leqslant c) = \bigcap_{k=1}^{\infty} X(f_k \leqslant c), \quad c \in \mathbb{R}. \tag{4.3.2}$$

Assume that $x \in X(f^* \leqslant c)$. Then $f^*(x) \leqslant c$, and from (4.3.1) it follows that $f_k(x) \leqslant c$ for any $k \in \mathbb{N}$. Therefore, $x \in \bigcap\limits_{k=1}^{\infty} X(f_k \leqslant c)$, so $X(f^* \leqslant c) \subset \bigcap\limits_{k=1}^{\infty} X(f_k \leqslant c)$.

Conversely, if $x \in \bigcap\limits_{k=1}^{\infty} X(f_k \leqslant c)$, then $f_k(x) \leqslant c$ for any $k \in \mathbb{N}$. Consequently,

$$\sup\{f_k(x) : \ k \in \mathbb{N}\} = f^*(x) \leqslant c,$$

that is, $x \in X(f^* \leqslant c)$, so $X(f^* \leqslant c) \supset \bigcap\limits_{k=1}^{\infty} X(f_k \leqslant c)$. Thus, the identity (4.3.2) is true, and the function f^* is measurable by Lemma 4.1.3.

The function $\inf\{f_k\}$ is measurable, since

$$\inf\{f_k\} = -\sup\{-f_k\}.$$

The function $\overline{f} = \overline{\lim} f_k$ is measurable, since

$$\overline{f}(x) = \overline{\lim} f_k(x) = \inf_{k \in \mathbb{N}} \left\{ \sup_{i \geqslant k} \{f_i(x)\} \right\}.$$

From the identity $\underline{\lim} f_k = -\overline{\lim}(-f_k)$, it follows that the function $\underline{f} = \underline{\lim} f_k$ is measurable.

If $\lim f_k$ exists, it is measurable, since in this case

$$\lim f_k = \underline{\lim} f_k = \overline{\lim} f_k.$$

$\square$

Definition 4.3.2. Two functions f and g defined on the set X are called *equivalent*, or *equal a.e.*, $f \overset{a.e.}{=} g$, if their values coincide a.e. on X, that is, if

$$\mu X(f \neq g) = 0.$$

Lemma 4.3.3. *If $f, g : X \mapsto \mathbb{R}$, $f \in S(X)$, and $f \overset{a.e.}{=} g$, then $g \in S(X)$.*

Proof. Let $X_0 := X(f \neq g)$. By assumption, $\mu X_0 = 0$. Consider the set $X_1 = X \setminus X_0$. The set X_1 is measurable as a difference of two measurable sets. We have $X = X_0 \dot\cup X_1$, and

$$X(g > c) = X_0(g > c) \dot\bigcup X_1(g > c) = X_0(g > c) \dot\bigcup X_1(f > c).$$

Since μ is assumed to be a complete measure, the set $X_0(g > c)$ is measurable as a subset of the null set X_0. At the same time, the set $X_1(f > c)$ is measurable by the property 8) of measurable functions. Therefore, the set $X(g > c)$ is measurable for any $c \in \mathbb{R}$. □

Definition 4.3.4. A sequences of functions $(f_k)_{k=1}^{\infty}$ is said to be convergent a.e. on the set X to a function f (we write $f_k \xrightarrow{a.e.} f$) if

$$\mu X(f_k \not\to f) = 0.$$

Example 4.3.5. *Let $X = [0, 1]$, μ be the Lebesgue measure, $f_k(x) = x^k$, and $f(x) \equiv 0$.*

Since $\lim\limits_{k\to+\infty} x^k = 0$ for $0 \leqslant x < 1$, and since $\lim\limits_{k\to+\infty} 1^k = 1$, we have $X(f_k \not\to f) = \{1\}$. Therefore, $f_k \xrightarrow{a.e.} f$, because $\mu(\{1\}) = 0$.

Theorem 4.3.6. *If $(f_k)_{k=1}^{\infty}$ is a sequence of measurable functions, and if $f_k \xrightarrow{a.e.} f$ on X, then $f \in S(X)$.*

Proof. Define $X_1 := X(f_k \to f)$ and $X_0 = X(f \not\to f)$. By assumption, $\mu X_0 = 0$, therefore, the set $X_1 = X \setminus X_0$ is measurable. Let us consider the following functions:

$$g_k(x) = \begin{cases} f_k(x), & x \in X_1, \\ 0, & x \in X_0, \end{cases} \qquad g(x) = \begin{cases} f(x), & x \in X_1, \\ 0, & x \in X_0. \end{cases}$$

Since $\mu X_0 = 0$, we get $f_k \overset{a.e.}{=} g_k$, $k \in \mathbb{N}$, and $f \overset{a.e.}{=} g$. At the same time, $g_k \to g$ on X pointwise. By Lemma 4.3.3, the functions g_k are measurable on X, so $g \in \mathcal{S}(X)$ as a pointwise limit of measurable functions, see Theorem 4.3.1. Now $f \in S(X)$ by Lemma 4.3.3, since $f \overset{a.e.}{=} g$. □

Theorem 4.3.7 (Egoroff)**.** *Let $\mu X < +\infty$, and let $(f_k)_{k=1}^{\infty}$ be a sequence of measurable functions that are a.e. finite[2] on X. If $f_k \xrightarrow{a.e.} f$, where f is a.e. finite on X, then for any $\delta > 0$ there exists a measurable set $X_\delta \subset X$ such that $\mu X_\delta < \delta$, and the sequences $(f_k)_{k=1}^{\infty}$ converges to f uniformly on $X \setminus X_\delta$.*

Proof. By Theorem 4.3.6, the function f is measurable on X. Let us introduce the following sets

$$X_0 = \left(\bigcup_{k=0}^{\infty} X(f_k = \pm\infty) \right) \bigcup X(f_k \not\to f), \qquad X_1 = X \setminus X_0.$$

The set X_0 is measurable as a countable union of measurable sets. Moreover, $\mu X_0 = 0$ due to σ-semi-additivity of the measure μ, since X_0 is a countable union of null sets. The set X_1 is measurable as difference of two measurable sets. Therefore, X_1 is measurable, and additionally, f, f_k, $k \in \mathbb{N}$, are finite on X_1, and $f_k \to f$ pointwise on X_1.

Consider the sequence

$$g_k(x) = |f_k(x) - f(x)|, \qquad k = 1, 2, 3, \ldots$$

The functions g_k are nonnegative and measurable on X_1, and $g_k(x) \to 0$ for any $x \in X_1$.

Let $(\varepsilon_j)_{j=1}^{\infty}$ be a sequence of positive numbers such that $\varepsilon_j \to 0$ as $j \to \infty$. Consider the following sets

$$X_{k,j} = \bigcap_{l=k}^{\infty} X_1(g_l < \varepsilon_j). \tag{4.3.3}$$

[2]Not necessary bounded! Consider e.g. the functions $f_k(x) = \dfrac{1}{x^k}$ that are finite a.e. on $X := [0, 1]$ but not bounded on X.

The sequences $(X_{k,j})$ increases as k grows, since

$$X_{k,j} = \bigcap_{l=k}^{\infty} X_1(g_l < \varepsilon_j) \subset \bigcap_{l=k+1}^{\infty} X_1(g_l < \varepsilon_j) = X_{k+1,j}.$$

Let us show that

$$\lim_{k\to\infty} X_{k,j} = X_1 \qquad j = 1, 2, 3, \dots \tag{4.3.4}$$

Let $x \in X_1$. Since $\lim\limits_{k\to+\infty} g_k = 0$, then for $\varepsilon_j > 0$ there exists a number k_0 such that $g_k(x) < \varepsilon_j$ for any $k \geqslant k_0$. This means that

$$x \in \bigcap_{l=k_0}^{\infty} X_1(g_l < \varepsilon_j) = X_{k_0,j},$$

therefore,

$$x \in \bigcup_{k=1}^{\infty} X_{k,j} = \lim_{k\to\infty} X_{k,j},$$

thus, $X_1 \subset \lim\limits_{k\to\infty} X_{k,j}$. The converse inclusion is obvious, so the identities (4.3.4) are true. Since μ is assumed to be σ-additive, it is continuous. Therefore, from (4.3.4), we obtain

$$\lim_{k\to\infty} \mu X_{k,j} = \mu X_1, \qquad j = 1, 2, 3, \dots,$$

So, for any $j \in \mathbb{N}$ there exists a number $k = k(j)$ such that

$$\mu(X_1 \setminus X_{k(j),j}) < \frac{\delta}{2^j},$$

where δ is an arbitrary positive number (here we use the fact that $\mu X_1 = \mu X < +\infty$).

The set

$$X_\delta = X_0 \bigcup \left(\bigcup_{j=1}^{\infty} (X_1 \setminus X_{k(j),j}) \right)$$

is the set defined in the assertion. Indeed,

$$\mu X_\delta \leqslant \mu X_0 + \sum_{j=1}^{\infty} \mu(X_1 \setminus X_{k(j),j}) < 0 + \sum_{j=1}^{\infty} \frac{\delta}{2^j} = \delta.$$

If $x \in X \setminus X_\delta$, then $x \notin X_0$, and $x \notin (X_1 \setminus X_{k(j),j})$ for any $j \in \mathbb{N}$. But since $x \notin X_0$, we have $x \in X_1$. And since $x \notin (X_1 \setminus X_{k(j),j})$ for any $j \in \mathbb{N}$, we obtain that $x \in X_{k(j),j}$ for all $j \in \mathbb{N}$. Thus, from (4.3.3) we obtain

$$|f_l(x) - f(x)| = g_l(x) < \varepsilon_j, \qquad l \geqslant k(j).$$

The number $k(j)$ does not depend on x, so f_k converges to f uniformly on $X \setminus X_\delta$. $\square$

Example 4.3.8. Consider the sequence $f_k(x) = \sin^k \pi x$, $k \in \mathbb{N}$ on the set $X = [0, 1]$ with the Lebesgue measure μ.

It is easy to see that for $x \in [0, 1]$, $x \neq \dfrac{1}{2}$, $\sin^k \pi x \to 0$, so

$$f_k \xrightarrow{a.e.} 0.$$

Choose δ such that $0 < \delta < \dfrac{1}{2}$, and define the set $X_\delta = \left(\dfrac{1}{2} - \delta, \dfrac{1}{2} + \delta\right)$, so $X \setminus X_\delta = \left[0, \dfrac{1}{2} - \delta\right] \bigcup \left[\dfrac{1}{2} + \delta, 1\right]$. Obviously, for any $x \in X \setminus X_\delta$, the inequality $0 \leqslant \sin^k \pi x \leqslant \sin^k \left(\pi\left(\dfrac{\pi}{2} - \delta\right)\right)$ holds, and $\sin^k \left(\pi\left(\dfrac{\pi}{2} - \delta\right)\right) \to 0$ as $k \to \infty$. Thus, the sequence $(f_k)_{k=1}^{\infty}$ converges uniformly to zero on the set $X \setminus X_\delta$.

For sequences of measurable functions one can define one more kind of convergence, the convergence in measure.

Definition 4.3.9. A sequences $(f_k)_{k=1}^{\infty}$ of functions measurable on X is said to be convergent to a measurable on X function f in measure (we write $f_k \xrightarrow{\mu} f$) if for any $\sigma > 0$,

$$\mu X\left(|f_k - f| \geqslant \sigma\right) \underset{k\to\infty}{\longrightarrow} 0.$$

Let us find interrelation between the convergence almost everywhere and the convergence in measure.

Theorem 4.3.10. *Let $\mu X < +\infty$. If a sequences $(f_k)_{k=1}^{\infty}$ of functions measurable on X converges a.e. to a function f, then $(f_k)_{k=1}^{\infty}$ converges to f on X in measure:*

$$f_k \xrightarrow{a.e.} f \implies f_k \xrightarrow{\mu} f.$$

Proof. By Theorem 4.3.6, the function f is measurable. On the contrary, suppose that $(f_k)_{k=1}^{\infty}$ does not converge to f in measure on the set X. Then for a certain $\sigma_0 > 0$ there exist $\delta_0 > 0$ and a sequence of indices $(k_j)_{j=1}^{\infty}$ such that

$$\mu X(|f_{k_j} - f| \geqslant \sigma_0) \geqslant \delta_0, \qquad j = 1, 2, 3, \ldots$$

Let

$$X' = \overline{\lim_{j}}\, X(|f_{k_j} - f| \geqslant \sigma_0).$$

Then by Theorem 2.2.6, $\mu X' \geqslant \delta_0$. So, if $x \in X'$, then according to the definition of the limit superior of a sequence of sets, among indices k_j, there exist infinitely many indices such that $x \in X(|f_{k_j} - f| \geqslant \sigma_0)$. This, in particular, means that for infinitely many indices we have $|f_{k_j}(x) - f(x)| \geqslant \sigma_0$, that is, $f_k(x) \not\longrightarrow f(x)$.

Thus, $f_k \not\longrightarrow f$ on the set X' of positive measure, a contradiction, since $f_k \xrightarrow{a.e.} f$ on X by assumption. □

Note that we cannot avoid the condition $\mu X < +\infty$ here.Consider, for example, the sequence $f_n(x) = \dfrac{x}{n}$ on $\mathbb{R}$.

However, the converse statement of Theorem 4.3.10 is not true.

Example 4.3.11. Let $X = [0, 1]$, and μ be the Lebesgue measure. Consider the functions

$$\varphi_{l,j}(x) = \begin{cases} 1, & x \in \left[\dfrac{j-1}{l}, \dfrac{j}{l}\right], \\ \\ 0, & x \notin \left[\dfrac{j-1}{l}, \dfrac{j}{l}\right], \end{cases}$$

where $l = 1, 2, \ldots$, $j = 1, 2, \ldots, l$.

Let us represent the functions $\varphi_{l,j}$ as one-index sequence as follows:

$$f_1(x) = \varphi_{11}(x), \quad f_2(x) = \varphi_{21}(x), \quad f_3(x) = \varphi_{22}(x), \quad f_4(x) = \varphi_{31}(x), \quad f_5(x) = \varphi_{32}(x), \ \ldots,$$

so

$$f_n(x) = \varphi_{l,j}(x), \qquad n = j + \sum_{k=1}^{l-1} k, \qquad l = 1, 2, \ldots, \quad j = 1, 2, \ldots, l.$$

The sequence f_n converges in measure on $[0, 1]$ to the function $f_0(x) \equiv 0$. Indeed, every function f_n belongs to a group of functions $\varphi_{l,j}$ with fixed index l each of which is nonzero only on the interval $\left[\dfrac{j-1}{l}, \dfrac{j}{l}\right]$ of length $\dfrac{1}{l}$. Thus, if we take $\sigma \leqslant 1$, we get

$$\mu X(|f_n - f_0| \geqslant \sigma) = \frac{1}{l} \underset{n\to\infty}{\longrightarrow} 0.$$

At the same time, for $l \geqslant 2$, and for any $x \in [0,1]$ in each group of functions $\varphi_{l,j}$ with fixed l, there exist functions that equal 1 at x and functions that equal 0 at x. Consequently, for any $x \in [0,1]$ the sequence $f_n(x)$ consists of infinitely many 1s and infinitely many 0s, so it is not convergent at all.

Thus, the sequence f_n converges in measure on $[0,1]$ to the function $f_0(x) \equiv 0$, but it does not converge at any point of the interval $[0,1]$.

However, the following theorem is true.

Theorem 4.3.12 (F. Riesz). *Any sequence $(f_n)_{n=1}^{\infty}$ of functions measurable on X convergent in measure on X to a measurable function f_0 contains a subsequence convergent a.e. on X to f_0.*

Proof. Let $(\varepsilon_j)_{j=1}^{\infty}$ and $(\eta_j)_{j=1}^{\infty}$ be two sequences of positive numbers such that

$$\varepsilon_j \underset{j\to\infty}{\longrightarrow} 0, \qquad \sum_{j=1}^{\infty} \eta_j < +\infty.$$

Consider the sequence of indices

$$n_1 < n_2 < \ldots < n_j < \ldots$$

constructed as follows: The index n_1 is chosen such that

$$\mu X(|f_{n_1} - f_0| \geqslant \varepsilon_1) < \eta_1.$$

Next we find an index $n_2 > n_1$ such that

$$\mu X(|f_{n_2} - f_0| \geqslant \varepsilon_2) < \eta_2.$$

If we already have the indices $n_1 < n_2 < \ldots < n_{j-1}$, then we take an index n_j such that

$$\mu X(|f_{n_j} - f_0| \geqslant \varepsilon_j) < \eta_j.$$

For any j we can always find an index n_j described above, because $f_n \xrightarrow{\mu} f_0$ by assumption. Since the process described can be continued indefinitely, as a result we obtain a subsequence $(f_{n_j})_{j=1}^{\infty}$. Let us show that $f_{n_j} \xrightarrow{a.e.} f_0$

For brevity sake, we denote $X_j = X(|f_{n_j} - f_0| \geqslant \varepsilon_j)$, $j \in \mathbb{N}$, and define the following set

$$X_0 = \overline{\lim}\, X_j = \bigcap_{l=1}^{\infty} \left(\bigcup_{j=l}^{\infty} X_j \right).$$

Since for any $l \in \mathbb{N}$, $X_0 \subset \bigcup\limits_{j=l}^{\infty} X_j$, we obtain

$$\mu X_0 \leqslant \mu \left(\bigcup_{j=l}^{\infty} X_j \right) \leqslant \sum_{j=l}^{\infty} \mu X_j < \sum_{j=l}^{\infty} \eta_j. \tag{4.3.5}$$

The series $\sum\limits_{j=1}^{\infty} \eta_j$ converges by assumption, so its reminder vanishes, therefore from (4.3.5) we get $\mu X_0 = 0$.

Let now $x \notin X_0$. Then there exists an index l such that $x \notin \bigcup\limits_{j=l} X_j$, that is, $x \notin X_j$ for all $j \geqslant l$. Consequently, for any $j \geqslant l$, one has

$$|f_{n_j}(x) - f_0(x)| < \varepsilon_j.$$

Since $\varepsilon_j \underset{j\to\infty}{\longrightarrow} 0$ by assumption, it follows that for any $x \in X \setminus X_0$, $f_{n_j}(x) \to f_0(x)$. □

In Example 4.3.11, we can take, for example, the subsequence of the sequence $(f_n)_{n=1}^{\infty}$ consisting only of functions $\varphi_{l,1}$. This subsequence tends to zero at any point of the interval $[0,1]$ but the point $x = 0$.

4.4 Approximation of measurable functions

In Example 4.1.5 we introduced simple functions that play a very important role in the theory of Lebesgue integral. We recall the definition of these function but using different notations.

The *characteristic function* (or *indicator function*) of a subset $A \subset X$ is the function $\chi_A : X \mapsto \mathbb{R}$ defined by

$$\chi_A = \begin{cases} 1 & \text{if} \quad x \in A \\ 0 & \text{if} \quad x \notin A \end{cases}$$

The function χ_A is measurable if and only if A is a measurable set.

Definition 4.4.1. Let $X = \dot{\bigcup}_{k=1}^{m} E_k$. Define a function $h : X \mapsto \mathbb{R}$ as follows: $h(x) = c_k$ whenever $x \in E_k$, where $c_k \in \mathbb{R}$. The function h is called a *simple function*. We will represent h in the form

$$h(x) = \sum_{k=1}^{m} c_k \chi_{E_k}(x). \tag{4.4.1}$$

Here we suppose that $c_k = 0$ if $\mu E_k = +\infty$.

In (4.4.1) some numbers c_k can coincide. However, for a given simple function h we can always find another representation of the form (4.4.1) with distinct numbers c_k. It can be done as follows. Let c'_1, c'_2, $\ldots$, c'_l be all distinct values of the function h. Define the sets $E'_j = \bigcup_{k:\ c_k = c'_j} E_k$, $j = 1, 2 \ldots, l$. The sets E'_j are disjoint, since the sets E_k are disjoint. So, we obtain a new (evidently, unique) representation of the function h:

$$h(x) = \sum_{j=1}^{l} c'_j \chi_{E'_j}(x), \tag{4.4.2}$$

where all c'_j are distinct.

As it was noticed in Example 4.1.5 if all the sets E_k are measurable, then the function h is measurable. Moreover, the sets E'_j are measurable if, and only if, the function h is measurable.

Theorem 4.4.2. *(on approximation) Any nonnegative function measurable on X can be represented as the limit of a non-decreasing sequence of nonnegative measurable simple functions.*

Proof. Let $f \in S^+(X)$, where $S^+(X) = S^+(X, \mathcal{M}, \mu)$ is the set of all nonnegative functions measurable on X. The measure μ is σ-finite on X, so

$$X = \bigcup_{l=1}^{+\infty} X_l,$$

where $X_1 \subseteq X_2 \subseteq X_3 \subseteq \cdots$.

For $l \in \mathbb{N}$, consider the truncations

$$F_l(x) = \begin{cases} f(x) & \text{if} \quad x \in X_l \quad \text{and} \quad f(x) \leqslant l, \\ l & \text{if} \quad x \in X_l \quad \text{and} \quad f(x) > l, \\ 0 & \text{otherwise.} \end{cases}$$

Then,

$$F_l(x) \nearrow f(x), \qquad \text{as} \quad l \to \infty \qquad \forall x \in X. \tag{4.4.3}$$

For a fixed $l \in \mathbb{N}$, we partition the range of $F_l(x)$, namely $[0, l]$, as follows. We divide the interval $[0, l]$ into subintervals

$$I_{l,k} = \left[\frac{k-1}{2^l}, \frac{k}{2^l}\right), \qquad k = 1, 2, \ldots, l \cdot 2^l,$$

and consider the sets

$$X_{l,0} = X_l(F_l = l) = X_l(f \geqslant l)$$

and

$$X_{l,k} = X_l\left(\frac{k-1}{2^l} \leqslant F_l < \frac{k}{2^l}\right) = X_l\left(\frac{k-1}{2^l} \leqslant f < \frac{k}{2^l}\right), \qquad k = 1, 2, \ldots, l \cdot 2^l.$$

Obviously,

$$X_l = X_{l,0} \dot{\bigcup} \left(\dot{\bigcup_{k=1}^{l\cdot 2^l}} X_{l,k}\right).$$

We consider the simple functions

$$h_l(x) = \begin{cases} l, & x \in X_{l,0}, \\ \dfrac{k-1}{2^l}, & x \in X_{l,k}, \end{cases}$$

and prove that the sequence $(h_l)_{l=1}^{\infty}$ is non-decreasing and convergent to f.

a) The sets $X_{l,k}$, $k = 1, 2, \ldots, l \cdot 2^l$, are measurable by Corollary 4.1.4. Therefore, for each $l \in \mathbb{N}$, the function $h_l(x)$ is a nonnegative measurable simple function.

b) Compare now the function h_l and h_{l+1}. If the point $x \in X_l$ is such that $f(x) < l$, then there exists an index k, $1 \leqslant k \leqslant l\cdot 2^l$ such that $x \in X_{l,k}$ and $h_l(x) = \dfrac{k-1}{2^l}$. When we consider the function h_{l+1}, we divide each interval $I_{l,k}$ into two parts. Therefore, either $f(x) \in \left[\dfrac{2k-2}{2^{l+1}}, \dfrac{2k-1}{2^{l+1}}\right)$ (so $h_{l+1}(x) = \dfrac{2k-2}{2^{l+1}} = \dfrac{k-1}{2^l}$), or $f(x) \in \left[\dfrac{2k-1}{2^{l+1}}, \dfrac{2k}{2^{l+1}}\right)$ (so $h_{l+1}(x) = \dfrac{2k-1}{2^{l+1}}$). In both cases we have $h_{l+1}(x) \geqslant h_l(x)$.

If $F_l(x) = l$ for a given $x \in X_l$ (that is, $f(x) \geqslant l$), then $h_l(x) = l$, and $h_{l+1}(x) \geqslant l = h_l(x)$. So, for any $x \in X_l$, $h_l(x) \leqslant h_{l+1}(x)$, and since $h_l(x) = 0$ for any $x \in X \setminus X_l$, one has $h_l(x) \leqslant h_{l+1}(x)$ for any $x \in X$. Thus, the sequence $(h_l)_{l=1}^{\infty}$ is non-decreasing on X.

c) Let us show now that $h_l(x) \nearrow f(x)$ as $l \to \infty$ for any $x \in \mathbb{X}$.

Let $x \in X$. For any $\varepsilon > 0$ there exists $l_0 \in \mathbb{N}$ such that for any $l \geqslant l_0$,

$$0 \leqslant f(x) - F_l(x) < \frac{\varepsilon}{2} \tag{4.4.4}$$

by (4.4.3). On the other hand, for $F_l(x) \leqslant l$, so there exists a number $k = k(l)$ such that $\dfrac{k-1}{2^l} \leqslant F_l(x) \leqslant \dfrac{k}{2^l}$. At the same time, $h_l(x) = \dfrac{k-1}{2^l}$, so

$$0 \leqslant F_l(x) - h_l(x) < \frac{1}{2^l}. \tag{4.4.5}$$

Thus, there exists $l_1 \in \mathbb{N}$ such that for any $l \geqslant l_1$ we have

$$0 \leqslant F_l(x) - h_l(x) < \frac{\varepsilon}{2}. \tag{4.4.6}$$

Now from (4.4.4) and (4.4.6) we obtain that

$$0 \leqslant f(x) - h_l(x) < \varepsilon$$

for any $l \geqslant \max\{l_0, l_1\}$, so $h_l(x) \nearrow f(x)$, as required. □

Corollary 4.4.3. *If f is a nonnegative bounded measurable function, then there exists a non-decreasing sequence of simple functions convergent to f uniformly on X if $\mu X < +\infty$.*

Proof. If $\mu X < +\infty$, then we can construct the sequence $h_l(x)$ without introducing the truncations $F_l(x)$.

Since f is bounded, there exists a number l_0 such that $f(x) \leqslant l_0$ on X. Then for the sequence $(h_l)_{l=1}^{\infty}$ constructed in the proof of Theorem 4.4.2 the inequality (4.4.5) holds for any $l \geqslant l_0$ and for any $x \in X$. □

Let us establish some additional approximation theorems for measurable functions.

Theorem 4.4.4. *Let $f \in S(X)$ and let f be a.e. finite on X. Then for any $\varepsilon > 0$ there exists a bounded function $g \in S(X)$ s.t. $\mu X(f \neq g) < \varepsilon$.*

Proof. Let $X_n = X(|f| > n)$, and $X_\infty = X(|f| = \infty)$. By assumption $\mu X_\infty = 0$. Moreover,

$$X_1 \supset X_2 \supset X_3 \supset \cdots, \qquad X_\infty = \bigcap_{n=1}^{\infty} X_n = \lim_{n\to\infty} X_n.$$

Since μ is σ-additive, by Theorem 2.2.5 we have

$$\mu X_\infty = \lim_{n\to\infty} \mu X_n.$$

So, given $\varepsilon > 0$, there exists $N > 0$ s.t. $\mu X_n < \varepsilon$ for any $n \geqslant N$. Now define the following function

$$g(x) = \begin{cases} f(x), & x \in X \setminus X_N, \\ 0, & x \in X_N. \end{cases}$$

This is a desirable function, since $g \in S(X)$, $|g(x)| \leqslant N$ on X, and the $X(f \neq g) = X_N$ is of measure not exceeding ε. □

Remark 4.4.5. Thus, every measurable a.e. finite function becomes bounded outside of a set of arbitrary small measure.

Measurable functions can also be approximated by a class of piece-wise constant functions called *step functions.*

Definition 4.4.6. A function f of the form

$$f(x) = \sum_{j=1}^{m} c_j \chi_{K_j}(x),$$

where K_j, $j = 1, \ldots, m$, are bricks in $\mathbb{R}^n$, and c_j are some real numbers, is called a step function.

Step functions are particular cases of simple functions. But unlike simple functions, every step function is measurable.

Theorem 4.4.7. *Let $f \in S(X)$. Then there exists a sequence of step functions $(\varphi_k)_{k=1}^{\infty}$ that converges to $f(x)$ a.e. on X.*

Proof. By Theorem 4.4.2, every nonnegative measurable function can be represented as a limit of a non-decreasing sequence of simple nonnegative functions. It is clear that an arbitrary measurable function can also be approximated by simple functions. Indeed, if $f \in S(X)$, then $f = f^+ - f^-$, where f^+ and f^- are nonnegative measurable functions. Then by Theorem 4.4.2 there exist two non-decreasing sequences of simple nonnegative functions $(h_n)_{n=1}^{\infty}$ and $(g_n)_{n=1}^{\infty}$ such that

$$f^+(x) = \lim_{n\to\infty} h_n(x) \qquad f^-(x) = \lim_{n\to\infty} g_n(x), \qquad x \in X.$$

Now for the functions $\psi_n(x) = h_n(x) - g_n(x)$ we obviously have

$$|\psi_n(x)| = |h_n(x)| + |g_n(x)| \leqslant |\psi_{n+1}(x)| \quad \text{and} \quad \lim_{n\to\infty} \psi_n(x) = f(x) \qquad \forall x \in X.$$

Thus, it is enough to establish that if A is a measurable set with finite measure, then $f(x) = \chi_A(x)$ can be approximated by step functions. By Definition 3.4.1 for any $\varepsilon > 0$ there exists an elementary set B s.t. $\mu(A \bigtriangleup B) < \varepsilon$. As we mentioned in Lecture **??**, the elementary set B can be represented as follows

$$B = \bigsqcup_{i=1}^{m} K_i,$$

where K_i are some bricks. Consider the step function

$$\psi(x) = \sum_{i=1}^{m} \chi_{K_i}(x).$$

Then we have that $\mu X(f \neq \psi) < \varepsilon$. Consequently, for every $k \in \mathbb{N}$, there exists a step function $\psi_k(x)$ s.t. if $E_k = X(f \neq \psi_k)$ then $\mu X_k \leqslant \frac{1}{2^k}$. Consider now the limit superior of the sequence $(E_k)_{k=1}^{\infty}$:

$$\overline{E} = \bigcap_{k=1}^{\infty} \bigcup_{l=k}^{\infty} E_l.$$

It is easy to see that $\mu\left(\bigcup\limits_{l=k}^{\infty} E_l\right) \leqslant 2^{-k}$, so $\mu\overline{E} = \lim\limits_{k\to\infty} \mu\left(\bigcup\limits_{l=k}^{\infty} E_l\right) = 0$. We thus obtain that

$$\chi_A(x) = \lim_{k\to\infty} \psi_k(x) \qquad \forall x \in X \setminus \overline{E},$$

and $\mu\overline{E} = 0$, as required. □

Now we are in a position to show that measurable functions can be approximated by continuous functions. First we prove the Luzin theorem stating that any a.e. finite measurable function on X is "nearly" continuous.

Theorem 4.4.8 (Luzin)**.** *Suppose f is measurable and a.e. finite on X, and $\mu X < +\infty$. Then for any $\varepsilon > 0$ there exists a closed set $F_\varepsilon \subset X$ such that*

$$\mu(X \setminus F_\varepsilon) < \varepsilon$$

and f is continuous on F_ε (that is, the restriction $f|_{F_\varepsilon}$ of the function f to the set F_ε is continuous on F_ε).

By $f\big|_{F_\varepsilon}$ we mean the restriction of f to the set F_ε. The conclusion of the theorem states that if f is viewed as a function defined only on F_ε, then f is continuous. However, the theorem does not make the stronger assertion that the function f defined on X is continuous at the points of F_ε.

Proof. By Theorem 4.4.7, there exists a sequence $(f_n)_{n=1}^{\infty}$ of step functions so that $f_n \xrightarrow{a.e.} f$ on X. Then for every $n \in \mathbb{N}$ we may find sets X_n so that $\mu X_n \leqslant \frac{1}{2^n}$ and f_n is continuous outside X_n. By Egorov's theorem 4.3.7, we may find a set $A_{\frac{\varepsilon}{3}}$ on which $f_n \to f$ uniformly and $\mu A_{\frac{\varepsilon}{3}} < \frac{\varepsilon}{3}$. Then we consider the set

$$F_0 = A \setminus \left(\bigcup_{n=N}^{\infty} X_n\right)$$

for N so large that $\sum\limits_{n=N}^{\infty} \mu X_n = \sum\limits_{n=N}^{\infty} \frac{1}{2^n} < \frac{\varepsilon}{3}$. Now for every $n \geqslant N$ the function f_n is continuous on F_0 thus f (being the uniform limit of $(f_n)_{n=1}^{\infty}$) is also continuous on F_0. To finish the proof, we merely need to approximate the set F_0 by a closed set $F_\varepsilon \subset F_0$ such that $\mu(F \setminus F_\varepsilon) < \frac{\varepsilon}{3}$, that is possible by Theorem 3.7.1. □

Thus, any measurable function is continuous outside of a set of arbitrary small measure. The next theorem gives us a tool for approximation of measurable functions by continuous functions.

4.5 Problems

Problem 4.1. Let $f \in S^+(X)$. Prove that $\sqrt{f} \in S^+(X)$.

Problem 4.2. Let $f \in S(X)$. Prove that $\sqrt[3]{f} \in S(X)$.

Problem 4.3. Let $f \in S(X)$. Is $\operatorname{sgn} f$ measurable?

Problem 4.4. Let $f \in S(X)$, and let $[f(x)]_n = \begin{cases} f(x) & \text{if} \quad |f(x)| \leqslant n \\ 0 & \text{if} \quad |f(x)| > n \end{cases}$. Is $[f(x)]_n$ measurable?

Problem 4.5. Prove that if the set $X(f > r)$ is measurable for any $r \in \mathbb{Q}$, then f is measurable. Is converse statement true? Explain the answer.

Problem 4.6. Prove that if for any $a, b \in \mathbb{R}$ either the sets $X(a \leqslant f \leqslant b)$, or the sets $X(a < f \leqslant b)$, or the sets $X(a \leqslant f < b)$, or the sets $X(a < f < b)$ are measurable, then $f \in S(X)$.

Problem 4.7. Let a function $f : \mathbb{R}^n \mapsto \mathbb{R}$ be continuous, and the functions $g_k : \mathbb{R} \mapsto \mathbb{R}$, $k = 1, \ldots, n$, be measurable on $\mathbb{R}$. Prove that the function $h : \mathbb{R} \mapsto \mathbb{R}$ defined as follows

$$h(x) = f(g_1(x), \ldots, g_n(x)),$$

is measurable.

Problem 4.8. Prove that if $\{f_k(x)\}_{k=1}^{+\infty} \in S(X)$, then the set of points $x \in X$ such that $\lim\limits_{k \to +\infty} f_k(x)$ exists is measurable.

Problem 4.9. Let the function $f : [0, 1] \mapsto \mathbb{R}$ be differentiable on $[0, 1]$. Prove that its derivative (the function $f' : [0, 1] \mapsto \mathbb{R}$) is Lebesgue measurable.

Problem 4.10. A complex-valued function $f : X \mapsto \mathbb{C}$ ($f(x) = u(x) + iv(x)$, $u, v : X \mapsto \mathbb{R}$) is called measurable if its real and imaginary parts u and v are measurable. Prove that the absolute value and the argument of a complex-valued measurable function are measurable functions.

Problem 4.11. Let $\{f_n\}_{n=1}^{+\infty} \in S(X)$, $f_n \xrightarrow{a.e.} f$ and $f_n \xrightarrow{a.e.} g$ on X. Prove that $f \stackrel{a.e.}{=} g$ on X.

Problem 4.12. Let $\{f_n\}_{n=1}^{+\infty} \in S(X)$, $f_n \xrightarrow{\mu} f$ and $f_n \xrightarrow{\mu} g$ on X. Prove that $f \stackrel{a.e.}{=} g$ on X.

Problem 4.13. Let $X = [0, 1]$, and μ be the Lebesgue measure on X. Let $f_n = e^{-nx}$, $n \in \mathbb{N}$, and $f_0 \equiv 0$. Prove that $f_n \xrightarrow{\mu} f_0$ on X.

Problem 4.14. Let $X = [0, \pi]$, and μ be the Lebesgue measure on X. Let $f_n = (\sin nx)^n$, $n \in \mathbb{N}$, and $f_0 \equiv 0$. Prove that $f_n \xrightarrow{\mu} f_0$ on X. Does $\{f_n\}_{n=1}^{+\infty}$ converge to $f_0 \equiv 0$ almost everywhere on X?

Hint: Use the followig Chebyshev's theorem:

Theorem 4.14.9. *For any irrational number α and for any real number β the inequality*

$$|\alpha q - p - \beta| < \frac{3}{p}$$

has infinitely many solutions (p, q), where $p \in \mathbb{N}$, $q \in \mathbb{Z}$.

Problem 4.15. Prove that the sequence $f_n(x) = |\sin nx|^{n^\alpha}$ does not converge to $f_0 \equiv 0$ a.e. on $X = [0, \pi]$ if $0 \leqslant \alpha \leqslant 2$. Does it have an a.e. limit if $\alpha > 2$? For what values of α does it converge in measure?

Problem 4.16. For any $\delta > 0$ find the "Egoroff set" X_δ ($\mu X_\delta < \delta$) such that the sequence $f_n(x) = \dfrac{n \sin x}{1 + n^2 \sin^2 x}$, $x \in [0, \pi]$, $n \in \mathbb{N}$, converges uniformly on $[0, \pi] \setminus X_\delta$.

Problem 4.17. For any $\delta > 0$ find the "Egoroff set" X_δ ($\mu X_\delta < \delta$) such that the sequence $f_n(x)$, $x \in [0,1]$, $n \in \mathbb{N}$, converges uniformly to $f \equiv 0$ on $[0,1] \setminus X_\delta$. Here

$$\text{a) } f_n(x) = e^{-n(1-x)}; \quad \text{b) } f_n(x) = \frac{2nx}{1+n^2x^2}; \quad \text{b) } f_n(x) = \frac{x^n}{1+x^n};$$

$$\text{d) } f_n(x) = \begin{cases} n^2 x, & 0 \leqslant x \leqslant \frac{1}{n}, \\ n^2\left(\frac{2}{n} - x\right), & \frac{1}{n} < x < \frac{2}{n}, \\ 0, & \frac{2}{n} \leqslant x \leqslant 1. \end{cases}$$

Problem 4.18. For any $\delta > 0$ find the "Egoroff set" X_δ ($\mu X_\delta < \delta$) such that the sequence $f_n = n\left(\sqrt{x + \frac{1}{n}} - \sqrt{x}\right)$, $x \in [0,1]$, $n \in \mathbb{N}$, converges uniformly on $[0,1] \setminus X_\delta$.

Problem 4.19 (Lusin's theorem)**.** Suppose f is measurable and finite-valued on X, and $\mu X < +\infty$. Then for any $\varepsilon > 0$ there exists a closed set $F_\varepsilon \subset X$ such that

$$\mu(X \setminus F_\varepsilon) < \varepsilon$$

and f is continuous on F_ε (that is, the restriction $f|_{F_\varepsilon}$ of the function f to the set F_ε is continuous on F_ε). *Hint*: Use approximation of measurable functions by simple functions and Egorov's theorem.

P.S. This problem actually says that every measurable and a.e. finite function is nearly continuous.

Problem 4.20. Let X be a compact set in $\mathbb{R}^n$, and let $F \subset X$ be closed. Suppose that ψ is defined and continuous on F. Prove that there exists a function $\varphi(x)$ defined on X satisfying the following properties

a) $\varphi(x)$ is continuous on X;

b) $\varphi(x) = \psi(x)$ for any $x \in F$;

c) $\max|\varphi(x)| = \max|\psi(x)|$.

Using this fact and Problem 4.19 prove that for any $f \in S(X)$ with $\mu X(f = \pm\infty) = 0$ and for any $\varepsilon > 0$, there exists a function $g(x)$ continuous on X such that $\mu X(f \neq g) < \varepsilon$, and if $|f(x)| \leqslant M$, then $|g(x)| \leqslant M$.

Problem 4.21. Construct a non-decreasing sequence of Lebesgue measurable simple functions $h_n(x)$ converging to a function $f(x)$ on X where

a) $f(x) = e^{-(3x+1)}$, $X = [0, +\infty)$;

b) $f(x) = e^{2x+1}$, $X = [1, +\infty)$;

c) $f(x) = \dfrac{1}{x-2}$, $X = (2, 20]$;

d) $f(x) = \dfrac{1}{(2x+1)(2x+3)}$, $X = [0, +\infty)$;

Problem 4.22. Prove that a function f monotone on an interval $[a,b] \subset \mathbb{R}$ is measurable (w.r.t. the Lebesgue measure).

Problem 4.23. Let $(X, \mathcal{M}, \mu)$ be a measure space, and $A \in \mathcal{M}$ with $\mu(A) < +\infty$. Suppose that f is a finite measurable function defined on A. Prove that the function $g(t) = \mu A(f > t)$ is non-strictly decreasing (non-increasing) and right-continuous on $\mathbb{R}$.

Chapter 5

Integration theory

In definition of the Lebesgue integral of a measurable function, we approximate the function by simple functions. By contrast, in definition of the Riemann integral of a function $f : [a, b] \mapsto \mathbb{R}$, we partition the domain $[a, b]$ into subintervals and approximate f by step functions that are constant on these subintervals. This difference is sometime expressed by saying that in the Lebesgue integral we partition the range, and in the Riemann integral we partition the domain. This trick allows to extend the set of integrable functions. Moreover, the Lebesgue integral is defined in $\mathbb{R}^n$ identically for any n, while the Riemann integral must be defined first on $\mathbb{R}$, and then extended to $\mathbb{R}^n$, $n \geqslant 2$. For functions defined on an abstract measure space, the Riemann integral is not defined at all.

Let $(X, \mathcal{M}, \mu)$ be a measure space, where the measure μ is supposed to be σ-additive and complete. And let $S(X, \mathcal{M}, \mu) = S(X)$ be the set of all functions μ-measurable on X. In what follows we assume all the considered sets to belong to the σ-algebra $\mathcal{M}$ and all the considered functions to be measurable (unless otherwise mentioned).

We construct the Lebesgue integral into three steps: First, we define the integral for simple functions, then we extend it to all nonnegative measurable functions, and then to all measurable functions.

5.1 Integral of simple functions

Let h be a measurable simple function such that all the sets in its representation (4.4.1) are measurable and of finite measure.

Definition 5.1.1. The expression

$$\int_X h(x)d\mu = \sum_{k=1}^{m} c_k \mu E_k \tag{5.1.1}$$

is called the *integral* of the simple function h over the set X w.r.t. the measure μ. Here we use the convention that $0 \cdot \infty = 0$.

One can verify that the value of the integral in (5.1.1) is independent on the representation (4.4.1). Indeed, among all the representations of the function h there exists a unique representation (4.4.2) in which all the numbers c'_j are distinct. Following the way of construction the representation (4.4.2), one has

$$\sum_{k=1}^{m} c_k \mu E_k = \sum_{j=1}^{l} c'_j \left(\bigcup_{k:\ c_k = c'_j} \mu E_k \right) = \sum_{j=1}^{l} c'_j \mu \left(\bigcup_{k:\ c_k = c'_j} E_k \right) = \sum_{j=1}^{\infty} c'_j \mu E'_j.$$

Thus, the integral of a simple function does not depend on its representation.

Let us study some properties of the integral of simple functions.

1) *If* $\mu X = 0$, *then* $\int\limits_X h(x)d\mu = 0$.

2) $\int\limits_X \alpha h(x)d\mu = \alpha \int\limits_X h(x)d\mu \qquad \forall \alpha \in \mathbb{R}$.

These two properties immediately follow from the definition of the integral.

3) $\int\limits_X (h_1(x) + h_2(x))d\mu = \int\limits_X h_1(x)d\mu + \int\limits_X h_2(x)d\mu$.

Let

$$h_1(x) = \sum_{k=1}^{m} c_k \chi_{E_k}(x), \qquad h_2(x) = \sum_{k=1}^{l} c'_j \chi_{E'_j}(x).$$

Consider the sets

$$E_{k,j} = E_k \bigcap E'_j, \qquad k = 1, 2, \dots, m \quad j = 1, 2, \dots, l.$$

Clearly,

$$E_k = \bigcupdot_{j=1}^{l} E_{k,j}, \qquad E'_k = \bigcupdot_{k=1}^{m} E_{k,j}, \qquad X = \bigcupdot_{j=1}^{l} \bigcupdot_{k=1}^{m} E_{k,j}.$$

It is also easy to see that $h_1(x)+h_2(x) = c_k+c'_j$ if $x \in E_{k,j}$, $k = 1, 2, \dots, m$, $j = 1, 2, \dots, l$. Therefore, the function $h_1(x) + h_2(x)$ is simple, so

$$\int\limits_X (h_1(x) + h_2(x)) = \sum_{k=1}^{m}\sum_{j=1}^{l}(c_k + c'_j)\mu E_{k,j} = \sum_{k=1}^{m} c_k \sum_{j=1}^{l} \mu E_{k,j} + \sum_{j=1}^{l} c'_j \sum_{k=1}^{m} \mu E_{k,j} =$$

$$= \sum_{k=1}^{m} c_k \mu E_k + \sum_{j=1}^{m} c'_j \mu E'_j = \int\limits_X h_1(x)d\mu + \int\limits_X h_2(x)d\mu.$$

4) *If* $X = X' \bigcupdot X''$, *then*

$$\int\limits_X h(x)d\mu = \int\limits_{X'} h(x)d\mu + \int\limits_{X''} h(x)d\mu.$$

Let $h(x) = \sum\limits_{k=1}^{m} c_k \chi_{E_k}(x)$. Consider the sets

$$E'_k = X' \bigcap E_k, \quad E''_k = X'' \bigcap E_k, \quad k = 1, 2, \dots, m.$$

Then we have $E_k = E'_k \bigcupdot E''_k$, $k = 1, 2, \dots, m$, and

$$\int\limits_X h(x)d\mu = \sum_{k=1}^{m} c_k \mu E_k = \sum_{k=1}^{m} c_k(\mu E'_k + \mu E''_k) = \sum_{k=1}^{m} c_k \mu E'_k + \sum_{k=1}^{m} c_k \mu E''_k = \int\limits_{X'} h(x)d\mu + \int\limits_{X''} h(x)d\mu.$$

5) *If* $h_1(x) \leqslant h_2(x)$ *for all* $x \in X$, *then*

$$\int\limits_X h_1(x)d\mu \leqslant \int\limits_X h_2(x)d\mu.$$

If $h(x) \geqslant 0$ on X, then, clearly, $\int\limits_X h(x)d\mu \geqslant 0$. Therefore, $\int\limits_X (h_1(x) - h_2(x))d\mu \geqslant 0$, and the property 5) follows now from the properties 2) and 3).

6) *If a non-increasing (i.e. non-strictly decreasing) sequence of nonnegative simple functions $(h_n)_{n=1}^{\infty}$ converges to 0 almost everywhere on X, then*

$$\lim_{n\to\infty}\int\limits_X h_n(x)d\mu = 0. \tag{5.1.2}$$

If $\mu X = 0$, then we are done, since all the integrals equal zero.

Suppose now that $0 < \mu X < +\infty$. The simple function $h_1(x)$ has only finitely many values (in its range), therefore, it is bounded, so there exists a number $M > 0$ such that $h_1(x) \leqslant M$ on X. Since the sequence $(h_n)_{n=1}^{\infty}$ is non-increasing and nonnegative, we have

$$0 \leqslant h_n(x) \leqslant M, \quad n \in \mathbb{N},\ x \in X. \tag{5.1.3}$$

Fix a number $\varepsilon > 0$ and define $\delta := \dfrac{\varepsilon}{2M}$. By Egoroff's theorem, Theorem 4.3.7, there exists a measurable set $X_\delta \subset X$ such that $\mu X_\delta < \delta$ and the sequence $(h_n)_{n=1}^{\infty}$ converges to 0 uniformly on $X \setminus X_\delta$. So, there exists a number $n_0 \in \mathbb{N}$ such that for any $n \geqslant n_0$

$$0 \leqslant h_n(x) < \frac{\varepsilon}{2\,\mu X}, \quad x \in X \setminus X_\delta. \tag{5.1.4}$$

Consider the following simple function

$$h_\varepsilon(x) = \begin{cases} M, & x \in X_\delta, \\ \dfrac{\varepsilon}{2\,\mu X}, & x \in X \setminus X_\delta. \end{cases}$$

From (5.1.3)–(5.1.4) it follows that $h_n(x) \leqslant h_\varepsilon(x)$ for any $x \in X$ whenever $n \geqslant n_0$. So by the property 5), one has for $n \geqslant n_0$

$$0 \leqslant \int\limits_X h_n(x)d\mu \leqslant \int\limits_X h_\varepsilon(x)d\mu = M \cdot \mu X_\delta + \frac{\varepsilon}{2\,\mu X} \cdot \mu(X \setminus X_\delta) <$$

$$< M \cdot \delta + \frac{\varepsilon}{2\,\mu X} \cdot \mu X = M \cdot \frac{\varepsilon}{2M} + \frac{\varepsilon}{2} = \varepsilon.$$

Finally, if $\mu X = +\infty$, then there exists $X_1 \subset X$ such that $\mu X_1 < +\infty$ and $h_1(x) = 0$ for any $x \in X \setminus X_1$. Since the sequence of non-negative simple functions $h_n(x)$ is non-increasing, we have that $h_n(x) = 0$, $n \in \mathbb{N}$, for any $x \in X \setminus X_1$. Now by the previous result, we have that (5.1.2) is true.

7) *If a sequence $(h_n)_{n=1}^{\infty}$ of nonnegative simple functions is non-increasing and*

$$\lim_{n\to+\infty}\int\limits_X h_n(x)d\mu = 0, \tag{5.1.5}$$

then $h_n \xrightarrow{a.e.} 0$ on X.

As before, suppose $\mu X > 0$ (otherwise, the assertion is trivial). Since the sequence $(h_n)_{n=1}^{\infty}$ is non-increasing and is bounded from below, it converges at any point x of the set X. Let

$$g(x) = \lim_{n\to+\infty} h_n(x), \qquad x \in X.$$

The functions h_n are nonnegative on X, so is the function $g(x)$, that is, $g(x) \geqslant 0$ for any $x \in X$. It suffices to show that $g(x) \overset{a.e.}{=} 0$ on X.

Indeed, it is clear that

$$X_0 = X(g \neq 0) = \bigcup_{k=1}^{\infty} X\left(g \geqslant \frac{1}{k}\right) = \bigcup_{k=1}^{\infty} X_k.$$

On the contrary, suppose that there exists an index k_0 such that $\mu X_{k_0} = \delta_0 > 0$, and consider a simple function h_0 defined as follows.

$$h_0(x) = \begin{cases} \dfrac{1}{k_0}, & x \in X_{k_0}, \\ \\ 0, & x \in X \setminus X_{k_0}. \end{cases}$$

For the function h_0, one has

$$h_n(x) \geqslant g(x) \geqslant h_0(x), \qquad n \in \mathbb{N}, \quad x \in X.$$

The property 5) now implies

$$\int_X h_n(x)d\mu \geqslant \int_X h_0(x)d\mu = \frac{1}{k_0} \cdot \mu X_0 + 0 \cdot \mu(X \setminus X_0) = \frac{1}{k_0} \cdot \delta_0 > 0, \qquad n \in \mathbb{N},$$

where we use our convention that $0 \cdot \infty = 0$. This contradicts with the condition (5.1.5), so $\mu X_k = 0$ for all $k \in \mathbb{N}$. From σ-additivity of the measure μ we have

$$\mu X_0 \leqslant \sum_{k=0}^{\infty} \mu X_k = 0.$$

5.2 Integral of nonnegative measurable functions

If $f \in S^+(X)$ (that is, f is nonnegative and measurable on X), then by Approximation Theorem 4.4.2 there exists a non-decreasing sequence of nonnegative simple functions $(h_n)_{n=1}^{\infty}$ convergent to f on the set X. By the property 5) of integral of simple functions the sequence of integrals

$$\left(\int_X h_n(x)d\mu\right)_{n\in\mathbb{N}}$$

is non-decreasing and has a limit (finite or infinite).

Definition 5.2.1. The expression

$$\int_X f(x)d\mu = \lim_{n\to+\infty} \int_X h_n(x)d\mu \tag{5.2.1}$$

is called the integral of the nonnegative function f over the set X w.r.t. measure μ.

If $\int_X f(x)d\mu$ is finite, then the function f is called *integrable* (or *summable*) on X. The set of all nonnegative integrable functions on X is denoted $L^+(X, \mathcal{M}, X)$, or, briefly, $L^+(X)$ if the measure μ is defined in advance.

Let us emphasize once again that the integral exists for **any** nonnegative measurable function but it can be infinite.

Before we start to study properties of the integral of nonnegative functions, we have to prove that the integral (5.2.1) is defined properly. Namely, we have to prove that it is independent on the sequence $(h_n)_{n=1}^{\infty}$ which we use to approximate f. We prove even a more general fact.

Theorem 5.2.2. *Let $f, g \in S^+(X)$ and $f(x) \leqslant g(x)$ for any $x \in X$. Suppose that $(h_n)_{n=1}^{\infty}$ and $(v_n)_{n=1}^{\infty}$ are non-decreasing sequences of nonnegative simple functions approximating the functions f and g, respectively. Then*

$$\lim_{n\to+\infty} \int_X h_n(x)d\mu \leqslant \lim_{n\to+\infty} \int_X v_n(x)d\mu. \tag{5.2.2}$$

Proof. Consider the difference $h_k - v_n$ for a fixed index k, $n \in \mathbb{N}$. Since the sequences $(v_n)_{n=1}^{\infty}$ is non-decreasing, the sequence $(h_k - v_n)_{n=1}^{\infty}$ is non-increasing, so it has limit as $n \to +\infty$

$$\lim_{n\to+\infty}(h_k(x) - v_n(x)) \leqslant \lim_{n\to+\infty}(f(x) - v_n(x)) = f(x) - g(x) \leqslant 0$$

for any $x \in X$.

Additionally, the sequence $(h_k - v_n)_{n\in\mathbb{N}}^{+}$ of positive parts of the sequence $(h_k - v_n)_{n\in\mathbb{N}}$ is non-increasing as well. Therefore,

$$\lim_{n\to+\infty}(h_k(x) - v_n(x))^+ = 0$$

for all $x \in X$.

By the property 6) of integrals of simple functions, we have

$$\lim_{n\to+\infty} \int_X (h_k(x) - v_n(x))^+ d\mu = 0.$$

From the inequality $(h_k(x) - v_n(x)) \leqslant (h_k(x) - v_n(x))^+$, it follows that

$$\lim_{n\to+\infty} \int_X (h_k(x) - v_n(x))d\mu \leqslant 0. \tag{5.2.3}$$

Here the integral (finite or equal $-\infty$) exists, since the sequence of integrals is non-increasing by the property 5). From (5.2.3) we obtain

$$\int_X h_k(x)d\mu \leqslant \lim_{n\to+\infty} \int_X v_n(x)d\mu.$$

From this inequality, we get (5.2.2) as $k \to +\infty$. $\square$

If we put $g(x) := f(x)$ and take two different non-decreasing sequences of nonnegative simple functions approximating f, we get by Theorem (5.2.2) that, on one hand,

$$\lim_{n\to+\infty} \int_X h_n(x)d\mu \leqslant \lim_{n\to+\infty} \int_X v_n(x)d\mu,$$

and, on the other hand, interchanging h_n and v_n (we can do this because of symmetry)

$$\lim_{n\to+\infty} \int_X v_n(x)d\mu \leqslant \lim_{n\to+\infty} \int_X h_n(x)d\mu.$$

Consequently,

$$\lim_{n\to+\infty} \int_X v_n(x)d\mu = \lim_{n\to+\infty} \int_X h_n(x)d\mu.$$

Thus, the definition of integral of nonnegative measurable functions does not depend on the non-decreasing sequences of simple nonnegative functions approximating these functions. Moreover, this fact implies that for any nonnegative measurable simple function $h(x)$, Definition 5.1.1 of the integral coincides with Definition 5.2.1, since we can put in Definition 5.2.1 $h_n(x) := h(x)$.

Now we are in a position to study properties of integral of nonnegative measurable functions.

1) *If $\mu X = 0$, then $\int\limits_X f(x)d\mu = 0$ for any nonnegative function f.*

On a null set any function is measurable, since the measure μ is assumed to be complete (see Example 4.1.2), so $f \in S^+(X)$. The property is obvious for simple functions, and can be proved for any function f in $S^+(X)$ by passage to the limit (when we approximate f by simple functions).

2) *If $\alpha \geqslant 0$, then*

$$\int\limits_X \alpha f(x)d\mu = \alpha \int\limits_X f(x)d\mu.$$

Moreover, if $f \in L^+(X)$, then $\alpha f \in L^+(X)$.

If $f(x) \geqslant 0$, then $\alpha f(x) \geqslant 0$ for any $x \in X$. If $h_n \nearrow f$, then $\alpha h_n \nearrow \alpha f$. Therefore,

$$\int\limits_X \alpha f d\mu = \lim_{n\to+\infty} \int\limits_X \alpha h_n d\mu = \alpha \lim_{n\to+\infty} \int\limits_X h_n d\mu = \alpha \int\limits_X f d\mu.$$

The second assertion is obvious.

3) *If $f, g \in S^+(X)$, then*

$$\int\limits_X (f(x) + g(x))d\mu = \int\limits_X f(x)d\mu + \int\limits_X g(x)d\mu,$$

and if $f, g \in L^+(X)$, then $f + g \in L^+(X)$.

Let $h_n \nearrow f$ and $v_n \nearrow g$. Then $h_n + v_n \nearrow f + g$, and

$$\int\limits_X (f(x) + g(x))d\mu = \lim_{n\to+\infty} \int\limits_X (h_n(x) + v_n(x))d\mu =$$

$$= \lim_{n\to+\infty} \int\limits_X h_n(x)d\mu + \lim_{n\to+\infty} \int\limits_X v_n(x)d\mu = \int\limits_X f(x)d\mu + \int\limits_X g(x)d\mu$$

The second assertion follows from the proved identity.

4) *If $X = X' \uplus X''$, where $X', X'' \in \mathcal{M}(X)$, then*

$$\int\limits_X f(x)d\mu = \int\limits_{X'} f(x)d\mu + \int\limits_{X''} f(x)d\mu,$$

and $f \in L^+(X)$ if, and only if, $f \in L^+(X')$ and $f \in L^+(X'')$.

If $h_n \nearrow f$ on X, then $h_n \nearrow f$ on both X' and X''. Conversely, if $h'_n \nearrow f$ on X' and $h''_n \nearrow f$ X'', then we set $h_n(x) := h'_n(x)$ for $x \in X'$, and $h_n(x) := h''_n(x)$ for $x \in X''$ to get $h_n(x) \nearrow f$ for $x \in X$. It is left to use the property 4) of integral of simple functions and pass to the limit as $n \to +\infty$.

5) *If $f(x) \leqslant g(x)$, $x \in X$, then*

$$\int\limits_X f(x)d\mu \leqslant \int\limits_X g(x)d\mu.$$

In particular, if $g \in L^+(X)$, then $f \in L^+(X)$.

This property is, in fact, Theorem 5.2.2.

6) *If $f \in L^+(X)$, then $\mu X(f = +\infty) = 0$. In other word, a nonnegative integrable function on X is finite almost everywhere on X.*

   Suppose that it is not true. Then $\mu X_0 = a > 0$, where $X_0 = X(f = +\infty)$. Consider the following sequence of simple functions:
$$h_n(x) = \begin{cases} n, & x \in X_0, \\ 0, & x \in X \setminus X_0. \end{cases}$$
   Since $f(x) \geqslant h_n(x)$ for $x \in X$, $n \in \mathbb{N}$, then by the property 5), one has
$$\int\limits_X f(x)d\mu \geqslant \int\limits_X h_n(x)d\mu = \int\limits_{X_0} nd\mu = n \cdot a \qquad \forall n \in \mathbb{N}.$$
   The number a is positive, therefore, the integral $\int\limits_X f(x)d\mu$ cannot be finite, a contradiction, since $f \in L^+(X)$ by assumption.

7) (Chebyshev's inequality) *If $f \in L^+(X)$, then*
$$\mu X(f \geqslant c) \leqslant \frac{1}{c} \int\limits_X f(x)d\mu, \qquad \forall c > 0.$$
   Let $c > 0$. Denote $X_c = X(f \geqslant c)$. Then we have
$$\int\limits_X f(x)d\mu = \int\limits_{X_c} f(x)d\mu + \int\limits_{X \setminus X_c} f(x)d\mu \geqslant \int\limits_{X_c} f(x)d\mu \geqslant c \cdot \mu X_c,$$
   as required.

8) *If $f \in S^+(X)$ and $\int\limits_X f(x)d\mu = 0$, then $f(x) = 0$ almost everywhere on X.*

   It is easy to check that $X(f > 0) = \bigcup\limits_{n=1}^{\infty} X\left(f \geqslant \frac{1}{n}\right)$. By Chebyshev's inequality we obtain
$$\mu X\left(f \geqslant \frac{1}{n}\right) \leqslant n \cdot \int\limits_X f(x)d\mu = 0.$$
   Due to σ-additivity of the measure μ, we have
$$\mu X(f > 0) \leqslant \sum_{n=1}^{\infty} \mu X\left(f \geqslant \frac{1}{n}\right) = 0.$$

9) *If $f \overset{a.e.}{=} g$ on X, then*
$$\int\limits_X f(x)d\mu = \int\limits_X g(x)d\mu.$$
   So $f \in L^+(X)$ if, and only, if $g \in L^+(X)$.

   Let $X' = X(f \neq g)$ and $X'' = X \setminus X''$. By assumption $\mu X' = 0$, and by construction $f(x) = g(x)$ for any $x \in X''$. From the properties 4) and 1) we have
$$\int\limits_X f(x)d\mu = \int\limits_{X'} f(x)d\mu + \int\limits_{X''} f(x)d\mu = \int\limits_{X'} g(x)d\mu + \int\limits_{X''} g(x)d\mu = \int\limits_X g(x)d\mu.$$

Theorem 5.2.3 (B. Levi)**.** *Let* $f(x) = \sum\limits_{k=1}^{\infty} f_k(x)$*, where* $f_k \in S^+(X)$*,* $k \in \mathbb{N}$*. Then* $f \in S^+(X)$*, and*

$$\int\limits_X f(x)d\mu = \sum_{k=1}^{\infty} \int\limits_X f_k(x)d\mu. \tag{5.2.4}$$

Proof. Let $s_n(x) = \sum\limits_{k=1}^{n} f_k(x)$. By the property 3) of measurable functions, we have $s_n(x) \in S^+(X)$. Therefore, $f \in S^+(X)$ by Theorem 4.3.1 as the limit of measurable functions. It is left to prove the identity (5.2.4).

Since $f(x) \geqslant s_n(x)$ for all $n \in \mathbb{N}$ and $x \in X$, by the properties 3) and 5) of integral of nonnegative functions, we have

$$\int\limits_X f(x)d\mu \geqslant \int\limits_X s_n(x)d\mu = \sum_{k=1}^{n} \int\limits_X f_k(x)d\mu.$$

If here $n \to +\infty$, one obtains

$$\int\limits_X f(x)d\mu \geqslant \sum_{k=1}^{\infty} \int\limits_X f_k(x)d\mu. \tag{5.2.5}$$

Let us now prove the opposite inequality. Let $(h_{k,j})_{j=1}^{\infty}$ be a sequence of nonnegative simple functions such that $h_{k,j} \nearrow f_k$ as $j \to \infty$, $k \in \mathbb{N}$, for any $x \in X$. Consider the functions

$$g_n(x) = \sum_{k=1}^{n} h_{k,n}(x).$$

The functions $g_n(x)$ are nonnegative measurable simple functions as sums of nonnegative measurable simple functions. Moreover, the sequence $(g_n)_{n=1}^{\infty}$ is non-decreasing, since

$$g_{n+1}(x) = \sum_{k=1}^{n+1} h_{k,n+1}(x) \geqslant \sum_{k=1}^{n} h_{k,n}(x) = g_n(x),$$

therefore, the sequence $(g_n)_{n=1}^{\infty}$ converges on X to a function g which is nonnegative and measurable by Theorem 4.3.1.

Let us fix $n \in \mathbb{N}$ and take an arbitrary $p \in \mathbb{N}$. Then

$$\sum_{k=1}^{n} h_{k,n+p}(x) \leqslant \sum_{k=1}^{n+p} h_{k,n+p}(x) = g_{n+p}(x) \leqslant \sum_{k=1}^{n+p} f_k(x) \leqslant f(x).$$

As $p \to +\infty$, we obtain

$$s_n(x) = \sum_{k=1}^{n} f_k(x) \leqslant g(x) \leqslant f(x). \tag{5.2.6}$$

Since $\lim\limits_{n\to+\infty} s_n(x) = f(x)$, then from (5.2.6) it follows that $g(x) = f(x)$ for $x \in X$. Consequently, $g_n(x) \nearrow f(x)$ for all $x \in X$. So we get by definition of the integral

$$\int\limits_X f(x)d\mu = \lim_{n\to+\infty} \int\limits_X g_n(x)d\mu = \lim_{n\to+\infty} \int\limits_X \sum_{k=1}^{n} h_{k,n}(x)d\mu \leqslant$$

$$\lim_{n\to+\infty} \int\limits_X \sum_{k=1}^{n} f_k(x)d\mu = \lim_{n\to+\infty} \sum_{k=1}^{n} \int\limits_X f_k(x)d\mu = \sum_{k=1}^{\infty} \int\limits_X f_k(x)d\mu.$$

This inequality together with (5.2.5) imply the identity (5.2.4). $\square$

Remark 5.2.4. Both parts of the identity (5.2.4) can be infinite.

Corollary 5.2.5. *If $(f_n)_{n=1}^{\infty}$ is a monotone non-decreasing sequence of nonnegative measurable functions, and $f(x) = \lim_{n\to+\infty} f_n(x)$, $x \in X$, then*

$$\int_X f(x)d\mu = \lim_{n\to+\infty} \int_X f_n(x)d\mu. \tag{5.2.7}$$

Proof. Consider the functions $\varphi_1(x) = f_1(x)$, $\varphi_k(x) = f_k(x) - f_{k-1}(x)$, $k \geqslant 2$. The functions $\varphi_k(x)$ are measurable and nonnegative on X, and $f(x) = \sum\limits_{k=1}^{\infty} \varphi_k(x)$. So by Levi's theorem, Theorem 5.2.3, we have

$$\int_X f(x)d\mu = \sum_{k=1}^{\infty} \int_X \varphi_k(x)d\mu = \lim_{n\to+\infty} \sum_{k=1}^{n} \int_X \varphi_k(x)d\mu = \lim_{n\to+\infty} \int_X \sum_{k=1}^{n} \varphi_k(x)d\mu = \lim_{n\to+\infty} \int_X f_k(x)d\mu.$$

□

Remark 5.2.6. Both parts of the identity (5.2.7) can be infinite.

Theorem 5.2.7 (Fatou)**.** *Let $(f_n)_{n=1}^{\infty}$ be a sequence of nonnegative measurable functions on X. Then*

$$\int_X \underline{\lim}\, f_n(x)d\mu \leqslant \underline{\lim} \int_X f_n(x)d\mu. \tag{5.2.8}$$

Proof. For any $x \in X$, we have

$$\underline{\lim}\, f_n(x) = \sup_{n\in\mathbb{N}} \left\{ \inf_{k\geqslant n} \{f_k(x)\} \right\}.$$

Introduce the functions $g_n(x) = \inf\limits_{k\geqslant n} \{f_k(x)\}$. Then it is clear that

$$g_n(x) \leqslant f_n(x), \qquad \forall x \in X,\ \forall n \in \mathbb{N}. \tag{5.2.9}$$

The sequence $g_n(x)$ is non-decreasing, since the number of terms in the sequence $\{f_n(x) : \ k \geqslant n\}$ decreases as n grows. Consequently, the limit $\lim\limits_{n\to+\infty} g_n(x)$ exists, and

$$\lim_{n\to+\infty} g_n(x) = \sup_{n\in\mathbb{N}}\{g_n(x)\} = \sup_{n\in\mathbb{N}} \left\{ \inf_{k\geqslant n} \{f_k(x)\} \right\} = \underline{\lim}\, f_n(x).$$

Now from Corollary 5.2.5 it follows that

$$\int_X \underline{\lim}\, f_n(x)d\mu = \int_X \lim_{n\to+\infty} g_n(x)d\mu = \lim_{n\to+\infty} \int_X g_n(x)d\mu. \tag{5.2.10}$$

The sequence $\left(\int_X f_n(x)d\mu \right)_{n=1}^{\infty}$ has the limit inferior (finite or equal to $+\infty$), since its terms are nonnegative. So there exists a subsequences of indices $(n_j)_{j=1}^{\infty}$ such that

$$\underline{\lim} \int_X f_n(x)d\mu = \lim_{j\to+\infty} \int_X f_{n_j}(x)d\mu \tag{5.2.11}$$

Now from (5.2.9)–(5.2.11) we obtain

$$\int_X \underline{\lim}\, f_n(x)d\mu = \lim_{n\to+\infty} \int_X g_n(x)d\mu = \lim_{j\to+\infty} \int_X g_{n_j}(x)d\mu \leqslant \lim_{j\to+\infty} \int_X f_{n_j}(x)d\mu = \underline{\lim} \int_X f_n(x)d\mu.$$

□

Corollary 5.2.8. *Let a sequence $(f_n)_{n=1}^{\infty}$ of nonnegative measurable functions on X converge to a nonnegative function f a.e. on X, and*

$$\int\limits_X f_n(x)d\mu \leqslant C, \qquad n \in \mathbb{N}.$$

Then

$$\int\limits_X f(x)d\mu \leqslant C$$

Proof. Indeed, the function $\underline{\lim}\, f_n$ is nonnegative by assumption and is measurable by Theorem 4.3.1. Since $f \overset{a.e.}{=} \underline{\lim}\, f_n (= \lim f_n)$ on X by assumption, the function f is measurable as well (by Lemma 4.3.3). From the property 9) of integral of nonnegative functions and from Theorem 5.2.7, we obtain

$$\int\limits_X f(x)d\mu = \int\limits_X \underline{\lim}\, f_n(x)d\mu \leqslant \underline{\lim} \int\limits_X f_n(x)d\mu \leqslant C, \qquad n \in \mathbb{N},$$

as required. □

Example 5.2.9. Note that in (5.2.8) strict inequality is possible. To show this, let us consider the Lebesgue measure μ on the interval $[0,1]$, and the sequence

$$f_n(x) = \begin{cases} n, & 0 \leqslant x \leqslant \dfrac{1}{n}, \\ 0, & \dfrac{1}{n} < x \leqslant 1. \end{cases}$$

Then we have

$$\int\limits_{[0,1]} f_n(x)d\mu = n \cdot \frac{1}{n} + 0 \cdot \left(1 - \frac{1}{n}\right) = 1, \qquad \forall n \in \mathbb{N},$$

therefore, $\underline{\lim} \int\limits_{[0,1]} f_n(x)d\mu = 1$. At the same time,

$$f(x) = \underline{\lim}\, f_n(x) = \begin{cases} +\infty, & x = 0, \\ 0, & 0 < x \leqslant 1. \end{cases}$$

so $f(x) \overset{a.e.}{=} g(x) \equiv 0$ on $[0,1]$. Thus, $\int\limits_{[0,1]} f(x)d\mu = \int\limits_{[0,1]} g(x)d\mu = 0$, and we get

$$0 = \int\limits_X \underline{\lim}\, f_n(x)d\mu < \underline{\lim} \int\limits_X f_n(x)d\mu = 1.$$

5.3 Integral of measurable functions

Let f be an arbitrary measurable function on a set X. Then it can be represented as follows

$$f(x) = f^+(x) - f^-(x),$$

where f^+ and f^- are defined in Definition 4.2.3.

The functions f^+ and f^- are nonnegative and measurable on X by the property 7) of measurable functions.

Definition 5.3.1. The expression

$$\int_X f(x)d\mu = \int_X f^+(x)d\mu - \int_X f^-(x)d\mu \tag{5.3.1}$$

is called the integral of the function f over the set X w.r.t. measure μ.

From the definition it follows that the integral (5.3.1) does not always exist. In fact, the following four cases are possible:

1) The integrals $\int_X f^+(x)d\mu$ and $\int_X f^-(x)d\mu$ are finite, so $\int_X f(x)d\mu$ is also finite.

2) The integral $\int_X f^+(x)d\mu$ is infinite but the integral $\int_X f^-(x)d\mu$ is finite. In this case, the integral $\int_X f(x)d\mu = +\infty$.

3) The integral $\int_X f^+(x)d\mu$ is finite but the integral $\int_X f^-(x)d\mu$ is infinite. In this case, the integral $\int_X f(x)d\mu = -\infty$.

4) The integrals $\int_X f^+(x)d\mu$ and $\int_X f^-(x)d\mu$ are infinite. In this case, the integral $\int_X f(x)d\mu$ *does not exist.*

Definition 5.3.2. A measurable function f is called (Lebesgue) *integrable* (or *summable*) on X w.r.t. measure μ if $\int_X f(x)d\mu$ exists and finite.

The set of all integrable functions on X is denoted $L(X, \mathcal{M}, X)$, or, briefly, $L(X)$ if the measure μ is defined in advance. It is clear that $L^+(X) \subset L(X)$, and that $f \in L(X)$ if, and only if, $f^+, f^- \in L^+(X)$.

Let us study properties of the integral of measurable functions.

1) *If $\mu X = 0$, then $\int_X f(x)d\mu = 0$ for any function f.*

   This property follows from the property 1) of integral of nonnegative functions.

2) *The function f is integrable on X if, and only if, the function $|f|$ is integrable on X. Moreover, if $f \in L(X)$, then*

$$\left|\int_X f(x)d\mu\right| \leqslant \int_X |f(x)|d\mu \tag{5.3.2}$$

   Indeed, let $f \in L(X)$, then $f^+, f^- \in L^+(X)$. Since $|f| = f^+ + f^-$, we obtain $|f| \in L^+(X)$ by the property 2) of integral of nonnegative functions.

   Conversely, let $|f| \in L^+(X)$. Since $f^+(x) \leqslant |f(x)|$ and $f^-(x) \leqslant |f(x)|$ for any $x \in X$, by the property 5) of integral of nonnegative functions we get $f^+, f^- \in L^+(X)$, so $f \in L(X)$. The inequality (5.3.2) follows from the fact that $|f| = f^+ + f^-$.

$$\left|\int\limits_X f(x)d\mu\right| = \left|\int\limits_X f^+(x)d\mu - \int\limits_X f^-(x)d\mu\right| \leqslant \left|\int\limits_X f^+(x)d\mu\right| + \left|\int\limits_X f^-(x)d\mu\right| =$$

$$= \int\limits_X f^+(x)d\mu + \int\limits_X f^-(x)d\mu = \int\limits_X |f(x)|d\mu.$$

3) *If* $f \overset{a.e.}{=} g$ *and the integral* $\int\limits_X f(x)d\mu$ *exists (but can be infinite), then the integral* $\int\limits_X g(x)d\mu$ *exists, and*

$$\int\limits_X f(x)d\mu = \int\limits_X g(x)d\mu.$$

Since $f \overset{a.e.}{=} g$, we have $f^+ \overset{a.e.}{=} g^+$ and $f^- \overset{a.e.}{=} g^-$, so $g^+, g^- \in L^+(X)$ by the property 9) of integral of nonnegative functions.

Corollary 5.3.3. *If* $f \in L(X)$, $f \overset{a.e.}{=} g$ *on* X, *then* $g \in L(X)$.

4) *If* $f \in L(X)$, *then* $\mu X(f = \pm\infty) = 0$.

If $f \in L(X)$, then $f^+, f^- \in L^+(X)$, and by the property 8) of integral of nonnegative functions we have $\mu X(f^+ = +\infty) = 0$ and $\mu X(f^- = +\infty) = 0$, so

$$\mu X(f = \pm\infty) = \mu X(f^+ = +\infty) \bigcup \mu X(f^- = +\infty) = 0$$

due to additivity of the measure μ.

5) *If* $X = X' \uplus X''$, *where* $X', X'' \in \mathcal{M}(X)$, *and the integral* $\int\limits_X f(x)d\mu$ *exists, then*

$$\int\limits_X f(x)d\mu = \int\limits_{X'} f(x)d\mu + \int\limits_{X''} f(x)d\mu. \tag{5.3.3}$$

If the integral $\int\limits_X f(x)d\mu$ exists, then one of the integrals $\int\limits_X f^+(x)d\mu$ or $\int\limits_X f^-(x)d\mu$ is finite. Without loss of generality we can suppose that the integral $\int\limits_X f^+(x)d\mu$ is finite. Then both integrals $\int\limits_{X'} f^+(x)d\mu$ and $\int\limits_{X''} f^+(x)d\mu$ are finite, so both integrals in the right-hand side of (5.3.3) exist. The identity (5.3.3) now follows from Definition 5.3.1 and from the property 4) of integral of nonnegative functions.

Remark 5.3.4. Existence of both integrals in the right-hand side of (5.3.3) does not imply the existence of the integral $\int\limits_X f(x)d\mu$.

Example 5.3.5. Let

$$f(x) = \begin{cases} +\infty, & x \in X', \\ -\infty, & x \in X'', \end{cases}$$

and $\mu X' > 0$, $\mu X'' > 0$. Then

$$\int\limits_{X'} f(x)d\mu = +\infty, \qquad \int\limits_{X''} f(x)d\mu = -\infty,$$

but the integral $\int\limits_X f(x)d\mu$ does not exists.

Corollary 5.3.6. *If $f \in L(X)$ and $X = X' \dot\cup X''$, where $X', X'' \in \mathcal{M}(X)$, then $f \in L(X')$ and $f \in L(X'')$. Conversely, if $f \in L(X')$ and $f \in L(X'')$, then $f \in L(X)$. In both cases the identity (5.3.3) holds.*

6) *If the integral $\int\limits_X f(x)d\mu$ exists and $\alpha \in \mathbb{R}$, then the integral $\int\limits_X \alpha f(x)d\mu$ exists, and*

$$\int\limits_X \alpha f(x)d\mu = \alpha \int\limits_X f(x)d\mu.$$

If the integral $\int\limits_X f(x)d\mu$ exists, then in the right-hand side of the identity

$$\int\limits_X f(x)d\mu = \int\limits_X f^+(x)d\mu - \int\limits_X f^-(x)d\mu.$$

one of integrals is finite. Without loss of generality, suppose that $\int\limits_X f^+(x)d\mu$ is finite. Furthermore, we have

$$\alpha f(x) = \begin{cases} \alpha f^+(x) - \alpha f^-(x), & \alpha \geqslant 0, \\ |\alpha| f^-(x) - |\alpha| f^+(x), & \alpha < 0. \end{cases}$$

Therefore, by Definition 5.3.1,

$$\int\limits_X \alpha f(x)d\mu = \begin{cases} \int\limits_X \alpha f^+(x)d\mu - \int\limits_X \alpha f^-(x)d\mu, & \alpha \geqslant 0, \\ \\ \int\limits_X |\alpha| f^-(x)d\mu - \int\limits_X |\alpha| f^+(x)d\mu, & \alpha < 0. \end{cases} \tag{5.3.4}$$

By the property 2) of integral of nonnegative functions, in (5.3.4), the integrals with f^+ are finite, so the integral in the left-hand side of (5.3.4) exists. Moreover, by the same property,

$$\int\limits_X \alpha f(x)d\mu = \begin{cases} \alpha \int\limits_X f^+(x)d\mu - \alpha \int\limits_X f^-(x)d\mu, & \alpha \geqslant 0, \\ \\ |\alpha| \int\limits_X f^-(x)d\mu - |\alpha| \int\limits_X f^+(x)d\mu, & \alpha < 0. \end{cases} =$$

$$= \alpha \left(\int\limits_X f^+(x)d\mu - \int\limits_X f^-(x)d\mu \right) = \alpha \int\limits_X f(x)d\mu.$$

Corollary 5.3.7. *If $f \in L(X)$, then $\alpha f \in L(X)$, $\alpha \in \mathbb{R}$.*

7) *If the integrals $\int\limits_X f(x)d\mu$ and $\int\limits_X g(x)d\mu$ exist, and if at least one of them is finite, then the integral $\int\limits_X (f(x)+g(x))d\mu$ exists, and*

$$\int\limits_X (f(x)+g(x))d\mu = \int\limits_X f(x)d\mu + \int\limits_X g(x)d\mu. \tag{5.3.5}$$

Without loss of generality, suppose that the integral $\int\limits_X f(x)d\mu$ is finite. Then the integrals $\int\limits_X f^+(x)d\mu$ and $\int\limits_X f^-(x)d\mu$, and one of the integrals $\int\limits_X g^+(x)d\mu$ and $\int\limits_X g^-(x)d\mu$ are also finite. Without loss of generality, suppose that the integral $\int\limits_X g^+(x)d\mu$. By the property 4) of integral of nonnegative functions, integrals of f^- and f^+ over any measurable subset of X are finite, as well. We must prove that one of the integrals $\int\limits_X (f(x)+g(x))^+d\mu$ or $\int\limits_X (f(x)+g(x))^-d\mu$ is finite, and then prove the identity (5.3.5).

We split the set X into the following six subsets

$$X_1 = X(f \geqslant 0, g \geqslant 0), \qquad X_4 = X(f < 0, g < 0),$$

$$X_2 = X(f \geqslant 0, g < 0, f+g \geqslant 0), \qquad X_5 = X(f \geqslant 0, g < 0, f+g < 0),$$

$$X_3 = X(f < 0, g \geqslant 0, f+g \geqslant 0), \qquad X_5 = X(f < 0, g \geqslant 0, f+g < 0).$$

It is clear that $X = \dot{\bigcup\limits_{j=1}^{6}} X_j$ and

$$(f+g)^+ = \begin{cases} f(x)+g(x), & X = \dot{\bigcup\limits_{j=1}^{3}} X_j, \\ 0, & X = \dot{\bigcup\limits_{j=4}^{6}} X_j, \end{cases}$$

$$(f+g)^- = \begin{cases} 0, & X = \dot{\bigcup\limits_{j=1}^{3}} X_j, \\ -(f(x)+g(x)), & X = \dot{\bigcup\limits_{j=4}^{6}} X_j. \end{cases}$$

On the set X_1 the functions f and g are nonnegative. Therefore, by the property 3) of integral of nonnegative functions, one has

$$\int\limits_{X_1} (f(x)+g(x))d\mu = \int\limits_{X_1} f(x)d\mu + \int\limits_{X_1} g(x)d\mu. \tag{5.3.6}$$

Moreover, all the integrals here are finite by assumption.

On the set X_2, we have $-g(x) \leqslant f(x)$, and the integral $\int\limits_{X_2} f(x)d\mu$ is finite by assumption. Therefore, the integral $\int\limits_{X_2} (-g(x))d\mu$ is also finite, so $\int\limits_{X_2} g(x)d\mu$ is finite by the property 6). Furthermore, in the identity

$$f(x) = (f(x) + g(x)) + (-g(x))$$

both functions in the right-hand side are nonnegative on X_2, so by the property 3) of integral of nonnegative functions and by the property 6) we get

$$\int\limits_{X_2} f(x)d\mu = \int\limits_{X_2} (f(x) + g(x))d\mu - \int\limits_{X_2} g(x)d\mu,$$

or

$$\int\limits_{X_2} (f(x) + g(x))d\mu = \int\limits_{X_2} f(x)d\mu + \int\limits_{X_2} g(x)d\mu. \tag{5.3.7}$$

All the integrals here are finite by assumption.

On the set X_3, we have $g(x) = (f(x) + g(x)) + (-f(x))$ where both functions in the right-hand side are nonnegative, so as above

$$\int\limits_{X_3} g(x)d\mu = \int\limits_{X_3} (f(x) + g(x))d\mu - \int\limits_{X_3} f(x)d\mu,$$

or

$$\int\limits_{X_3} (f(x) + g(x))d\mu = \int\limits_{X_3} f(x)d\mu + \int\limits_{X_3} g(x)d\mu. \tag{5.3.8}$$

Moreover, all the integrals in (5.3.8) are finite by assumption.

According to the property 4) of integral of nonnegative functions, we now obtain

$$\int\limits_{X} (f(x) + g(x))^{+}d\mu = \int\limits_{\dot{\bigcup}_{j=1}^{3} X_j} (f(x) + g(x))^{+}d\mu + \int\limits_{\dot{\bigcup}_{j=4}^{6} X_j} 0 \cdot d\mu = \sum_{j=1}^{3} \int\limits_{X_j} (f(x) + g(x))d\mu. \tag{5.3.9}$$

On the set X_4, we have $-(f(x) + g(x)) = (-f(x)) + (-g(x))$ where both functions in the right-hand side are nonnegative, so by the property 3) of integral of nonnegative functions and by the property 6) we get

$$\int\limits_{X_4} (f(x) + g(x))d\mu = \int\limits_{X_4} f(x)d\mu + \int\limits_{X_4} g(x)d\mu. \tag{5.3.10}$$

By assumption the integral $\int\limits_{X_4} f(x)d\mu$ is finite, but $\int\limits_{X_4} g(x)d\mu$ can be equal to $-\infty$, so the integral in the left-hand side of the identity (5.3.10) is either finite, or equal to $-\infty$.

Analogously, for the sets X_5 and X_6 we obtain

$$\int\limits_{X_5} (f(x) + g(x))d\mu = \int\limits_{X_5} f(x)d\mu + \int\limits_{X_5} g(x)d\mu \tag{5.3.11}$$

$$\int\limits_{X_6} (f(x) + g(x))d\mu = \int\limits_{X_6} f(x)d\mu + \int\limits_{X_6} g(x)d\mu \tag{5.3.12}$$

Moreover, the integral in the left-hand side of (5.3.12) is finite by assumption, while the integral in the left-hand side of (5.3.11) can equal $-\infty$.

Now by the property 4) of integral of nonnegative functions and by the property 6) we have

$$\int_X (f(x)+g(x))^- d\mu = \int_{\overset{3}{\underset{j=1}{\dot\bigcup}} X_j} 0\cdot d\mu + \int_{\overset{6}{\underset{j=4}{\dot\bigcup}} X_j} [-(f(x)+g(x))]\, d\mu = -\sum_{j=4}^{6} \int_{X_j} (f(x)+g(x))d\mu, \quad (5.3.13)$$

where the integral on the left-hand side can be either finite or equal to $+\infty$.

Thus, we proved that the integral $\int_X (f(x)+g(x))d\mu$ exists. Now from the property 5) and from the identities (5.3.6)–(5.3.13) one obtains

$$\int_X (f(x)+g(x))d\mu = \int_X (f(x)+g(x))^+ d\mu - \int_X (f(x)+g(x))^- d\mu = \sum_{j=1}^{6} \int_{X_j} (f(x)+g(x))d\mu =$$

$$= \sum_{j=1}^{6} \left(\int_{X_j} f(x)d\mu + \int_{X_j} g(x)d\mu \right) = \int_X f(x)d\mu + \int_X g(x)d\mu.$$

Other situations (when integral of f is infinite but the integral of g is finite and so on) can be proved analogously.

Remark 5.3.8. In the conditions of the property 7) there can appear such a situation when the sum $f(x)+g(x)$ is not defined at some points of the set (for instance, when $f(x) = +\infty$ but $g(x) = -\infty$). However, by the property 4) the set of such points is a null set (for example, in the proof of the property 7) we had $\mu X(f = \pm\infty) = 0$), therefore, by the property 3), we can change the function whose integral is finite to another function which is equal to the function we change a.e. on X and is everywhere finite on X (leaving the same notation for the new function). After that the sum $f(x)+g(x)$ is defined everywhere on X but the value of the integral $\int_X (f(x)+g(x))d\mu$ remains unchanged.

Corollary 5.3.9. *If $f, g \in L(X)$, then $f+g \in L(X)$.*

8) *If $f(x) \leqslant g(x)$ a.e. on X, then*

$$\int_X f(x)d\mu \leqslant \int_X g(x)d\mu, \quad (5.3.14)$$

provided both integrals exist.

Without loss of generality we can assume that $f(x) \leqslant g(x)$ everywhere on X. Otherwise, we redefine either one or both functions on the null-set where the inequality $f(x) \leqslant g(x)$ fails. According to the property 3), this operation does not affect existence and values of the corresponding integrals. It is clear that

$$f^+(x) \leqslant g^+(x), \qquad f^-(x) \geqslant g^-(x), \qquad x \in X.$$

By the property 5) of integral of nonnegative functions, we have

$$\int_X f^+(x)d\mu \leqslant \int_X g^+(x)d\mu, \qquad \int_X f^-(x)d\mu \geqslant \int_X g^-(x)d\mu.$$

Now we subtract the second inequality from the first one to obtain (5.3.14).

9) *If* $|f(x)| \leqslant g(x)$ *a.e. on* X, *and* $g \in L(X)$, *Then* $f \in L(X)$, *and*

$$\left|\int_X f(x)d\mu\right| \leqslant \int_X g(x)d\mu \tag{5.3.15}$$

Without loss of generality we can assume that $|f(x)| \leqslant g(x)$ everywhere on X. Otherwise, we redefine either one or both functions on the null-set where the inequality $|f(x)| \leqslant g(x)$ fails. According to the property 3), this operation does not affect existence and values of the corresponding integrals. The inequality (5.3.15) follows from the property 5) of nonnegative functions and from the property 2).

Corollary 5.3.10. *If a function* f *is measurable and a.e. bounded on* X, *and* $\mu X < +\infty$, *then* $f \in L(X)$. *Moreover, if* $m \leqslant f(x) \leqslant M$ *a.e. on* X, *then*

$$m \cdot \mu X \leqslant \int_X f(x)d\mu \leqslant M \cdot \mu X. \tag{5.3.16}$$

Proof. Since $|f(x)| \leqslant \max\{|m|, M\} =: C$ a.e. on X, and the simple function $h(x) \equiv C$ is integrable on X, the function f is integrable on X, as well. To prove (5.3.16), consider the simple functions $h_1(x) \equiv m$ and $h_2(x) \equiv M$ on X. We have $h_1(x) \leqslant f(x) \leqslant h_2(x)$ a.e. on X. Now (5.3.16) follows from (5.3.14). □

Remark 5.3.11. By Theorem 4.4.4 every a.e. bounded measurable functions is Lebesgue integrable. So every a.e. finite measurable function is integrable out of a set of arbitrary small measure.

10) (Absolute continuity of integral) *If* $f \in L(X)$, *then for any* $\varepsilon > 0$, *there exists* $\delta > 0$ *such that*

$$\left|\int_{X_\delta} f(x)d\mu\right| < \varepsilon$$

for any subset $X_\delta \subset X$ *with* $\mu X_\delta < \delta$.

Due to the property 2), it is enough to prove the absolute continuity of integral of nonnegative functions. So, suppose that $f \in L^+(X)$, and fix $\varepsilon > 0$. By Definition 5.2.1 of integral of nonnegative functions, there exists a nonnegative simple function h such that $h(x) \leqslant f(x)$ on X and

$$\int_X f(x)d\mu - \int_X h(x)d\mu = \int_X (f(x) - h(x))d\mu < \frac{\varepsilon}{2}$$

Let $h(x) = \sum\limits_{k=1}^{m} c_k \chi_{E_k}(x)$, and $M = \max\{c_k : \ k = 1, 2, \ldots, m\}$. Take the number $\delta := \dfrac{\varepsilon}{2M}$, and let $X_\delta \subset X$ be such that $\mu X_\delta < \delta$. Put $h_\delta \equiv M$ on X_δ, then $h(x) \leqslant h_\delta(x)$ for all $x \in X_\delta$, so by the property 4) of integral of simple functions, we have

$$\int_{X_\delta} h(x)d\mu \leqslant \int_{X_\delta} h_\delta(x)d\mu = M \cdot \mu X_\delta = M \cdot \frac{\varepsilon}{2M} = \frac{\varepsilon}{2}.$$

Thus, we obtain

$$\int_{X_\delta} f(x)d\mu = \int_{X_\delta} f(x)d\mu - \int_{X_\delta} h(x)d\mu + \int_{X_\delta} h(x)d\mu = \int_{X_\delta} (f(x)d\mu - h(x))d\mu + \int_{X_\delta} h(x)d\mu \leqslant$$

$$\leqslant \int_X (f(x) - h(x))d\mu + \int_{X_\delta} h(x)d\mu < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

as required.

11) (Complete additivity of integral) *Let $f \in L(X)$ and the set X can be represented as $X = \dot{\bigcup}_{k=1}^{\infty} X_k$, where all the sets X_k are measurable. Then*

$$\int_X f(x)d\mu = \sum_{k=1}^{\infty} \int_{X_k} f(x)d\mu.$$

Suppose first that $f \in L^+(X)$ and set

$$f_k(x) = \begin{cases} f(x), & x \in X_k, \\ 0, & x \in X \setminus X_k. \end{cases}$$

All the nonnegative functions f_k are measurable on X by the property 8) of measurable functions, and

$$f(x) = \sum_{k=1}^{\infty} f_k(x), \qquad x \in X.$$

Now by Levi's Theorem 5.2.3, we have

$$\int_X f(x)d\mu = \sum_{k=1}^{\infty} \int_X f_k(x)d\mu = \sum_{k=1}^{\infty} \int_{X_k} f(x)d\mu.$$

If f is an arbitrary integrable function on X, then $f = f^+ - f^-$, and the functions f^+ and f^- are integrable and nonnegative on X. Therefore,

$$\int_X f^+(x)d\mu = \sum_{k=1}^{\infty} \int_{X_k} f^+(x)d\mu, \qquad \int_X f^-(x)d\mu = \sum_{k=1}^{\infty} \int_{X_k} f^-(x)d\mu. \tag{5.3.17}$$

Since both series in the right-hand side of the identities (5.3.17) converge, we obtain

$$\int_X f(x)d\mu = \int_X f^+(x)d\mu - \int_X f^-(x)d\mu = \sum_{k=1}^{\infty} \left(\int_{X_k} f^+(x)d\mu - \int_{X_k} f^-(x)d\mu \right) = \sum_{k=1}^{\infty} \int_{X_k} f(x)d\mu.$$

Proposition 5.3.12. *Let $X = \dot{\bigcup}_{k=1}^{\infty} E_k$, $E_k \in \mathcal{M}(X)$, $k \in \mathbb{N}$, and*

$$f(x) = \sum_{n=1}^{\infty} a_n \chi_{E_n}(x).$$

The function $f \in L(X)$ if, and only if

$$\sum_{n=1}^{\infty} |a_n| \mu E_n < +\infty$$

Proof. It is sufficient to prove the proposition for the case when $f \geqslant 0$ on X. So $a_k \geqslant 0$, $k \in \mathbb{N}$. Consider the simple functions $h_n(x) = \sum_{k=1}^{n} a_k \chi_{E_k}(x)$. Clearly, the sequence $(h_n)_{n=1}^{\infty}$ is non-decreasing and

$$\lim_{n \to +\infty} h_n(x) = f(x).$$

Now by Definition 5.2.1, we have

$$\int_X f(x)d\mu = \lim_{n \to +\infty} \int_X h_n(x)d\mu = \lim_{n \to +\infty} \sum_{k=1}^{n} \int_X a_k \chi_{E_k}(x)d\mu = \lim_{n \to +\infty} \sum_{k=1}^{n} a_k \mu E_k = \sum_{k=1}^{\infty} a_k \mu E_k.$$

Thus, $f \in L^+(X)$ if, and only if, the series $\sum_{k=1}^{\infty} a_k \mu E_k$ converges. $\square$

Theorem 5.3.13 (Lebesgue's Dominated Convergence Theorem). *Let a sequence $(f_n)_{n=1}^{\infty}$ of integrable functions on X converge to a function f almost everywhere on X. If the inequality $|f_n(x)| \leqslant g(x)$, $n \in \mathbb{N}$, holds almost everywhere on X, and $g(x) \in L(X)$, then $f \in L(X)$, and*

$$\int_X f(x)d\mu = \lim_{n\to+\infty} \int_X f_n(x)d\mu. \tag{5.3.18}$$

Proof. From the conditions of the theorem it follows that $|f(x)| \leqslant g(x)$ a.e. on X, so $f \in L(X)$ by the property 9). If necessary, we redefine the functions f_n and f on sets of measure zero, so that f_n converges to f everywhere on X (this operation does not affect the value of the integrals of these functions), and the inequality $|f_n(x)| \leqslant g(x)$ holds everywhere on X.

Applying Fatou's Theorem 5.2.7 to the sequence $(g + f_n)_{n=1}^{\infty}$ of nonnegative functions on X, we get

$$\int_X \underline{\lim}\,(g(x) + f_n(x))d\mu \leqslant \underline{\lim} \int_X (g(x) + f_n(x))d\mu. \tag{5.3.19}$$

Now since $f_n(x) \underset{n\to+\infty}{\longrightarrow} f(x)$ on X, we have

$$\underline{\lim}(g(x) + f_n(x)) = \lim(g(x) + f_n(x)) = g(x) + \lim f_n(x) = g(x) + f(x),$$

so the left-hand side of (5.3.19) has the form

$$\int_X g(x)d\mu + \int_X f(x)d\mu. \tag{5.3.20}$$

At the same time, the right-hand side of (5.3.19) have the form

$$\int_X g(x)d\mu + \underline{\lim} \int_X f_n(x)d\mu. \tag{5.3.21}$$

From (5.3.19)–(5.3.21) we obtain

$$\int_X f(x)d\mu \leqslant \underline{\lim} \int_X f_n(x)d\mu. \tag{5.3.22}$$

Now let us consider the sequence $(g - f_n)_{n=1}^{\infty}$ of nonnegative functions on X. In the same way as above, one can prove that

$$-\int_X f(x)d\mu \leqslant \underline{\lim} \left(-\int_X f_n(x)d\mu \right),$$

or

$$\int_X f(x)d\mu \geqslant \overline{\lim} \int_X f_n(x)d\mu. \tag{5.3.23}$$

From (5.3.22) and (5.3.23) it follows that $\lim\limits_{n\to+\infty} \int\limits_X f_n(x)d\mu$ exists and the identity (5.3.18) holds. □

5.3.1 Invariance Properties

Definition 5.3.14. Let f be a function defined on $\mathbb{R}^n$. The function f_c, defined as $f_c(x) = f(x - c)$ is called the *translation* of f by a vector $c \in \mathbb{R}^n$.

The following theorem is true.

Theorem 5.3.15. *Let $f \in L(\mathbb{R}^n)$. Then $f_c \in L(\mathbb{R}^n)$, and*

$$\int\limits_{\mathbb{R}^n} f(x-c)d\mu = \int\limits_{\mathbb{R}^n} f(x)d\mu. \tag{5.3.24}$$

Proof. Suppose first that $f(x) = \chi_A(x)$, where A is a measurable set. Then obviously $f_c(x) = \chi_{A_c}$, where $A_c = \{x + c : \ x \in A\}$. In this case, the assertion of the theorem holds, since $\mu A = \mu A_c$ as we proved in Section 3.8, and

$$\int\limits_{\mathbb{R}^n} f(x-c)d\mu = \mu A_c = \mu A = \int\limits_{\mathbb{R}^n} f(x)d\mu.$$

As a result of linearity, the identity (5.3.24) holds for all simple functions.

Now if $f(x)$ is non-negative and $(h_n(x))_{n=1}^{\infty}$ is a sequence of simple functions that increase pointwise a.e to f (such a sequence exists by Theorem 4.4.2), then $(h_n(x-c))_{n=1}^{\infty}$ is a sequence of simple functions that increase to $f_c(x)$ pointwise a.e, and Corollary 5.2.5 implies (5.3.24) in this special case. In general case, we represent the function as $f = f^+ - f^-$ where the nonnegative functions f^+ and f^- are defined in Definition 4.2.3. Since the assertion of the theorem holds for f^+ and f^-, it holds for the function f. □

In the same way one can prove the following theorem.

Theorem 5.3.16. *Let $f \in L(\mathbb{R}^n)$. Then $f(ax) \in L(\mathbb{R}^n)$ for $a \in \mathbb{R} \setminus \{0\}$, and*

$$\int\limits_{\mathbb{R}^n} f(ax)d\mu = |a|^{-n} \int\limits_{\mathbb{R}^n} f(x)d\mu. \tag{5.3.25}$$

5.4 Difference between Riemann and Lebesgue definite integrals

One can compare the Riemann and Lebesgue integrals only if both are defined, that is, if X is rectifiable set in $\mathbb{R}^n$ and μ is the Lebesgue measure. For the sake of simplicity, let us consider the case $X = [a, b]$. In what follow, we denote Lebesgue integral as $(L)\int\limits_a^b f(x)dx$ instead of $\int\limits_{[a,b]} f(x)d\mu$, omitting the symbol (L) whenever it is clear that we use Lebesgue integration. As well, we denote Riemann integral as $(R)\int\limits_a^b f(x)dx$.

To define Riemann integral, one needs to divide the integration domain into l parts I_k. Then, in each part we choose a point and consider the integral sum $\sigma = \sum\limits_{k=1}^{l} f(\xi_k)|I_k|$. The Riemann integral exists, that is, the integral sums converge if the arbitrariness of choosing points ξ_k does not affect the value of the integral sum, that is, if the integrand f is not "too discontinuous".

Lebesgue's approach to integration is different in essence. Let us state it briefly. Suppose that f be a *bounded* Lebesgue measurable function on $[a, b]$, and set

$$m := \inf\{f(x) : \ x \in [a, b]\}, \qquad M := \sup\{f(x) : \ x \in [a, b]\}.$$

Split the interval $[m, M]$ into several parts by points $m = y_0 < y_1 < y_2 < \ldots < y_l = M$, and put $X_k := \{x \in [a, b] : \ y_{k-1} < f(x) < y_k\}$, $k = 1, \ldots, l$. Then we choose points $\xi_k \in X_k$ and construct the following integral sum $\sigma = \sum\limits_{k=1}^{l} f(\xi_k)\mu X_k$.

Furthermore, one introduces the lower and upper integral sums, $s = \sum\limits_{k=1}^{l} y_{k-1} \cdot \mu X_k$ and $S = \sum\limits_{k=1}^{l} y_k \cdot \mu X_k$ whose properties are similar to the ones of the lower and upper Darboux sums. It can be proved that for any bounded Lebesgue measurable function on $[a, b]$, the following identity holds

$$\lim_{l \to +\infty} s = \lim_{l \to +\infty} S = (L) \int_a^b f(x) dx.$$

Thus, any bounded Lebesgue measurable function is Lebesgue integrable.

In fact, Definition 4.2 of integral of nonnegative functions defines Lebesgue integral as the limit of lower integral sums.

Thus, in Lebesgue's approach we divide the integration domain into parts where *the values* of the integrand are close to each other, while in Riemann integration the integration domain is split into small parts regularly. So Lebesgue's approach allows us to substantially extend the class of integrable function. For example, Dirichlet's function

$$D(x) = \begin{cases} 1, & x \in \mathbb{Q} \bigcap [a, b], \\ 0, & x \in [a, b] \setminus \mathbb{Q}, \end{cases}$$

is non-Riemann integrable (see Example 5.4.8 below) but it is a simple function from Lebesgue's point of view, so

$$(L) \int_a^b D(x) dx = 1 \cdot \mu \left(\mathbb{Q} \bigcap [a, b] \right) + 0 \cdot \mu \left([a, b] \setminus \mathbb{Q} \right) = 1 \cdot 0 + 0 \cdot 1 = 0.$$

The following theorem shows that the class of Lebesgue integrable functions contains the class of Riemann integrable functions.

Theorem 5.4.1. *If a function is Riemann integrable on $[a, b]$, then it is Lebesgue integrable on $[a, b]$, and*

$$(L) \int_a^b f(x) dx = (R) \int_a^b f(x) dx.$$

Proof. Let $(P_n)_{n \in \mathbb{N}}$ be a partition of the interval $[a, b]$ such that $a = x_0^{(n)} < x_1^{(n)} < x_2^{(n)} < \ldots < x_{k_n}^{(n)} = b$. Moreover, let the partition P_{n+1} be a refinement of P_n for any $n \in \mathbb{N}$, and $\Delta_n \to 0$ as $n \to +\infty$, where $\Delta_n = \max\{|x_j^{(n)} - x_{j-1}^{(n)}| : \ j = 1, \ldots, k_n\}$ is the norm of the partition P_n. Let us set $I_i^{(n)} := [x_{i-1}^{(n)}, x_i^{(n)}]$, $|I_i^{(n)}| = x_i^{(n)} - x_{i-1}^{(n)}$, and

$$m_i^{(n)} = \inf\{f(x) : \ x \in I_i^{(n)}\}, \qquad M_i^{(n)} = \sup\{f(x) : \ x \in I_i^{(n)}\},$$

where $i = 1, \ldots, k_n$, $n \in \mathbb{N}$. Introduce the lower and the upper Darboux sums

$$s_n = \sum_{i=1}^{k_n} m_i^{(n)} |I_i^{(n)}|, \qquad S_n = \sum_{i=1}^{k_n} M_i^{(n)} |I_i^{(n)}| \tag{5.4.1}$$

The Riemann integrability of the function f means that

$$\lim_{n \to +\infty} s_n = \lim_{n \to +\infty} S_n = (R) \int_a^b f(x) dx. \tag{5.4.2}$$

Consider two sequences of simple functions

$$h_n(x) = m_i^{(n)}, \quad x \in \left(x_{i-1}^{(n)}, x_i^{(n)} \right), \quad i = 1, \ldots, k_n, \quad n \in \mathbb{N},$$

$$\varkappa_n(x) = M_i^{(n)}, \quad x \in \left(x_{i-1}^{(n)}, x_i^{(n)} \right), \quad i = 1, \ldots, k_n, \quad n \in \mathbb{N}.$$

At the points $x_i^{(n)}$, these functions can be defined arbitrarily. Indeed, for each n we have finitely many such points, and for all $n \in \mathbb{N}$ we get countably many points where we define the functions h_n and $\varkappa_n$ arbitrarily. But any countable set is a null-set, so it is not important for constructing the Lebesgue integral of the function f how we define the functions h_n and $\varkappa_n$ at the points $x_i^{(n)}$.

The functions h_n and $\varkappa_n$ are measurable on $[a,b]$. Moreover, since each partition P_n is a refinement of the previous partition P_{n-1}, the functions h_n cannot decrease (a.e. on $[a,b]$), while the functions $\varkappa_n$ cannot increase (a.e. on $[a,b]$). Consequently, the following limits exist a.e. on $[a,b]$

$$g_1(x) = \lim_{n\to+\infty} h_n(x), \qquad g_2(x) = \lim_{n\to+\infty} \varkappa_n(x),$$

where the functions $g_1(x)$ and $g_2(x)$ are measurable on $[a,b]$ by Theorem 4.3.6.

Since the following holds

$$h_n(x) \leqslant f(x) \leqslant \varkappa_n(x), \qquad \text{a.e. on } [a,b],$$

in the limit case we have

$$g_1(x) \leqslant f(x) \leqslant g_2(x), \qquad \text{a.e. on } [a,b]. \tag{5.4.3}$$

Consider now the sequence $(\varkappa_n - h_n)_{n=1}^{\infty}$. This sequence converges to $g_2(x) - g_1(x)$ a.e. on $[a,b]$, and is majorized by the integrable function $\varkappa_1 - h_1$. Consequently, we can apply Lebesgue's Dominated Convergence Theorem 5.3.13 to the sequence $(\varkappa_n - h_n)_{n=1}^{\infty}$:

$$\int_a^b (g_2(x) - g_1(x))dx = \lim_{n\to+\infty} \int_a^b (\varkappa_n(x) - h_n(x))dx =$$

$$= \lim_{n\to+\infty} \int_a^b \varkappa_n(x)dx - \lim_{n\to+\infty} \int_a^b h_n(x)dx = \lim_{n\to+\infty} S_n - \lim_{n\to+\infty} s_n = 0.$$

The function $g_2(x) - g_1(x) \geqslant 0$ a.e. on $[a,b]$, so from the property 8) of integral of nonnegative functions it follows that $g_2(x) - g_1(x) = 0$ a.e. on $[a,b]$. Now the inequalities (5.4.3) imply $f \overset{a.e.}{=} g_1$ on $[a,b]$. Thus, f is measurable on $[a,b]$ by Lemma 4.3.3.

Since f is bounded and measurable on $[a,b]$, it is Lebesgue integrable on $[a,b]$ by Corollary 5.3.10, so according to Corollary 5.2.5 we obtain

$$(L)\int_a^b f(x)dx = (L)\int_a^b g_1(x)dx = \lim_{n\to+\infty} \int_a^b h_n(x)dx = \lim_{n\to+\infty} s_n = (R)\int_a^b f(x)ds,$$

as required. □

5.4.1 Discontinuities of Riemann integrable functions

It is very well known that continuous, piecewise continuous, and monotone functions are Riemann integrable (see e.g. [9, Chapter 6]). In this section we completely describe the class of Riemann integrable functions in terms of their discontinuities. To do this, let us introduce the following objects.

Let $J := [a,b]$ and let $I(c,r) := (c-r, c+r)$ be the open interval centered at c of radius $r > 0$. Suppose that $f : [a,b] \mapsto \mathbb{R}$ is a *bounded* function.

Definition 5.4.2. The **oscillation of** f **on** $I(c,r)$ is the following value

$$\operatorname{osc}(f,c,r) = \sup |f(x) - f(y)|$$

where the supremum is taken over all $x, y \in J \bigcap I(c,r)$. This quantity exists since f is bounded.

Definition 5.4.3. The **oscillation of** f **at** c is the value

$$\operatorname{osc}(f,c) = \lim_{r\to 0} \operatorname{osc}(f,c,r).$$

This limit exists because $\operatorname{osc}(f,c,r) \geqslant 0$ is a non-strictly decreasing function of r.

It is clear from Definitions 5.4.2 and 5.4.3 that f is continuous at c if, and only if, $\operatorname{osc}(f,c) = 0$ (This is a Baire theorem, see [8, Ch. V, §4]). For each $\varepsilon > 0$ we define a set A_ε by

$$A_\varepsilon = \{c \in J : \ \operatorname{osc}(f,c) \geqslant \varepsilon\}.$$

It is easy to see that the set of points in J where f is discontinuous is $\bigcup\limits_{\varepsilon>0} A_\varepsilon$.

Lemma 5.4.4. *If $\varepsilon > 0$, then the set A_ε is closed (and therefore compact).*

Proof. Suppose that $c_n \in A_\varepsilon$ converges to c and assume that $c \notin A_\varepsilon$, say, $\operatorname{osc}(f,c) = \varepsilon - \delta$ where $\delta > 0$. Select r so that $\operatorname{osc}(f,c,r) < \varepsilon - \dfrac{\delta}{2}$, and choose n with $|c_n - c| < \dfrac{r}{2}$. So if $x, y \in J \bigcap I\left(c_n, \dfrac{r}{2}\right)$, then $x, y \in J \bigcap I\,(c, r)$, therefore, $\operatorname{osc}\left(f, c_n, \dfrac{r}{2}\right) < \varepsilon$ which implies $\operatorname{osc}(f, c_n) < \varepsilon$, a contradiction. □

Now we are in s position to describe the class of Riemann integrable functions in terms of their discontinuities.

Theorem 5.4.5 (Lebesgue). *A bounded function f on $[a,b]$ is Riemann integrable if, and only if, it is continuous almost everywhere on $[a,b]$.*

Proof. By assumption, there exists a number $M > 0$ such that $|f(x)| \leqslant M$ on $[a,b]$.

Suppose that the set D of discontinuities of f has Lebesgue measure 0, and let $\varepsilon > 0$. Since $A_\varepsilon \subset D$, we have $\mu(A_\varepsilon) = 0$, since Lebesgue measure μ is complete by Theorem 3.4.13. The set A_ε is measurable, and its Lebesgue measure coincides with its outer measure. So by Definition 3.3.1, given $\varepsilon > 0$ there exists a cover of A_ε by intervals (bricks in $\mathbb{R}$), $A_\varepsilon \subset \bigcup\limits_j I_j$, such that

$$\sum_j |I_j| \leqslant \mu^*(A_\varepsilon) + \frac{\varepsilon}{2} = \frac{\varepsilon}{2}.$$

Moreover, we can find open intervals $\widetilde{I}_j \supset I_j$ such that

$$|\widetilde{I}_j| < |I_j| + \frac{\varepsilon}{2^{j+1}},$$

so we have

$$\sum_j |\widetilde{I}_j| < \sum_j |I_j| + \frac{\varepsilon}{2} \leqslant \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

The cover $\bigcup\limits_j \widetilde{I}_j \supset A_\varepsilon$ of the closed set A_ε by open sets $\widetilde{I}_j$ contains a finite subcover $A_\varepsilon \subset \bigcup\limits_{j=1}^{N} \widetilde{I}_j =: I$ such that

$$|I| = \sum_{j=1}^{N} |\widetilde{I}_j| < \varepsilon.$$

The complement $CI = [a,b] \setminus I$ of I is compact, and around each point z in this complement we can find an interval F_z with $\sup\limits_{x,y\in F_z} |f(x) - f(y)| < \varepsilon$, since $z \notin A_\varepsilon$. We may now choose a finite subcover of $\bigcup\limits_{z\in CI} F_z$, which we denote by $I_{N+1}, ..., I_{N'}$. Now, taking all the end points of the intervals $I_1, I_2, ..., I_{N'}$ we obtain a partition P of $[a,b]$ with

$$S_{N'} - s_{N'} \leqslant 2M \sum_{j=1}^{N} |I_j| + \varepsilon(b-a) \leqslant C\varepsilon,$$

where the Darboux sums $s_{N'}$ and $S_{N'}$ are defined in (5.4.1). Hence f is integrable on $[a, b]$ by (5.4.2), as required.

Conversely, suppose that f is integrable on $[a, b]$, and let D be its set of discontinuities. Since D equals $\bigcup\limits_{n=1}^{\infty} A_{\frac{1}{n}}$, it suffices to prove that each $A_{\frac{1}{n}}$ has measure 0 according to Corollary 3.4.11. Let $\varepsilon > 0$ and choose a partition $P = \{x_0, x_1, ..., x_N\}$ so that $S_N - s_N < \dfrac{\varepsilon}{n}$. Then, if $A_{\frac{1}{n}}$ intersects $I_j = (x_{j-1}, x_j)$ we must have $\sup\limits_{x\in I_j} f(x) - \inf\limits_{x\in I_j} f(x) \geqslant \dfrac{1}{n}$, and this shows that

$$\frac{1}{n} \sum_{\{j:\ I_j \bigcap A_{\frac{1}{n}} \neq \varnothing\}} |Ij| \leqslant \sum_{\{j:\ I_j \bigcap A_{\frac{1}{n}} \neq \varnothing\}} [\sup_{x\in I_j} f(x) - \inf_{x\in I_j} f(x)]|Ij| \leqslant S_N - s_N < \frac{\varepsilon}{n}.$$

So by taking intervals intersecting $A_{\frac{1}{n}}$ and making them slightly larger, we can cover $A_{\frac{1}{n}}$ with open intervals of total length less than 2ε. Therefore, $A_{\frac{1}{n}}$ has measure 0, and we are done. $\square$

Corollary 5.4.6. *Continuous, piece-wise continuous, and monotone bounded functions on $[a, b]$ are Riemann integrable on $[a, b]$.*

Proof. Indeed, the set of discontinuity is empty for continuous functions, finite for piece-wise continuous functions, and at most countable for monotone functions (see Homework), so its Lebesgue measure is 0 for all these classes of functions. $\square$

Example 5.4.7. Let us consider the Dirichlet function

$$\mathcal{D}(x) = \begin{cases} 1, & x \in \mathbb{Q} \bigcap [0, 1], \\ 0, & x \in [0, 1] \setminus \mathbb{Q}. \end{cases}$$

The function $\mathcal{D}(x)$ is discontinuous at any point of $[0, 1]$, since for any $0 < \varepsilon < 1$, $\delta > 0$, and for any $x_0 \in [0, 1]$ there exist $x \in [0, 1] \bigcap (x_0 - \delta, x_0 + \delta)$ such that $|\mathcal{D}(x) - \mathcal{D}(x_0)| = 1 > \varepsilon$. Thus, $\mathcal{D}(x) \notin \mathcal{R}[0, 1]$ by Theorem 5.4.5, since its set of discontinuities is the whole interval $[0, 1]$ whose Lebesgue measure is positive (equals 1).

However, the set $X_0 := \mathbb{Q} \bigcap [0, 1]$ is a subset of the countable set $\mathbb{Q}$. Since $\mu\mathbb{Q} = 0$ and the Lebesgue measure is complete, we have $\mu X_0 = 0$. Therefore, $\mathcal{D} \overset{a.e.}{=} 0$ on $[0, 1]$, and the function $g(x) \equiv 0$ is measurable and integrable on $[0, 1]$. Thus, $\mathcal{D} \in L^+[0, 1]$ and by property 3) of integral of measurable functions

$$\int\limits_{[0,1]} \mathcal{D}(x) d\mu = \int\limits_{[0,1]} 0 \cdot d\mu = 0.$$

Example 5.4.8. Consider the Riemann function

$$f(x) = \begin{cases} 0, & x \in [0, 1] \setminus \mathbb{Q}, \\ \\ \dfrac{1}{n}, & x = \dfrac{m}{n} \quad \text{where } m \text{ and } n \text{ are coprime, } m < n. \end{cases}$$

The set of discontinuities of the function f is $X_0 := \mathbb{Q} \bigcap [0, 1]$. Indeed, if $c \in X_0$, then there exists a number n such that $f(c) = \dfrac{1}{n}$. Let us take an arbitrary number $0 < \varepsilon_0 < \dfrac{1}{n}$. Then $\forall \delta > 0$ there exists $x \in [0, 1] \bigcap (c - \delta, c + \delta)$ such that $f(x) = 0$, so we have

$$|f(x) - f(c)| = \frac{1}{n} > \varepsilon_0.$$

Consequently, X_0 is a subset of the set of discontinuities of the function f.

Let us show that f is continuous on $[0,1] \setminus X_0$. In fact, suppose that $a \in [0,1] \setminus X_0$. For any $\varepsilon > 0$, there exist only *finitely* many positive integers not exceeding $\frac{1}{\varepsilon}$. So there are only finitely many rational numbers in X_0 such that $f(m/n) \geqslant \varepsilon$. One can choose $\delta > 0$ so small that the interval $(a-\delta, a+\delta)$ does not contain these rational numbers, we have

$$|f(x) - f(a)| < \varepsilon$$

for any $x \in (a-\delta, a+\delta) \bigcap [0,1]$. Thus, f is continuous on $[0,1] \setminus X_0$.

The set X_0 is dense in $[0,1]$ but $\mu X_0 = 0$, so $f \in \mathcal{R}[0,1]$ by Theorem 5.4.5. Moreover, $f \overset{a.e.}{=} 0$ on $[0,1]$, so

$$\int\limits_{[0,1]} f(x) d\mu = \int\limits_{[0,1]} 0 \cdot d\mu = 0.$$

Example 5.4.9. Let $X = [0,1]$. For $n \in \mathbb{N}$, consider the sets

$$A_n = \left[\frac{1}{n+1}, \frac{1}{n}\right), \quad B_n = \left[\frac{1}{n+1}, \frac{1}{2}\left(\frac{1}{n+1} + \frac{1}{n}\right)\right), \quad C_n = \left[\frac{1}{2}\left(\frac{1}{n+1} + \frac{1}{n}\right), \frac{1}{n}\right),$$

and construct the function

$$f(x) = \sum_{n=1}^{\infty} (n+1) \left[\chi_{C_n}(x) - \chi_{B_n}(x)\right].$$

Note that $A_n = B_n \mathbin{\dot\cup} C_n$, $n \in \mathbb{N}$, $f \in L(A_n)$, and

$$\int\limits_{A_n} f(x) d\mu = (n+1)[\mu C_n - \mu B_n] = 0,$$

since $\mu B_n = \mu C_n = \dfrac{1}{2n(n+1)}$. However, $f \notin L[0,1]$ by Proposition 5.3.12, since

$$\sum_{n=1}^{\infty} (n+1)[\mu B_n + \mu C_n] = \sum_{n=1}^{\infty} \frac{n+1}{n(n+1)} = \sum_{n=1}^{\infty} \frac{1}{n} = +\infty.$$

Consider now a real number $a \in (0,1)$. There exists a natural number $n \in \mathbb{N}$ such that $\dfrac{1}{n+1} \leqslant a < \dfrac{1}{n}$. Consequently, on the interval $[a,1]$ the function f is bounded and has finitely many points of discontinuities. Thus, by Theorem 5.4.5, the function f is Riemann integrable on $[a,1]$, and by Theorem 5.4.1

$$\left|(\mathcal{R}) \int_a^1 f(x) dx\right| = \left|(L) \int\limits_{(a,1)} f(x) d\mu\right| = \left|\sum_{k=1}^{n-1} \int\limits_{\left(\frac{1}{k+1}, \frac{1}{k}\right)} f(x) d\mu + \int\limits_{\left(a, \frac{1}{n}\right)} f(x) d\mu\right| =$$

$$= \left|\int\limits_{\left(a, \frac{1}{n}\right)} f(x) d\mu\right| \leqslant (n+1)\left(\frac{1}{n} - \frac{1}{n+1}\right) = \frac{1}{n}.$$

Clearly, $a \to +0$ if, and only if, $n \to +\infty$, so

$$\lim_{a \to +0} \left|(\mathcal{R}) \int_a^1 f(x) dx\right| \leqslant \lim_{n \to +\infty} \frac{1}{n} = 0,$$

thus, $f \in \mathcal{R}(+0,1]$, and

$$(\mathcal{R}) \int_{+0}^{1} f(x)dx = 0.$$

This example show that even if a function is non-Lebesgue integrable, there can exist an improper Riemann integral.

5.4.2 Approximation of Riemann integrable functions

Theorem 5.4.10. *Suppose f is Riemann integrable on $[a,b]$, and is bounded by M. Then there exists a sequence $(f_k)_{k=1}^{\infty}$ of continuous functions on $[a,b]$ so that*

$$\sup_{x\in[a,b]} |f_k(x)| \leqslant M, \qquad k = 1, 2, \dots,$$

and

$$\int_a^b |f(x) - f_k(x)|dx \to 0 \quad as \quad k \to \infty.$$

Proof. Given $\varepsilon > 0$, we may choose a partition $a = x_0 < x_1 < \cdots < x_N = b$ of the interval $[a,b]$ so that the upper and lower Darboux sums of f differ by at most ε. Denote by f^* the step function defined by

$$f^*(x) = \sup_{x_{j-1}\leqslant y\leqslant x_j} f(y) \quad \text{if} \quad x \in [x_{j-1}, x_j) \quad \text{for} \quad 1 \leqslant j \leqslant N.$$

By construction we have $|f^*| \leqslant M$, and moreover

$$\int_a^b |f^*(x) - f(x)|dx = \int_a^b (f^*(x) - f(x))dx < \varepsilon. \tag{5.4.4}$$

Now we can modify f^* to make it continuous and still approximate f in the sense of the lemma. For small $\delta > 0$, let $\widetilde{f}(x) = f^*(x)$ when the distance of x from any of the division points $x_0, \dots, x_N$ exceeds δ. In the δ-neighborhood of x_j for $j = 1, \dots, N-1$, define $\widetilde{f}(x)$ to be the linear function for which $\widetilde{f}(x_j \pm \delta) = f^*(x_j \pm \delta)$. Near $x_0 = a$, $\widetilde{f}$ is linear with $\widetilde{f}(a) = 0$ and $\widetilde{f}(a+\delta) = f^*(a+\delta)$. Similarly, near $x_N = b$ the function $\widetilde{f}$ is linear with $\widetilde{f}(b) = 0$ and $\widetilde{f}(b-\delta) = f^*(b-\delta)$. The absolute value of this extension is also bounded by M. Moreover, $\widetilde{f}$ differs from f^* only in the N intervals of length 2δ surrounding the division points. Thus

$$\int_a^b |f^*(x) - \widetilde{f}(x)|dx \leqslant 2MN \cdot 2\delta.$$

If we choose δ sufficiently small, we get

$$\int_a^b |f^*(x) - \widetilde{f}(x)|dx < \varepsilon. \tag{5.4.5}$$

As a result, equations (5.4.4)–(5.4.5), and the triangle inequality yield

$$\int_a^b |\widetilde{f}(x) - f(x)|dx < 2\varepsilon.$$

Denoting by f_k the $\widetilde{f}$ so constructed, when $2\varepsilon = \dfrac{1}{k}$, we see that the sequence $(f_k)_{k=1}^{\infty}$ has the properties required by the lemma. $\square$

Remark 5.4.11. We construct $\widetilde{f}(x)$ such that $\widetilde{f}(a) = \widetilde{f}(b)$, so we may extend $\widetilde{f}$ to a continuous and $(b-a)$- periodic function on $\mathbb{R}$. If the function f satisfies the condition $f(a) = f(b)$, we also can consider it to be $(b-a)$-periodic on the real line. Thus, our approximating function $\widetilde{f}(x)$ can be used to approximate periodic functions.

5.4.3 Approximation of Lebesgue integrable functions

To be written soon...

5.5 Fubini's theorem and its applications

In elementary calculus integrals of continuous functions of several variables are often calculated by iterating one-dimensional integrals. We shall now examine this important analytic device from the general point of view of Lebesgue integration in $\mathbb{R}^n$, and we shall see that a number of interesting issues arise. In general, we can represent $\mathbb{R}^n$ as a product $\mathbb{R}^n = \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$ where $n = n_1 + n_2$, and $n_1, n_2 \geqslant 1$. A point in $\mathbb{R}^n$ then takes the form (x, y), where $x \in \mathbb{R}^{n_1}$ and $y \in \mathbb{R}^{n_2}$. With such a decomposition of $\mathbb{R}^n$ in mind, the general notion of a slice, formed by fixing one variable, becomes natural.

Definition 5.5.1. Let f be a function defined on $\mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$. The *slice* of f corresponding to $y \in \mathbb{R}^{n_2}$ is the function f^y of the $x \in \mathbb{R}^{n_1}$ variable, given by

$$f^y(x) = f(x, y).$$

Similarly, the slice of f for a fixed $x \in \mathbb{R}^{n_1}$ is $f_x(y) = f(x, y)$.

In the case of a set $A \subset \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$ we define its slices by

$$A^y = \{x \in \mathbb{R}^{n_1} : \ (x, y) \in A\} \qquad \text{and} \qquad A_x = \{y \in \mathbb{R}^{n_2} : \ (x, y) \in A\}.$$

Recall that even if f is measurable on $\mathbb{R}^n$, it is not necessarily true that the slice f^y is measurable on $\mathbb{R}^{n_2}$ for each y; nor does the corresponding assertion necessarily hold for a measurable set: the slice E^y may not be measurable for each y. An easy example arises in $\mathbb{R}^2$ by placing a one-dimensional non-measurable set on the x-axis; the set A in $\mathbb{R}^2$ has measure zero, but E^y is not measurable for $y = 0$. What saves us is that, nevertheless, measurability holds for almost all slices.

5.5.1 Statement and proof of Fubini's theorem

Theorem 5.5.2. *Suppose $f(x, y)$ is integrable on $\mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$. Then for almost every $y \in \mathbb{R}^{n_2}$:*

(i) The slice $f^y(x)$ is integrable on $\mathbb{R}^{n_1}$.

(ii) The function $\int\limits_{\mathbb{R}^{n_1}} f^y(x)\mu(dx)$ is integrable on $\mathbb{R}^{n_2}$, and

(iii) The following identity holds:

$$\int\limits_{\mathbb{R}^{n_2}} \left(\int_{\mathbb{R}^{n_1}} f(x, y)\mu(dx) \right) \mu(dy) = \int\limits_{\mathbb{R}^n} f(x, y) d\mu.$$

Clearly, the theorem is symmetric in x and y so that we also may conclude that the slice $f_x(y)$ is integrable on $\mathbb{R}^{n_2}$ for a.e. x. Moreover, $\int\limits_{\mathbb{R}^{n_2}} f_x(y)\mu(dy)$ is integrable, and

$$\int\limits_{\mathbb{R}^{n_1}} \left(\int_{\mathbb{R}^{n_2}} f(x, y)\mu(dy) \right) \mu(dx) = \int\limits_{\mathbb{R}^n} f(x, y) d\mu.$$

In particular, Fubini's theorem states that the integral of f on $\mathbb{R}^n$ can be computed by iterating lower-dimensional integrals, and that the iterations can be taken in any order

$$\int\limits_{\mathbb{R}^{n_2}} \left(\int_{\mathbb{R}^{n_1}} f(x,y)\mu(dx)\right)\mu(dy) = \int\limits_{\mathbb{R}^{n_1}} \left(\int_{\mathbb{R}^{n_2}} f(x,y)\mu(dy)\right)\mu(dx) = \int\limits_{\mathbb{R}^{n}} f(x,y)d\mu.$$

The proof of Fubini's theorem which we give next consists of a sequence of six steps. We begin by letting $\mathcal{F}$ denote the set of integrable functions on $\mathbb{R}^n$ which satisfy all three conclusions in the theorem, and set out to prove that $L(\mathbb{R}^n) \subset \mathcal{F}$.

We proceed by first showing that $\mathcal{F}$ is closed under operations such as linear combinations (Step 1) and limits (Step 2). Then we begin to construct families of functions in $\mathcal{F}$. Since any integrable function is the "limit" of simple functions, and simple functions are themselves linear combinations of sets of finite measure, the goal quickly becomes to prove that the function $\chi_A(x)$ belongs to $\mathcal{F}$ whenever A is a measurable subset of $\mathbb{R}^n$ with finite measure. To achieve this goal, we begin with bricks and work our way up to sets of type $G_\delta(\mathbb{R}^n)$ (Step 3), and sets of measure zero (Step 4). Finally, a limiting argument shows that all integrable functions are in $\mathcal{F}$. This will complete the proof of Fubini's theorem.

Proof of Theorem 5.5.2.

1) *Any finite linear combination of functions in $\mathcal{F}$ also belongs to $\mathcal{F}$.*

   Indeed, let $(f_m)_{m=1}^N \subset \mathcal{F}$. For each m there exists a set $A_m \subset \mathbb{R}^{n_2}$ of measure 0 so that f_m^y is integrable on $\mathbb{R}^{n_1}$ whenever $y \notin A_m$. Then, if $A = \bigcup\limits_{m=1}^{N} A_m$, the set A has measure 0, and in the complement of A, the y-slice corresponding to any finite linear combination of the f_m is measurable, and also integrable. By linearity of the integral, we then conclude that any linear combination of the f_m's belongs to $\mathcal{F}$.

2) *Suppose $(f_m)_{m=1}^\infty$ is a sequence of measurable functions in $\mathcal{F}$ so that $f_m \nearrow f$ or $f_m \searrow f$, where f is integrable (on $\mathbb{R}^n = \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$). Then $f \in \mathcal{F}$.*

   By taking $-f_m$ instead of f_m if necessary, we note that it suffices to consider the case of an increasing sequence. Also, we may replace f_m by $f_m - f_1$ and assume that the f_m's are non-negative. Now, we observe that an application of Corollary 5.2.5 yields

$$\lim_{m\to\infty} \int\limits_{\mathbb{R}^n} f_m(x,y)\mu(dx)\mu(dy) = \int\limits_{\mathbb{R}^n} f(x,y)\mu(dx)\mu(dy). \tag{5.5.1}$$

   By assumption, for each m there exists a set $A_m \subset \mathbb{R}^{n_2}$, so that f_m^y is integrable on $\mathbb{R}^{n_1}$ whenever $y \notin A_m$. If $A = \bigcup\limits_{m=1}^{\infty} A_m$, then $\mu A = 0$ in $\mathbb{R}^{n_2}$, and if $y \notin A$, then f_m^y is integrable on $\mathbb{R}^{n_1}$ for all m, and, by Corollary 5.2.5, we find that

$$g_m(y) = \int\limits_{\mathbb{R}^{n_1}} f_m^y(x)\mu(dx) \nearrow g(y) = \int\limits_{\mathbb{R}^{n_1}} f^y(x)\mu(dx)$$

   as $m \to \infty$. By assumption, each $g_m(y)$ is integrable, so that another application of Corollary 5.2.5 yields

$$\int\limits_{\mathbb{R}^{n_2}} g_m(y)\mu(dy) \longrightarrow \int\limits_{\mathbb{R}^{n_2}} g(y)\mu(dy) \quad \text{as} \quad m \to \infty. \tag{5.5.2}$$

   By the assumption that $f_m \in \mathcal{F}$ we have

$$\int\limits_{\mathbb{R}^{n_2}} g_m(y)\mu(dy) = \int\limits_{\mathbb{R}^n} f_m(x,y)\mu(dx)\mu(dy),$$

and combining this fact with (5.5.1) and (5.5.2), we conclude that

$$\int_{\mathbb{R}^{n_2}} g(y)\mu(dy) = \int_{\mathbb{R}^n} f(x,y)\mu(dx)\mu(dy).$$

Since f is integrable, the right-hand integral is finite, and this proves that g is integrable. Consequently $g(y)$ is finite-valued a.e. on $\mathbb{R}^{n_2}$ in variable y, hence f^y is integrable for a.e. $y \in \mathbb{R}^{n_2}$, and

$$\int_{\mathbb{R}^{n_2}} \left(\int_{\mathbb{R}^{n_1}} f(x,y)\mu(dx) \right) \mu(dy) = \int_{\mathbb{R}^n} f(x,y)d\mu.$$

This proves that $f \in \mathcal{F}$ as desired.

3) *Any characteristic function of a set E that is a $G_\delta(\mathbb{R}^n)$ and of finite measure belongs to $\mathcal{F}$.*

We proceed in stages of increasing order of generality.

(a) First suppose E is a bounded open cube in $\mathbb{R}^n$, such that $X = Q_1 \times Q_2$, where Q_1 and Q_2 are open cubes in $\mathbb{R}^{n_1}$ and $\mathbb{R}^{n_2}$, respectively. Then, for each y the function $\chi_E(x,y)$ is measurable in x, and integrable with

$$g(y) = \int_{\mathbb{R}^{n_1}} \chi_E(x,y)\mu(dx) = \begin{cases} \mu(Q_1), & y \in Q_2, \\ 0, & y \notin Q_2. \end{cases}$$

Consequently, $g(y) = \mu(Q_1)\chi_{Q_2}(y)$ is also measurable and integrable, with

$$\int_{\mathbb{R}^{n_2}} g(y)\mu(dy) = \mu(Q_1)\mu(Q_2).$$

Since we initially have $\displaystyle\int_{\mathbb{R}^n} \chi_E(x,y)\mu(dy) = \mu(E) = \mu(Q_1)\mu(Q_2)$, we deduce that $\chi_E \in \mathcal{F}$.

(b) Now suppose E is a subset of the boundary of some closed cube. Then, since the boundary of a cube has measure 0 in $\mathbb{R}^n$, we have $\displaystyle\int_{\mathbb{R}^n} \chi_E(x,y)\mu(dx)\mu(dy) = 0$.

Next, we note, after an investigation of the various possibilities, that for almost every y, the slice E^y has measure 0 in $\mathbb{R}^{n_1}$, and therefore if $g(y) = \displaystyle\int_{\mathbb{R}^{n_1}} \chi_E(x,y)\mu(dx)$ we have $g(y) = 0$ for a.e. $y \in \mathbb{R}^{n_2}$. As a consequence, $\displaystyle\int_{\mathbb{R}^{n_2}} g(y)\mu(dy) = 0$, and therefore $\chi_E \in \mathcal{F}$.

(c) Suppose now E is a finite union of closed cubes whose interiors are disjoint, $E = \displaystyle\bigcup_{k=1}^{N} Q_k$. Then, if Q_k° denotes the interior of Q_k, we may represent $\chi_E(x)$ as a linear combination of the the functions $\chi_{Q_k^\circ}(x)$ and $\chi_{A_k}(x)$ where A_k is a subset of the boundary of Q_k for $k = 1, \ldots, N$. By our previous analysis, we know that $\chi_{Q_k^\circ}(x)$ and $\chi_{A_k}(x)$ belong to $\mathcal{F}$ for all k, and since Step 1 guarantees that $\mathcal{F}$ is closed under finite linear combinations, we conclude that $\chi_E \in \mathcal{F}$, as desired.

(d) Next, we prove that if E is open and of finite measure, then $\chi_E \in \mathcal{F}$. This follows from taking a limit in the previous case. Indeed, by Theorem 3.5.9, we may represent E as a countable union of almost disjoint closed cubes

$$E = \bigcup_{j=1}^{\infty} Q_j.$$

Consequently, if we let $f_m(x) = \bigcup_{j=1}^{m} \chi_{Q_j}(x)$, then we note that the functions f_m increase to $f(x) = \chi_E(x)$, which is integrable since μE is finite. Therefore, we may conclude by Step 2 that $f \in \mathcal{F}$.

(e) Finally, if $E \in G_\delta(\mathbb{R}^n)$ of finite measure, then $\chi_E \in \mathcal{F}$. Indeed, by definition, there exist open sets $\widetilde{G}_1, \widetilde{G}_2, \ldots$, such that

$$E = \bigcap_{k=1}^{\infty} \widetilde{G}_k.$$

Since E has finite measure, there exists an open set $\widetilde{G}_0$ of finite measure with $E \subset \widetilde{G}_0$. If we set

$$G_k = \widetilde{G}_0 \bigcap \left(\bigcap_{j=1}^{k} \widetilde{G}_j \right)$$

then $(G_k)_{k=1}^{\infty}$ is a decreasing sequence of open sets of finite measure such that

$$G_1 \supset G_2 \supset \cdots$$

and

$$E = \bigcap_{k=1}^{\infty} G_k.$$

Therefore, the sequence of functions $f_k(x) = \chi_{G_k}(x)$ decreases to $f(x) = \chi_E(x)$, and since $\chi_{G_k} \in \mathcal{F}$ for all k by (d) above, we conclude by Step 2 that $\chi_E(x)$ belongs to $\mathcal{F}$.

4) *If E has measure 0, then $\chi_E(x)$ belongs to $\mathcal{F}$.*

Indeed, since E is measurable, we may choose a set $G \in G_\delta(\mathbb{R}^n)$ with $E \subset G$ and $\mu G = 0$ (see Theorem 3.7.5). Since $\chi_G \in \mathcal{F}$ (by the previous step) we find that

$$\int_{\mathbb{R}^{n_2}} \left(\int_{\mathbb{R}^{n_1}} \chi_G(x,y)\mu(dx) \right) \mu(dy) = \int_{\mathbb{R}^n} \chi_G(x,y) d\mu = 0.$$

Therefore

$$\int_{\mathbb{R}^{n_1}} \chi_G(x,y)\mu(dx) = 0 \qquad \text{for a.e.} \quad y \in \mathbb{R}^{n_2}.$$

Consequently, the slice G^y has measure 0 for a.e. $y \in \mathbb{R}^{n_2}$. The simple observation that $E^y \subset G^y$ then shows that E^y has measure 0 for a.e. $y \in \mathbb{R}^{n_2}$, and $\int_{\mathbb{R}^{n_1}} \chi_E(x,y)\mu(dx) = 0$ for a.e. $y \in \mathbb{R}^{n_2}$.

Therefore,

$$\int_{\mathbb{R}^{n_2}} \left(\int_{\mathbb{R}^{n_1}} \chi_E(x,y)\mu(dx) \right) \mu(dy) = 0 = \int_{\mathbb{R}^n} \chi_E(x,y) d\mu$$

and thus $\chi_E \in \mathcal{F}$, as was to be shown.

5) *If E is any measurable subset of $\mathbb{R}^n$ with finite measure, then $\chi_E(x)$ belongs to $\mathcal{F}$.*

To prove this, recall first that by Theorem 3.7.5 there exists a set of finite measure $G \in G_\delta(\mathbb{R}^n)$, with $E \subset G$ and $\mu(G \setminus E) = 0$. Since

$$\chi_E(x) = \chi_G(x) - \chi_{G \setminus E}(x),$$

and $\mathcal{F}$ is closed under linear combinations, we find that $\chi_E \in \mathcal{F}$, as desired.

6) *If $f \in L(\mathbb{R}^n)$, then $f \in \mathcal{F}$.*

We note first that f has the decomposition $f = f^+ - f^-$, where non-negative and integrable functions f^+ and f^- are defined in Definition 4.2.3. So by Step 1 we may assume that f is itself non-negative. By Theorem 4.4.2, there exists a sequence $(h_m)_{m=1}^{\infty}$ of simple functions that increase to f. Since each $h_m(x)$ is a finite linear combination of characteristic functions of sets with finite measure, we have $h_m \in \mathcal{F}$ by Steps 5 and 1, hence $f \in \mathcal{F}$ by Step 2.

□

5.5.2 Applications of Fubini's theorem

Theorem 5.5.3 (Tonelly)**.** *Let $f(x,y) \in S^+(\mathbb{R}^{n_1} \times \mathbb{R}^{n_2})$. Then for almost every $y \in \mathbb{R}^{n_2}$:*

(i) The slice $f^y(x)$ is measurable on $\mathbb{R}^{n_1}$.

(ii) The function defined by $\int\limits_{\mathbb{R}^{n_1}} f^y(x)\mu(dx)$ is measurable on $\mathbb{R}^{n_2}$.

Moreover:

(iii)

$$\int\limits_{\mathbb{R}^{n_2}} \left(\int_{\mathbb{R}^{n_1}} f(x,y)\mu(dx) \right) \mu(dy) = \int\limits_{\mathbb{R}^n} f(x,y) d\mu,$$

where the integrals exist (but can be infinite).

In practice, this theorem is often used in conjunction with Fubini's theorem. Indeed, suppose we are given a measurable function f on $\mathbb{R}^n$ and asked to compute $\int\limits_{\mathbb{R}^n} f d\mu$. To justify the use of iterated integration, we first apply the present theorem to $|f|$. Using it, we may freely compute (or estimate) the iterated integrals of the non-negative function $|f|$. If these are finite, Theorem 5.5.3 guarantees that f is integrable, that is, $\int\limits_{\mathbb{R}^n} |f| d\mu < +\infty$. Then the hypothesis in Fubini's theorem is verified, and we may use that theorem in the calculation of the integral of f.

Proof of Theorem 5.5.3. Consider the truncations

$$f_l(x,y) = \begin{cases} f(x,y), & \text{if } \ |(x,y)| < l \ \text{ and } \ f(x,y) < l, \\ 0, & \text{otherwise,} \end{cases}$$

where $|(x,y)| < l$ means the cube $\widehat{Q}_l$ in $\mathbb{R}^n$ (see (3.5.1))

Each f_l is integrable, and by part (i) in Fubini's Theorem 5.5.2 there exists a set $E_l \subset \mathbb{R}^{n_2}$ of measure 0 such that the slice $f_l^y(x)$ is measurable for all $y \in E_l^c$. Then, if we set $E = \bigcup\limits_l E_l$, we find that $f_l^y(x)$ is measurable for all $y \in E^c$ and all l. Moreover, $\mu(E) = 0$. Since $f_l^y \nearrow f^y$, Corollary 5.2.5 implies that if $y \notin E$, then

$$\int\limits_{\mathbb{R}^{n_1}} f_l(x,y)\mu(dx) \nearrow \int\limits_{\mathbb{R}^{n_1}} f(x,y)\mu(dx) \quad \text{as} \quad l \to \infty.$$

Again by Fubini's Theorem 5.5.2, $\int\limits_{\mathbb{R}^{n_1}} f_l(x,y)\mu(dx)$ is measurable for all $y \in E^c$, hence so is $\int\limits_{\mathbb{R}^{n_1}} f(x,y)\mu(dx)$. Another application of Corollary 5.2.5 then gives

$$\int\limits_{\mathbb{R}^{n_2}} \left(\int_{\mathbb{R}^{n_1}} f_l(x,y)\mu(dx) \right) \mu(dy) \to \int\limits_{\mathbb{R}^{n_2}} \left(\int_{\mathbb{R}^{n_1}} f(x,y)\mu(dx) \right) \mu(dy), \tag{5.5.3}$$

By part (iii) in Fubini's Theorem 5.5.2 we know that

$$\int_{\mathbb{R}^{n_2}} \left(\int_{\mathbb{R}^{n_1}} f_l(x,y)\mu(dx)\right)\mu(dy) = \int_{\mathbb{R}^n} f_l(x,y)d\mu, \tag{5.5.4}$$

A final application of Corollary 5.2.5 directly to f_l also gives

$$\int_{\mathbb{R}^n} f_l(x,y)d\mu \to \int_{\mathbb{R}^n} f(x,y)d\mu. \tag{5.5.5}$$

Combining (5.5.3), (5.5.4), and (5.5.5) completes the proof of the theorem. □

Corollary 5.5.4. *If E is a measurable set in $\mathbb{R}^{n_1}\times\mathbb{R}^{n_2}$, then for almost every $y\in\mathbb{R}^{n_2}$ the slice*

$$E^y = \{x\in\mathbb{R}^{n_1}:\ (x,y)\in E\}$$

is a measurable subset of $\mathbb{R}^{n_1}$. Moreover $\mu(E^y)$ is a measurable function of y and

$$\mu(E) = \int_{\mathbb{R}^{n_2}} \mu(E^y)\mu(dy).$$

This is an immediate consequence of the first part of Theorem 5.5.3 applied to the function $\chi_E(x)$. Clearly, a symmetric result holds for the x-slices in $\mathbb{R}^{n_2}$.

We have thus established the basic fact that if E is measurable on $\mathbb{R}^{n_1}\times\mathbb{R}^{n_2}$, then for almost every $y\in\mathbb{R}^{n_2}$ the slice E^y is measurable in $\mathbb{R}^{n_1}$ (and also the symmetric statement with the roles of x and y interchanged). One might be tempted to think that the converse assertion holds. To see that this is not the case let us recall an example considered for non-Borel measurable sets.

Let $\mathcal{N}$ denote a non-measurable subset of $\mathbb{R}$, and then define

$$E = [0,1]\times\mathcal{N}\subset\mathbb{R}\times\mathbb{R},$$

we see that

$$E^y = \begin{cases} [0,1], & y\in\mathcal{N},\\ 0, & y\notin\mathcal{N}.\end{cases}$$

Thus E^y is measurable for every y. However, if E were measurable, then the corollary would imply that $E_x = \{y\in\mathbb{R}:\ (x,y)\in E\}$ is measurable for almost every $x\in\mathbb{R}$, which is not true since E_x is equal to $\mathcal{N}$ for all $x\in[0,1]$.

Let us now study measurability and measures of product sets of the form $E_1\times E_2$ where $E_k\in\mathbb{R}^{n_k}$, $k=1,2$.

Theorem 5.5.5. *If $E = E_1\times E_2$ is a measurable subset of $\mathbb{R}^n$, and $\mu^*(E_2)>0$, then E_1 is measurable.*

Proof. By Corollary 5.5.4, we have that for a.e. $y\in\mathbb{R}^{n_2}$, the slice function

$$(\chi_{E_1\times E_2})^y(x) = \chi_{E_1}(x)\chi_{E_2}(y)$$

is measurable as a function of x. In fact, we claim that there is some $y\in E_2$ such that the above slice function is measurable in x; for such a y we would have $\chi_{E_1\times E_2}(x,y) = \chi_{E_1}(x)$, and this would imply that E_1 is measurable.

To prove the existence of such a y, we use the assumption that $\mu^*(E_2)>0$. Indeed, let F denote the set of $y\in\mathbb{R}^{n_2}$ such that the slice E^y is measurable. Then $\mu(F^c)=0$ according Corollary 5.5.4. However, $E_2\bigcap F$ is not empty because $\mu^*(E_2\bigcap F)>0$. To see this, note that $E_2 = (E_2\bigcap F)\bigcup(E_2\bigcap F^c)$, hence

$$0<\mu^*(E_2)\leqslant\mu^*\left(E_2\bigcap F\right)+\mu^*\left(E_2\bigcap F^c\right) = \mu^*\left(E_2\bigcap F\right),$$

because $E_2\bigcap F^c$ is a subset of a set of measure zero. □

To deal with a converse of the above result, we need the following lemma.

Lemma 5.5.6. *If $E_1 \in \mathbb{R}^{n_1}$ and $E_2 \in \mathbb{R}^{n_2}$, then*

$$\mu^*(E_1 \times E_2) \leqslant \mu^*(E_1)\mu^*(E_2),$$

and if one of the sets E_k has outer measure zero, then $\mu^*(E_1 \times E_2) = 0$.

Proof. Let $\varepsilon > 0$. By Definition 3.3.1, we there exist bricks $\{K_j^{(1)}\}_{j=1}^\infty$ in $\mathbb{R}^{n_1}$ and $\{K_j^{(2)}\}_{j=1}^\infty$ in $\mathbb{R}^{n_2}$ such that

$$E_1 \subset \bigcup_{j=1}^\infty K_j^{(1)} \quad \text{and} \quad E_2 \subset \bigcup_{j=1}^\infty K_j^{(2)}$$

and

$$\sum_{j=1}^\infty mK_j^{(1)} \leqslant \mu^*(E_1) + \varepsilon \quad \text{and} \quad \sum_{j=1}^\infty mK_j^{(2)} \leqslant \mu^*(E_2) + \varepsilon.$$

Since $E_1 \times E_2 \subset \bigcup\limits_{j,l=1}^\infty K_j^{(1)} \times K_l^{(2)}$, the semi-additivity of the outer measure yields

$$\mu^*(E_1 \times E_2) \leqslant \sum_{j,l=1}^\infty m(K_j^{(1)} \times K_l^{(2)}) = \left(\sum_{j=1}^\infty mK_j^{(1)}\right)\left(\sum_{l=1}^\infty mK_l^{(2)}\right) \leqslant (\mu^*(E_1)+\varepsilon)(\mu^*(E_2)+\varepsilon).$$

If neither E_1 nor E_2 has outer measure 0, then from the above we find

$$\mu^*(E_1 \times E_2) \leqslant \mu^*(E_1)\mu^*(E_2) + o(\varepsilon),$$

and since ε is arbitrary, we have $\mu^*(E_1 \times E_2) \leqslant \mu^*(E_1)\mu^*(E_2)$.

If for instance $\mu^*(E_1) = 0$, consider for each positive integer m the set $E_2^m = E_2 \bigcap \{y \in \mathbb{R}^{n_2} : |y| \leqslant m\}$. Then, by the above argument, we find that $\mu^*(E_1 \times E_2^m) = 0$. Since $(E_1 \times E_2^m) \nearrow (E_1 \times E_2)$ as $m \to \infty$, we conclude that $\mu^*(E_1 \times E_2) = 0$. □

Theorem 5.5.7. *Suppose E_1 and E_2 are measurable subsets of $\mathbb{R}^{n_1}$ and $\mathbb{R}^{n_2}$, respectively. Then $E = E_1 \times E_2$ is a measurable subset of $\mathbb{R}^n$. Moreover,*

$$\mu(E) = \mu(E_1)\mu(E_2),$$

and if one of the sets E_k, $k = 1, 2$, has measure zero, then $\mu(E) = 0$.

Proof. It suffices to prove that E is measurable, because then the assertion about $\mu(E)$ follows from Corollary 5.5.4. Since each set E_k, $k = 1, 2$, is measurable, there exist sets $G_k \subset \mathbb{R}^{n_k}$ of type $G_\delta(\mathbb{R}^{n_k})$ such that $E_k \subset G_k$ and $\mu^*(G_k \setminus E_k) = 0$ for each $k = 1, 2$ according to Theorem 3.7.5. Clearly, $G = G_1 \times G_2$ is measurable in $\mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$ and

$$(G_1 \times G_2) \setminus (E_1 \times E_2) \subset ((G_1 \setminus E_1) \times G_2) \bigcup (G_1 \times (G_2 \setminus E_2)).$$

By Lemma 5.5.6 we conclude that $\mu^*(G \setminus E) = 0$, hence E is measurable. □

As a consequence of this proposition we have the following.

Theorem 5.5.8. *Suppose f is a measurable function on $\mathbb{R}^{n_1}$. Then the function $\widehat{f}$ defined by $\widehat{f}(x,y) = f(x)$ is measurable on $\mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$.*

Proof. Recall first that if $c \in \mathbb{R}$ and $E_1 = \{x \in \mathbb{R}^{n_1} : f(x) < c\}$, then E_1 is measurable by definition. Since

$$\{(x,y) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} : \widehat{f}(x,y) < c\} = E_1 \times \mathbb{R}^{n_2},$$

Theorem 5.5.7 shows that $\{\widehat{f}(x,y) < c\}$ is measurable for each $c \in \mathbb{R}$. Thus $\widehat{f}(x,y)$ is a measurable function on $\mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$, as desired. □

This theorem will be of use when we will study the Fourier transform.

Theorem 5.5.9. *If f is a measurable function on $\mathbb{R}^n$, then the function $\widehat{f}(x,y) = f(x-y)$ is measurable on $\mathbb{R}^n \times \mathbb{R}^n$.*

Proof. By picking $E = \{z \in \mathbb{R}^n : \ f(z) < c\}$, we see that it suffices to prove that whenever E is a measurable subset of $\mathbb{R}^n$, then $\widetilde{E} = \{(x,y) : \ x - y \in E\}$ is a measurable subset of $\mathbb{R}^n \times \mathbb{R}^n$.

Note first that if G is an open set, then $\widetilde{G}$ is also open. Taking countable intersections shows that if $E \in G_\delta(\mathbb{R}^n)$, then so is $\widetilde{E}$. Assume now that $\mu(\widetilde{E}_j) = 0$ for each j, where $\widetilde{E}_j = \widetilde{E} \bigcap \widehat{Q}_j$ and $\widehat{Q}_j = \{|y| < j\}$. Again, take G to be open in $\mathbb{R}^n$, and let us calculate $\mu\left(\widetilde{G} \bigcap \widehat{Q}_j\right)$. We have that

$$\chi_{\widetilde{G} \bigcap \widehat{Q}_j}(x,y) = \chi_G(x-y)\chi_{\widehat{Q}_j}(y).$$

Hence

$$\mu\left(\widetilde{G} \bigcap \widehat{Q}_j\right) = \int\limits_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} \chi_G(x-y)\chi_{\widehat{Q}_j}(y)\mu(dy) \right) \mu(dx) =$$

$$= \int\limits_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} \chi_G(x-y)\mu(dx) \right) \chi_{\widehat{Q}_j}(y)\mu(dy) = \mu(G)\mu(\widehat{Q}_j),$$

by the translation-invariance of the measure. Now if $\mu(E) = 0$, there is a sequence of open sets G_m such that $E \subset G_m$ and $\mu(G_m) \to 0$. It follows from the above that $\widetilde{E}_j \subset \widetilde{G}_m \bigcap \widehat{Q}_j$ and $\mu\left(\widetilde{G}_m \bigcap \widehat{Q}_j\right) \to 0$ as $m \to \infty$ for each fixed j. This shows $\mu(\widetilde{E}_j) = 0$, and hence $\mu(\widetilde{E}) = 0$. The proof of the proposition is concluded once we recall that any measurable set E can be written as the difference of a $G_\delta(\mathbb{R}^n)$ and a set of measure zero. $\square$

Finally, let us establish one of the fundamental facts of the integration theory, which is similar to a correspondent property of the Riemann integral.

Theorem 5.5.10. *Suppose $f(x)$ is a non-negative function on $\mathbb{R}^n$, and let*

$$A = \{(x,y) \in \mathbb{R}^n \times \mathbb{R} \ : \ 0 \leqslant y \leqslant f(x)\}.$$

Then:

1) f is measurable on $\mathbb{R}^n$ if and only if A is measurable in $\mathbb{R}^{n+1}$.

2) If the conditions in 1) *hold, then*

$$\int\limits_{\mathbb{R}^n} f(x)dx = \mu(A).$$

Proof. If f is measurable on $\mathbb{R}^n$, then by Theorem 5.5.8 the function

$$F(x,y) = y - f(x)$$

is measurable on $\mathbb{R}^{n+1}$, so we immediately see that $A = \{y \geqslant 0\} \bigcap \{F \leqslant 0\}$ is measurable.

Conversely, suppose that A is measurable. We note that for each $x \in \mathbb{R}^n$ the slice $A_x = \{y \in \mathbb{R} : (x,y) \in A\}$ is a closed segment, namely $A_x = [0, f(x)]$. Consequently Corollary 5.5.4 (with the roles of x and y interchanged) yields the measurability of $\mu(A_x) = f(x)$. Moreover

$$\mu(A) = \int\limits_{\mathbb{R}^{n+1}} \chi_A(x,y)dxdy = \int\limits_{\mathbb{R}^n} \mu(A_x)dx = \int\limits_{\mathbb{R}^n} f(x)dx,$$

as was to be shown. $\square$

Later we will prove a somewhat converse statement by introducing the so-called Radon-Nikodym derivative.

5.6 Problems

5.1.1 Basic properties of Lebesgue integral

Problem 5.1. Prove that the function

$$f(x) = \begin{cases} x^2, & x \in \mathbb{Q} \bigcap [0,1], \\ \sqrt{x}, & x \in [0,1] \setminus \mathbb{Q}, \end{cases}$$

is Lebesgue integrable on $[0,1]$, find its Lebesgue integral over $[0,1]$. Is $f(x)$ Riemann integrable on $[0,1]$?

Problem 5.2. Calculate the integral $\int\limits_{[0,1]} f(x)d\mu$, where μ is the Lebesgue measure and

$$f(x) = \begin{cases} \sin x, & x \in \mathbb{Q}, \\ \\ \cos x, & x \notin \mathbb{Q}. \end{cases}$$

Problem 5.3. Calculate the integral $\int\limits_{[0,1]} f(x)d\mu$, where μ is the Lebesgue measure and

$$f(x) = \begin{cases} \sin x, & x \in \mathbb{Q}, \\ \\ \sin^2 x, & x \notin \mathbb{Q}. \end{cases}$$

Problem 5.4. Calculate the integrals

$$a) \int\limits_{(0,+\infty)} e^{-[x]}d\mu, \qquad b) \int\limits_{(3,+\infty)} e^{-[3x+1]}d\mu, \qquad c) \int\limits_{(4,+\infty)} e^{-[2x+2]}d\mu.$$

where $[x]$ is the integer part of x (not exceeding x).

Problem 5.5. Calculate

$$a) \int\limits_{(0,+\infty)} \frac{d\mu}{[x+1]\cdot[x+2]}, \qquad b) \int\limits_{(0,+\infty)} \frac{d\mu}{[2x+1]\cdot[2x+3]}, \qquad c) \int\limits_{(3,+\infty)} \frac{d\mu}{[3x+1]},$$

$$d) \int\limits_{(3,+\infty)} \frac{(-1)^{[x]}}{[x+1]\cdot[x+2]}d\mu, \qquad e) \int\limits_{(0,+\infty)} \frac{(-1)^{[x]}}{[2x+1]\cdot[2x+2]}d\mu,$$

where $[x]$ is the integer part of x (not exceeding x).

Hint: In the problem e) use the fact that

$$\sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} = \frac{\pi}{4}.$$

Problem 5.6. Does the integral

$$\int\limits_{(0,+\infty)} \frac{(-1)^{[x]}}{[x+1]}d\mu,$$

exist? Here μ is the Lebesgue measure, and $[x]$ is the integer part of x (not exceeding x).

Problem 5.7. Does the integral

$$\int\limits_{(0,+\infty)} \frac{(-1)^{[2x+1]}}{[x+1]} d\mu,$$

exist? Here μ is the Lebesgue measure, and $[x]$ is the integer part of x (not exceeding x).

Problem 5.8. Calculate the integral

$$\int\limits_{\mathcal{D}} f(x,y) d\mu,$$

where $\mathcal{D} = \{(x,y) :\ 0 \leqslant x \leqslant 1,\ 0 \leqslant y \leqslant 1\}$, μ is the Lebesgue measure, and

$$f(x,y) = \begin{cases} 1, & \text{where} \quad xy \notin \mathbb{Q} \\ 0, & \text{where} \quad xy \in \mathbb{Q}. \end{cases}$$

Problem 5.9. Let $f_n \in S^+(X)$, and

$$\int\limits_X f_n(x) d\mu \underset{n\to\infty}{\longrightarrow} 0.$$

Prove that $f_n \overset{\mu}{\underset{n\to\infty}{\longrightarrow}} 0$. Does this sequence converge to 0 a.e. on X?

Problem 5.10. Prove that the condition

$$\int\limits_X \frac{|f_n(x)|}{1+|f_n(x)|} d\mu \underset{n\to\infty}{\longrightarrow} 0$$

is equivalent to the fact $f_n \overset{\mu}{\underset{n\to\infty}{\longrightarrow}} 0$ whenever $\mu X < +\infty$. Construct a counterexample in the case when $X = \mathbb{R}$.

Problem 5.11. Let $f \in L(\mathbb{R}^n)$, and let $X_y = \{x \in \mathbb{R}^n :\ |f(x)| > y\}$. Prove that

$$\int\limits_{\mathbb{R}^n} |f(x)| d\mu = \int\limits_0^{+\infty} \mu(X_y) dy.$$

Problem 5.12. Let $f, g \in S^+(X)$. Prove that

$$\int\limits_X f(x)g(x) d\mu = \int\limits_0^{+\infty} \Phi(y) dy,$$

where $\Phi(y) = \int\limits_{X_y} f(x) dx$ and $X_y = X(g > y)$.

Hint: Use approximation of nonnegative measurable functions and Levi's theorem.

Problem 5.13. Let $f \in L(X)$ and for any set $E \subset X$, $E \in \mathcal{M}(X)$, the identity

$$\int\limits_E f(x) d\mu = 0$$

holds. Prove that $f(x) \overset{a.e.}{=} 0$ on X.

Problem 5.14. Let $f \in S(X)$, $\mu X < +\infty$, and $X_n = X(n-1 \leqslant f < n)$. Prove that $f \in L(X)$ if, and only if, the series $\sum\limits_{n=-\infty}^{+\infty} |n| \cdot \mu X_n$ converges.

Problem 5.15. Construct a finite-valued function $f \in S^+(\mathbb{R})$ such that

$$\sum_{n=1}^{+\infty} n \cdot \mu X_n < +\infty,$$

where $X_n = X(n \leqslant f < n+1)$, but such that $f \notin L^+(\mathbb{R})$.

Problem 5.16. Let $\mu X < +\infty$, $f \in S(X)$, and there exist constants $A > 0$ and $\alpha > 1$ such that for any $\varepsilon > 0$ the following inequality holds

$$\mu\{x \in X : \ |f(x)| > \varepsilon\} < \frac{A}{\varepsilon^\alpha}$$

Prove that $f \in L(X)$ w.r.t. the measure μ.

Problem 5.17. Prove that if $f \in L(X)$ and $X_n = X(|f| \geqslant n)$, then $\lim\limits_{n\to+\infty} n \cdot \mu X_n = 0$.

Problem 5.18. Let $\mu X < +\infty$ and $f \in S(X)$. Define the sets $X_n = \{x \in X : \ |f(x)| > n\}$, $n = 0, 1, 2, \ldots$. Prove that $f \in L(X)$ if, and only if,

$$\sum_{n=1}^{\infty} \mu X_n < \infty.$$

Problem 5.19. Construct an example of a function such that $f \in S^+(\mathbb{R})$, and

$$\sum_{n=1}^{\infty} \mu X_n < \infty.$$

where X_n are defined in Problem 5.18, and $f \notin L^+(\mathbb{R})$.

Problem 5.20. Let $f \in S^+(x)$, $\mu X < +\infty$, and $X_n = X(n-1 \leqslant f < n)$, $n = 1, 2, \ldots$. Prove that $f^m \in L^+(X)$ if, and only if, the series $\sum\limits_{n=1}^{+\infty} n^m \cdot \mu X_n$ converges, $m \in \mathbb{N}$.

Problem 5.21. Let $f \in S^+(x)$, $\mu X < +\infty$, and $B_n = X(f \geqslant n-1)$, $n = 1, 2, \ldots$. Prove that $f^m \in L^+(X)$ if, and only if, the series $\sum\limits_{n=1}^{+\infty} n^{m-1} \cdot \mu B_n$ converges, $m \in \mathbb{N}$.

Problem 5.22. Prove that the function $f(x) = \dfrac{1}{x} \cos \dfrac{1}{x}$ is non-Lebesgue integrable on $[0, 1]$.

Problem 5.23. Let f be unbounded *integrable* function on X. Put

$$[f(x)]_n = \begin{cases} f(x), & |f(x)| \leqslant n, \\ n \quad , & |f(x)| > n, \end{cases} \qquad [f(x)]_n^0 = \begin{cases} f(x), & |f(x)| \leqslant n, \\ 0 \quad , & |f(x)| > n. \end{cases}$$

Prove that

$$\int_X f(x) d\mu = \lim_{n\to+\infty} \int_X [f(x)]_n d\mu = \lim_{n\to+\infty} \int_X [f(x)]_n^0 d\mu.$$

Problem 5.24. Let $f, f_n \in L^+(X)$, $n \in \mathbb{N}$. Suppose that $f_n \xrightarrow{\mu} f$ on X, and

$$\lim_{n\to\infty} \int\limits_X f_n(x)d\mu = \int\limits_X f(x)d\mu.$$

Prove that for any $A \subset X$, $A \in \mathcal{M}(X)$,

$$\lim_{n\to\infty} \int\limits_A f_n(x)d\mu = \int\limits_A f(x)d\mu.$$

Hint: Use Riesz's Theorem on subsequences of sequences of measurable functions convergent in measure and the Fatou theorem.

Problem 5.25. Construct an example showing that in Problem 5.24 the condition $f_n(x) \geqslant 0$ on X is substantial.

Problem 5.26. Let $f \in L^+(X)$, and let $g \in S(X)$, $|g(x)| \leqslant M$ on X. Prove that there exists a number K, $\inf\{g(x) :\ x \in X\} \leqslant K \leqslant \sup\{g(x) :\ x \in X\}$, such that

$$\int_X f(x)g(x)d\mu = K \int_X f(x)d\mu.$$

Construct an example showing that the nonnegativity of the function f on X is crucial.

Problem 5.27. Let $\mu X < +\infty$ and $f, f_n \in L(X)$, $n \in \mathbb{N}$. Suppose that $f_n \xrightarrow{\mu} f$ on X, the functions $(f_n)_{n=1}^{+\infty}$ have uniformly absolute continuous integrals over X, that is, for any $\varepsilon > 0$ there exists $\delta > 0$ such that for any $A \subset X$, $A \in \mathcal{M}(X)$, $\mu A < \delta$, one has

$$\left|\int\limits_A f_n(x)d\mu\right| < \varepsilon \qquad \forall n \in \mathbb{N}.$$

Prove that

$$\lim_{n\to\infty} \int\limits_X f_n(x)d\mu = \int\limits_X f(x)d\mu.$$

Problem 5.28. Let $\{f_n\}_{n=1}^{+\infty} \in L(X)$, and $f_n \xrightarrow{a.e.} f$ on X as $n \to +\infty$. Suppose that $|f_n| \leqslant g(x)$, $n \in \mathbb{N}$ almost everywhere on X, where $g(x) \in L(X)$ and $|g(x)| \leqslant M$ almost everywhere on X. Prove that

$$\lim_{n\to+\infty} \int\limits_X f_n(x)g(x)d\mu = \int_X f(x)g(x)d\mu.$$

Problem 5.29. Let $f \in L(X)$, and let $(X_n)_{n=1}^{+\infty} \subset X$ be a monotone decreasing sequence of measurable sets, and $X_0 := \lim\limits_{n\to\infty} X_n$. Is the identity

$$\int\limits_{X_0} f(x)d\mu = \lim_{n\to+\infty} \int_{X_n} f(x)d\mu.$$

true?

Problem 5.30. Let $\mu X < +\infty$. Prove that if the integral $\int\limits_X f(x)g(x)d\mu$ exists and finite for any $f \in L(X)$, then $g(x)$ is bounded almost everywhere on X.

Problem 5.31. Prove that if

$$\int_X f(x)g(x)d\mu \geqslant 0$$

for any f such that $f(x) \geqslant 0$ on X, then $g(x) \geqslant 0$ almost everywhere on X.

Problem 5.32. Let $f(x) = \dfrac{1}{x^2}$ and μ be the Lebesgue measure on $(0,1)$. Prove that

$$\int_{(0,1)} f(x)d\mu = +\infty$$

using only the definition of Lebesgue integral for nonnegative functions.

Problem 5.33. Let $f(x) = \dfrac{1}{x}$ and μ be the Lebesgue measure on $(0,1)$. Prove that

$$\int_{(0,1)} f(x)d\mu = +\infty$$

using only the definition of Lebesgue integral for nonnegative functions.

Problem 5.34. Construct a sequence of functions $(f_n)_{n=1}^{+\infty}$, $f_n \in L^+([0,1])$ such that $f_n(x) \to 0$ for any $x \in [0,1]$ but

$$\int_{[0,1]} f_n(x)d\mu \not\to 0 \quad \text{as} \quad n \to \infty.$$

Problem 5.35. Show that there exists a sequence of functions $(f_n)_{n=1}^{+\infty}$, $f_n \in L^+([0,1])$ such that $f_n(x) \to 0$ for any $x \in [0,1]$ and

$$\int_{[0,1]} f_n(x)d\mu \to 0 \quad \text{as} \quad n \to \infty,$$

but $F(x) = \sup\limits_n f_n(x) \notin L^+([0,1])$.

Hint: Consider the sequence $f_n(x) = \dfrac{1}{n}2^n \chi_{A_n}(x)$, where $A_n = (2^{-n}, 2^{-n+1})$, $n \in \mathbb{N}$.

Problem 5.36. Construct an example of a sequence $(f_n)_{n=1}^{\infty} \in L(\mathbb{R})$ such that $f_n \underset{n\to\infty}{\longrightarrow} 0$ **uniformly** on $\mathbb{R}$ but

$$\int_{\mathbb{R}} f_n(x)d\mu \underset{n\to\infty}{\longrightarrow} \infty.$$

Problem 5.37. Let $(X, \mathcal{M}, \mu)$ be a measure set, and let $f \in L^+(X)$. Define

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

Prove that

1) ν is a σ-additive measure;

2) if g is integrable w.r.t. ν, then fg is integrable w.r.t. μ, and

$$\int_X g(x)d\nu(x) = \int_X f(x)g(x)d\mu(x).$$

5.1.2 Difference between Lebesgue and Riemann definite integrals

In this section, μ is the Lebesgue measure.

Problem 5.38. Consider the functions

$$f(x) = \begin{cases} 0, & x = 0, \\ \ln \dfrac{1}{x}, & x \in (0,1], \end{cases}$$

and

$$f_n(x) = \begin{cases} 0, & x \in \left[0, \frac{1}{n}\right), \\ \ln \dfrac{1}{x}, & x \in \left[\frac{1}{n}, 1\right]. \end{cases}$$

Prove that $f_n \xrightarrow{a.e.} f$ on $[0,1]$. Is $f(x)$ Riemann integrable on the interval $[0,1]$? Is it Lebesgue integrable on $[0,1]$? If so, find the Lebesgue integral of this function on $[0,1]$.

Problem 5.39. Is the function

$$f(x) = \begin{cases} 1, & x \in \mathbb{Q} \bigcap [0,1], \\ x^3, & x \in [0,1] \setminus \mathbb{Q}, \end{cases}$$

Riemann integrable on the interval $[0,1]$? Is it Lebesgue integrable on $[0,1]$? If so, find the Lebesgue integral of this function on $[0,1]$.

Problem 5.40. Let $K \in \mathbb{R}^n$ be a brick, and $f \in \mathcal{R}(K)$. Prove that $f \in L(K)$, and

$$(L) \int\limits_K f(x) d\mu = (\mathcal{R}) \int\limits_K f(x) dx^n.$$

Problem 5.41. Let $f \in \mathcal{R}(a+0, b]$, that is, let the improper integral

$$(\mathcal{R}) \int\limits_{a+0}^{b} f(x)dx := \lim_{c \to a+0} (\mathcal{R}) \int\limits_{c}^{b} f(x)dx$$

be finite. Suppose that $f(x) \geqslant 0$ on $(a,b]$. Prove that $f \in L^+(a,b)$ and

$$(L) \int\limits_{(a,b)} f(x) d\mu = (\mathcal{R}) \int\limits_{a+0}^{b} f(x)dx.$$

Problem 5.42. Let $f \in \mathcal{R}[a', b]$ for any $a' \in (a,b)$. Prove that $|f| \in \mathcal{R}(a+0, b]$ if, and only if, $f \in L(a,b)$. And if f is integrable, then

$$(L) \int\limits_{(a,b)} f(x) d\mu = (\mathcal{R}) \int\limits_{a+0}^{b} f(x)dx.$$

Hint: Putting $a_0 := b$ and considering some sequence $a_n \searrow a$, prove that

$$(\mathcal{R}) \int\limits_{a+0}^{b} f(x)dx = \lim_{n \to \infty} (\mathcal{R}) \int\limits_{a_n}^{b} f(x)dx,$$

then, use the complete additivity of the Lebesgue integral.

Problem 5.43. Let $\alpha \in \mathbb{R}$, and

$$f(x) = \begin{cases} x^\alpha, & x \in (0,1], \\ 0, & x = 0. \end{cases}$$

Find all α such that

a) $f \in \mathcal{R}[0,1]$;

b) $f \in \mathcal{R}(+0,1]$;

c) $f \in L[0,1]$.

Problem 5.44. Let $\alpha \in \mathbb{R}$, and $f(x) = x^\alpha$ on $[1,+\infty)$. Find all α such that

a) $f \in \mathcal{R}[1,+\infty)$;

b) $f \in L[1,+\infty)$.

Problem 5.45. Let $\alpha \in \mathbb{R}$, $\beta \geqslant 0$, and

$$f(x) = \begin{cases} x^\alpha \sin x^\beta, & x \in (0,1], \\ 0, & x = 0. \end{cases}$$

Find all pairs (α,β) such that

a) $f \in \mathcal{R}[0,1]$;

b) $f \in \mathcal{R}(+0,1]$;

c) $f \in L[0,1]$.

Problem 5.46. Let $\alpha \in \mathbb{R}$, $\beta < 0$, and

$$f(x) = \begin{cases} x^\alpha \sin x^\beta, & x \in (0,1], \\ 0, & x = 0. \end{cases}$$

Find all pairs (α,β) such that

a) $f \in \mathcal{R}[0,1]$;

b) $f \in \mathcal{R}(+0,1]$;

c) $f \in L[0,1]$.

Problem 5.47. Let $\alpha,\beta \in \mathbb{R}$, and $f(x) = x^\alpha \sin x^\beta$ on $[1,+\infty)$. Find all pairs (α,β) such that

a) $f \in \mathcal{R}[1,+\infty)$;

b) $f \in L[1,+\infty)$.

Problem 5.48. Find the limit

$$\lim_{n\to\infty} (L) \int\limits_{(0,+\infty)} \frac{d\mu(x)}{x^{\frac{1}{n}} \left(1+\frac{x}{n}\right)^n}.$$

Hint: To prove that integrands $\dfrac{1}{x^{\frac{1}{n}} \left(1+\frac{x}{n}\right)^n}$ are Lebesgue integrable on $(0,+\infty)$, estimate them by Lebesgue integrable functions on $(0,1)$ and $[1,+\infty)$ separately. Then use Lebesgue's Dominated Convergence Theorem.

Problem 5.49. Let $\mathbb{Q}\bigcap[0,1] = \{r_n\}_{n=1}^{\infty}$ and $\varepsilon > 0$. Prove that the series

$$\sum_{n=1}^{\infty} \frac{1}{n^{1+\varepsilon}\sqrt{|x-r_n|}}.$$

converges a.e. on $[0,1]$.

Hint: Using Problem 5.42 and Levi's theorem, prove that the sum of the series is Lebesgue integrable on $(0,1)$.

Problem 5.50. Let $X = (0,1)$. Construct nonnegative functions $(f_n)_{n=1}^{\infty}$ and f measurable w.r.t. the Lebesgue measure μ such that $f(x) = \overline{\lim} f_n(x)$ on X,

$$\int_X f_n(x)d\mu \leqslant C \qquad \forall n \in \mathbb{N}$$

for some constant $C > 0$, but $f \notin L(X)$.

Problem 5.51. Let $\mathcal{C}$ be the Cantor set on $[0,1]$, and

$$f(x)\begin{cases} 1 & \text{if} \quad x \in \mathcal{C}, \\ x & \text{if} \quad x \in \left[0, \dfrac{1}{2}\right] \setminus \mathcal{C}, \\ x^2 & \text{if} \quad x \in \left(\dfrac{1}{2}, 1\right] \setminus \mathcal{C}. \end{cases}$$

Find

$$(L)\int_{(0,1)} f(x)d\mu.$$

Problem 5.52. Let $\mathcal{C}$ be the Cantor set on $[0,1]$ and let

$$[0,1] \setminus \mathcal{C} = \dot{\bigcup_{n=1}^{\infty}} (a_n, b_n).$$

Define the functions f_n as follows. For any $n \in \mathbb{N}$, we have $f_n(x) = 0$ for any $x \in [0,1] \setminus (a_n, b_n)$, $f_n\left(\dfrac{a_n+b_n}{2}\right) = 1$, and $f_n(x)$ is continuous and linear on the intervals $\left[a_n, \dfrac{a_n+b_n}{2}\right]$, $\left[\dfrac{a_n+b_n}{2}, b_n\right]$. Let

$$f(x) = \sum_{n=1}^{\infty} f_n(x).$$

Find

$$(L)\int_{(0,1)} f(x)d\mu.$$

Problem 5.53. Let $\mathcal{C}$ be the Cantor set on $[0,1]$ and let

$$[0,1] \setminus \mathcal{C} = \dot{\bigcup_{n=1}^{\infty}} (a_n, b_n).$$

Define the functions f_n as follows. For any $n \in \mathbb{N}$, we have $f_n(x) = 0$ for any $x \in [0,1] \setminus (a_n, b_n)$, $f_n\left(\dfrac{a_n+b_n}{2}\right) = \dfrac{b_n-a_n}{2}$, and $f_n(x)$ is continuous and linear on the intervals $\left[a_n, \dfrac{a_n+b_n}{2}\right]$, $\left[\dfrac{a_n+b_n}{2}, b_n\right]$. Let

$$f(x) = \sum_{n=1}^{\infty} f_n(x).$$

Find

$$(L) \int_{(0,1)} f(x) d\mu.$$

Problem 5.54. Let

$$x = \sum_{i=1}^{\infty} \frac{x_i}{2^i}$$

be the binary representation of a number $x \in [0, 1)$, where $x_i \in \{0, 1\}$, and $\lim_{i \to \infty} x_i \neq 1$.

Let $f_k(x) = 2x_k - 1$ for $k \in \mathbb{N}$ and $x \in (0, 1)$. Prove that the sequence $\{f_k(x)\}_{k=1}^{\infty}$ is orthonormal on $(0, 1)$, that is, prove that

$$(L) \int_{(0,1)} f_k(x) f_j(x) d\mu = 0, \qquad \text{for} \quad j \neq k,$$

and

$$(L) \int_{(0,1)} f_k^2(x) d\mu = 1, \qquad k \in \mathbb{N}.$$

Problem 5.55. Find functions $f, g \in \mathcal{R}((+0, 1])$ such that

$$h(x) = \max\{f(x), g(x)\} \notin \mathcal{R}((+0, 1]).$$

5.1.3 Fubini's theorem

Let $X_1 \subset \mathbb{R}^{n_1}$ and $X_2 \subset \mathbb{R}^{n_2}$. Let μ_1 and μ_2 be two Lebesgue-Stieltjes σ-finite (or finite) measures defined on $\mathcal{M}_1 = \mathcal{M}(X_1)$ and $\mathcal{M}_2 = \mathcal{M}(X_2)$, respectively. The Descartes product of these two measures spaces is the following space $(X, \mathcal{M}, \mu)$, where $X = X_1 \times X_2$, $\mathcal{M} = \mathcal{M}_1 \times \mathcal{M}_2$, that is, any $A \in \mathcal{M}$ can be represented as $A = A_1 \times A_2$ with $A_1 \in \mathcal{M}_1$ and $A_2 \in \mathcal{M}_2$. The measure μ of $A \in \mathcal{M}$ is defined as $\mu(A) = \mu_1(A_1) \cdot \mu_2(A_2)$. it is easy to see that μ is a Lebesgue-Stieltjes measure defined on $\mathcal{M} = \mathcal{M}(X)$.

In this section, we consider such three measure spaces.

Problem 5.56. Let $A \in \mathcal{M}$, $\mu A < \infty$. Prove that for μ_1-a.e. x in X_1, the slice A_x is μ_2-measurable, that is, $A_x \in \mathcal{M}_2$. Also prove that the function $\mu_2(A_x)$ is μ_1-measurable on X_1, and

$$\mu(A) = \int_{X_1} \mu_2(A_x) d\mu_2(x).$$

Problem 5.57. Let $A \in \mathcal{M}$, and $f(x, y)$ is integrable on A w.r.t the measure μ. Prove that

1) f_x is μ_2-integrable on A_x for μ_1-a.e. x in X_1, where A_x is the x slice of A;

2) the function

$$\int_{A_x} f_x(y) d\mu_2(y)$$

is μ_1-integrable on X_1;

3) and

$$\int_A f(x, y) d\mu = \int_{X_1} \int_{A_x} f_x(y) d\mu_2(y) d\mu_1(x).$$

Problem 5.58. Let $A \in \mathcal{M}$, and $f(x, y) \in S^+(A)$. Prove that

1) f_x is μ_2-measurable for μ_1-a.e. x in X_1;

2) the function
$$\int_{A_x} f_x(y)d\mu_2(y)$$
is μ_1-measurable, where A_x is the x slice of A;

3) and
$$\int_A f(x,y)d\mu = \int_{X_1}\int_{A_x} f_x(y)d\mu_2(y)d\mu_1(x).$$

Problem 5.59. Construct a function $f(x,y) \notin L((0,1)\times(0,1))$ such that $f(x_0,y)$ is integrable on $(0,1)$ for any fixed $x_0 \in (0,1)$, and $f(x,y_0)$ is integrable on $(0,1)$ for any fixed $y_0 \in (0,1)$, and such that
$$\int_{(0,1)} f(x_0,y)d\mu(y) = \int_{(0,1)} f(x,y_0)d\mu(x) = 0.$$

Problem 5.60. Construct a finite function $f(x,y)$ measurable on $(0,1)\times(0,1)$ such that
$$\int_{(0,1)}\left(\int_{(0,1)} f(x,y)d\mu(x)\right)d\mu(y) = 0,$$
and
$$\int_{(0,1)}\left(\int_{(0,1)} f(x,y)d\mu(y)\right)d\mu(x) = 1.$$

Hint: Consider the function
$$f(x,y) = \sum_{n=0}^{\infty} 4^n\left[\chi_{A_n\times A_n}(x,y) - 2\chi_{A_{n+1}\times A_n}(x,y)\right],$$
where $A_n = (2^{-n}, 2^{-n+1})$, $n \in \mathbb{N}$.

Problem 5.61. Let $a, b \in \mathbb{R}$, $a \neq b$. Construct a finite function $f(x,y)$ measurable on $(0,1)\times(0,1)$ such that
$$\int_{(0,1)}\left(\int_{(0,1)} f(x,y)d\mu(x)\right)d\mu(y) = a,$$
and
$$\int_{(0,1)}\left(\int_{(0,1)} f(x,y)d\mu(y)\right)d\mu(x) = b.$$

Problem 5.62. Prove that the function $f(x,y) = e^{-xy}\sin x \sin y$ is integrable on $(0,+\infty)\times(0,+\infty)$ (w.r.t. Lebesgue measure).

Problem 5.63. Let $f(x) \in L_{\mu_1}(X_1)$, $g(x) \in L_{\mu_2}(X_2)$. Prove that $f(x)\cdot g(y) \in L_\mu(X)$. Here L_ν means integrability w.r.t. the measure ν.

Problem 5.64. Prove that if $A = A_1 \times A_2$ ($A_1 \subset X_1$ and $A_2 \subset X_2$) is measurable (w.r.t. μ) and has positive measure, then the sets A_1 and A_2 are measurable w.r.t measures μ_1 and μ_2, respectively.

Problem 5.65. Let a function f be measurable on $X = X_1 \times X_2$ (w.r.t. μ), and let the following integral exist

$$I = \int\limits_{X_1} \left(\int\limits_{X_2} |f(x,y)| d\mu_2(y) \right) d\mu_1(x) < +\infty.$$

Prove that $f \in L(X, \mathcal{M}, \mu)$.

Problem 5.66. Let $f, g \in L(\mathbb{R})$. Prove that the function $h(x,t) = f(t)g(x-t) \in L(\mathbb{R})$ for a.e. $x \in \mathbb{R}$ and that the convolution

$$(f * g)(x) = \int\limits_{\mathbb{R}} f(t)g(x-t) d\mu(t)$$

belongs to $L(\mathbb{R})$.

Chapter 6

Differentiation of integrable functions

I cover pages 98–108, 115–136 and 285–292 from the book [11].

6.1 Problems

6.1.1 Differentiation

Problem 6.1. Prove that if f is integrable on $\mathbb{R}^n$, and f is not identically zero, then

$$f^*(x) \geqslant \frac{c}{|x|^n}, \quad \text{for some } c > 0 \text{ and all } |x| \geqslant 1,$$

where $|x| = \sqrt{x_1^2 + \cdots + x_n^2}$. Conclude that f^* is not integrable on $\mathbb{R}^n$. Then, show that the weak type estimate

$$\mu(\{x \ : \ F^* > \alpha\}) \leqslant \frac{c}{\alpha}$$

for all $\alpha > 0$ whenever $\int |f| d\mu = 1$, is the best possible in the following sense: if f is supported in the unit ball with $\int |f| d\mu = 1$, then

$$\mu(\{x \ : \ F^* > \alpha\}) \geqslant \frac{c'}{\alpha}$$

for some $c' > 0$ and all sufficiently small α.

Hint: For the first part, use the fact that $\int\limits_B |f| d\mu > 0$ for some ball B.

Problem 6.2. Consider the function on $\mathbb{R}$ defined by

$$f(x) = \begin{cases} \dfrac{1}{|x|(\ln \frac{1}{|x|})^2} & \text{if } \ |x| \leqslant \dfrac{1}{2}, \\ \\ 0 & \text{otherwise.} \end{cases}$$

(a) Verify that f is integrable.

(b) Establish the inequality

$$f^*(x) \geqslant \frac{c}{|x|(\ln \frac{1}{|x|})} \quad \text{for some } c > 0 \text{ and all } |x| \leqslant \frac{1}{2},$$

to conclude that the maximal function f^* is not locally integrable.

6.1.2 Functions of bounded variation

By $V[a,b]$ we denote the set of all functions of bounded variation defined on the interval $[a,b]$. By $V_a^b(f)$ we denote the (total) variation of the function f on $[a,b]$:

$$V_a^b(f) = \sup_T \sum_{k=1}^N |F(t_k) - F(t_{k-1})|,$$

where the supremum is taken over all partitions of the interval $[a,b]$:

$$T := \{t_k\}_{k=0}^N, \qquad a = t_0 < t_1 < \ldots < t_{N-1} < t_N = b.$$

Problem 6.3. Let $c \in (a,b)$. Prove that $f \in V[a,b]$ if, and only if, $f \in V[a,c] \cap V[c,b]$, and

$$V_a^b(f) = V_a^c(f) + V_c^b(f).$$

Problem 6.4. Let $f \in V[a,b]$. Prove that f can be represented as a difference of *strictly increasing* functions.

Problem 6.5. Let

$$f(x) = \begin{cases} 1 & \text{for} \quad x = a \in [0,1], \\ \\ 0 & \text{for} \quad x \in [0,1] \setminus \{a\}. \end{cases}$$

Represent f as a difference of two non-strictly increasing functions.

Problem 6.6. Let $f(x) = \sin x$ defined on $[0, 2\pi]$. Represent f as a difference of two non-strictly increasing functions.

Problem 6.7. Let $f \in V[a,b]$. Prove that f is non-strictly increasing on $[a,b]$ if, and only if, $V_a^b(f) = f(b) - f(a)$.

Problem 6.8. Let f be finite on $[a,b]$, and let φ be a strictly increasing continuous function on $[a,b]$ with $\varphi(a) = a$ and $\varphi(b) = b$. Prove that $f \in V[a,b]$ if, and only if, $g(x) = f(\varphi(x)) \in V[a,b]$. Show by an example that continuity of φ is crucial in this result. Find $f \notin V[0,1]$ and strictly increasing function φ with $\varphi(0) = 0$, $\varphi(1) = 1$ such that $g(x) = f(\varphi(x)) \in V[0,1]$.

Problem 6.9. Let $f, g \in V[a,b]$. Prove that f and g are bounded, $f \cdot g \in V[a,b]$, and

$$V_a^b(f \cdot g) \leqslant \sup_{x \in [a,b]} |g(x)| \cdot V_a^b(f) + \sup_{x \in [a,b]} |f(x)| \cdot V_a^b(g).$$

Problem 6.10. Let $f, g \in V[a,b]$, and $|f(x)| \geqslant C > 0$ for $x \in [a,b]$. Prove that $\dfrac{g}{f} \in V[a,b]$.

Problem 6.11. Let $f, g \in V[a,b]$. Prove that $\max\{f, g\} \in V[a,b]$.

Problem 6.12. Let $f \in V[a,b]$. Prove that $|f| \in V[a,b]$, and $V_a^b(|f|) \leqslant V_a^b(f)$.

Problem 6.13. Find all real α and β such that the function

$$f(x) = \begin{cases} x^\alpha \sin x^\beta & \text{for} \quad x \in (0,1], \\ 0 & \text{for} \quad x = 0, \end{cases}$$

is of bounded variation on $[0,1]$.

Problem 6.14. Let $f \in V[a,b]$. Prove that $f \in C[a,b]$ if, and only if, $V_a^x(f) \in C[a,b]$.

Problem 6.15. Let a set $A \subset [a,b]$ be of Lebesgue measure zero, $\mu A = 0$. Find a continuous (non-strictly) increasing function on $[a,b]$ s.t. $f'(x) = \infty$ for any $x \in A$.

Hint: To construct a necessary function, use the following auxiliary function

$$g(x) = \sum_{n=1}^{\infty} n^2 \chi_{G_n}(x),$$

where $G_1 \supset G_2 \supset \ldots \supset A$, and $\mu G_n < 2^{-n}$.

Problem 6.16 (Fubini's Little Theorem)**.** Let $\{f_n(x)\}_{n=1}^{\infty}$ be a sequence of non-strictly increasing functions on $[a,b]$. Let also the series

$$\sum_{n=1}^{\infty} f_n(x)$$

converge everywhere on $[a,b]$, and let $f(x)$ be its sum. Prove that

$$f'(x) = \sum_{n=1}^{\infty} f_n'(x), \qquad \text{a.e. on} \quad [a,b].$$

Problem 6.17. Construct a sequence $\{f_n(x)\}_{n=1}^{\infty}$ of non-strictly increasing continuous functions on $[-1,1]$ such that the series

$$\sum_{n=1}^{\infty} f_n(x)$$

converges uniformly on $[-1,1]$ to a function $f(x)$, and

$$\sum_{n=1}^{\infty} f_n(0) = 0,$$

but $f'(0) = \infty$.

Hint: To construct a necessary function, use the following auxiliary functions

$$g_n(x) = \begin{cases} 0, & 0 \leqslant x \leqslant 2^{-n-1}, \\ 2^{n+1}\sqrt{x}(x - 2^{-n-1}), & 2^{-n-1} < x < 2^{-n}, \\ \sqrt{x}, & 2^{-n} \leqslant x \leqslant 1. \end{cases}$$

Problem 6.18. Given a sequence $\{f_n(x)\}_{n=1}^{\infty} \subset V[a,b]$ such that

$$\sum_{n=1}^{\infty} |f_n(a)| < \infty \qquad \text{and} \qquad \sum_{n=1}^{\infty} V_a^b(f_n) < \infty,$$

prove that the series

$$\sum_{n=1}^{\infty} f_n(x)$$

converges to some function $f(x) \in V[a,b]$ uniformly on $[a,b]$, and that

$$V_a^b(f) \leqslant \sum_{n=1}^{\infty} V_a^b(f_n).$$

Problem 6.19. Consider the following jump function

$$h(x) = \sum_{k=1}^{\infty} a_k g_k(x),$$

where

$$\alpha = \sum_{k=1}^{\infty} |a_k| < \infty$$

and $g_k(x) = \chi_{(c_k,b]}$ with $c_k \in [a,b]$, $c_k \neq c_j$ if $k \neq j$. Prove that $V_a^b(h) = \alpha < \infty$.

Problem 6.20. Construct a *strictly* increasing function $f(x) \in C[a,b] \cap V[a,b]$ such that $f'(x) = 0$ a.e. on $[a,b]$.

Problem 6.21. Let $\{f_n(x)\}_{n=1}^{\infty}$ be a sequence of non-strictly increasing functions on $[a,b]$. Suppose that there exists a constant $C > 0$ such that $|f_n(x)| \leqslant C$ for any $x \in [a,b]$ and for any $n \in \mathbb{N}$. Prove that there exists a subsequence $\{f_{n_k}\}_{k=1}^{\infty}$ converging everywhere on $[a,b]$ to a non-strictly increasing function on $[a,b]$.

Problem 6.22. Let $\{f_n(x)\}_{n=1}^{\infty} \subset V[a,b]$. Suppose that there exists a constant $C > 0$ such that $V_a^b(f_n)$ and $|f_n(x)| \leqslant C$ for any $x \in [a,b]$ and for any $n \in \mathbb{N}$. Prove that there exists a subsequence $\{f_{n_k}\}_{k=1}^{\infty}$ converging everywhere on $[a,b]$ to a function $f \in V[a,b]$.

Problem 6.23. Let $f \in C[a,b]$. For a given partition $T = \{a = t_0 < t_1 < \cdots < t_{n-1} < t_n = b\}$ of the interval $[a,b]$ we define the intervals $\Delta_k = [t_{k-1}, t_k]$ and the magnitudes $M_k = \max\limits_{t \in \Delta_k} f(t)$ and $m_k = \min\limits_{t \in \Delta_k} f(t)$ for $k = 1, \ldots, n$. Let

$$\Omega(T) = \sum_{k=1}^{n} (M_k - m_k).$$

Prove that

$$V_a^b(f) = \lim_{\lambda(T) \to 0} V_T(f) = \lim_{\lambda(T) \to 0} \Omega(T),$$

where $\lambda(T) = \max\limits_{1 \leqslant k \leqslant n} |\Delta_k|$. Here the value $V_a^b(f)$ can be finite or infinite.

Construct a function $f \notin V[0,1]$ and a sequence $\{T_i\}_{i=1}^{\infty}$ of partitions of the interval $[0,1]$ such that $\lambda(T_i) \to 0$ as $i \to \infty$ and $V_{T_i}(f) = 0$.

Hint: Show that $\liminf\limits_{\lambda(T) \to 0} V_T(f) \geqslant \alpha$ for any $\alpha < V_a^b(f)$.

Problem 6.24. Let $f \in C[a,b]$. The function $g(y)$ equal to the number of solutions of the equation $f(x) = y$ is called *Banach indicatrix* ($g(y) = \infty$ is possible). Prove that g is measurable and Lebesgue integrable, and

$$V_a^b(f) = \int\limits_{\mathbb{R}} g(y) d\mu(y).$$

Hint: Approximate $g(y)$ by an increasing sequence of some measurable functions, and use Levi's theorem and Problem 6.23.

Problem 6.25. Let $f \in V[a,b]$ and $f(x) = \psi(x) + j(x)$, where $j(x)$ is the jump function related to $f(x)$. Prove that $V_a^b(f) = V_a^b(\psi) + V_a^b(j)$.

Problem 6.26. Let $f \in C[a,b] \cap V[a,b]$. Prove that there exist nonnegative non-strictly increasing continuous function $f_1(x)$ and $f_2(x)$ such that $f(x) = f_1(x) - f_2(x)$ and $V_a^b(f) = V_a^b(f_1) + V_a^b(f_2)$.

Problem 6.27. Let $f \in V[a,b]$. Prove that there exist nonnegative non-strictly increasing function $f_1(x)$ and $f_2(x)$ such that $f(x) = f_1(x) - f_2(x)$ and $V_a^b(f) = V_a^b(f_1) + V_a^b(f_2)$.

6.1.3 Absolutely continuous functions

If a function f is absolutely continuous on $[a,b]$, then we write $f \in AC[a,b]$.

Problem 6.28. Let $f, g \in AC[a,b]$. Prove that $f \cdot g \in AC[a,b]$, and if $g(x) \neq 0$ on $[a,b]$, then $\dfrac{f}{g} \in AC[a,b]$.

Problem 6.29. Prove that if $f \in AC[a,b]$, then $|f| \in AC[a,b]$.

And conversely, prove that if $|f| \in AC[a,b]$ and $f \in C[a,b]$, then $f \in AC[a,b]$.

Problem 6.30. Prove that if $f \in AC[a,b]$, then $f \in V[a,b]$.

Problem 6.31. Construct a function $f \in C[0,1] \cap V[0,1]$ such that $f \notin AC[0,1]$.

Problem 6.32. Prove that if $f \in AC[a,b]$, then $V_a^x(f) \in AC[a,b]$.

Problem 6.33. Prove that $f \in AC[a,b]$ if, and only if, for any $\varepsilon > 0$ there exists $\delta > 0$ such that for any countable system $S = \{(a_k, b_k)\}_{k=1}^{\infty}$ of disjoint open intervals in $[a,b]$ with

$$\sum_{k=1}^{\infty} |b_k - a_k| < \delta$$

one has

$$\sum_{k=1}^{\infty} |f(b_k) - f(a_k)| < \varepsilon$$

Problem 6.34. Let $f \in AC[a,b]$, and suppose that $g(x)$ is a non-strictly increasing on $[c,d]$, $g(x) \in AC[c,d]$, $g(c) = a$, $g(d) = b$. Prove that $f(g(x)) \in AC[c,d]$.

Problem 6.35. A function f defined on $[a,b]$ is said to possess *Luzin N-property* if for any Lebesgue measurable set E, $E \subset [a,b]$, with $\mu E = 0$, the set $f(E)$ is Lebesgue measurable and $\mu\left(f(E)\right) = 0$.

Prove that if $f \in AC[a,b]$, then f possesses Luzin N-property.

Hint: Use the result of Problem 6.33.

Problem 6.36. Let $f \in C[a,b]$. Prove that $f(E)$ is Lebesgue measurable for any Lebesgue measurable set $E \subset [a,b]$ if, and only if, $f(x)$ possesses Luzin N-property.

Hint: Use the fact that continuous functions transfers compact sets to compact sets and the result of Problem 3.22.

Problem 6.37 (Banach–Zarecki)**.** Let $f \in C[a,b] \cap V[a,b]$, and suppose that f possesses Luzin N-property. Prove that $f \in AC[a,b]$.

Problem 6.38. Let $f \in AC[a,b]$ be non-strictly increasing on $[a,b]$, and let a set $A \subseteq [a,b]$ be Lebesgue measurable. Prove that

$$\mu(f(A)) = \int\limits_{A} f'(t)dt,$$

where μ is the Lebesgue measure.

Hint: Prove first that the Lebesgue-Stieltjes measure μ_f constructed with the function f is absolutely continuous w.r.t. the Lebesgue measure μ.

Problem 6.39. Let $F, G \in AC[a,b]$. Prove that

$$\int\limits_{[a,b]} F(x)G'(x)d\mu = F(b)G(b) - F(a)G(a) - \int\limits_{[a,b]} F'(x)G(x)d\mu.$$

Problem 6.40. Let $f \in L[a,b]$, and suppose that the function $G \in AC[c,d]$ is *strictly* increasing on $[c,d]$. Moreover, let $G(c) = a$, $G(d) = b$, and the inverse function $G^{-1}(y)$ belong to $AC[a,b]$. Prove that $f(G(y))G'(y) \in L[c,d]$, and

$$\int\limits_{[a,b]} f(x)d\mu(x) = \int\limits_{[c,d]} f(G(y))G'(y)d\mu(y).$$

Problem 6.41. Let $f(x)$ be Lipschitz continuous on $[a,b]$. Prove that $f \in AC[a,b]$.

Prove also that the function $f(x) = \dfrac{1}{\ln \frac{2}{x}}$ on $(0,1]$, $f(0) = 0$, is absolutely continuous on $[0,1]$, but not Lipschitz continuous.

Problem 6.42. Find all real α and β such that the function

$$f(x) = \begin{cases} x^\alpha \sin x^\beta & \text{for} \quad x \in (0,1], \\ 0 & \text{for} \quad x = 0, \end{cases}$$

is absolutely continuous on $[0,1]$.

Hint: Use Problems 6.41, 6.35, 6.13, and 6.37.

Problem 6.43. Let $f \in C[a,b] \cap V[a,b]$. Prove that there exists a unique representation $f(x) = g(x) + h(x)$, where $g \in AC[a,b]$, $g(a) = f(a)$, and $h(x) \equiv 0$ on $[a,b]$ or $h(x) \in C[a,b] \cap V[a,b]$ is such that $h(x) \not\equiv const$ and $h'(x) = 0$ a.e. on $[a,b]$.

Chapter 7

Linear spaces

7.1 Hölder and Minkowski inequalities

On the plane (ξ, η), consider the line $\eta = \xi^{p-1}$ for $\xi \geqslant 0$, $p > 1$. Let the number q be such that

$$\frac{1}{p} + \frac{1}{q} = 1. \tag{7.1.1}$$

From this identity it follows that $q - 1 = \dfrac{1}{p-1}$, so $q > 1$, since $p > 1$ by assumption. Moreover, the equation $\eta = \xi^{p-1}$ is equivalent to the equation $\xi = \eta^{q-1}$.

For $a, b > 0$, let S_1 be the square of the set bounded by the axis $O\xi$, by the graph of the function $\eta = \xi^{p-1}$, and by the line $\xi = a$, and let S_2 be the square of the set bounded by the axis $O\eta$, by the graph of the function $\eta = \xi^{p-1}$, and by the line $\eta = b$. Obviously,

$$S_1 + S_2 \geqslant ab, \tag{7.1.2}$$

where equality holds if, and only if, $b = a^{p-1}$. The squares S_1 and S_2 can be calculated as follows

$$S_1 = \int_0^a \xi^{p-1} d\xi = \frac{a^p}{p}, \quad S_2 = \int_0^b \eta^{q-1} d\eta = \frac{b^q}{q}. \tag{7.1.3}$$

Substituting these values of S_1 and S_2 into (7.1.2), we obtain the following inequality

$$ab \leqslant \frac{a^p}{p} + \frac{b^q}{q}, \tag{7.1.4}$$

called the *Young inequality*. Note that the equality in (7.1.4) is possible if, and only if, $b = a^{p-1}$, that is, if, and only if, $b^q = a^p$.

Hölder inequality

Let $(X, \mathcal{M}(X), \mu)$ be a measure space. Suppose that the measure μ is complete and σ-additive. Suppose also that $f, g \in S(X)$ such that $|f|^p, |g|^q \in L(X)$, where p and q are related as in (7.1.1). Put now in the Young inequality (7.1.4)

$$a := \frac{|f(x)|}{\left(\int\limits_X |f(x)|^p d\mu\right)^{\frac{1}{p}}}, \qquad b := \frac{|g(x)|}{\left(\int\limits_X |g(x)|^q d\mu\right)^{\frac{1}{q}}},$$

then we have

$$\frac{|f(x)g(x)|}{\left(\int\limits_X |f(x)|^p d\mu\right)^{\frac{1}{p}} \left(\int\limits_X |g(x)|^q d\mu\right)^{\frac{1}{q}}} \leqslant \frac{|f(x)|^p}{p\left(\int\limits_X |f(x)|^p d\mu\right)} + \frac{|g(x)|^q}{q\left(\int\limits_X |g(x)|^q d\mu\right)}.$$

The right-hand side of this inequality is Lebesgue integrable on X by assumption. Therefore, the left-hand side is also Lebesgue integrable on X by property 8) of integral of measurable functions. Thus, we have

$$\frac{\int\limits_X |f(x)g(x)|d\mu}{\left(\int\limits_X |f(x)|^p d\mu\right)^{\frac{1}{p}} \left(\int\limits_X |g(x)|^q d\mu\right)^{\frac{1}{q}}} \leqslant \frac{\int\limits_X |f(x)|^p d\mu}{p\left(\int\limits_X |f(x)|^p d\mu\right)} + \frac{\int\limits_X |g(x)|^q d\mu}{q\left(\int\limits_X |g(x)|^q d\mu\right)} = \frac{1}{p} + \frac{1}{q} = 1,$$

so

$$\int\limits_X |f(x)g(x)|d\mu \leqslant \left(\int\limits_X |f(x)|^p d\mu\right)^{\frac{1}{p}} \left(\int\limits_X |g(x)|^q d\mu\right)^{\frac{1}{q}}. \tag{7.1.5}$$

This inequality is called the *Hölder inequality*. For the case $p = q = 2$ we get

$$\int\limits_X |f(x)g(x)|d\mu \leqslant \left(\int\limits_X |f(x)|^2 d\mu\right)^{\frac{1}{2}} \left(\int\limits_X |g(x)|^2 d\mu\right)^{\frac{1}{2}}.$$

This inequality is called Cauchy–Schwarz–Bunyakovsky inequality.

Remark 7.1.1. We prove the Hölder inequality under implicit assumption that the integrals $\int\limits_X |f(x)|^p d\mu$, $\int\limits_X |g(x)|^q d\mu$ are non-zero. However, if, for instance, $\int\limits_X |f(x)|^p d\mu = 0$, then $f(x) \overset{a.e.}{=} 0$ on X by the property 8) of integral of nonnegative functions. But in this case, the left-hand side of (7.1.5) is zero together with its right-hand side, so the inequality (7.1.5) holds.

Remark 7.1.2. In the proof of the Hölder inequality we assumed that both integrals in the right-hand side of (7.1.5) are finite. In fact, the inequality (7.1.5) holds if one or both integrals are infinite. But in this case it becomes useless. The main meaning of the Hölder inequality is that from the integrability of the functions $|f|^p$ and $|g|^q$ we get the integrability of the product $|fg|$.

Remark 7.1.3. Since the equality in the Young inequality appears if, and only if, $a^p = b^q$, we have that in the Hölder inequality the equality appears if, and only if, $|f(x)|^p \overset{a.e.}{=} C|g(x)|^q$ on X for some constant $C > 0$.

Remark 7.1.4. In the same way, from the Young inequality, one can prove the following finite sum form of the Hölder inequality for real x_k, y_k, $k = 1, \dots, n$,

$$\left|\sum_{k=1}^n x_k y_k\right| \leqslant \left(\sum_{k=1}^n |x_k|^p\right)^{\frac{1}{p}} \left(\sum_{k=1}^n |y_k|^q\right)^{\frac{1}{q}}, \tag{7.1.6}$$

where equality holds if, and only if,

$$\frac{|x_j|^p}{\sum\limits_{k=1}^n |x_k|^p} = \frac{|y_j|^q}{\sum\limits_{k=1}^n |y_k|^q}, \quad \operatorname{sgn} x_j y_j = const, \qquad j = 1, \dots, n.$$

Remark 7.1.5. From (7.1.6), one can easily deduce the following form of the Hölder inequality

$$\left|\sum_{k=1}^{\infty} x_k y_k\right| \leqslant \left(\sum_{k=1}^{\infty} |x_k|^p\right)^{\frac{1}{p}} \left(\sum_{k=1}^{\infty} |y_k|^q\right)^{\frac{1}{q}},$$

provided the series in the right hand-side of this inequality converge.

Minkowski inequality

Let now $f, g \in S(X)$ and $|f|^p, |g|^p \in L(X)$, $p > 1$. Consider two sets $X_1 = X(|f| \geqslant |g|)$ and $X_2 = X(|f| < |g|)$. Then we have

$$\int\limits_X |f(x)+g(x)|^p d\mu = \int\limits_{X_1} |f(x)+g(x)|^p d\mu + \int\limits_{X_2} |f(x)+g(x)|^p d\mu \leqslant 2^p \int\limits_{X_1} |f(x)|^p d\mu + 2^p \int\limits_{X_2} |g(x)|^p d\mu,$$

so $|f+g|^p \in L(X)$. Moreover,

$$\int\limits_X |f(x)+g(x)|^p d\mu = \int\limits_X |f(x)+g(x)|^{p-1} \cdot |f(x)+g(x)| d\mu \leqslant$$

$$\leqslant \int\limits_X |f(x)+g(x)|^{p-1} \cdot |f(x)| d\mu + \int\limits_X |f(x)+g(x)|^{p-1} \cdot |g(x)| d\mu$$

Applying Hölder inequality to both summands in the right-hand side of the inequality above and recalling that $(p-1)q = p$, we get

$$\int\limits_X |f(x)+g(x)|^p d\mu \leqslant \left(\left(\int\limits_X |f(x)|^p d\mu\right)^{\frac{1}{p}} + \left(\int\limits_X |g(x)|^p d\mu\right)^{\frac{1}{p}}\right) \left(\int\limits_X |f(x)+g(x)|^p d\mu\right)^{\frac{1}{q}}$$

Now we divide both parts of the inequality by $\left(\int\limits_X |f(x)+g(x)|^p d\mu\right)^{\frac{1}{q}}$ to obtain the Minkowski inequality

$$\left(\int\limits_X |f(x)+g(x)|^p d\mu\right)^{\frac{1}{p}} \leqslant \left(\int\limits_X |f(x)|^p d\mu\right)^{\frac{1}{p}} + \left(\int\limits_X |g(x)|^p d\mu\right)^{\frac{1}{p}}. \tag{7.1.7}$$

Here we used the identity $1 - \dfrac{1}{q} = \dfrac{1}{p}$.

Remark 7.1.6. We prove the Minkowski under implicit assumption that the integral $\int\limits_X |f(x)+g(x)|^p d\mu$ is non-zero. However, if it is zero then the inequality (7.1.7) is obvious.

Remark 7.1.7. We deduce the Minkowski inequality from the Hölder inequality which holds for $p > 1$. However, the Minkowski holds for $p \geqslant 1$, since for $p = 1$ it is obvious.

Remark 7.1.8. Since the equality in the Hölder inequality appears if, and only if, $|f(x)|^p \overset{a.e.}{=} C|g(x)|^q$ on X for some constant $C > 0$, we have that in the Minkowski inequality the equality appears if, and only if, either $f(x) \overset{a.e.}{=} Cg(x)$ on X for some constant $C \geqslant 0$, or $g(x) \overset{a.e.}{=} 0$ on X.

Remark 7.1.9. In the same way, from the Hölder inequality, one can prove the following finite sum form of the Minkowski inequality for real x_k, y_k, $k = 1, \dots, n$,

$$\left(\sum_{k=1}^{n} |x_k + y_k|^p\right)^{\frac{1}{p}} \leqslant \left(\sum_{k=1}^{n} |x_k|^p\right)^{\frac{1}{p}} + \left(\sum_{k=1}^{n} |y_k|^p\right)^{\frac{1}{p}}. \tag{7.1.8}$$

where equality holds if, and only if,

$$x_j = \lambda y_j,\ \lambda \geqslant 0, \quad \text{or} \quad y_j = 0, \qquad j = 1, \dots, n.$$

Remark 7.1.10. As above, from (7.1.8), one can easily deduce the series form of the Minkowski inequality

$$\left(\sum_{k=1}^{\infty} |x_k + y_k|\right)^{\frac{1}{p}} \leqslant \left(\sum_{k=1}^{\infty} |x_k|^p\right)^{\frac{1}{p}} + \left(\sum_{k=1}^{\infty} |y_k|^p\right)^{\frac{1}{p}}, \tag{7.1.9}$$

provided the series in the right hand-side of this inequality converge.

7.2 Linear normed spaces

Definition 7.2.1. A non-empty set V is called a *linear* or *vector space* over a field F if it satisfies the following conditions.

I. For any $x, y \in V$, there is a uniquely determined element $z \in V$ called *the sum* of x and y, and denoted as $x + y$. Moreover,

   1) $x + y = y + x$ (commutativity),
   2) $x + (y + z) = (x + y) + z$ (associativity),
   3) There exists an element $0 \in V$ such that $x + 0 = x$ for all $x \in V$ (existence of zero),
   4) For any $x \in V$ there exists in element $-x \in V$ such that $x + (-x) = 0$ (existence of the opposite element),

II. For any number $\alpha \in F$ and for any element $x \in V$, the element $\alpha x \in V$ is defined (the product of the element x and the number α). Moreover,

   1) $\alpha(\beta x) = (\alpha\beta)x$,
   2) There exists a number 1 in the field F such that $1 \cdot x = x$ for any $x \in V$,
   3) $(\alpha + \beta)x = \alpha x + \beta x$,
   4) $\alpha(x + y) = \alpha x + \alpha y$.

Example 7.2.2.

1) The real line $\mathbb{R}$ with ordinary operations of addition and product by a real number is a linear space over the field $F = \mathbb{R}$.

2) The collection of all possible n-tuple of real numbers $x = (x_1, \dots, x_n)$ with the following product and addition

$$(x_1, \dots, x_n) + (y_1, \dots, y_n) = (x_1 + y_1, \dots, x_n + y_n),$$

$$\alpha(x_1, \dots, x_n) = (\alpha x_1, \dots, \alpha x_n),$$

is a linear space over the field $F = \mathbb{R}$. This space is denoted $\mathbb{R}^n$. Analogously, one can define the space $\mathbb{C}^n$ over the field $F = \mathbb{C}$.

3) The set of all (real or complex) continuous functions on $[a,b]$ with ordinary addition and multiplication by a number (real or complex) is a linear space over the field $\mathbb{R}$ or $\mathbb{C}$, denoted $C[a,b]$.

4) The space l_p, $p>1$, whose elements are sequences of numbers (real or complex)

$$x=(x_1,\dots,x_n,\dots)$$

satisfying the condition

$$\sum_{k=1}^{\infty}|x|^p<+\infty,$$

with operations

$$(x_1,\dots,x_n,\dots)+(y_1,\dots,y_n,\dots)=(x_1+y_1,\dots,x_n+y_n,\dots),$$

$$\alpha(x_1,\dots,x_n,\dots)=(\alpha x_1,\dots,\alpha x_n,\dots),$$

is a linear space over the field $\mathbb{R}$ or $\mathbb{C}$. The fact that the sum of two element in l_p belongs l_p, follows from the Minkowski inequality (7.1.9).

5) The set of all convergent (real or complex) sequences with coordinate-wise summation and multiplication by a (real or complex) number is a linear space.

6) The set of all (real or complex) sequences convergent to 0 with coordinate-wise summation and multiplication by a (real or complex) number is a linear space.

7) The set of all bounded (real or complex) sequences with coordinate-wise summation and multiplication by a (real or complex) number is a linear space.

8) The set of all (real or complex) sequences ($\mathbb{R}^\infty$ or $\mathbb{C}^\infty$) with coordinate-wise summation and multiplication by a (real or complex) number is a linear space.

Definition 7.2.3. Two linear spaces V and V' over a field $\mathbb{F}$ are called *isomorphic* if there exists a one-to-ne correspondence between their elements which agrees with the linear operations in V and V'. This means that from $x\leftrightarrow x'$ and $y\leftrightarrow y'$ $(x,y\in V,\ x',y'\in V')$ it follows that

$$x+y\leftrightarrow x'+y'$$

and

$$\alpha x\leftrightarrow \alpha x'\qquad \forall\alpha\in\mathbb{F}.$$

Isomorphic spaces can be considered as different realizations of *the same linear space*. For example, the space $\mathbb{R}^n$ (respectively, $\mathbb{C}^n$) is isomorphic to the space $\mathbb{R}^{n-1}[x]$ (respectively, $\mathbb{C}^{n-1}[x]$) of all polynomials of degree $n-1$ with real (complex) coefficients.

Definition 7.2.4. Elements x_1, x_2, …, x_m of a linear space V are called *linearly dependent* if there exist numbers $\alpha_k\in\mathbb{F}$, $k=1,\dots,m$, such that $\sum\limits_{k=1}^{m}|\alpha_k|\neq 0$ and

$$\alpha_1x_1+\alpha_2x_2+\dots+\alpha_mx_m=0. \tag{7.2.1}$$

Otherwise, the elements x_1, x_2, …, x_m are called *linearly independent*. In other word, the elements x_1, x_2, …, x_m are linearly independent if the identity (7.2.1) implies $\alpha_1=\alpha_2=\ldots=\alpha_m=0$.

Definition 7.2.5. An *infinite* system of elements x_1, x_2, …, of a linear space V is called *linearly independent* if any its finite subsystem is linearly independent.

If in the linear space V one can find n linearly independent elements, and any $n+1$ its elements are linearly dependent, then the space V is said to have *dimension* n. If in the space V one can find arbitrary many linearly independent elements, then the space V is called *infinitely dimensional*. Any linearly independent system with n elements of n-dimensional space V is called a *basis* of the space V.

It is easy to show that the spaces $\mathbb{R}^n$ and $\mathbb{C}^n$ are n-dimensional.

Definition 7.2.6. A linear space V over a field $\mathbb{F}$ is called *normed* if for any $x \in V$ there exists the number $\|x\|$ called the *norm* of the element x, satisfying the following conditions:

1) $\|x\| \geqslant 0$ ($= 0$ if, and only if $x = 0$),
2) $\|x + y\| \leqslant \|x\| + \|y\|$ (triangle inequality),
3) $\|\alpha x\| = |\alpha| \cdot \|x\|$,

for all $x, y \in V$, and for any $\alpha \in \mathbb{F}$.

Note that the function $\rho(x, y) = \|x - y\|$, $x, y \in V$, is a metric, since it satisfies the following properties

1) $\rho(x, y) \geqslant 0$ ($= 0$ if, and only if, $x = y$),
2) $\rho(x, y) = \rho(y, x)$,
3) $\rho(x, z) \leqslant \rho(x, y) + \rho(y, z)$.

Thus, any linear normed space is a metric space!

Theorem 7.2.7. *Let V be a linear normed space, and let a sequence $\{x_n\}_{n=1}^{\infty} \subset V$ is convergent to $x \in V$ in the norm of the space V. Then $\|x_n\| \to \|x\|$ as $n \to \infty$.*

Proof. From the triangle inequality we obtain the inequalities

$$\|x_n\| \leqslant \|x_n - x\| + \|x\|, \qquad \|x\| \leqslant \|x_n - x\| + \|x_n\|,$$

which imply

$$|\|x_n\| - \|x\|| \leqslant \|x_n - x\|,$$

as required, since $\|x_n - x\| \to 0$ as $n \to \infty$ by assumption. □

Converse, of course, is not true. For example, the sequence of real numbers $x_n = (-1)^{n-1}$, $n \in \mathbb{N}$, has no limit but the sequence $|x_n| = 1$, $n \in \mathbb{N}$, converges to 1.

Definition 7.2.8. A linear normed space V is called *complete* w.r.t. its norm if any Cauchy sequence of its elements converges, that is, from $\|x_n - x_m\| \underset{n,m\to\infty}{\longrightarrow} 0$ it follows that the sequence $(x_n)_{n=1}^{\infty}$ converges.

Definition 7.2.9. Linear normed spaces complete w.r.t. their norm are called *Banach spaces*.

Example 7.2.10.

1) The real line $\mathbb{R}$ is a complete linear normed space with the norm $\|x\| = |x|$.
2) The spaces $\mathbb{R}^n$ and $\mathbb{C}^n$ become complete linear normed spaces of elements $x = (x_1, \dots, x_n)$ if we set

$$\|x\|_p = \left(\sum_{k=1}^{n} |x_k|^p\right)^{\frac{1}{p}}, \qquad 1 \leqslant p < +\infty.$$

We can also introduce the norm

$$\|x\|_\infty = \max_{1 \leqslant k \leqslant n} |x_k|.$$

It is easy to show that all the introduced norms are equivalent, that is for any $p_1, p_2 \in [1, +\infty]$ there exist numbers c_{p_1,p_2} and C_{p_1,p_2} such that

$$c_{p_1,p_2}\|x\|_{p_1} \leqslant \|x\|_{p_2} \leqslant C_{p_1,p_2}\|x\|_{p_1}.$$

3) The set $C[a,b]$ of (real or complex) continuous functions on $[a,b]$ is a complete normed linear space with the norm
$$\|f\| = \max_{a\leqslant x\leqslant b} |f(x)|.$$
Indeed, $\|f\| \geqslant 0$ for any $f \in C[a,b]$. Moreover, if $\|f\| = 0$, it is clear that $f(x) \equiv 0$ on $[a,b]$. Furthermore, for $f, g \in C[a,b]$ and for any $x \in [a,b]$, one has
$$|f(x) + g(x)| \leqslant |f(x)| + |g(x)| \leqslant \|f\| + \|g\|.$$
Since it is true for any $x \in [a,b]$, this is true for $\|f + g\|$. Thus, $\|f\|$ satisfies the property 2) of norms. The property 3) is obvious in this case.

Furthermore, if $(f_n)_{n=1}^{\infty} \subset C[a,b]$ is a Cauchy sequence: $\|f_n - f_m\| \underset{n,m\to\infty}{\longrightarrow} 0$, then we have that for any $c \in [a,b]$ the sequence $(f_n(c))_{n=1}^{\infty}$ is a Cauchy sequence of real (complex) numbers, so it is convergent as $\mathbb{R}$ and $\mathbb{C}$ are complete spaces. Thus, the sequence $(f_n(x))_{n=1}^{\infty}$ converges pointwise to a function $f(x)$ on $[a,b]$. Since $(f_n(x))_{n=1}^{\infty}$ is a Cauchy sequence, we have that
$$\forall\varepsilon > 0 \ \exists n_0 \in \mathbb{N} : \ \forall n \geqslant n_0 \ \forall p \in \mathbb{N} \ \max_{x\in[a,b]} |f_{n+p}(x) - f_n(x)| < \varepsilon.$$
If $p \to \infty$, we obtain
$$\forall\varepsilon > 0 \ \exists n_0 \in \mathbb{N} : \ \forall n \geqslant n_0 \ \max_{x\in[a,b]} |f_n(x) - f(x)| \leqslant \varepsilon.$$
Thus, the sequence $(f_n(x))_{n=1}^{\infty}$ *uniformly* converges to $f(x)$ on $[a,b]$, so $f \in C[a,b]$ as a uniform limit of a sequence of continuous functions. Consequently, $C[a,b]$ is complete.

4) The space l_p, $p \geqslant 1$, is a complete linear normed space with the norm
$$\|x\|_p = \left(\sum_{k=1}^{\infty} |x|^p\right)^{\frac{1}{p}}.$$
Indeed, it is clear that the function $\|x\|_p : l_p \mapsto [0, +\infty)$ satisfies the properties 1) and 3) of norms. Moreover, by Minkowsky's inequality (7.1.9) it satisfies the property 2), so it is a norm by Definition 7.2.6.

Suppose $(x^{(n)})_{n=1}^{\infty}$ is a Cauchy sequence of elements l_p, that is,
$$\sum_{k=1}^{\infty} |x_k^{(n)} - x_k^{(m)}|^p \to 0 \qquad \text{as} \qquad n, m \to +\infty. \tag{7.2.2}$$
This, in particular, means that for any $k \in \mathbb{N}$
$$|x_k^{(n)} - x_k^{(m)}| \to 0 \qquad \text{as} \qquad n, m \to +\infty,$$
so for any $k \in \mathbb{N}$ the sequence $(x_k^{(n)})_{n=1}^{\infty}$ converges to a number x_k. Thus, we have an infinite sequence $x = (x_k)_{k=1}^{\infty}$. It is left to prove that $x \in l_p$. In its turn it follows from Minkowski inequality. In fact, from (7.2.2) we obtain that for any $M \in \mathbb{N}$ and for any $\varepsilon > 0$, there exists a number $n_0 \in \mathbb{N}$ such that
$$\sum_{k=1}^{M} |x_k^{(n)} - x_k^{(m)}|^p < \varepsilon \qquad n, m \geqslant n_0.$$
As $m \to \infty$, we obtain for any $M \in \mathbb{N}$,
$$\sum_{k=1}^{M} |x_k^{(n)} - x_k|^p \leqslant \varepsilon \qquad n \geqslant n_0.$$

So when $M \to \infty$, one has
$$\sum_{k=1}^{\infty} |x_k^{(n)} - x_k|^p \leqslant \varepsilon \qquad n \geqslant n_0.$$
Now by Minkowski's inequality (7.1.9), we get
$$\left(\sum_{k=1}^{\infty} |x_k|^p\right)^{\frac{1}{p}} \leqslant \left(\sum_{k=1}^{\infty} |x_k^{(n)} - x_k|^p\right)^{\frac{1}{p}} + \left(\sum_{k=1}^{\infty} |x_k^{(n)}|^p\right)^{\frac{1}{p}}.$$
Since both series in the right-hand side of this inequality are convergent, the series in its left-hand side is convergent as well, so $x \in l_p$.

However, not every linear normed space is complete.

Example 7.2.11. Consider the set of all functions continuous on $[-1,1]$, and introduce the norm
$$\|f\| = (L) \int\limits_{[-1,1]} |f(x)|d\mu = (\mathcal{R}) \int\limits_{-1}^{1} |f(x)|dx. \tag{7.2.3}$$
It is clear that the number $\|f\|$ satisfies all the properties of norms. Thus, we get a linear normed space that we denote as $\mathfrak{R}[-1,1]$.

Let $(f_n(x))_{n=1}^{\infty} \subset \mathfrak{R}[-1,1]$ be the following sequence
$$f_n(x) = \begin{cases} -1 & \text{if} \quad -1 \leqslant x \leqslant -\frac{1}{n}, \\ nx & \text{if} \quad -\frac{1}{n} \leqslant x \leqslant \frac{1}{n}, \\ 1 & \text{if} \quad \frac{1}{n} \leqslant x \leqslant 1. \end{cases}$$
It is easy to see that
$$\|f_n(x) - f_m(x)\| = \int\limits_{-1}^{1} |f_n(x) - f_m(x)|dx = \left|\frac{1}{m} - \frac{1}{n}\right| \to 0 \quad \text{as} \quad n, m \to \infty,$$
and
$$\|f_n(x) - \operatorname{sgn}(x)\| = \int\limits_{-1}^{1} |f_n(x) - \operatorname{sgn}(x)|dx \to 0 \quad \text{as} \quad n \to \infty.$$
At the same time, for any function f continuous on $[-1,1]$ one has
$$0 < \|f(x) - \operatorname{sgn}(x)\| \leqslant \|f(x) - f_n(x)\| + \|f_n(x) - \operatorname{sgn}(x)\|,$$
so the sequence $(f_n)_{n=1}^{\infty}$ does not have continuous limit on $[-1,1]$. Thus, the space $\mathfrak{R}[-1,1]$ is not complete.

As we will se below, the completion of the space $\mathfrak{R}[-1,1]$ (that is, roughly speaking, adding all the limits of sequences of elements of the space $\mathfrak{R}[-1,1]$) is exactly the space $L[-1,1]$ with the norm (7.2.3).

7.3 Spaces L_p

Let $(X, \mathcal{M}(X), \mu)$ be a measure space, and the measure μ is supposed to be σ-additive and complete. Recall that two functions $f, g \in S(X))$ are called equivalent on X and denoted $f \sim g$ if $f \overset{a.e.}{=} g$ on X. Let us split the set $S(X)$ into equivalence classes. It is possible since the equivalence relation $f \sim g$ is reflexive ($f \sim f$), symmetric ($f \sim g \Longrightarrow g \sim f$), and transitive ($f \sim g,\ g \sim h \Longrightarrow f \sim h$). In what follows, we denote the class of functions equivalent to a function f by the same letter f, and do not distinguish functions that belong to the same equivalence class. So dealing with an equivalence class we pick a representative function of this class (it may be any function in the class) and deal with this function.

Definition 7.3.1. The collection of all equivalence classes of measurable on X functions satisfying the inequality

$$\int_X |f(x)|^p d\mu < +\infty, \qquad p \geqslant 1, \tag{7.3.1}$$

is defined to be a normed linear space $L_p(X, \mathcal{M}, \mu)$ with the norm

$$\|f\|_p := \left(\int_X |f(x)|^p d\mu \right)^{\frac{1}{p}} \tag{7.3.2}$$

Remark 7.3.2. The norm defined in (7.3.2) is finite and does not depend on the representative functions of equivalence classes, since all function from one equivalence class have the same integral over X.

Remark 7.3.3. In what follows we denote the class $L_p(X, \mathcal{M}, \mu)$ as $L_p(X)$ if it does not lead to any disambiguations.

It is clear that $L_p(X)$ is a linear space. Indeed, for any number λ (real or complex) and for any $f \in L_p(X)$, the class of functions λf belongs to the space $L_p(X)$ by the property 6) of integral of measurable functions. Moreover, from the Minkowski inequality (7.1.7). it follows that $f + g \in L_p$ whenever $f, g \in L_p(X)$.

It is also easy to check that the number $\|f\|_p$ satisfies all the properties of norms. Indeed, $\|f\|_p \geqslant 0$ for any $f \in L_p(X)$. Moreover, if $\|f\|_p = 0$, then $f \overset{a.e.}{=} 0$ on X, so it is belong to the class of functions which has the function $f(x) \equiv 0$ as a representative function. Thus, the norm satisfies the property 1) of norms. The property 3) is obvious, and the property 2) follows from the Minkowski inequality (7.1.7).

Definition 7.3.4. The convergence in the norm of the space $L_p(X)$ is called the *convergence in the* p^{th} *mean*. For $p = 1$, it is usually called the *mean convergence*, and for $p = 2$ it is called the *mean-square convergence*. If a sequence $(f_n)_{n=1}^\infty$ of functions in $L_p(X)$ converges to $f_0 \in L_p(X)$, we denote $f_n \xrightarrow{L_p} f_0$.

Theorem 7.3.5 (Uniqueness of limit)**.** *Any convergent sequences in $L_p(X)$, $p \geqslant 1$, has a unique limit.*

Proof. Let $(f_n)_{n=1}^\infty$ be such that $f_n \xrightarrow[n\to\infty]{L_p} f$ and $f_n \xrightarrow[n\to\infty]{L_p} g$. Then by Minkowski's inequality (7.1.7) we have

$$\|f - g\|_p = \|(f - f_n) + (f_n - g)\|_p \leqslant \|f - f_n\|_p + \|g - f_n\|_p \to 0 \quad \text{as} \quad n \to \infty,$$

so $\|f - g\|_p = 0$. Thus, $f \overset{a.e.}{=} g$ on X, therefore, they belong to one class and are considered as one element in the space $L_p(X)$ by Definition 7.3.1. □

By Theorem 7.2.7, we have

Theorem 7.3.6 (Continuity of norm)**.** *If $f_n \xrightarrow[n\to\infty]{L_p} f_0$, $p \geqslant 1$, then $\|f_n\|_p \xrightarrow[n\to\infty]{} \|f\|_p$.*

Theorem 7.3.7. *If $f_n \xrightarrow[n\to\infty]{L_p} f_0$, $p \geqslant 1$, then $(f_n)_{n=1}^\infty$ is a Cauchy sequence w.r.t. the norm $\|\cdot\|_p$.*

Proof. By assumption, for any $\varepsilon > 0$, there exists a number $n_0 \in \mathbb{N}$ such that for any $n \geqslant n_0$ one has

$$\|f_n - f\|_p < \frac{\varepsilon}{2}.$$

Let $n, m \geqslant n_0$. Then from Minkowski's inequality (7.1.7), it follows that

$$\|f_n - f_m\|_p = \|(f_n - f) + (f - f_m)\|_p \leqslant \|f_n - f\|_p + \|f - f_m\|_p < \varepsilon,$$

as required. □

The converse statement is not so easy to prove and means that the space $L_p(X)$ is Banach, provided $\mu X < +\infty$. To prove this statement we need the following lemma which can also be helpful in some other cases.

Lemma 7.3.8. *If $f \in S(X)$ and $f \in L_p(X)$ for some number $p > 1$, then $f \in L_{p_1}(X)$ for any $p_1 \in [1, p)$. Moreover, the following inequality holds*

$$\left(\int\limits_X |f(x)|^{p_1} d\mu\right)^{\frac{1}{p_1}} \leqslant (\mu X)^{\frac{p-p_1}{p_1 p}} \left(\int\limits_X |f(x)|^p d\mu\right)^{\frac{1}{p}} \tag{7.3.3}$$

Proof. Let us introduce the numbers p' and q' as follows

$$p' = \frac{p}{p_1}, \qquad q' = \frac{p}{p - p_1}.$$

It is easy to see that $\frac{1}{p'} + \frac{1}{q'} = 1$.

Applying the Hölder inequality with powers p' and q' to the integral $\int\limits_X |f(x)|^{p_1} d\mu$, we obtain

$$\int\limits_X |f(x)|^{p_1} d\mu = \int\limits_X |f(x)|^{p_1} \cdot 1 d\mu \leqslant \left(\int\limits_X |f(x)|^{p_1 \cdot \frac{p}{p_1}}\right)^{\frac{p_1}{p}} \cdot \left(\int\limits_X 1^{\frac{p}{p-p_1}}\right)^{\frac{p-p_1}{p}} = (\mu X)^{\frac{p-p_1}{p}} \left(\int\limits_X |f(x)|^p\right)^{\frac{p_1}{p}},$$

that gives us (7.3.3) after taking powering in $\frac{1}{p_1}$. $\square$

Theorem 7.3.9 (Riesz-Fischer)**.** *The space $L_p(X)$, $p \geqslant 1$, $\mu X < +\infty$, is a complete linear normed space, that is, it is Banach w.r.t. the norm $\|\cdot\|_p$.*

Proof. Let $(f_n)_{n=1}^{+\infty}$ be a Cauchy sequence of functions in the space L_p. For any $\varepsilon > 0$, there exists a number $n_0 \in \mathbb{N}$ such that

$$\int\limits_X |f_m(x) - f_n(x)|^p d\mu < \varepsilon^p, \qquad m, n \geqslant n_0. \tag{7.3.4}$$

If $p > 1$,then from (7.3.3) with $p_1 = 1$ we obtain

$$\int\limits_X |f_m(x) - f_n(x)| d\mu < (\mu X)^{\frac{p-1}{p}} \cdot \varepsilon, \qquad m, n \geqslant n_0.$$

Consequently, if $(f_n)_{n=1}^{\infty}$ is a Cauchy sequence in $L_p(X)$, $p \geqslant 1$, it is a Cauchy sequence in $L_1(X)$.

Furthermore, for the number $\varepsilon = \frac{1}{2^k}$ there exists a number $n_0(k)$ such that

$$\int\limits_X |f_m(x) - f_n(x)| d\mu < \frac{1}{2^k}, \qquad m, n \geqslant n_0(k).$$

For $k = 1$, we take $n_1 > n_0(1)$, then for $k = 2$, we take $n_2 > \max\{n_1, n_0(2)\}$, etc. When we have the numbers $n_1 < n_2 < \cdots < n_{k-1}$, we take $n_k > \max\{n_{k-1}, n_0(k)\}$, etc. Thus, we obtain a subsequence $(f_{n_k})_{k=1}^{\infty}$ such that

$$\int\limits_X |f_{n_{k+1}}(x) - f_{n_k}(x)| d\mu < \frac{1}{2^k}, \qquad k = 1, 2, \ldots.$$

Introduce the following function

$$g(x) = |f_{n_1}(x)| + \sum_{k=2}^{\infty} |f_{n_k}(x) - f_{n_{k-1}}(x)|. \tag{7.3.5}$$

By Lemma 7.3.8, all the terms of this series are integrable functions. According to Levy's Theorem 5.2.3, one has

$$\int_X g(x)d\mu = \int_X |f_{n_1}(x)|d\mu + \sum_{k=2}^{\infty} \int_X |f_{n_k}(x) - f_{n_{k-1}}(x)|d\mu < \int_X |f_{n_1}(x)|d\mu + \sum_{k=2}^{\infty} \frac{1}{2^{k-1}} < +\infty$$

By the property 6) of integral of nonnegative functions, the function $g(x)$ is finite a.e. on X. In other word, the series (7.3.5) converges a.e. on X. Therefore, the series

$$f_{n_1}(x) + \sum_{k=2}^{\infty} (f_{n_k}(x) - f_{n_{k-1}}(x))$$

also converges a.e. on X, and its sum $f_0(x)$ is a measurable and a.e. finite on X. Moreover, we have

$$f_0(x) = \lim_{k\to+\infty} \left(f_{n_1}(x) + \sum_{j=2}^{k} (f_{n_j}(x) - f_{n_{j-1}}(x)) \right) = \lim_{k\to+\infty} (f_{n_k}(x)).$$

Thus, the subsequence $(f_{n_k})_{k=1}^{\infty}$ converges to $f_0(x)$ a.e. on X.

Put now $m = n_k$ in the inequality (7.3.4). Since

$$|f_{n_k}(x) - f_n(x)|^p \underset{k\to\infty}{\longrightarrow} |f_0(x) - f_n(x)|^p \qquad \text{a.e. on} \quad X,$$

by Corollary 5.2.8

$$\int_X |f_0(x) - f_n(x)|^p d\mu < \varepsilon^p, \qquad n \geqslant n_0. \tag{7.3.6}$$

Now we prove that $f_0 \in L_p(X)$ and that $f_n \xrightarrow{L_p} f_0$. Indeed, for any number $n \geqslant n_0$, by the Minkowski inequality (7.1.7) and by (7.3.6), we have

$$\left(\int_X |f_0(x)|^p d\mu \right)^{\frac{1}{p}} = \left(\int_X |(f_0(x) - f_n(x)) + f_n(x)|^p d\mu \right)^{\frac{1}{p}} \leqslant$$

$$\leqslant \left(\int_X |f_0(x) - f_n(x)|^p d\mu \right)^{\frac{1}{p}} + \left(\int_X |f_n(x)|^p d\mu \right)^{\frac{1}{p}} \leqslant \varepsilon + M_n < +\infty,$$

where $M_n = \left(\int_X |f_n(x)|^p d\mu \right)^{\frac{1}{p}}$. Thus, $f_0 \in L_p(X)$.

Now from (7.3.6) it follows that

$$\|f_n - f_0\|_p < \varepsilon, \qquad n \geqslant n_0,$$

which means that the sequence $(f_n)_{n=1}^{\infty}$ converges to f_0 in p^{th} mean.

So, we get that a Cauchy sequence in L_p converges, as required. $\square$

Note that the condition $\mu X < +\infty$ was used only to establish that any Cauchy sequence in $L_p(X)$, $p > 1$, is a Cauchy sequence in $L_1(X)$. This means that we can omit this condition for the space $L_1(X)$.

Corollary 7.3.10. *Let X be the space $\mathbb{R}^n$ or any subset of $\mathbb{R}^n$ with finite or infinite measure. Then the space $L_1(X)$ is complete w.r.t. the norm $\|\cdot\|_1$.*

From the proof of Theorem 7.3.9 we obtain the following fact.

Corollary 7.3.11. *Let X be the space $\mathbb{R}^n$ or any subset of $\mathbb{R}^n$ with finite or infinite measure. If $(f_n)_{n=1}^{\infty}$ converges to f in $L_1(X)$, then there exists a subsequence $(f_{n_k})_{k=1}^{\infty}$ such that*

$$f_{n_k} \xrightarrow{a.e.} f \quad on \quad X.$$

7.4 Relations between different types of convergence

Let $(X, \mathcal{M}, \mu)$ be a measure space with $\mu X < +\infty$, and let $(f_n)_{n=1}^{\infty}$ be a sequence of measurable functions on X, and $f_0 \in S(X)$. We know the following kinds of convergence.

1) Uniform convergence on X,

2) Convergence a.e. on X (pointwise convergence),

3) Convergence in the space L_p (convergence in p^{th} mean),

4) Convergence in measure.

Let us consider interrelations between these kinds of convergence.

I. *Uniform convergence implies all other kinds of convergence.*

a) If $(f_n(x))_{n=1}^{\infty}$ converges to $f_0(x)$ uniformly on X, then it converges to $f_0(x)$ for any $x \in X$, so it converges pointwisely (almost everywhere) on X.

b) By definition of the uniform convergence, for any $\varepsilon > 0$ there exists a number $n_0 \in \mathbb{N}$ such that for any $n \geqslant n_0$ and for any $x \in X$ the following inequality holds

$$|f_n(x) - f_0(x)| < \varepsilon.$$

This implies

$$\|f_n - f_0\|_p = \left(\int_X |f_n(x) - f_0(x)|^p \right)^{\frac{1}{p}} < \varepsilon(\mu X)^{\frac{1}{p}},$$

so $f_n \xrightarrow{L_p} f_0$.

c) From the definition of the uniform convergence we get

$$X(|f_n - f_0| \geqslant \varepsilon) = \varnothing$$

for any $n \geqslant n_0$. Therefore, $f_n \xrightarrow{\mu} f_0$.

However, convergence in measure, converges in p^{th} mean, and point-wise convergence do not imply the uniform convergence, generally speaking.

Example 7.4.1. Let $X = [0, 1]$, and μ be the Lebesgue measure. Consider the sequence $f_n(x) = x^n$.

This sequence converges to $f_0(x) \equiv 0$ a.e. on X (everywhere but the point $x = 1$). Moreover, for any $p \geqslant 1$,

$$\|f_n - f_0\|_p = \left(\int\limits_X |x^n - 0|^p \right)^{\frac{1}{p}} = \left(\frac{1}{np+1} \right)^{\frac{1}{p}} \underset{n\to\infty}{\longrightarrow} 0$$

so $f_n \xrightarrow{L_p} f_0$. Additionally, for any $0 < \sigma < 1$, one has

$$\mu X(|f_n - f_0| \geqslant \sigma) = 1 - \sqrt[n]{\sigma} \underset{n\to\infty}{\longrightarrow} 0,$$

therefore, $f_n \xrightarrow{\mu} f_0$. However, $f_n(x)$ does not converge to $f_0(x)$ uniformly on X, since

$$\sup_{x\in[0,1]} |f_n(x) - f_0(x)| = \sup_{x\in[0,1]} |x^n| = 1 \underset{n\to\infty}{\not\longrightarrow} 0.$$

II. *Pointwise convergence and convergence in p^{th} mean are incomparable.*

Example 7.4.2. Let $X = [0, 1]$, and μ be the Lebesgue measure, and let

$$f_n(x) = \begin{cases} n^{\frac{1}{p}}, & 0 \leqslant x \leqslant \dfrac{1}{n}, \\ 0, & \dfrac{1}{n} < x \leqslant 1. \end{cases}$$

Clearly, $f_n \xrightarrow{a.e.} f_0$ where $f_0(x) \equiv 0$. However,

$$\|f_n - f_0\|_p = \left(\int\limits_X |f_n(x) - f_0(x)|^p d\mu \right)^{\frac{1}{p}} = \left(\int\limits_0^{\frac{1}{n}} n dx \right)^{\frac{1}{p}} = 1 \underset{n\to\infty}{\not\longrightarrow} 0,$$

so $f_n \not\to f_0$ in $L_p(X)$.

Example 7.4.3. Let $X = [0, 1]$, and μ be the Lebesgue measure, and let

$$\varphi_{km}(x) = \begin{cases} 1, & x \in \left[\dfrac{m-1}{k}, \dfrac{m}{k} \right], \\ 0, & x \notin \left[\dfrac{m-1}{k}, \dfrac{m}{k} \right], \end{cases}$$

where $k = 1, 2, \ldots$, and $m = 1, 2, \ldots, k$. Let us represent the functions $\phi_{l,j}$ as one sequence as follows:

$$f_1(x) = \varphi_{11}(x), \quad f_2(x) = \varphi_{21}(x), \quad f_3(x) = \varphi_{22}(x), \quad f_4(x) = \varphi_{31}(x), \quad f_5(x) = \varphi_{32}(x), \ \ldots,$$

so

$$f_n(x) = \varphi_{l,j}(x), \qquad n = j + \sum_{k=1}^{l-1} k, \qquad l = 1, 2, \ldots, \quad j = 1, 2, \ldots, l.$$

The sequence f_n converges in p^{th} mean on $[0, 1]$ to the function $f_0(x) \equiv 0$, since

$$\|f_n - f_0\|_p = \left(\int\limits_{[0,1]} |f_n(x) - f_0(x)|^p d\mu \right)^{\frac{1}{p}} = \left(\int\limits_{\frac{m-1}{k}}^{\frac{m}{k}} 1 \cdot dx \right)^{\frac{1}{p}} = \left(\frac{1}{k} \right)^{\frac{1}{p}} \underset{n\to\infty}{\longrightarrow} 0.$$

Obviously, $k \to \infty \iff n \to \infty$. Thus, the sequence f_n converges in p^{th} mean on $[0, 1]$ to the function $f_0(x) \equiv 0$, but it does not converge at any point of the interval $[0, 1]$ as it was established in Example 4.3.11.

III. *Pointwise convergence implies convergence in measure. Converse is not true, generally speaking.*

This was established in Theorem (4.3.10) and Example 4.3.11.

IV. *Convergence in p^{th} mean with exponent p_2 implies convergence in p^{th} mean with exponent p_1 for any $1 \leqslant p_1 < p_2$.*

This follows from the estimate (7.3.3) proved in Lemma 7.3.8.
The converse statement is not true, generally speaking, as the following example shows.

Example 7.4.4. Let $X = [0,1]$, and μ be the Lebesgue measure, and let

$$f_n(x) = \begin{cases} \left(\dfrac{n}{\ln n}\right)^{\frac{1}{p}}, & 0 \leqslant x \leqslant \dfrac{1}{n}, \\ 0, & \dfrac{1}{n} < x \leqslant 1, \end{cases}$$

for some $p \geqslant 1$. Put $f_0(x) \equiv 0$.
Since we have

$$\|f_n - f_0\|_p = \left(\int\limits_{[0,1]} |f_n(x)|^p d\mu \right)^{\frac{1}{p}} = \left(\int\limits_0^{\frac{1}{n}} \frac{n}{\ln n} dx \right)^{\frac{1}{p}} = \left(\frac{1}{\ln n}\right)^{\frac{1}{p}} \underset{n\to\infty}{\longrightarrow} 0,$$

so $f_n \xrightarrow{L_p} f_0$, but $f_n(x) \not\to f_0$ in p_1^{th} mean for any $p_1 > p$, because

$$\|f_n - f_0\|_{p_1} = \left(\int\limits_0^{\frac{1}{n}} \left(\frac{n}{\ln n}\right)^{\frac{p_1}{p}} dx \right)^{\frac{1}{p_1}} = n^{\frac{p_1 - p}{p_1 p}} \cdot \left(\frac{1}{\ln n}\right)^{\frac{1}{p}} \underset{n\to\infty}{\longrightarrow} +\infty.$$

V. *Convergence in p^{th} mean for any $p \geqslant 1$ implies convergence in measure.*

Since convergence in p^{th} mean for any $p \geqslant 1$ implies convergence in mean (that is, with the exponent $p = 1$), it is sufficient to show that $f_n \xrightarrow{L_1} f_0$ implies $f_n \xrightarrow{\mu} f_0$.

So let $f_n \xrightarrow{L_1} f_0$. For any $\sigma > 0$, $\delta > 0$, and $\varepsilon = \sigma\delta$, there exists a natural number $n_0 = n_0(\varepsilon)$ such that

$$\|f_n - f_0\|_1 = \int\limits_X |f_n(x) - f_0(x)| d\mu < \varepsilon.$$

By Chebyshev's inequality (see the property 7) of integral of nonnegative functions), one has for any $n \geqslant n_0$

$$\mu X(|f_n - f_0| \geqslant \sigma) \leqslant \frac{1}{\sigma} \int\limits_X |f_n(x) - f_0(x)| d\mu < \frac{\varepsilon}{\sigma} = \delta,$$

so $f_n \xrightarrow{\mu} f_0$.

The following example shows that converse is not true.

Example 7.4.5. Let $X = [0,1]$ and μ be the Lebesgue measure. Suppose that

$$f_n(x) = \begin{cases} n, & 0 \leqslant x \leqslant \dfrac{1}{n}, \\ 0, & \dfrac{1}{n} < x \leqslant 1. \end{cases}$$

The sequence f_n converges to the function $f_0(x) \equiv 0$ in measure on X, since for any $\sigma > 0$

$$\mu X(|f_n - f_0| \geqslant \sigma) \leqslant \frac{1}{n}.$$

However, for any $n \in \mathbb{N}$

$$\|f_n - f_0\|_1 = \int\limits_X |f_n(x) - f_0(x)| d\mu = 1,$$

so $f_n \not\longrightarrow 0$ in p^{th} mean for any $p \geqslant 1$.

7.4.1 Approximation of integrable functions

We say that a family $\mathcal{G}$ of integrable functions is dense in $L_p(X)$, $p \geqslant 1$, if for any $f \in L_p(X)$ and $\varepsilon > 0$, there exists $g \in \mathcal{G}$ so that $\|f - g\|_p < \varepsilon$. Fortunately, we are familiar with many families that are dense in $L_p(X)$, and we describe some in the theorem that follows. These are useful when one is faced with the problem of proving some fact or identity involving integrable functions. In this situation a general principle applies: the result is often easier to prove for a more restrictive class of functions (like the ones in the theorem below), and then a density (or limiting) argument yields the result in general.

Theorem 7.4.6. *Let X be a compact set in $\mathbb{R}^n$, and let $F \subset X$ be closed. If ψ is defined and continuous on F, then there exists a function $\varphi(x)$ defined on X satisfying the following properties*

a) $\varphi(x)$ *is continuous on* X*;*

b) $\varphi(x) = \psi(x)$ *for any* $x \in F$*;*

c) $\max |\varphi(x)| = \max |\psi(x)|$.

Proof. For the case $X = [a, b]$, a proof can be found in the book [8, Chapter 4, § 4, Lemma 2]. A similar proof for arbitrary compact set $X \subset \mathbb{R}^n$ is more difficult.

Another way (more topological) is to take the function

$$\varphi(x) = \begin{cases} \inf\limits_{y \in F} \left(\psi(y) + \dfrac{\rho(x, y)}{\rho(x, F)} - 1 \right), & x \in X \setminus F, \\ \psi(x), & x \in F, \end{cases} \tag{7.4.1}$$

which satisfies all the conditions of the theorem. Here $\rho(x, y) = \|x - y\|_2$ and $\rho(x, F) = \inf\limits_{y \in F} \rho(x, y)$. □

Problem. Prove that the function (7.4.1) is continuous on X.

From Theorems 4.4.8 and 7.4.6 it immediately follows the following fact.

Theorem 7.4.7 (Luzin)**.** *Let $X \subset \mathbb{R}^n$ be a compact set. Then for any $f \in S(X)$ and for any $\varepsilon > 0$ there exists a function g continuous on X and such that $\mu X(f \neq g) < \varepsilon$.*

If $|f(x)| \leqslant M$, then $|g(x)| \leqslant M$, as well.

Corollary 7.4.8 (Frechet)**.** *Let $X \subset \mathbb{R}^n$ be a compact set. Then for any $f \in S(X)$ there exists sequence $(f_n)_{n=1}^\infty$ of continuous functions converging to f a.e. on X.*

Corollary 7.4.9 (Borel)**.** *Let $X \subset \mathbb{R}^n$ be a compact set. Then for any $f \in S(X)$ there exists sequence $(f_n)_{n=1}^\infty$ of continuous functions converging to f in measure on X.*

Now we can study function families dense in the space of all integrable functions.

Theorem 7.4.10. *Let $X \subset \mathbb{R}^n$ be a compact set. The following families of functions are dense in $L_1(X)$:*

(i) The bounded measurable functions.

(ii) The simple functions.

(iii) The step functions.

(iv) The continuous functions of compact support.

Here X is supposed to be the space $\mathbb{R}^n$ or a (bounded or unbounded) subset of the space $\mathbb{R}^n$.

Recall that the set $\mathrm{supp} f = \overline{\{x \in X : \ f(x) \neq 0\}}$ is called the *support* of the function f (note that support of a function is always a closed set by definition).

Proof. Suppose that $f \in L_1(X)$. By the absolute continuity of Lebesgue integral (see the property 10) of integral of measurable functions), for any $\varepsilon > 0$ there exists $\delta > 0$ such that

$$\int\limits_{X_\delta} |f| d\mu < \varepsilon$$

for any X_δ with $\mu X_\delta < \delta$.

Now by the property 4) of integral of measurable functions, we have $\mu X(f = \pm\infty) = 0$. By Theorem 4.4.4, for any $\delta > 0$ there exists a bounded function $g \in S(X)$ such that $\mu X(f \neq g) < \delta$, and we set $g(x) = 0$ for any $x \in X(f \neq g)$. So finally we get

$$\int\limits_X |f - g| d\mu = \int\limits_{X(f \neq g)} |f| d\mu < \varepsilon.$$

Thus, $\|f - g\|_1 < \varepsilon$, as required.

Let again f be an integrable function on X. We can always represent f as follows $f = f^+ - f^-$, where f^+ and f^- are nonnegative functions defined in Definition 4.2.3, and it now suffices to prove the theorem when $f \geqslant 0$.

For (ii), Theorem 4.4.2 guarantees the existence of a sequence $(h_k)_{k=1}^\infty$ of non-negative simple functions that increase to f pointwise. By the Lebesgue Dominated Convergence Theorem 5.3.13, we then have that

$$\int\limits_X f(x) d\mu = \lim_{n \to \infty} \int\limits_X h_n(x) d\mu,$$

so

$$\lim_{n \to \infty} \int\limits_X (f(x) - h_n(x)) d\mu = 0,$$

that is equivalent to

$$\|f - h_n\|_1 \to 0 \qquad \text{as} \qquad n \to \infty,$$

since $f - h_n \geqslant 0$ for any $n \in \mathbb{N}$. Thus, there are simple functions that are arbitrarily close to f in the L_1 norm.

For (iii), we first note that by (ii) it suffices to approximate simple functions by step functions. Then, we recall that a simple function is a finite linear combination of characteristic functions of sets of finite measure, so it suffices to show that if A is such a measurable set, then there is a step function φ so that $\|\chi_A - \varphi\|_1$ is small. However, since A is measurable, then by Definition 3.4.1 for any $\varepsilon > 0$ there exists an elementary set B such that

$$\mu^*(A \Delta B) = \mu(A \Delta B) < \varepsilon.$$

The elementary set $B = \sum\limits_{l=1}^m K_l$ is a finite sum of bricks that we can always consider as almost disjoint (see Definition 3.5.8). Thus $\chi_A(x)$ and $\varphi(x) = \sum\limits_{k=1}^m \chi_{K_l}(x)$ differ at most on a set of measure ε, and as a result we find that $\|\chi_A - \varphi\|_1 < \varepsilon$.

By (iii), it suffices to establish (iv) when f is the characteristic function of a brick. In the one-dimensional case, where f is the characteristic function of an interval $[a, b]$, we may choose a continuous piecewise linear function g defined by

$$g(x) = \begin{cases} 1, & x \in [a, b], \\ 0, & x \notin [a - \varepsilon, b + \varepsilon], \end{cases}$$

and with g linear on the intervals $[a - \varepsilon, a]$ and $[b, b + \varepsilon]$. Then $\|f - g\|_1 < \varepsilon$. In n dimensions, it suffices to note that the characteristic function of a rectangle is the product of characteristic functions of intervals. Then, the desired continuous function of compact support is simply the product of functions like g defined above. □

Remark 7.4.11. Note that there exist some other ways to approximate the characteristic function of measurable set by continuous functions with compact support. For further details, see [5, 4].

Remark 7.4.12. In fact, bounded measurable functions are dense in the space $L_p(X)$. To prove this it is sufficient to note that if $f \in L_p(X)$, then $f^p \in L_1(X)$, and use absolute continuity of Lebesgue integral for the function f^p, at the same time, approximating the function f by a bounded measurable function.

As an application of Theorem 7.4.10, we now examine how continuity properties of f are related to the way the translations f_c vary with c (see Definition 5.3.14). Note that for any given $x \in \mathbb{R}^n$, the statement that $f_c(x) \to f(x)$ as $c \to 0$ is the same as the continuity of f at the point x.

However, a general $f \in L_1(\mathbb{R}^n)$ may be discontinuous at every x, even when corrected on a set of measure zero. Nevertheless, there is an overall continuity that an arbitrary $f \in L_1(\mathbb{R}^n)$ enjoys, one that holds in the norm.

Proposition 7.4.13. *Let $f \in L_1(\mathbb{R}^n)$. Then*

$$\|f - f_c\|_1 \to 0 \quad \textit{as} \quad c \to 0.$$

The proof is a simple consequence of the approximation of integrable functions by continuous functions of compact support as given in Theorem 7.4.10.

Proof. By Theorem 7.4.10 for any $\varepsilon > 0$, one can find a continuous function g of compact support such that $\|f - g\|_1 < \varepsilon$. Now

$$f_c - f = (g_c - g) + (f_c - g_c) + (g - f).$$

Since $\|f_c - g_c\|_1 = \|f - g\|_1 < \varepsilon$ and since g is continuous of compact support we have that clearly

$$\|g_c - g\|_1 = \int\limits_{\mathbb{R}^n} |g(x - c) - g(x)| dx \to 0 \quad \text{as} \quad c \to 0.$$

So if $|c| < \delta$, where δ is sufficiently small, then $\|g_c - g\|_1 < \varepsilon$, and as a result $\|f_c - f\|_1 < 3\varepsilon$, whenever $|c| < \delta$. □

The results above for $L_1(\mathbb{R}^n)$ lead immediately to an extension in which $\mathbb{R}^n$ can be replaced by any fixed subset X of positive measure. In fact, if X is such a subset, we can define $L_1(X)$ and carry out the arguments that are analogous to $L_1(\mathbb{R}^n)$. Better yet, we can proceed by extending any function f on X by setting $\widetilde{f} = f$ on X and $\widetilde{f} = 0$ on $\mathbb{R}^n \setminus X$, and defining $\|f\|_{L_1(X)} = \|\widetilde{f}\|_{L_1(\mathbb{R}^n)}$. Moreover, for measure spaces $(X, \mathcal{M}, \mu)$ with compact X, the statement of Theorem 7.4.10 can be partially established for the spaces $L_p(X)$ with $p > 1$.

Theorem 7.4.14. *The following families of functions are dense in $L_p(X)$, $p \geqslant 1$, where $X \subset \mathbb{R}^n$ is a compact set:*

(i) The bounded measurable functions.

(ii) The continuous functions.

Proof. The proof of the statement (i) is similar to the one in Theorem 7.4.10. Indeed, if $f \in L_p(X)$, then by the absolute continuity of Lebesgue integral (see the property 10) of integral of measurable functions), for any $\varepsilon > 0$ there exists $\delta > 0$ such that

$$\int_{X_\delta} |f|^p d\mu < \varepsilon^p$$

for any X_δ with $\mu X_\delta < \delta$.

Now by the property 4) of integral of measurable functions, we have $\mu X(f = \pm\infty) = 0$. By Theorem 4.4.4, for any $\delta > 0$ there exists a bounded function $g \in S(X)$ such that $\mu X(f \neq g) < \delta$, and we set $g(x) = 0$ for any $x \in X(f \neq g)$. So finally we get

$$\int_X |f - g|^p d\mu = \int_{X(f \neq g)} |f|^p d\mu < \varepsilon^p.$$

Thus, $\|f - g\|_1 < \varepsilon$, as required.

Now it is enough to approximate a bounded measurable function in the norm L_p. Let $f \in S(X)$ be bounded, so $|f(x)| \leqslant M$. By Theorem 7.4.7, for any $\varepsilon > 0$ there exists a continuous function g s.t.

$$\mu X_0 = \mu X(f \neq g) < \frac{\varepsilon^p}{(2M)^p}.$$

Then we have

$$\|f - g\|_p^p = \int_X |f - g|^p d\mu = \int_{X_0} |f - g|^p d\mu \leqslant (2M)^p \cdot \mu X_0 < \varepsilon^p.$$

□

From this theorem and from Weierstrass' theorems it easy to obtain the following facts.

Corollary 7.4.15. *The families of polynomials and trigonometric polynomials are dense in $L_p[-\pi, \pi]$, $p \geqslant 1$.*

Recall that trigonometric polynomials are functions of the form

$$T_n(x) = \frac{a_0}{2} + \sum_{m=1}^{n} (a_m \cos mx + b_m \sin mx).$$

7.5 Basic theory of inner product spaces

In this section we introduce a more narrow class of linear normed spaces, namely, the inner product spaces.

Let X be a linear space over the field[1] $\mathbb{F} = \mathbb{C}$.

Definition 7.5.1. The function $(\cdot, \cdot) : X \times X \mapsto \mathbb{C}$ is called *inner (scalar) product* of the elements of the space X if it possesses the following properties:

1) for any $x \in X$, $(x, x) \geqslant 0$ $(= 0 \iff x = 0)$;

2) $(x, y) = \overline{(y, x)}$ for any[2] $x, y \in X$;

3) $(\alpha x + \beta y, z) = \alpha(x, z) + \beta(y, z)$ for any $x, y, z \in X$ and $\alpha, \beta \in \mathbb{C}$.

[1] In fact, the theory discussed here can be easily transferred to the case $\mathbb{F} = \mathbb{R}$.
[2] In the case $\mathbb{F} = \mathbb{R}$, this property looks as $(x, y) = (y, x)$

From the properties of the inner product it follows that[3]

$$(x, \alpha y) = \overline{(\alpha y, x)} = \overline{\alpha(y, x)} = \overline{\alpha} \cdot \overline{(y, x)} = \overline{\alpha}(x, y).$$

Definition 7.5.2. A linear space X over a field $\mathbb{F}$ with an inner product is called *Euclidean* if $\mathbb{F} = \mathbb{R}$, and is called *unitary* if $\mathbb{F} = \mathbb{C}$.

Let X be a Euclidean (unitary) space. For any $x \in X$, introduce the following number

$$\|x\| = \sqrt{(x, x)}. \tag{7.5.1}$$

This number is denoted as norm. And we have to prove that this is a norm, indeed. To do this, we need the following inequality:

$$|(x, y)| \leqslant \|x\| \cdot \|y\|, \qquad \forall x, y \in X. \tag{7.5.2}$$

This inequality is called *Cauchy–(Bunyakovsky)–Schwarz* inequality. For finite sums it was proved by Cauchy in 1821. For integrals, it was proved by Bunyakovsky in 1859 and by Schwarz in 1888.

To prove the inequality (7.5.2), let us represent the inner product as follows $(x, y) = |(x, y)|e^{i\theta}$. Then for any $t \in \mathbb{R}$ we have

$$(te^{-i\theta}x + y, te^{-i\theta}x + y) = (te^{-i\theta}x, te^{-i\theta}x) + (te^{-i\theta}x, y) + (y, te^{-i\theta}x) + (y, y) =$$

$$= te^{i\theta} \cdot te^{-i\theta} \cdot (x, x) + te^{-i\theta}(x, y) + t^{i\theta}(y, x) + \|y\|^2 = t^2\|x\|^2 + \|y\|^2 + t \cdot |(x, y)| + t \cdot |(y, x)| =$$

$$= t^2\|x\|^2 + 2t \cdot |(x, y)| + \|y\|^2 \geqslant 0.$$

The quadratic polynomial is nonnegative if, and only if, its discriminant is non-positive. In our case, the discriminant has the form $|(x, y)|^2 - \|x\|^2 \cdot \|y\|^2$, and its non-positivity implies (7.5.2).

Prove now that the number (7.5.1) is a norm.

1) $\|x\| \geqslant 0$ $(= 0 \Longleftrightarrow x = 0)$.

   This property follows from the first property of inner products.

2) *The triangle inequality*

   By the Cauchy–Bunyakovsky-Schwarz inequality, we have

$$\|x + y\|^2 = (x + y, x + y) = \|x\|^2 + (x, y) + \overline{(x, y)} + \|y\|^2 = \|x\|^2 + 2\operatorname{Re}(x, y) + \|y\|^2 \leqslant$$

$$\leqslant \|x\|^2 + 2|(x, y)| + \|y\|^2 \leqslant \|x\|^2 + 2\|x\| \cdot \|y\| + \|y\|^2 = (\|x\| + \|y\|)^2.$$

   Consequently,

$$\|x + y\| \leqslant \|x\| + \|y\|, \tag{7.5.3}$$

   as required.

3) $\|\alpha x\| = \sqrt{(\alpha x, \alpha x)} = \sqrt{\alpha \cdot \overline{\alpha}(x, x)} = \sqrt{|\alpha| \cdot \|x\|^2} = |\alpha| \cdot \|x\|$.

Thus, we have that any Euclidean (unitary) space is a linear normed space.

Proposition 7.5.3. *The inner product of a Euclidean (unitary) space X is a continuous function of its variables.*

Proof. Indeed, let $x_n \to x$ and $y_n \to y$ by norm in X. Then by Theorem 7.2.7, $\|y_n\| \to \|y\|$ as $n \to \infty$, so there exists a constant $C > 0$ such that $\|y_n\| \leqslant C$ for all $n \in \mathbb{N}$. Thus, by Cauchy-Bunyakovsky-Schwarz inequality (7.5.2) we have

$$|(x_n, y_n) - (x, y)| \leqslant |(x_n, y_n) - (x, y_n)| + |(x, y_n) - (x, y)| = |(x_n - x, y_n)| + |(x, y_n - y)| \leqslant$$

$$\leqslant \|x_n - x\| \cdot \|y_n\| + \|y_n - y\| \cdot \|x\| \leqslant C\|x_n - x\| + \|y_n - y\| \cdot \|x\| \underset{n\to\infty}{\longrightarrow} 0.$$

Thus, $\lim\limits_{n\to\infty} (x_n, y_n) = (x, y)$, as required. □

[3] If $\mathbb{F} = \mathbb{R}$, then we have $(x, \alpha y) = \alpha(x, y)$.

Definition 7.5.4. For any two non-zero elements x, y of a Euclidean space X, cosine of the angle α between x and y is defined by the following formula

$$\cos\alpha = \frac{(x, y)}{\|x\| \cdot \|y\|} \ .$$

It is clear that if $\alpha = \dfrac{\pi}{2}$, then $(x, y) = 0$. Thus, we can introduce for both Euclidean and unitary spaces the following notion.

Definition 7.5.5. Two non-zero elements x and y of a Euclidean (unitary) space are called *orthogonal* if $(x, y) = 0$. We denote this as $x \perp y$

For orthogonal elements of Euclidean and inner product spaces there exists the Pythagorean theorem.

Theorem 7.5.6 (Pythagorean theorem)**.** *If $x \perp y$, then*

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

Proof. Indeed, we have

$$\|x + y\|^2 = (x + y, x + y) = \|x\|^2 + (x, y) + (y, x) + \|y\|^2 = \|x\|^2 + \overline{(x, y)} + \|y\|^2 = \|x\|^2 + \|y\|^2.$$

□

Definition 7.5.7. Let X be a Euclidean or unitary space. A system $(e_n)_{n=1}^{\omega}$ of elements of the space X is called *orthogonal* if $(e_n, e_m) = 0$ whenever $n \neq m$. Here ω can be finite or infinite (if the space X is infinite-dimensional).

The system $(e_n)_{n=1}^{\omega}$ is called *orthonormal* if

$$(e_n, e_m) = \begin{cases} 1 & \text{if} \quad m = n, \\ 0 & \text{if} \quad m \neq n. \end{cases}$$

Definition 7.5.8. Let $(e_n)_{n=1}^{\omega}$, $1 \leqslant \omega \leqslant \infty$, be an orthonormal system of elements of a Euclidean (unitary) space X. For any $x \in X$, then numbers

$$c_n(x) = (x, e_n), \qquad n = 1, 2, \ldots, \omega,$$

are called *Fourier coefficients* of the element x with respect to the system $(e_n)_{n=1}^{\omega}$.

Theorem 7.5.9 (On projection)**.** *Let $(e_k)_{k=1}^{n}$ be an orthonormal system of elements of a Euclidean (unitary) space X. Suppose that $x \in X$ and*

$$y = \sum_{k=1}^{n} c_k(x) e_k.$$

Then

$$x - y \perp y.$$

Proof. Note first that for any $k = 1, \ldots, n$,

$$c_k(x) = c_k(y),$$

so that for $1 \leqslant k \leqslant n$ we have

$$(x - y, e_k) = (x, e_k) - (y, e_k) = c_k(x) - c_k(y) = 0.$$

Therefore,

$$(x - y, y) = (x - y, \sum_{k=1}^{n} c_k(x) e_k) = \sum_{k=1}^{n} \overline{c_k(x)} (x - y, e_k) = 0,$$

as required. □

Corollary 7.5.10. *If $(e_k)_{k=1}^n$ is an orthonormal system of elements of a Euclidean (unitary) space X, then*

$$\left\|\sum_{k=1}^n a_k e_k\right\|^2 = \sum_{k=1}^n |a_k|^2.$$

Proof.

$$\left\|\sum_{k=1}^n a_k e_k\right\|^2 = \left\|a_1 e_1 + \sum_{k=2}^n a_k e_k\right\|^2 = \|a_1 e_1\|^2 + \left\|\sum_{k=2}^n a_k e_k\right\|^2 = |a_1|^2 \cdot \|e_1\|^2 + \left\|\sum_{k=2}^n a_k e_k\right\|^2 =$$

$$= |a_1|^2 + |a_2|^2 + \cdots + |a_n|^2 = \sum_{k=1}^n |a_k|^2.$$

□

Theorem 7.5.11 (Bessel's inequality)**.** *Let X be an infinitely dimensional Euclidean or unitary space. If $(e_n)_{n=1}^\infty$ is an orthonormal system of elements of the space X, then*

$$\sum_{n=1}^\infty |c_n(x)|^2 \leqslant \|x\|^2 \qquad \forall x \in X. \tag{7.5.4}$$

Proof. Let $y = \sum\limits_{n=1}^m c_n(x) e_n$. Then by Theorems 7.5.6 and 7.5.9, we have

$$\|x\|^2 = \|x-y\|^2 + \|y\|^2. \tag{7.5.5}$$

Since $\|x-y\|^2 \geqslant 0$, from (7.5.5) we obtain

$$\sum_{n=1}^m |c_n(x)|^2 = \|y\|^2 \leqslant \|x\|^2 \qquad \forall m \in \mathbb{N}.$$

Tending now m to $+\infty$ we get (7.5.4), as required. □

By the Bessel inequality (7.5.4) for any $x \in X$ the series $\sum\limits_{n=1}^\infty |c_n(x)|^2$ converges, thus we have the following corollary.

Corollary 7.5.12 (Riemann–Lebesgue's Lemma)**.** *Let X be an infinitely dimensional Euclidean or unitary space, and let $(e_n)_{n=1}^\infty$ be an orthonormal system of elements of the space X. Then for any $x \in X$ its Fourier coefficients $c_n(x)$ satisfy the following*

$$c_n(x) \underset{n\to\infty}{\longrightarrow} 0.$$

Theorem 7.5.13. *Let $(e_k)_{k=1}^n$ is an orthonormal system of elements of a unitary space X. Suppose that $x \in X$, and $c_k(x)$, $k = 1, \ldots, n$, are its Fourier coefficients with respect to the system $(e_k)_{k=1}^n$. Then the functional*

$$\Delta(\lambda_1, \lambda_2, \ldots, \lambda_n) = \left\|x - \sum_{k=1}^n \lambda_k e_k\right\|, \qquad \lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{C}$$

achieves its minimal value for

$$\lambda_k = c_k(x), \qquad k = 1, \ldots, n.$$

This property is called the *minimization property* of Fourier coefficients.

Proof. Let

$$y = \sum_{k=1}^{n} c_k(x) e_k.$$

Since $(x - y)$ By Corollary 7.5.10 and by Theorem 7.5.6, we have

$$\left\| x - \sum_{k=1}^{n} \lambda_k e_k \right\|^2 = \left\| x - y + \sum_{k=1}^{n} (c_k(x) - \lambda_k) e_k \right\|^2 = \|x - y\|^2 + \left\| \sum_{k=1}^{n} (c_k(x) - \lambda_k) e_k \right\|^2 =$$

$$= \|x - y\|^2 + \sum_{k=1}^{n} |\lambda_k - c_k|^2.$$

It is clear now that

$$\min_{\lambda_1,\ldots,\lambda_n} \left\| x - \sum_{k=1}^{n} \lambda_k e_k \right\|^2 = \|x - y\|^2 = \|x\|^2 - \|y\|^2 = \|x\|^2 - \sum_{k=1}^{n} |c_k(x)|^2.$$

□

The geometric meaning of this theorem is that the shortest distance from a point to its orthogonal projection to a subspace is the perpendicular from the point to the projection.

7.5.1 Gram-Schmidt orthogonalization method

Let X be an infinitely dimensional unitary (Euclidean) space, and let $(x_n)_{n=1}^{\infty} \subset X$ be a countable system of linearly independent elements of X. We show that in X there exists a countable orthonormal system $(e_n)_{n=1}^{\infty}$ of elements of X such that for any $n \in \mathbb{N}$ the element e_n is a linear combination of elements x_1, x_2, …, x_n. Our proof of this fact is constructive.

Since the system $(x_n)_{n=1}^{\infty}$ is linearly independent, $\|x_1\| \neq 0$. Set now $e_1 := \dfrac{x_1}{\|x_1\|}$. Then we set $y_2 := \alpha_2^{(1)} e_1 - x_2$, and find $\alpha_2^{(1)}$ from the condition $(e_1, y_2) = 0$. Then we have

$$\alpha_2^{(1)} = (x_2, e_1).$$

Since the system $(x_n)_{n=1}^{\infty}$ is linearly independent, $y_2 \neq 0$, so we can set

$$e_2 := \frac{y_2}{\|y_2\|},$$

and $(e_1, e_2) = 0$. Furthermore, let $y_3 = \alpha_3^{(1)} e_1 + \alpha_3^{(2)} e_2 - x_3$. We find $\alpha_3^{(1)}$ and $\alpha_3^{(2)}$ from the conditions $(y_3, e_1) = (y_3, e_2) = 0$. Obviously,

$$\alpha_3^{(1)} = (x_3, e_1), \qquad \alpha_3^{(2)} = (x_3, e_2),$$

and $y_3 \neq 0$, so we set

$$e_3 := \frac{y_3}{\|y_3\|}.$$

Let we already have an orthonormal system $(e_n)_{n=1}^{m}$ such that each e_n is a linear combination of elements x_1, x_2, …, x_n. We set $y_{n+1} = \alpha_{n+1}^{(1)} e_1 + \alpha_{n+1}^{(2)} e_2 + \cdots + \alpha_{n+1}^{(n)} e_n - x_n$, and find the coefficients $\alpha_{n+1}^{(k)}$, $k = 1, \ldots, n$, from the conditions

$$(y_{n+1}, e_k) = 0, \quad k = 1, \ldots, n.$$

Then
$$\alpha_{n+1}^{(k)} = (x_{n+1}, e_k), \quad k = 1, \dots, n,$$
and $y_{n+1} \neq 0$, so
$$e_{n+1} := \frac{y_{n+1}}{\|y_{n+1}\|}.$$
Continuing in the same manner, we will get a countable orthonormal system $(e_n)_{n=1}^{\infty}$ of elements of X such that for any $n \in \mathbb{N}$ the element e_n is a linear combination of elements $x_1, x_2, \dots, x_n$. This process is called the *Gram–Schmidt orthogonalization* process.

7.5.2 Orthonormal systems in infinitely dimensional spaces

Let again X be an infinitely dimensional unitary (Euclidean) space, and let $(e_n)_{n=1}^{\infty}$ be an orthonormal system of elements of X (that always exists as we established above).

Definition 7.5.14. For any $x \in X$ the series
$$\sum_{n=1}^{\infty} c_n(x) e_n,$$
where $c_n(x) = (x, e_n)$, $n \in \mathbb{N}$, is called the *Fourier series* of the element x.

From the identity (7.5.5) it follows that the Fourier series of any element $x \in X$ converges to x (in the norm of the space X induced by the inner product) *if and only if* the following identity holds:
$$\|x\|^2 = \sum_{n=1}^{\infty} |c_n(x)|^2. \tag{7.5.6}$$
This identity is called *Parseval's* identity.

Definition 7.5.15. An orthonormal system $(e_n)_{n=1}^{\infty}$ of elements of X is called *closed (in Steklov's sense)* if for any $x \in X$ Parseval's identity (7.5.6) holds.

Thus, we obtain the following fact.

Theorem 7.5.16. *Any element of an infinitely dimensional unitary (Euclidean) space X can be expanded into its Fourier series w.r.t. an orthonormal system $(e_n)_{n=1}^{\infty} \subset X$ if, and only if, the system $(e_n)_{n=1}^{\infty}$ is closed (in Steklov's sense).*

Let now
$$x_\alpha, \quad \alpha \in A,$$
be a system of linearly independent elements of a linear *normed* space X over a field $\mathbb{F}$. Here A is the set of indices (finite, countable or of higher cardinality).

Definition 7.5.17. The collection of all elements of X of the form $\sum\limits_{\alpha \in A} \lambda_\alpha x_\alpha$, $\lambda_k \in \mathbb{F}$, is called the *linear span* of this system, and denoted $\operatorname*{span}\limits_{\alpha \in A} (x_\alpha)$.

Definition 7.5.18. The system[4] $(x_\alpha)_{\alpha \in A}$ of elements of a linear normed space X is called *complete* in X if its linear span is dense in X, that is, if for any $x \in X$ the following conditions hold:

> for any $\varepsilon > 0$ there exist elements $x_{\alpha_1}, x_{\alpha_2}, \dots, x_{\alpha_n}$ of the system $(x_\alpha)_{\alpha \in A}$ and numbers $\lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{F}$ such that
> $$\left\| x - \sum_{k=1}^{n} \lambda_k x_{\alpha_k} \right\| < \varepsilon.$$

[4] It is supposed to be linearly independent.

Example 7.5.19. By Theorem 7.4.10 the following systems of elements of $L_p[a,b]$, $p>1$, are complete in $L_p[a,b]$: bounded measurable functions, step functions, simple functions, continuous functions of compact support.

Theorem 7.5.20. *Let X be in infinitely dimensional unitary (Euclidean) space, and let $(e_n)_{n=1}^{\infty}$ be an orthonormal system of elements of X. If the system $(e_n)_{n=1}^{\infty}$ is complete, then every element $x \in X$ can be expanded into its Fourier series w.r.t. the system $(e_n)_{n=1}^{\infty}$.*

Proof. Since the system $(e_n)_{n=1}^{\infty}$ is complete, for any $x \in X$ and for any $\varepsilon > 0$ there exist numbers α_1, α_2, $\ldots$, $\alpha_{N_\varepsilon} \in \mathbb{F}$ such that

$$\left\| x - \sum_{n=1}^{N_\varepsilon} \alpha_n e_n \right\|^2 < \varepsilon.$$

From Theorem (7.5.13) and from (7.5.5), we have

$$\|x\|^2 - \sum_{n=1}^{N_\varepsilon} |(x,e_n)|^2 = \left\| x - \sum_{n=1}^{N_\varepsilon} (x,e_n) e_n \right\|^2 \leqslant \left\| x - \sum_{n=1}^{N_\varepsilon} \alpha_n e_n \right\|^2 < \varepsilon.$$

This means that

$$\|x\|^2 = \sum_{n=1}^{\infty} |(x,e_n)|^2$$

So the system $(e_n)_{n=1}^{\infty}$ is closed (in Steklov's sense), therefore, every element $x \in X$ can be expanded into its Fourier series w.r.t. to the system $(e_n)_{n=1}^{\infty}$. □

7.6 Hilbert spaces

This section is devoted to complete inner product spaces.

Definition 7.6.1. A Euclidean (unitary) space is called a real (complex) *Hilbert* space if it is complete with respect to the norm induced by the inner product (that is, defined by the formula (7.5.1)).

Example 7.6.2. The (complex) space l_2 is a (complex) Hilbert space with inner product

$$(x,y) = \sum_{n=1}^{\infty} x_n \overline{y}_n, \qquad \forall x, y \in l_2,$$

since l_2 is complete with respect to the norm

$$\|x\| = \left(\sum_{n=1}^{\infty} |x|^2 \right)^{\frac{1}{2}} = \sqrt{(x,x)},$$

as we proved in Example 7.2.10.

Example 7.6.3. Another example is the space $L_2[a,b]$ with inner product

$$(f,g) = \int_a^b f(x)\overline{g(x)}dx, \qquad \forall f, g \in L_2[a,b],$$

which is complete w.r.t. the inner product induced by this norm. We will study this space in detail in Section 7.6.5.

Theorem 7.6.4. *If $a = (a_n)_{n=1}^{\infty} \in l_2$, then for any orthonormal system $(e_n)_{n=1}^{\infty}$ of elements of a Hilbert space H, the series $\sum_{n=1}^{\infty} a_n e_n$ converges. Moreover, there exists an element $x \in H$ such that this series is the Fourier series of x.*

Proof. From the Pythagorean theorem 7.5.6 it follows that for any natural $p < m$,

$$\left\| \sum_{n=m}^{p} a_n e_n \right\|^2 = \sum_{n=m}^{p} |a_n|^2.$$

Since the series $\sum_{n=1}^{\infty} |a_n|^2$ converges (because $a \in l_2$), we obtain that the sequence of partial sums $S_m = \sum_{n=1}^{m} a_n e_n$ of the series $\sum_{n=1}^{\infty} a_n e_n$ is a Cauchy sequence:

$$\|S_p - S_m\| = \left\| \sum_{n=m}^{p} a_n e_n \right\| \to 0 \quad \text{as} \quad p, m \to \infty.$$

Since the space H is complete, the sequence S_n has a limit $x \in H$ which is the sum of the considered series

$$x = \sum_{n=1}^{\infty} a_n e_n$$

Furthermore, by Proposition 7.5.3, the inner product is a continuous functional of its variables, therefore,

$$(x, e_n) = \lim_{m \to \infty} \left(\sum_{k=1}^{m} a_k e_k, e_n \right) = \lim_{m \to \infty} \sum_{k=1}^{m} (a_k e_k, e_n) = a_n.$$

Thus, the series $\sum_{n=1}^{\infty} a_n e_n$ is the Fourier series of the element x, as required. □

Theorem 7.6.5. *Let H be a Hilbert space (real or complex). An orthonormal system $(e_n)_{n=1}^{\infty} \subset H$ is complete if, and only if, the only element of H orthogonal to any e_n, $n \in \mathbb{N}$, is the zero element.*

Proof. If the system $(e_n)_{n=1}^{\infty}$ is complete, then by Theorem 7.5.20 Parseval's identity (7.5.6) holds. So if $(x, e_n) = c_n(x) = 0$, $n \in \mathbb{N}$, then $\|x\| = 0$, so $x = 0$ by the first property of the inner product.

Conversely, suppose that the system $(e_n)_{n=1}^{\infty} \subset H$ is such that the only element of H orthogonal to all elements e_n is zero. Consider an element $x \in H$ and show that if a series

$$\sum_{n=1}^{\infty} a_n e_n$$

is the Fourier series of x, then it converges to x in the norm of the space H induced by its inner product.

Indeed, from Bessel's inequality (7.5.4) it follows that the series $\sum_{n=1}^{\infty} |a_n|^2$ converges. So by Theorem 7.6.4, there exists an element $y \in H$ such that

$$y = \sum_{n=1}^{\infty} a_n e_n,$$

and $\sum_{n=1}^{\infty} a_n e_n$ is the Fourier series for y. Thus, we have

$$(x - y, e_n) = 0, \qquad \forall n \in \mathbb{N},$$

so $x - y = 0$, as required. □

Remark 7.6.6. Some authors call an orthonormal system in a Hilbert space H complete if only if the zero element of H is orthogonal to all the elements of the system. Theorem 7.6.5 says that this definition of complete systems is equivalent to our Definition 7.6.5.

Proposition 7.6.7. *An orthonormal system $(e_n)_{n=1}^{\infty}$ of elements of a Hilbert space X is closed (in Steklov's sense) if, and only if, it is complete.*

Proof. Let $(e_n)_{n=1}^{\infty}$ be closed. From Parseval's identity (7.5.6) it immediately follows that if for some $x \in X$ we have $c_n(x) = 0$, $n \in \mathbb{N}$, then $\|x\| = 0$, so $x = 0$. Therefore, the system $(e_n)_{n=1}^{\infty}$ is complete by Theorem 7.6.5.

If $(e_n)_{n=1}^{\infty}$ is complete, then by Theorem 7.5.20 any element $x \in X$ can be expanded into its Fourier series. Now Theorem 7.5.16 implies that the system $(e_n)_{n=1}^{\infty}$ is closed. $\square$

The following theorem will allow us to study closed systems of functions in the space of integrable functions.

Theorem 7.6.8. *Let a set $L \subset H$ is dense in a Hilbert space H, and let an orthonormal system $(e_n)_{n=1}^{\infty} \subset L$ is closed in L. Then it is closed in H.*

Proof. Let $x \in H$. Consider the partial sums

$$S_n(x) = \sum_{k=1}^{n} c_k(x) e_k.$$

It is clear that $S_n(A_1x_1 + A_2x_2) = A_1S_n(x_1) + A_2S_n(x_2)$ for any $x_1, x_2 \in H$ and for any $A_1, A_2 \in \mathbb{F}$. Moreover, by Bessel's inequality (7.5.4) we have $\|S_n(x)\| \leqslant \|x\|$.

By assumption, for any $\varepsilon > 0$ there exists an element $y \in L$ such that $\|x - y\| < \dfrac{\varepsilon}{3}$. Thus, we obtain

$$\|x - S_n(x)\| \leqslant \|x - y\| + \|S_n(y) - S_n(x)\| + \|y - S_n(y)\|. \tag{7.6.1}$$

Note that by Bessel's inequality (7.5.4) and by assumption

$$\|S_n(y) - S_n(x)\| = \|S_n(y - x)\| \leqslant \|y - x\| < \frac{\varepsilon}{3}. \tag{7.6.2}$$

Moreover, by assumption the system $(e_n)_{n=1}^{\infty}$ is closed in L, so for the element y Parseval's identity (7.5.6) holds:

$$\|y\|^2 = \sum_{k=1}^{\infty} |c_k(y)|^2,$$

and by (7.5.5) we have

$$\|y - S_n(y)\|^2 = \|y\|^2 - \sum_{k=1}^{n} |c_k(y)|^2 < \frac{\varepsilon}{3} \tag{7.6.3}$$

for sufficiently large n. Now from (7.6.1)–(7.6.3) it follows that

$$\|x - S_n(x)\| < \varepsilon$$

for sufficiently large n. This means that the system $(e_n)_{n=1}^{\infty}$ is complete, so by Theorems 7.5.16 and 7.5.20 it is closed, as required. $\square$

7.6.1 Isomorphism of separable Hilbert space

Definition 7.6.9. Two Euclidean (unitary) spaces X_1 and X_2 are called *isomorphic* if there exists a bijection (one-to-one correspondence) $F : X_1 \mapsto X_2$ such that

$$(F(x), F(y))_2 = (x, y), \qquad x, y \in X_1,$$

where $(\cdot,\cdot)_1$ and $(\cdot,\cdot)_2$ are the inner products of the spaces X_1 and X_2, respectively. F is called an isomorphism.

Definition 7.6.10. A subset L' of a linear space L is called a subspace of L if L' is a linear space. A subset is called *closed* if for any sequence $(x_n)_{n=1}^{\infty} \subset L'$ convergent to an element $x \in L$, we have that $x \in L'$.

Definition 7.6.11. A real (complex) Hilbert space H is a *completion* of a Euclidean (unitary) space X if there is a subspace X' in H isomorphic to X and dense in H.

Theorem 7.6.12. *Any Euclidean (unitary) space has a completion.*

This is a particular case of the corresponding theorem for metric spaces and we leave a proof of this theorem for the class of "Functional Analysis".

Definition 7.6.13. A linear normed space X is called *separable* if there exists a countable complete system of elements of the space X.

Definition 7.6.14. A countable system $(e_n)_{n=1}^{\infty}$ of elements of a linear normed space X is called a *basis of the space* X if every element x of the space X has a unique expansion w.r.t. this system, that is, there exists a unique series sequence $(\lambda_n)_{n=1}^{\infty}$ such that

$$x = \sum_{n=1}^{\infty} \lambda_n e_n$$

Here the series converges to x w.r.t. the norm of the space X, that is,

$$\forall \varepsilon > 0 \quad \exists N : \quad \forall n \geqslant N, \quad \left\| x - \sum_{k=1}^{n} \lambda_k e_k \right\| < \varepsilon.$$

Example 7.6.15. By the 1$^{\text{st}}$ Weierstrass theorem, the system $(x^n)_{n=1}^{\infty}$ complete in $C[a,b]$, but it is not a basis in $C[a,b]$. Later we will show that the system

$$1, \sin x, \cos x, \sin 2x, \cos 2x, \ldots, \sin nx, \cos nx, \ldots$$

is a basis in $L_2[-\pi,\pi]$. We will also show that this system is not a basis in $C[-\pi,\pi]$.

Theorem 7.6.16. *In any separable Euclidean (unitary) space there exists an orthonormal basis.*

Proof. Let X be a separable Euclidean (unitary) space. Then there exists a countable complete linearly independent system $(x_n)_{n=1}^{\infty}$ of elements of the space X. Using the Gram–Schmidt orthogonalization process, from $(x_n)_{n=1}^{\infty}$ we construct an orthonormal system $(e_n)_{n=1}^{\infty}$ which is complete in X. By Theorem 7.5.20 any element of X can be expanded into its Fourier series w.r.t. this system, so $(e_n)_{n=1}^{\infty}$ is an orthonormal basis in X. $\square$

Theorem 7.6.17. *Any separable real (or complex) Hilbert space H is isomorphic to the space l_2 over $\mathbb{R}$ (or $\mathbb{C}$).*

Proof. By Theorem 7.6.16, in H there exists an orthonormal basis $(e_n)_{n=1}^{\infty}$. Then to each element $x \in H$ we can correspond the sequence $c_n(x) = (x, e_n)$, $n \in \mathbb{N}$, of its Fourier coefficients. By Theorem 7.5.20, the system $(e_n)_{n=1}^{\infty}$ is closed in Steklov's sense, so the Fourier series of the element x satisfies Parseval's identity 7.5.6, so $(c_n(x))_{n=1}^{\infty} \in l_2$.

Conversely, let $a = (a_n)_{n=1}^{\infty} \in l_2$. Then by Theorem 7.6.4, the numbers a_n, $n \in \mathbb{N}$, are the Fourier coefficients of an element x of the space H.

Thus, between the spaces H and l_2 there exists a one-to-one correspondence. Furthermore, it is obvious that if $x \sim (a_n)$ and $y \sim (b_n)$, then $(\alpha x+\beta y) \sim \alpha(a_n)+\beta(b_n)$, where $\alpha, \beta \in \mathbb{F}$. Moreover, by Theorem 7.5.20 we have

$$\|x\|^2 = \sum_{n=1}^{\infty} |a_n|^2, \qquad \|y\|^2 = \sum_{n=1}^{\infty} |b_n|^2,$$

and from the continuity of the inner product (see Proposition 7.5.3) it follows that

$$(x, y)_H = \sum_{n=1}^{\infty} a_n \bar{b}_n = (a, b)_{l_2}. \tag{7.6.4}$$

□

It is clear that the space l_2 is separable, thus, from Theorem 7.6.17 we obtain the following fact.

Theorem 7.6.18. *All separable infinitely dimensional Hilbert spaces are isomorphic to each other.*

Proof. Let H and H' be separable infinitely dimensional Hilbert spaces. According to Theorem 7.6.16 there exist orthonormal bases $(e_n)_{n=1}^{\infty} \subset H$ and $(e'_n)_{n=1}^{\infty} \subset H'$. Then to each element $x = \sum_{n=1}^{\infty} a_n e_n \in H$ we correspond the element $x' = \sum_{n=1}^{\infty} a_n e'_n \in H'$. Clearly, this is the required isomorphism. □

From the facts established above it follows that if $(e_n)_{n=1}^{\infty}$ is a basis in a separable Euclidean (unitary) space X, then the completion of the space X is a Hilbert space consisting of all the series of the form

$$\sum_{n=1}^{\infty} a_n e_n$$

where $(a_n) \in l_2$, with the inner product of the form

$$(x, y) = \sum_{n=1}^{\infty} a_n \bar{b}_n,$$

whenever

$$x = \sum_{n=1}^{\infty} a_n e_n, \qquad y = \sum_{n=1}^{\infty} b_n e_n.$$

Note that not every Hilbert space is separable.

Example 7.6.19. Let $H_{\mathbb{R}}$ be the set of all functions defined on $\mathbb{R}$ and having at most countable number of non-zero (real) values satisfying

$$\sum_{t \in \mathbb{R}} x^2(t) < \infty, \qquad \forall x(t) \in H_{\mathbb{R}}.$$

This is a Euclidean space with the inner product

$$(x(t), y(t)) = \sum_{t \in \mathbb{R}} x(t) y(t). \tag{7.6.5}$$

For every function $f(t)$ defined on $\mathbb{R}$ the set $\overline{\{t \in \mathbb{R} \ : \ f(t) \neq 0\}}$ is called the *support* of f and is denoted as supp$f$. Thus, the space $H_{\mathbb{R}}$ is the space of functions with countable support.

This space is complete. Indeed, let $\{x_n(t)\}_1^\infty$ is a Cauchy sequence in the norm induced by the inner product (7.6.5): for any $\varepsilon > 0$ there exists a number $N_\varepsilon > 0$ such that $\forall n, m \geqslant N_\varepsilon$

$$\sum_{k=1}^{\omega} |x_n(t_k) - x_m(t_k)|^2 < \varepsilon, \qquad 1 \leqslant \omega \leqslant \infty,$$

where $t_k \in T$, and $T = \bigcup\limits_{n=1}^{\infty} \operatorname{supp}[x_n(t)]$. It is clear that T is at most countable as at most countable union of at most countable sets. Thus, we have that for any $t_k \in T$,

$$|x_n(t_k) - x_m(t_k)|^2 < \varepsilon,$$

so the sequence $x_n(t_k)$ is a Cauchy sequence for every $t_k \in T$, thus $x_n(t_k)$ converges to $x(t_k)$. The function $x(t)$ has at most countable support. Moreover, for any finite number $M \leqslant \omega$, we obtain

$$\sum_{k=1}^{M} |x_n(t_k) - x_m(t_k)|^2 < \varepsilon,$$

and running m to infinity

$$\sum_{k=1}^{M} |x_n(t_k) - x(t_k)|^2 \leqslant \varepsilon,$$

so

$$\|x_n - x\|^2 = \sum_{k=1}^{\omega} |x_n(t_k) - x(t_k)|^2 \leqslant \varepsilon.$$

Now for sufficiently large n we get

$$\|x\|^2 \leqslant \|x_n - x\|^2 + \|x_n\|^2 < \infty.$$

Therefore, $x \in H_{\mathbb{R}}$, so $H_{\mathbb{R}}$ is a real Hilbert space.

However, $H_{\mathbb{R}}$ is not separable. Indeed, if

$$x_\tau(t) = \begin{cases} 1, & t = \tau, \\ 0, & t \neq \tau, \end{cases}$$

then $\|x_\tau - x_{\tau'}\| = \sqrt{2}$ whenever $\tau \neq \tau'$. There exists a bijection between the system $\{x_\tau(t)\}_{\tau \in \mathbb{R}}$ and the real line $\mathbb{R}$.

If $H_{\mathbb{R}}$ were separable, there would exist a countable dense system $(e_n)_{n=1}^\infty$. However, $\bigcup\limits_{n=1}^{\infty} \operatorname{supp}[e_n(t)]$ is countable, so for any linear combination of any elements of this system there exists a point $\tau \in \mathbb{R}$ that do not belong the $\bigcup\limits_{n=1}^{\infty} \operatorname{supp}[e_n(t)]$, so there are always exist functions x_τ that cannot be approximated by any countable system of elements of $H_{\mathbb{R}}$, so this space is not separable.

7.6.2 Orthogonal projections

Let L be a subspace of a Euclidean (unitary) subspace X.

Definition 7.6.20. An element $y_0 \in L$ is called the *orthogonal projection of an element* $x_0 \in X$ *onto the subspace* L if

$$(x_0 - y_0, y) = 0 \qquad \forall y \in L. \tag{7.6.6}$$

Clearly, every element $x \in X$ can have only one orthogonal projection onto L. Indeed, if there exist $y_1, y_2 \in L$ such that

$$(x - y_1, y) = 0, \qquad (x - y_2, y) = 0 \qquad \forall y \in L,$$

then

$$\|y_1 - y_2\|^2 = (y_1 - y_2, y_1 - x + x - y_2) = (y_1 - y_2, y_1 - x) + (y_1 - y_2, x - y_2) = 0,$$

since $y_1 - y_2 \in L$. Thus, $y_1 = y_2$.

Theorem 7.6.21. *An element $y_0 \in L$ is the orthogonal projection of an element $x_0 \in X$ if, and only if,*

$$\|x_0 - y_0\| = \inf_{y \in L} \|x_0 - y\|. \tag{7.6.7}$$

Proof. If $y_0 \in L$ satisfies the condition (7.6.6), then for any $y \in L$

$$\|x_0 - y\|^2 = ((x_0 - y_0) + (y_0 - y), (x_0 - y_0) + (y_0 - y)) = \|x_0 - y_0\|^2 + \|y_0 - y\|^2,$$

so the condition (7.6.7) holds.

Conversely, let the condition (7.6.7) holds for an element $y_0 \in L$. Let us consider the function

$$f(t) = \|x_0 - y_0 + ty\|^2, \qquad t \in \mathbb{R},$$

where y is an arbitrary element of L.

If X is Euclidean, then we have

$$f(t) = \|x_0 - y_0\|^2 + 2t(x_0 - y_0, y) + t^2\|y\|^2.$$

Since this function has the minimal value at the point $t = 0$ by assumption, we obtain $f'(0) = 0$ that implies the condition (7.6.6).

If X is a unitary space, then

$$f(t) = \|x_0 - y_0\|^2 + 2t\,\mathrm{Re}(x_0 - y_0, y) + t^2\|y\|^2,$$

so analogously,

$$\mathrm{Re}(x_0 - y_0, y) = 0 \qquad \forall y \in L.$$

And considering the function

$$g(t) = \|x_0 - y_0 + ity\|^2, \qquad t \in \mathbb{R},$$

in the same way we obtain

$$\mathrm{Im}(x_0 - y_0, y) = 0 \qquad \forall y \in L,$$

therefore, the condition (7.6.6) holds in the case of unitary spaces too. □

Theorem 7.6.22. *If a subspace L of a Euclidean (unitary) space X is complete, then for any $x \in X$, there exists an orthogonal projection onto the subspace L.*

Proof. According to Theorem 7.6.21 it suffices to prove that for any $x_0 \in X$ there exists $y_0 \in L$ such that

$$\|x_0 - y_0\| = \inf_{y \in L} \|x_0 - y\|.$$

Given $x_0 \in X$, define the number

$$d := \inf_{y \in L} \|x_0 - y\|^2.$$

Then there exists a sequence $(y_n)_{n=1}^{\infty} \subset L$ such that

$$\lim_{n \to \infty} \|x_0 - y_n\|^2 = d. \tag{7.6.8}$$

It is easy to see that this system is Cauchy. Indeed, from the Apollonius identity

$$\|y_m - y_n\|^2 = 2\|y_n - x_0\|^2 + 2\|y_m - x_0\|^2 - 4\left\|x_0 - \frac{y_m + y_n}{2}\right\|^2$$

which is true for any $n, m \in \mathbb{N}$, it follows that

$$\left\|x_0 - \frac{y_m + y_n}{2}\right\|^2 \geqslant d,$$

since $\frac{y_m + y_n}{2} \in L$. So we have

$$\|y_m - y_n\|^2 \leqslant 2\|y_n - x_0\|^2 + 2\|y_m - x_0\|^2 - 4d, \tag{7.6.9}$$

and (7.6.8) implies that for any $\varepsilon > 0$ there exists a number $N \in \mathbb{N}$ such that for any $n \geqslant N$,

$$\|x_0 - y_n\| < d + \frac{\varepsilon^2}{4}.$$

Thus from (7.6.9) we get

$$\|y_n - y_m\| < \varepsilon \quad \forall n, m \geqslant N.$$

Since the subspace L is a complete linear normed space, there exists an element $y_0 \in L$ such that $\|y_n - y_0\| \to 0$ as $n \to \infty$. From the continuity of norm we obtain

$$\lim_{n\to\infty} \|x_0 - y_n\| = \|x_0 - y_0\|,$$

as required. □

7.6.3 Linear functionals and the Riesz representation theorem

Let again X be a Euclidean or unitary space.

Definition 7.6.23. Any number-valued mapping f is called a functional. That is, if $f : X \mapsto \mathbb{C}$ or $\mathbb{R}$, then f is a functional.

Definition 7.6.24. A functional $f : X \mapsto \mathbb{F}$ defined on X is called linear if

$$f(c_1x_1 + c_2x_2) = c_1f(x_1) + c_2f(x_2), \quad \forall c_1, c_2 \in \mathbb{F}, \quad x_1, x_2 \in X,$$

where $\mathbb{F} = \mathbb{C}$ or $\mathbb{R}$.

Definition 7.6.25. The set $K = \{x \in X : \ f(x) = 0\}$ is called the *kernel* of the functional f.

Definition 7.6.26. A functional $f : X \mapsto \mathbb{C}$ or $\mathbb{R}$ is called bounded if there exists a constant $M > 0$ such that

$$|f(x)| \leqslant M\|x\| \quad \forall x \in X.$$

Lemma 7.6.27. *A linear functional defined on a linear normed space X is continuous on X if it is continuous at* 0*.*

Proof. Let $f(x_n) \to 0$ as $n \to \infty$ whenever $x_n \to 0$ in X as $n \to \infty$, that is $\|x_n\| \to 0$.

Suppose that $\|x_n - x\| \underset{n\to\infty}{\longrightarrow} 0$, that is, $x_n \to x$ in X. Then $f(x_n) - f(x) = f(x_n - x) \to 0$ as $n \to \infty$, so $f(x_n) \to f(x)$ as $n \to \infty$ for any $x \in X$. □

Theorem 7.6.28. *In a normed space a linear functional is continuous if and only if it is bounded.*

Proof. Let f is bounded, and let $\|x_n\| \to 0$ as $n \to \infty$. Then

$$|f(x_n)| \leqslant C \cdot \|x_n\| \to 0 \quad \text{as} \quad n \to \infty.$$

So f is continuous at zero, therefore, is continuous by Lemma 7.6.27.

Conversely, suppose that f is continuous but unbounded. Then there exists a sequence $(x_n)_{n=1}^{\infty} \subset X$ such that $|f(x_n)| > n\|x_n\|$. Introduce a new sequence

$$\xi_n = \frac{x_n}{n\|x_n\|}.$$

Since f is linear, we have $|f(\xi_n)| > 1$. At the same time, $\|\xi_n\| = \dfrac{1}{n} \to 0$ as $n \to \infty$. Thus, we have $\xi_n \to 0$ as $n \to \infty$ in the norm of X, so by continuity $f(\xi_n) \to 0$, a contradiction. □

Definition 7.6.29. Let f be a linear continuous functional defined on a normed space X. The number $\|f\| := \inf M$ where inferior is taken over all numbers $M > 0$ satisfying the inequality

$$|f(x)| \leqslant M\|x\| \quad \forall x \in X,$$

is called the *norm* of the functional f.

Lemma 7.6.30. *The norm of a linear continuous functional defined on a normed space X can be defined by the formula*

$$\|f\| = \sup_{\substack{x\in X,\\ \|x\|\neq 0}} \frac{|f(x)|}{\|x\|} = \sup_{\substack{x\in X,\\ \|x\|\leqslant 1}} |f(x)| = \sup_{\substack{x\in X,\\ \|x\|=1}} |f(x)|. \tag{7.6.10}$$

Proof. It is easy to see that the norm $\|f\|$ is the inferior of the numbers $M > 0$ satisfying the inequality

$$\frac{|f(x)|}{\|x\|} \leqslant M, \qquad \forall x \in X,\ x \neq 0.$$

This means that for any $\varepsilon > 0$ there exists an element $\widehat{x} \in X$, $\|\widehat{x}\| \neq 0$, such that

$$\frac{|f(\widehat{x})|}{\|\widehat{x}\|} > \|f\| - \varepsilon.$$

By definition of supremum, we have

$$\|f\| = \sup_{\substack{x\in X,\\ \|x\|\neq 0}} \frac{|f(x)|}{\|x\|}.$$

Furthermore, since $\dfrac{|f(x)|}{\|x\|} = f\left(\dfrac{x}{\|x\|}\right)$, so

$$\|f\| = \sup_{\substack{x\in X,\\ \|x\|=1}} |f(x)| \leqslant \sup_{\substack{x\in X,\\ \|x\|\leqslant 1}} |f(x)|.$$

Now if $\|x\| \leqslant 1$, then

$$|f(x)| \leqslant \|f\| \cdot \|x\| \leqslant \|f\|,$$

so

$$\sup_{\|x\|\leqslant 1} |f(x)| \leqslant \|f\|.$$

Consequently,

This implies that

$$\sup_{\|x\|\leqslant 1} |f(x)| = \|f\|.$$

□

It is clear that the norm of lineal continuous functionals possesses all three properties of norms. So we can introduce the linear normed space of linear bounded functional defined on X. This space is called *dual* to the space X and is usual denoted X^*.

Example 7.6.31. Let $t_k \in [a,b]$, and $c_k \in \mathbb{R}$, $k = 1,\dots,n$, $a < t_1$ and $t_n < b$. In $C[a,b]$ consider the following functional

$$f(x) \stackrel{def}{=} \sum_{k=1}^{n} c_k x(t_k), \qquad x \in C[a,b].$$

This functional is obviously linear.Moreover,

$$|f(x)| \leqslant \sum_{k=1}^{n} |c_k| \cdot |x(t_k)| \leqslant \left(\sum_{k=1}^{n} |c_k|\right) \cdot \|x\|, \qquad \forall x \in C[a,b],$$

so

$$\|f\| = \sup_{\substack{x \in C[a,b],\\ \|x\| \neq 0}} \frac{|f(x)|}{\|x\|} \leqslant \sum_{k=1}^{n} |c_k|.$$

Thus, f is a bounded linear functional.

Consider now the function $y(t) \in C[a,b]$ which satisfies $y(t_k) = \operatorname{sgn} c_k$, $k = 1,\dots,n$, and is linear between the points $a, t_1, \dots, t_n, b$, and $y(a) = y(b) = 0$. It is clear that $\|y\| = 1$. Then we have

$$f(y) = \sum_{k=1}^{n} c_k \cdot \operatorname{sgn} c_k = \sum_{k=1}^{n} |c_k|.$$

So we found a function in $C[a,b]$ on which the functional f achieves the value equal to a bound of its norm, so the norm is equal to this bound:

$$\|f\| = \sum_{k=1}^{n} |c_k|.$$

In Hilbert spaces, it is possible to find a general form of any linear bounded functional.

Lemma 7.6.32. *For any element $a \in X$, where X is a Euclidean or unitary space, the functional*

$$f(x) = (x,a) \quad x \in X \tag{7.6.11}$$

is linear and bounded. Moreover,

$$\|f\| = \|a\|.$$

Proof. The linearity of f is obvious, and the boundness follows from the Cauchy-Bunyakovsky-Schwarz inequality:

$$|f(x)| \leqslant \|a\| \cdot \|x\| \quad \forall x \in X.$$

Since $f(a) = \|a\|^2$, we have $\|f\| = \|a\|$. □

Lemma 7.6.33. *If a functional defined on X is given by the formula* (7.6.11)*, then the element a is uniquely defined.*

Proof. Indeed, let

$$f(x) = (x,a) \qquad \text{and} \qquad f(x) = (x,b),$$

then $(x, b-a) = 0$ for any $x \in X$. In particular, $(b-a, b-a) = \|b-a\| = 0$, so $a = b$. □

Theorem 7.6.34 (Riesz)**.** *Any linear continuous functional defined on a Hilbert space H can be uniquely represented as in* (7.6.11) *for a certain element $a \in H$.*

Proof. Let K be the kernel of the functional f. From the continuity of the functional f it follows that K is a closed subspace of the space H: if $f(y_n) = 0$, $y_n \in Y$, $n \in \mathbb{N}$, and $y_n \underset{n\to\infty}{\longrightarrow} y$, then $f(y) = 0$. Moreover, since the space H is complete, the subspace K us complete as well (any Cauchy sequence in K converges, and the limit must belong to K).

If $K = H$, then $f(x) \equiv 0$ on H, so clearly

$$f(x) = (0, x) \qquad x \in H.$$

Suppose now that $K \neq H$. Then there exists an element $x_0 \in H$ such that $f(x_0) \neq 0$. Let $y_0 \in K$ be the orthogonal projection of the element x_0 onto the subspace K which exists according to Theorem 7.6.22. We set

$$z_0 = x_0 - y_0.$$

Then we have $f(z_0) = f(x_0) \neq 0$, and $(z_0, y) = 0$ for any $y \in K$.

Since

$$f\left(x - \frac{f(x)}{f(z_0)} z_0\right) = 0 \qquad x \in H,$$

it follows that

$$x - \frac{f(x)}{f(z_0)} z_0 \in K \qquad \forall x \in H.$$

Therefore,

$$\left(x - \frac{f(x)}{f(z_0)} z_0, z_0\right) = 0 \qquad \forall x \in H.$$

Thus, we obtain

$$(x, z_0) = \frac{\|z_0\|^2}{f(z_0)} \cdot f(x).$$

Note that $f(z_0) \neq 0$ implies $\|z_0\| \neq 0$. Consequently, the element $a = \dfrac{\overline{f(z_0)}}{\|z_0\|^2} z_0$ exists, and

$$f(x) = (x, a), \qquad a \in H,$$

as required. □

This theorem and Lemma 7.6.32 show that the space H^* dual to a Hilbert space H is isomorphic to H. This is principal difference between Hilbert and Banach spaces that isomorphic to a subspace of their dual space.

Example 7.6.35. In l_2 let us consider the following functional

$$f(x) = \sum_{n=1}^{\infty} \frac{x_n + x_{n+1}}{2^n},$$

where $x = (x_1, x_2, \ldots, x_n, \ldots) \in l_2$. This functional is bounded. Indeed, if $y \in l_2$ is such that $y_n = \dfrac{1}{2^n}$, then by Cauchy-Bunyakovski-Schwarz inequality we obtain

$$|f(x)| \leqslant \sum_{n=1}^{\infty} \frac{1}{2^n} \cdot |x_n| + 2\sum_{n=1}^{\infty} \frac{1}{2^{n+1}} \cdot |x_{n+1}| \leqslant 3\|y\|_{l_2} \cdot \|x\|_{l_2}.$$

To find the norm we will use Theorem 7.6.34. The functional f can be rewritten in the form

$$f(x) = \frac{x_1}{2} + 3\sum_{n=2}^{\infty} \frac{x_n}{2^n} = (a, x),$$

where $a = (a_1, a_2, \dots, a_n, \dots)$ with

$$a_1 = \frac{1}{2}, \quad a_n = \frac{3}{2^n}, \qquad n = 2, 3, \dots.$$

Thus, by Lemma 7.6.32, we have

$$\|f\| = \|a\|_{l_2} = \frac{1}{4} - 9 - \frac{9}{4} + 9\sum_{n=0}^{\infty}\frac{1}{4^n} = \frac{1}{4} - 9 - \frac{9}{4} + 12 = 1.$$

7.6.4 Weak convergence

Definition 7.6.36. A sequence $(x_n)_{n=1}^{\infty}$ of elements of a linear normed space X is called *weakly convergent* to an element $x \in X$ if

$$\lim_{n\to\infty} f(x_n) = f(x) \qquad \forall f \in X^*.$$

Due to the Riesz Theorem 7.6.34, we can define the weak convergence in Hilbert spaces in the inner product form.

Definition 7.6.37. A sequence $(x_n)_{n=1}^{\infty}$ of elements of a Hilbert space H is called *weakly convergent* to an element $x \in H$ if

$$\lim_{n\to\infty} (x_n, y) = (x, y) \qquad \forall y \in H.$$

This definition will allow us later to understand the notion of weak solutions of differential equations.

Theorem 7.6.38. *If the sequences $(x_n)_{n=1}^{\infty}$ of elements of a linear normed space X converges to $x \in X$ in the norm of X, then it is weakly convergent to x.*

Proof. By linearity of the inner product and by the boundness of functionals in the space X^* we have for any $y \in X$

$$|f(x_n) - f(x)| = |f(x_n - x)| \leqslant \|f\| \cdot \|x_n - x\| \to 0 \qquad \text{as} \quad n \to \infty.$$

$\square$

Since any Hilbert space is a normed space, we have the corresponding theorem for Hilbert spaces which we state separately for the sake of convenience in the future.

Theorem 7.6.39. *If the sequences $(x_n)_{n=1}^{\infty}$ of elements of a Hilbert space H converges to $x \in H$ in the norm of H, then it is weakly convergent to x.*

Generally speaking, converse is not true.

Example 7.6.40. Consider the sequence of coordinate vectors $(e_n)_{n=1}^{\infty}$ in l_2, where $e_n = (\underbrace{0, \dots, 0}_{n-1}, 1, 0, \dots)$.
This sequence weakly convergent in l_2. Indeed, for any $a \in l_2$ we have

$$(e_n, a) = a_n \to 0 \qquad \text{as} \quad n \to \infty,$$

since the series $\sum\limits_{n=1}^{\infty} |a_n|^2$ converges. However, the sequence $(e_n)_{n=1}^{\infty}$ is not convergent in l_2, because for any $a \in l_2$

$$\|e_n - a\|^2 = \|e_n\|^2 - 2\,\mathrm{Re}(e_n, a) + \|a\|^2 \underset{n\to\infty}{\longrightarrow} 1 + \|a\|^2 > 1.$$

However, in finite-dimensional case the weak convergence is equivalent to the convergent in norm.

Example 7.6.41. Consider the space $\mathbb{C}^n$ with standard inner product. Let $(e_k)_{k=1}^n$ be an arbitrary orthonormal basis in $\mathbb{C}^n$, and suppose that a sequence $(x_m)_{m=1}^\infty$ weakly converges to an element $x \in \mathbb{C}^n$. Then we have

$$x_m = \sum_{k=1}^n x_m^{(k)} e_k \qquad \text{and} \qquad x = \sum_{k=1}^n x^{(k)} e_k,$$

and

$$x_m^{(k)} = (x_m, e_k) \underset{m\to\infty}{\longrightarrow} (x, e_k) = x^{(k)}, \qquad k = 1, \dots, n.$$

This means that the sequence $(x_m)_{m=1}^\infty$ converges to x entrywise. This implies

$$\|x_m - x\|^2 = \sum_{k=1}^n |x_m^{(k)} - x^{(k)}|^2 \to 0 \qquad \text{as} \quad m \to \infty.$$

This example gives us an idea that in l_2 the coefficient-wise convergence does not imply the convergence in norm. Indeed, if $x^{(n)} \in l_2$ is such that $x_k^{(n)} = \dfrac{1}{k^{\frac{1}{2}+\frac{1}{n}}}$, then

$$\|x^{(n)}\|_{l_2}^2 = \sum_{k=1}^\infty \left(\frac{1}{k^{\frac{1}{2}+\frac{1}{n}}}\right)^2 = \sum_{k=1}^\infty \frac{1}{k^{1+\frac{2}{n}}} < +\infty \qquad \forall n \in \mathbb{N}.$$

The coefficient-wise limit of the sequence $x^{(n)}$ is the sequence x whose coefficients are $x_k = \dfrac{1}{k^{\frac{1}{2}}}$, so

$$\|x\|_{l_2}^2 = \sum_{k=1}^\infty \left(\frac{1}{k^{\frac{1}{2}}}\right)^2 = \sum_{k=1}^\infty \frac{1}{k} = +\infty,$$

thus, $x \notin l_2$.

With an additional property the weak convergence might imply the convergence in norm. In fact, by Theorem 7.2.7 it follows that if $\|x_n - x\| \to 0$ as $n \to \infty$, then $\|x_n\| \to \|x\|$. And converse is not true.

Theorem 7.6.42. *Let a sequence $(x_n)_{n=1}^\infty$ of elements of Hilbert space H weakly converges to $x \in H$, and let $\|x_n\| \to \|x\|$ as $n \to \infty$. Then $(x_n)_{n=1}^\infty$ converges to x in norm of H.*

Proof. Since $(x_n)_{n=1}^\infty$ weakly converges to x, we have $(x_n, x) \to \|x\|^2$ and $(x, x_n) \to \|x\|^2$. Thus, we have

$$\|x_n - x\|^2 = \|x_n\|^2 - (x_n, x) - (x, x_n) + \|x\|^2 =$$

$$= \left[\|x_n\|^2 - \|x\|^2\right] + \left[\|x\|^2 - (x_n, x)\right] + \left[\|x\|^2 - (x, x_n)\right] \to 0 \qquad \text{as} \quad n \to \infty.$$

□

7.6.5 The space L_2

Consider now the (complex) space $L_2[-\pi, \pi]$. According to Theorem 7.3.9, the space is complete w.r.t. the norm[5]

$$\|f\|_2 = \left(\int_{-\pi}^{\pi} |f(x)|^2 dx\right)^{\frac{1}{2}}, \qquad f \in L_2[-\pi, \pi].$$

[5]In spite of simplicity, in what follows we use the standard notation $\int_{-\pi}^{\pi} g(x)dx$ for the Lebesgue integral $\int_{[-\pi,\pi]} g(x)d\mu$.

Moreover, we can introduce the following function of two variables

$$(f, g) = \left(\int_{-\pi}^{\pi} f(x)\overline{g(x)}dx \right)^{\frac{1}{2}}, \qquad f, g \in L_2[-\pi, \pi]. \tag{7.6.12}$$

It is easy to see that this function defines an inner product on the space $L_2[-\pi, \pi]$ since it satisfies all three properties of inner products. Indeed, for $f \in L_2[-\pi, \pi]$, $(f, f) = \|f\|_2^2 \geqslant 0$ $(= 0 \iff f = 0)$. This is the property the norm $\|\cdot\|_2$ which we established earlier. The linearity of the function (7.6.12) follows from the linearity of Lebesgue integral (see Properties 6 and 7 of Lebesgue integral). Finally, the property $(f, g) = \overline{(g, f)}$ is obvious.

Thus, the space $L_2[-\pi, \pi]$ is a Hilbert space with inner product defined by (7.6.12).

Theorem 7.6.43. *The space $L_2[-\pi, \pi]$ is separable.*

Proof. Let $f \in L_2[-\pi, \pi]$. By Theorem 7.4.14, for any $\varepsilon > 0$ there exists a function $g \in C[-\pi, \pi]$ such that $\|f - g\| < \frac{\varepsilon}{3}$. In it its turn, the function g can be approximated by a polynomial p by the first Weierstrass theorem: $\max\limits_{x \in [-\pi, \pi]} |g(x) - p(x)| < \frac{\varepsilon}{6\pi}$. This implies that

$$\|g - p\|_2 = \left(\int_{-\pi}^{\pi} |g(x) - p(x)|^2 dx \right)^{\frac{1}{2}} \leqslant \max_{x \in [-\pi, \pi]} |g(x) - p(x)| \int_{-\pi}^{\pi} dx < \frac{\varepsilon}{6\pi} \cdot 2\pi = \frac{\varepsilon}{3}.$$

Furthermore, every polynomial p with complex coefficients can be approximated by a polynomial q with complex rational coefficients (that is, the real and imaginary part of every coefficient is rational). So if $p(x) = \sum\limits_{k=0}^{n} a_k x^k$, $a_k \in \mathbb{C}$, and $q(x) = \sum\limits_{k=0}^{n} b_k x^k$, $\operatorname{Re} b_k, \operatorname{Im} b_k \in \mathbb{Q}$, then

$$\|p - q\|_2 \leqslant 2\pi \cdot \max_{x \in [-\pi, \pi]} |p(x) - q(x)| < 2\pi \sum_{k=0}^{n} |a_k - b_k| \pi^k < \frac{\varepsilon}{3},$$

since we can choose a polynomial q such that $|a_k - b_k| < \frac{\varepsilon}{6(n+1)\pi^{k+1}}$. Combining all the previous we get

$$\|f - q\|_2 \leqslant \|f - g\|_2 + \|g - p\|_2 + \|p - q\|_2 < \varepsilon.$$

Thus, the system of all polynomials with complex rational coefficients is complete in the space $L_2[-\pi, \pi]$. By Example 1.3.10 the set of polynomials with rational coefficients is countable. But every polynomial Q with complex rational coefficients has the form $Q(x) = \operatorname{Re} Q(x) + i \operatorname{Im} Q(x)$. So the set of all polynomials with complex rational coefficients is the union of two countable sets, so it is countable by Theorem 1.3.5. Since all the polynomials are Lebesgue integrable on $[-\pi, \pi]$, we get that the space $L_2[-\pi, \pi]$ contains a countable complete system, so it is separable. □

Note that this theorem can be proved much simpler. Indeed, according to Corollary 7.4.15, the countable system $(x^n)_{n=1}^{\infty}$ is complete in $L_p[-\pi, \pi]$, $p > 1$, so even $L_p[-\pi, \pi]$ is separable. But this system is not orthonormal in $L_2[-\pi, \pi]$ (for $p \neq 2$, $L_p[a, b]$ has no inner product). So, the proof of Theorem 7.6.43 gives an idea that the orthonormal polynomial basis in $L_2[-\pi, \pi]$ can be a system of polynomials with rational coefficients. In fact, such a system exists, and the polynomials are called *Legendre polynomials.* These polynomials have the form

$$L_n(x) = \frac{1}{2^n n!} \cdot \frac{d^n}{dx^n}(x^2 - 1)^n.$$

Corollary 7.6.44. *If Parseval's identity w.r.t. a system $(e_n)_{n=1}^{\infty}$ holds for all the functions x^n, $n = 0, 1, 2, \ldots$, then the system $(e_n)_{n=1}^{\infty}$ is closed.*

Proof. Indeed, consider the polynomial

$$P(x) = \sum_{k=0}^{m} a_k x^k.$$

Then

$$S_n(P) = \sum_{k=0}^{m} a_k S_k(x^k),$$

so

$$\|P - S_n(P)\|_2 \leqslant \sum_{k=0}^{m} |a_k| \cdot \|x_k - S_n(x^k)\|_2 \to 0 \qquad \text{as} \quad n \to \infty.$$

Thus, Parseval's identity holds for any polynomials, and the set of all polynomials is dense in $L_2[-\pi, \pi]$ by Theorem 7.4.14. □

But do orthonormal bases exist in $L_2[-\pi, \pi]$? Yes, as we mentioned above. Legendre polynomials form such a basis. The existence of such a basis is provided by Theorems 7.6.43 and 7.6.16.

Let us consider the following countable system of continuous functions

$$1,\ \sin x,\ \cos x,\ \ldots, \sin nx,\ \cos nx,\ \ldots \tag{7.6.13}$$

it is easy to see that this system is orthonormal in $L_2[-\pi, \pi]$.

Lemma 7.6.45. *If $n \in \mathbb{N} \cup \{0\}$, then*

$$\int_{-\pi}^{\pi} \sin nx dx = 0.$$

Proof. If $n = 0$, then the statement is obvious. Let $n > 0$.. Then we have

$$\int_{-\pi}^{\pi} \sin nx dx = \left.\frac{-\cos nx}{n}\right|_{-\pi}^{\pi} = 0.$$

□

Analogously we have the following

Lemma 7.6.46. *If $n \in \mathbb{N}$, then*

$$\int_{-\pi}^{\pi} \cos nx dx = 0.$$

Lemma 7.6.47. *For any $n \in \mathbb{N}$,*

$$\int_{-\pi}^{\pi} \cos^2 nx dx = \int_{-\pi}^{\pi} \sin^2 nx dx = \pi.$$

Proof. The formulæ follow from the well-known identities

$$\cos^2 x = \frac{1 + \cos 2x}{2}, \qquad \sin^2 x = \frac{1 - \cos 2x}{2}$$

and from Lemmata 7.6.45–7.6.46. □

So we get

Lemma 7.6.48. *The system* (7.6.13) *is orthogonal in* $L_2[-\pi, \pi]$.

Proof. This fact follows from the Lemmata 7.6.45–7.6.47 and the standard formulae of sines and cosines sums of sums of angles:

$$\sin nx \sin mx = \frac{\cos(n-m)x - \cos(n+m)x}{2},$$

$$\cos nx \cos mx = \frac{\cos(n-m)x + \cos(n+m)x}{2},$$

$$\sin nx \cos mx = \frac{\sin(n-m)x + \sin(n+m)x}{2},$$

□

Corollary 7.6.49. *The trigonometric system* (7.6.13) *is an orthogonal basis in* $L_2[-\pi,\pi]$.

Proof. Indeed, the trigonometric polynomials

$$T_n(x) = \frac{a_0}{2} + \sum_{m=1}^{n} (a_m \cos mx + b_m \sin mx)$$

are dense in $L_2[-\pi,\pi]$ by Corollary 7.4.15. But Parseval's identity holds for any trigonometric polynomial:

$$\int_{-\pi}^{\pi} |T_n(x)|^2 dx = \frac{\pi |a_0|^2}{2} + \pi \sum_{m=1}^{n} (|a_m|^2 + |b_m|^2).$$

Consequently, the system (7.6.13) is closed in $L_2[-\pi,\pi]$ by Theorem 7.6.8. By Proposition 7.6.7 the system (7.6.13) is complete, so it is a basis in $L_2[-\pi,\pi]$ by Theorem 7.5.20. □

Thus, any function f from the space $L_2[-\pi,\pi]$ can be expanded in to the series

$$f(x) = \frac{a_0}{2} + \sum_{m=1}^{\infty} (a_m \cos mx + b_m \sin mx)$$

convergent in the norm of $L_2[-\pi,\pi]$, where the coefficients a_m and b_m have the form

$$a_m = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos mx dx, \qquad b_m = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin mx dx, \qquad m = 0, 1, 2, \dots \tag{7.6.14}$$

are the Fourier coefficients of the function f w.r.t. the system (7.6.13).

From the properties of separable Hilbert spaces, we have following fact.

Proposition 7.6.50. *A function f belongs to the space $L_2[-\pi,\pi]$ if, and only if, the sequences of its Fourier coefficients defined by* (7.6.14) *belong to the space l_2. Moreover, the Parseval's identity holds:*

$$\|f\|_2^2 = \int_{-\pi}^{\pi} |f(x)|^2 dx = \frac{\pi |a_0|^2}{2} + \pi \sum_{m=1}^{\infty} (|a_m|^2 + |b_m|^2). \tag{7.6.15}$$

From the proof of Theorem 7.6.17 (see formula (7.6.4)) it follows that for any two functions $f, g \in L_2[-\pi,\pi]$ with Fourier coefficients

$$a_m = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos mx dx, \qquad b_m = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin mx dx, \qquad m = 0, 1, 2, \dots$$

and

$$c_m = \frac{1}{\pi}\int\limits_{-\pi}^{\pi} g(x)\cos mx dx, \qquad d_m = \frac{1}{\pi}\int\limits_{-\pi}^{\pi} g(x)\sin mx dx, \qquad m = 0, 1, 2, \dots$$

one has

$$(f, g) = \int\limits_{-\pi}^{\pi} f(x)\overline{g(x)}dx = \frac{a_0\overline{c_0}}{4} + \sum_{m=1}^{\infty}(a_m\overline{c_m} + b_m\overline{d_m}). \tag{7.6.16}$$

This formula implies the following interesting fact.

Proposition 7.6.51. *The Fourier series of a function $f \in L_2[-\pi, \pi]$ w.r.t. the system* (7.6.13) *can be integrated term by term over any measurable set $A \subset [-\pi, \pi]$, i.e.*

$$\int\limits_A f(x)d\mu = a_0\mu A + \sum_{n=1}^{\infty}\left(a_n\int\limits_A \cos nx d\mu + b_n\int\limits_A \sin nx d\mu\right). \tag{7.6.17}$$

Proof. In fact, let $g(x) = \chi_A(x)$, the characteristic function of the set A. It is bounded and measurable, since A is measurable. Substituting $g(x)$ into the formula (7.6.16) we get (7.6.17), as required. □

Remark 7.6.52. Note that if $(e_n)_{n=1}^{\infty}$ is an arbitrary orthonormal basis in $L_2[-\pi, \pi]$, then Propositions 7.6.50–7.6.51 are also true for the corresponding Fourier series w.r.t. this basis.

7.7 Problems

7.7.1 Linear normed spaces. Spaces L_p.

Problem 7.1. Prove that any normed linear space is a metric (linear) space with the metric $\rho(f,g) = \|f-g\|$.

Problem 7.2. Prove that the space $C^1[a,b]$ of continuously differentiable functions on $[a,b]$ with norm

$$\|f\| = \max_{x\in[a,b]} |f(x)| + \max_{x\in[a,b]} |f'(x)|$$

is a complete linear normed space. (Prove first that $\|f\|$ is a norm.)

Problem 7.3. Prove that the space of all bounded (real or complex) sequences is a *complete* linear normed space with norm

$$\|x\|_\infty = \sup_{k\in\mathbb{N}} |x_k|.$$

Problem 7.4. Let $f_n \to f$ in $L_1(X)$. Prove that $|f_n| \to |f|$ in $L_1(X)$.

Problem 7.5. Let $f_n \to f$ in $L_1(X)$, and $f_n \to g$ almost everywhere on X. Prove that $f \overset{a.e.}{=} g$ on X.

Problem 7.6. Let $\{f_n(x)\}_{n=1}^{+\infty} \in S(X)$ and $f_0(x) \in S(X)$, where $\mu X < +\infty$. Prove that if $f_n \xrightarrow{\mu} f_0$ on X, and the functions $f_n(x)$, $n \geqslant 1$, are bounded almost everywhere on X as a whole (that is, there is one constant for all functions for almost all points x), then $f_n \xrightarrow{L_1} f_0$ on X.

Problem 7.7. Prove that the sequence $f_n = n^2xe^{-nx}$, $n \in \mathbb{N}$, converges everywhere on $[0,1]$ to the function $f_0(x) \equiv 0$, but f_n does not converge in $L_2[0,1]$. Does it converge in $L_1[0,1]$ and $C[0,1]$?

Problem 7.8. Does the sequence $f_n(x) = (\sin nx)^n$ converge to $f_0 \equiv 0$ in L_p norm, $p \geqslant 1$, on $[0,\pi]$?

Hint: Use Stirling's formula:

$$n! = \sqrt{2\pi n}\left(\frac{n}{e}\right)^n \left(1 + O\left(\frac{1}{n}\right)\right) \quad \text{as} \quad n \to \infty.$$

Problem 7.9. Do the following expressions define norms?

1) $x \mapsto |\arctan x|$ for $x \in \mathbb{R}$;

2) $x \mapsto \left(\sum\limits_{k=1}^{n} |x_k|^p\right)^{\frac{1}{p}}$ for $x = (x_1,\dots,x_n) \in \mathbb{R}^n$ if $0 < p < 1$ and $n \geqslant 2$;

3) $x \mapsto \max\limits_{a\leqslant t\leqslant \frac{a+b}{2}} |x(t)|$ for $x \in C[a,b]$;

4) $x \mapsto |x(a)| + \max\limits_{a\leqslant t\leqslant b} |x'(t)|$ for $x \in C^1[a,b]$;

5) $x \mapsto |x(b) - x(a)| + \max\limits_{a\leqslant t\leqslant b} |x'(t)|$ for $x \in C^1[a,b]$.

Problem 7.10. Prove that in the space $C[0,\pi]$ (with maximum norm) the functions $1, \cos t, \cos^2 t$ are linearly independent but the functions $1, \cos 2t, \cos^2 t$ are linearly dependent.

Problem 7.11. Consider the space B of all n times continuously differentiable functions on $[a,b]$ with the norm

$$\|x\|_s = \max_{0\leqslant j\leqslant n} \{ \sup_{t\in[a,b]} p_j(t)|x^{(j)}(t)|\},$$

where $p_j(t) \in C[a,b]$, $j = 0,1,\dots,n$, are some positive functions.

Prove that $\|\cdot\|_s$ defines a norm and that B is a complete space.

Problem 7.12 (Extended Hölder inequality)**.** Let $f \in L_1(X)$ and $g \in L_\infty(X)$. Prove that $fg \in L_1(X)$ and

$$\int\limits_X |f(x)g(x)|d\mu \leqslant \|f\|_1 \cdot \|g\|_\infty.$$

Problem 7.13. Let $1 \leqslant s < p < \infty$. Show that there exists a function $f \in L_r(0, +\infty)$ for any $r \in [s, p]$ such that $f \notin L_r(0, +\infty)$ for any $r \notin [s, p]$.

Problem 7.14. Let $f \in L_p(X)$ for all $p \geqslant p_0 \geqslant 1$. Prove that there exists a finite or infinite limit of the magnitudes $\|f\|_p$ as $p \to \infty$. Prove also that if the limit is finite, then $f \in L_\infty(X)$ and

$$\lim_{p\to\infty} \|p\| = \|f\|_\infty.$$

If the limit is infinite, prove that $f \notin L_\infty(X)$.

Problem 7.15. Let $(f_n)_{n=1}^\infty \subset L_p(X)$ and $f_0 \in L_p(X)$ for some $p \geqslant 1$. Suppose that $f_n \xrightarrow{\mu} f_0$, and $|f_n(x)| \leqslant |g(x)|$ on X, $\forall n \in \mathbb{N} \cup \{0\}$, where $g \in L_p(X)$. Prove that $f_n \xrightarrow{L_p} f$.

Hint: Prove first that Lebesgue's Dominated Convergence Theorem is true if we change the convergence a.e. to convergence in measure.

Problem 7.16. Let $\mu X < +\infty$ and $1 \leqslant r < p < \infty$. Given a sequence $(f_n)_{n=1}^\infty \subset L_p(X)$ such that $f_n \underset{n\to\infty}{\xrightarrow{\mu}} f$, suppose that there exists a number $C > 0$ so that $\|f_n\|_p \leqslant C$ for any $n \in \mathbb{N}$. Prove that $f_n \underset{n\to\infty}{\xrightarrow{L_r(X)}} f$.

Problem 7.17. Let $1 \leqslant p < \infty$. Given a sequence $\{f_n(x)\}_{n=1}^\infty \subset L_p(X)$, suppose that

$$\sum_{n=1}^\infty \|f_n\|_p < +\infty.$$

Prove that the series

$$\sum_{n=1}^\infty f_n(x)$$

converges absolutely a.e. in $L_p(X)$, and

$$\left\| \sum_{n=1}^\infty f_n(x) \right\|_p \leqslant \sum_{n=1}^\infty \|f_n\|_p.$$

Problem 7.18. Does the sequence $\{x_n\}_{n=1}^\infty$ converge in the space E? Here

1) $E = l_1, \qquad x_n = \left(\underbrace{0, \ldots, 0}_{n-1}, \frac{1}{n^\sigma}, \frac{1}{(n+1)^\sigma}, \frac{1}{(n+2)^\sigma}, \ldots \right), \qquad \sigma > 1;$

2) $E = l_2, \qquad x_n = \left(\frac{1}{n}, \underbrace{0, \ldots, 0}_{n-2}, 1, 0, 0, \ldots \right);$

3) $E = C^1[0,1], \qquad x_n(t) = \frac{t^{n+1}}{n+1} - \frac{t^{n+2}}{n+2};$

4) $E = L_1[0,1], \qquad x_n(t) = \begin{cases} e^{-\frac{t}{n}}, & t \in \mathbb{R} \setminus \mathbb{Q}, \\ 0, & t \in \mathbb{Q}, \end{cases};$

$$5)\ E = L_2[0,1], \qquad x_n(t) = \begin{cases} \sqrt{n} - n\sqrt{n}t, & t \in \left[0, \dfrac{1}{n}\right], \\ \\ 0, & t \in \left(\dfrac{1}{n}, 1\right], \end{cases}.$$

Problem 7.19. Let $[a,b] \subset \mathbb{R}$, $1 \leqslant p < \infty$, and let $f \in L_p[a,b]$. Define its *modulus of continuity* as follows

$$\omega(f;\delta)_p = \sup_{0 \leqslant h \leqslant \delta} \left(\int\limits_{[a,b-h]} |f(x+h) - f(x)|^p d\mu \right)^{\frac{1}{p}}.$$

Prove that $\omega(f;\delta)_p \to 0$ as $\delta \to +0$. Here μ is the Lebesgue measure.

Problem 7.20. Let $f, g \in L_1(\mathbb{R})$ and $f(x)$ is bounded. Prove that the convolution

$$(f * g)(x) = \int\limits_{\mathbb{R}} f(t) g(x-t) d\mu(t)$$

is continuous on $\mathbb{R}$.

Problem 7.21. Let $f \in L_p(\mathbb{R})$, $1 \leqslant p < +\infty$ and $h > 0$. Define

$$f_h(x) = \frac{1}{2h} \int\limits_{(x-h,x+h)} f(t) d\mu(t).$$

Prove that $\|f_h\|_{L_p} \leqslant \|f\|_{L_p}$, and $\|f - f_h\|_{L_p} \to 0$ as $h \to +0$.

Problem 7.22. Prove that a subspace L_0 of a Banach space L is a Banach space if and only if L_0 is closed.

Problem 7.23. Is the set $L = \left\{ x = (x_n)_{n=1}^{\infty} \in l_p : \sum\limits_{k=1}^{\infty} x_k = 0,\ x_k \in \mathbb{R} \right\}$ a linear closed subspace of the spaces l_p, $p \geqslant 1$? Remind that in l_1 the norm is $\|x\|_1 = \sum\limits_{n=1}^{\infty} |x_n|$.

Hint: Consider the cases $p = 1$ and $p > 1$ separately. Consider the sequence $x^{(n)} = \left(1, \underbrace{-\frac{1}{n}, -\frac{1}{n}, \ldots, -\frac{1}{n}}_{n}, 0, 0, \ldots \right)$.

7.7.2 Linear inner product spaces

Problem 7.24. Let X be a ***complex*** normed space. Prove that X is a unitary space if and only if the parallelogram identity holds:

$$\|x+y\|^2 + \|x-y\|^2 = 2(\|x\|^2 + \|y\|^2). \tag{7.24.1}$$

Hint: As an inner product, consider the construction

$$(x,y) = \frac{1}{4}\{\|x+y\|^2 - \|x-y\|^2\} + \frac{i}{4}\{\|x+iy\|^2 - \|x-iy\|^2\}, \tag{7.24.2}$$

where $i = \sqrt{-1}$. Prove first that (x,y) is continuous w.r.t. its both variables. Then use that fact that any real number can be approximated by a sequence of rational numbers.

Prove the same for real normed linear spaces.

Problem 7.25. Prove that in l_p, $p \neq 2$, it is impossible to introduce an inner product agreed with the norm

$$\|x\|_p = \left(\sum_{k=1}^{+\infty} |x_k|^p\right)^{\frac{1}{p}}.$$

Hint: Use the parallelogram identity.

Problem 7.26. Prove that in $L_p[a,b]$, $p \neq 2$, it is impossible to introduce an inner product agreed with the norm

$$\|f\|_p = \left(\int_a^b |f(x)|^p d\mu\right)^{\frac{1}{p}}.$$

Problem 7.27. Prove that the space $C[a,b]$ of continuous functions on $[a,b]$ with norm

$$\|f\| = \max_{x\in[a,b]} |f(x)|$$

cannot be a Hilbert space.

Problem 7.28. Prove that in a Euclidean (unitary) space X the following identity holds

$$\|z-x\|^2 + \|z-y\|^2 = \frac{1}{2}\|x-y\|^2 + 2\left\|z - \frac{x+y}{2}\right\|^2$$

for any $x, y, z \in X$. This identity is called the Apollonius identity.

Problem 7.29. Let X be the linear space of functions $x(t)$ defined on $(-\infty, \infty)$ and satisfying the condition

$$\int_{-\infty}^{+\infty} |x(t)|^2 e^{-t^2} dt.$$

Prove that the space X with the inner product

$$\int_{-\infty}^{+\infty} x(t)\overline{y(t)} e^{-t^2} dt, \qquad x, y \in X$$

is a Hilbert space.

Problem 7.30. Prove that in an infinite dimensional Hilbert space H any closed unit ball contains infinitely many *non-intersecting* closed balls with radius $\dfrac{1}{4}$.

A closed ball $B(a,r)$ in H of radius r with the centre at a is the set $B(a,r) = \{x \in H \ : \ \|x-a\| \leqslant r\}$.

Hint: Use the fact that any infinite dimensional Hilbert space contains an infinite orthonormal system of elements (due to Gram-Schmidt process).

Problem 7.31. Let L be a subspace of a Hilbert space H. The subspace $L^\perp = \{x \in H \ : \ (x,y) = 0 \ \ \forall y \in L\}$ of H is called the orthogonal complement of L. Prove that $L^\perp$ is closed.

Problem 7.32. Let L be a *closed* subspace of a Hilbert space H. Prove that $H = L \oplus L^\perp$, that is, any $x \in H$ can be *uniquely* represented in the form $x = u + v$ where $u \in L$ and $v \in L^\perp$ and $\|x-u\| = \|v\|$. In this case, u is the projection of x onto L.

Problem 7.33. Let L be a subspace of a Hilbert space H. Prove that L is dense in H (i.e. the closure of L coincides with H: $\overline{L} = H$) if, and only if, $L^\perp = \{0\}$.

Remind that the closure of L is the set consisting of all points of L and all limit points of L.

Problem 7.34. In the space $L_2[-1,1]$, construct projections of any function $x \in L_2[-1,1]$ onto the subspaces of even and odd functions.

Problem 7.35. Prove that for a fixed $n \in \mathbb{N}$ the set

$$L_n = \left\{ x \in l_2, x = (\xi_1, \xi_2, \ldots) \ : \ \sum_{k=1}^{n} \xi_k = 0 \right\}$$

is a closed subspace of l_2. Describe $L_n^{\perp}$ and find the distance between $e_1 = (1,0,0,\ldots)$ and L:

$$\rho(e_1, L_n) = \inf_{y \in L_n} \|y - e_1\|_{l_2}.$$

Find also $\lim\limits_{n\to\infty} \rho(e_1, L_n)$.

Problem 7.36. For the function e^t, find the polynomial $p(t)$ of degree 2 such that the norm $\|e^t - p(t)\|$ is minimal in $L_2[-1,1]$.

Problem 7.37. Prove that the system $(t^n)_{n=1}^{\infty}$ is not a basis in $L_2[0,1]$.

Hint: Prove that if $(t^n)_{n=1}^{\infty}$ were a basis in $L_2[0,1]$, then any function $f(t) \in L_2[0,1]$ would be infinitely differentiable. To do this, consider the integral $g(s) = \displaystyle\int_0^s f(t)dt$ as a function on $[0,1]$, where f is an arbitrary function on $L_1[0,1]$, and prove that $g(s)$ is an infinitely differentiable function whenever $(t^n)_{n=1}^{\infty}$ is a basis.

Problem 7.38. Prove that the system $e_n(t) = \dfrac{\sqrt{2}}{\sin \mu_n} \sin \mu_n t$, where μ_n are *positive* solutions of the equation $\tan \mu = \mu$, is an orthogonal system in $L_2[0,1]$.

Problem 7.39. Let $\alpha = (\alpha_1, \alpha_2, \ldots)$ be a sequence of *positive* numbers: $\alpha_n > 0$, $n \in \mathbb{N}$, and let $l_{2,\alpha}$ be a space of complex sequences $x = (x_1, x_2, \ldots)$ satisfying the condition

$$\sum_{k=1}^{\infty} \alpha_k |x_k|^2 < +\infty.$$

Prove that the space $l_{2,\alpha}$ with the inner product

$$\sum_{k=1}^{\infty} \alpha_k x_k \overline{y_k} < +\infty$$

is a complex separable Hilbert space. Construct an orthonormal basis in this space if

1) $\alpha_n = e^{-n}$, $n \in \mathbb{N}$;
2) $\alpha_n = n$, $n \in \mathbb{N}$;
3) $\alpha_n = n^2$, $n \in \mathbb{N}$.

Problem 7.40. Does the mapping f acting on the space X define a linear and/or continuous functional:

1)
$$f(x) = x(2) + \int_{-2}^{2} t x(t) dt, \qquad X = C[-2,2];$$

2)
$$f(x) = x_1 - 4x_3, \qquad X = l_3;$$

3)
$$f(x) = \int_0^1 x^2(t)dt, \qquad X = C[0,1];$$

4)
$$f(x) = \int_0^1 x^2(t)dt, \qquad X = L_2[0,1];$$

5)
$$f(x) = \int_0^1 x(t)\sin^2 t dt, \qquad X = L_2[0,1];$$

Problem 7.41. Does the mapping $f(x) = x'(0)$ acting on the set L of differentiable functions in $C[-1,1]$ define a linear continuous functional?

Problem 7.42. Does the mapping $f(x) = \sum_{k=1}^{\infty} x_k$ acting on the subspace $L = \{x \in l_2 \ : \ \left|\sum_{k=1}^{\infty} x_k\right| < \infty\}$ define a linear continuous functional?

Hint: Consider the sequence $x_n = \left(\underbrace{\frac{1}{n}, \frac{1}{n}, \dots, \frac{1}{n}}_{n}, 0, 0, \dots\right)$.

Problem 7.43. Let $\mathbb{R}^n_p$ be the linear normed space of n-dimensional vectors $x = (x_1, \dots, x_n)$ with the norm
$$\|x\|_p = \begin{cases} \left(\sum\limits_{k=1}^{n} |x_k|^p\right)^{\frac{1}{p}} & \text{if} \quad 1 \leqslant p < +\infty; \\ \max\limits_{1 \leqslant k \leqslant n} |x_k| & \text{if} \quad p = +\infty. \end{cases}$$
Find the general representation of linear functionals in $\mathbb{R}^n_p$.

Problem 7.44. Prove that any linear functional defined on a finite-dimensional space is continuous.

Problem 7.45. Prove that a linear functional in a normed space is continuous if and only if its kernel is closed.

Problem 7.46. Let f be a linear functional defined on a linear normed space L. Prove that f is continuous if and only if the sets $\{x \in L \ : \ f(x) < c\}$ and $\{x \in L \ : \ f(x) > c\}$ are open in L.

Problem 7.47. Find the norm of the following functionals

1)
$$f(x) = \frac{x(\varepsilon) + x(-\varepsilon) - 2x(0)}{\varepsilon^2} \qquad \text{on} \quad C[-1,1], \qquad \varepsilon \in (0,1) \quad \text{is fixed;}$$

2)
$$f(x) = \int_{-1}^{1} x(t)dt - \frac{1}{2n+1}\sum_{k=-n}^{n} x\left(\frac{k}{n}\right) \qquad \text{on} \quad C[-1,1], \qquad n \in \mathbb{N} \quad \text{is fixed;}$$

3)
$$f(x) = \sum_{k=1}^{\infty} [1 - (-1)^k]\frac{k-1}{k} x_k \qquad \text{where} \quad x = (x_k)_{k=1}^{\infty} \in l_1;$$

4)
$$f(x) = \sum_{k=1}^{\infty} \frac{x_k}{\sqrt{k(k+1)}} \qquad \text{where} \quad x = (x_k)_{k=1}^{\infty} \in l_2.$$

5)
$$f(x) = \int_0^{\frac{1}{2}} \sqrt{t}x(t^2)dt \qquad \text{on} \quad L_2[0,1];$$

6)
$$f(x) = \int_0^{\frac{1}{2}} x(t)\,\mathrm{sgn}\left(t - \frac{1}{2}\right) dt \qquad \text{on} \quad L_2[0,1];$$

Hint: To prove that a number M satisfying $|f(x)| \leqslant M\|x\|$ is the norm of the functional f, try to find an element x ($\|x\| = 1$) such that $f(x) = M$ or a sequence $(x_n)_{n=1}^{\infty}$ such that $\dfrac{f(x_n)}{\|x_n\|} > M - \varepsilon_n$, and $\varepsilon_n \to 0$ as $n \to \infty$.

Chapter 8

Divergent series. Cesàro and Abel methods of summation.

8.1 Divergent series.

By definition, a number series

$$S = \sum_{m=0}^{\infty} a_m$$

converges to the number S if its partial sums

$$S_n = \sum_{m=0}^{n-1} a_m$$

have *finite* limit as $n \to \infty$ which is equal S.

According to this definition the series

$$1 - 1 + 1 - 1 + 1 - 1 + \cdots \tag{8.1.1}$$

diverges. Indeed, its partial sums S_n

$$S_1 = 1, \quad S_2 = 0, \quad S_3 = 1, \quad S_4 = 0, \ldots,$$

that is,

$$S_{2k+1} = 1, \quad S_{2k} = 0, \quad k \in \mathbb{N}.$$

However, if we follow Euler and set (formally)

$$a = 1 - 1 + 1 - 1 + 1 - 1 + \cdots\,,$$

then we obtain

$$a = 1 - (1 - 1 + 1 - 1 + 1 - 1 + \cdots) = 1 - a,$$

so $2a = 1$, and $a = \dfrac{1}{2}$.

We can get the same result in another way. To do this consider the function

$$f(z) = \sum_{m=0}^{\infty} z^m = \frac{1}{1-z}. \tag{8.1.2}$$

This series converges for $|z| < 1$ but the function $f(z)$ exists for $z = -1$, so we get

$$a = f(-1) = \sum_{m=0}^{\infty} (-1)^m = \frac{1}{1-(-1)} = \frac{1}{2}.$$

However, the situation is not so good as it seems from the first view, because it can be shown that the "sum" a of the series (8.1.1) can be a number different from $\frac{1}{2}$. Indeed, let us represent the series (8.1.1) as follows

$$a = 1 - 0 - 1 + 1 - 0 - 1 + 1 - 0 - 1 + \cdots \tag{8.1.3}$$

and consider the function

$$g(z) = \sum_{m=1}^{\infty} z^{3m} - \sum_{m=1}^{\infty} z^{3m+2} = \frac{1}{1-z^3} - \frac{z^2}{1-z^3} = \frac{1-z^2}{1-z^3} = \frac{1+z}{1+z+z^2}. \tag{8.1.4}$$

From (8.1.3)–(8.1.4) it follows that

$$a = g(1) = \frac{2}{3}.$$

Moreover, it is clear that for any $m, n \in \mathbb{N}$, $m < n$,

$$g_{mn}(x) = \frac{1 + x + \cdots + x^m}{1 + x + \cdots + x^n} = \frac{1-x^m}{1-x^n} = 1 - x^m + x^n - x^{m+n} + x^{2n} - \cdots ,$$

so

$$a = g_{mn}(1) = \frac{m}{n}.$$

Additionally, if we substitute $z = -2$ into (8.1.2), then we get

$$f(-2) = 1 - 2 + 2^2 - 2^3 + \cdots = \sum_{n=0}^{+\infty} (-2)^n = \frac{1}{1+2} = \frac{1}{3}. \tag{8.1.5}$$

Here the "sum" of a series with integer terms is fractional. There is no paradox here, since $\frac{1}{3}$ is not the sum in usual sense. This is just a number functionally dependent on the series (8.1.5). Moreover, if we take $z = 2$ in (8.1.2), then we obtain a really surprising thing

$$f(2) = 1 + 2 + 2^2 + 2^3 + \cdots = \sum_{n=0}^{+\infty} 2^n = \frac{1}{1-2} = -1. \tag{8.1.6}$$

Thus, without specific rules regarding "summing" of divergent series, we can obtain a lot of paradoxical results.

8.1.1 General methods of summing

Let $\mathcal{D}$ be the set of all number series, $\mathcal{D}_c (\subset \mathcal{D})$ be the set of all convergent series, and $\mathcal{D}_a (\subset \mathcal{D}_c)$ be the set of all absolute convergent series. Let $S : \mathcal{D}_c \mapsto \mathbb{C}$ denote the ordinary summation operator defined on the set $\mathcal{D}_c$. To every convergent series S corresponds its sum: $\forall \sigma \in \mathcal{D}_c \ \exists S(\sigma) \in \mathbb{C}$, where $S(\sigma)$ is the sum of the series σ.

Consider now some other summation operator S^* defined on $\mathcal{B}^* \subset \mathcal{D}$. It is natural to assume that $\mathcal{D}_c \subset \mathcal{B}^*$. Then the operator S^* must possess the following properties:

1) *Regularity*:

$$\forall \sigma \in \mathcal{D}_c \qquad S^*(\sigma) = S(\sigma).$$

2) *Shift invariance*:

$$S^*\left(a_0 + \sum_{m=1}^{\infty} a_n\right) = a_0 + S^*\left(\sum_{m=1}^{\infty} a_n\right).$$

3) *Linearity*: $\forall \alpha, \beta \in \mathbb{C}$, and $\forall \sigma_1 \sigma_2 \in \mathcal{B}$

$$S^*(\alpha\sigma_1 + \beta\sigma_2) = \alpha S^*(\sigma_1) + \beta S^*(\sigma_2).$$

The second condition is less important, and some significant methods, such as Borel summation (see below), do not possess it.

Sometimes S^* possess an additional property

4) $S^* : \mathcal{B}^* \mapsto \mathbb{C}$ *is a homomorphism.*

This means that S^* preserves the product of series. For example, the operator S of standard summation is a homomorphism on the set $\mathcal{D}_a$, since the sum of the product of two absolutely convergent series is equal to the product of their sums.

Two summation methods (operators) S_1 and S_2 defined on sets $\mathcal{B}_1$ and $\mathcal{B}_2$ are called *consistent* and S_2 is stronger than S_1 if $\mathcal{B}_1 \subset \mathcal{B}_2 \subset \mathcal{D}$, and

$$S_2(\sigma) = S_1(\sigma) \qquad \forall \sigma \in \mathcal{B}_1.$$

In fact, it is possible to construct a hierarchy of **some** *regular* summation methods. That is, we can find a sequence of summation operators S_k defined on sets $\mathcal{B}_k$ such that

$$\mathcal{D}_a \subset \mathcal{D}_c \subset \mathcal{B}_1 \subset \mathcal{B}_2 \subset \cdots \subset \mathcal{D}$$

and

$$S_{k+1}(\sigma) = S_k(\sigma) \qquad \forall \sigma \in \mathcal{B}_k$$

for $k = 0, 1, 2, \ldots$. Here $S_0 = S$ is the standard summation method.

Of course, it is possible to construct summation operators that are incomparable. In what follows, we will consider a hierarchy of a sequence of summation operators.

Let us study two the most important for us summability methods.

8.1.2 Cesàro summability method

Let $(s_n)_{n=1}^{\infty}$ be a sequence of numbers. Then the numbers

$$\sigma_N = \frac{s_1 + s_2 + \cdots + s_N}{N} \tag{8.1.7}$$

are called *Cesàro means*. The sequence $(s_n)_{n=1}^{\infty}$ is called *Cesàro convergent* if the sequence $(\sigma_N)_{N=1}^{\infty}$ of Cesàto means has a finite limit.

Respectively, the series

$$\sum_{m=0}^{\infty} a_n$$

are called *Cesàro summable* if the sequence of its partial sums is Cesàro convergent.

For instance, for the series (8.1.1), Cesàro means of the sequence of its partial sums have the form

$$\sigma_{2k} = \frac{1}{2}, \qquad \sigma_{2k-1} = \frac{k}{2k-1}, \qquad k = 1, 2, 3, \ldots.$$

It is clear that $\lim_{N\to\infty} \sigma_N = \dfrac{1}{2}$, so the sum $a = \dfrac{1}{2}$ of the series (8.1.1) obtained by Euler is the Cesàro sum of this series.

Note that the Cesàro summability method possesses all three properties of generalized summation operators.

1) *Cesàro summation method is regular.*

   Indeed, let a series $\sum\limits_{m=0}^{\infty} a_m$ converges to a number A. Suppose first that $A = 0$, then for any $\delta > 0$ there exists a number $M_1 > 0$ such that for any $n > M_1$ one has $|S_n| < \delta$, where $S_n = \sum\limits_{m=0}^{n-1} a_m$. Moreover, since the sequence S_n is convergent, it is bounded, so there exists a number $B > 0$ such that that $|S_n| < B$ for all $n \in \mathbb{N}$. Choose $\varepsilon > \delta > 0$. Then there exists a number $M_2 > 0$ such that $\forall n > M_2$

$$\frac{M_1(B-\delta)}{n} < \varepsilon - \delta.$$

   Let $N > \max\{M_1, M_2\}$, then we have

$$|\sigma_N| = \left|\frac{S_1 + S_1 + \cdots + S_N}{N}\right| \leqslant \frac{\sum\limits_{n=1}^{M_1} |S_n|}{N} + \frac{\sum\limits_{n=M_1+1}^{N} |S_n|}{N} \leqslant \frac{M_1(B-\delta)}{N} + \delta < \varepsilon.$$

   If $A \neq 0$, we can consider a new series $\sum\limits_{m=0}^{\infty} \widetilde{a}_m$ such that $\widetilde{a}_0 = a_0 - A$, and $\widetilde{a}_m = a_m$, $m \geqslant 1$. Then we have that the sum $\widetilde{A}$ of this series equals zero, so $\widetilde{\sigma}_N \to 0$ as $N \to \infty$. But clearly $\widetilde{S}_n = S_n - A$ for all n, so $\widetilde{\sigma}_N = \sigma_N - A$. Therefore, $\sigma_N \to A$ as $N \to \infty$.

2) *Cesàro summation method is shift invariant.*

   This property was, in fact, established in the previous item.

3) *Cesàro summation method is linear.*

   This property is obvious because of linearity of the Cesàro means (8.1.7).

However, Cesàro summation operator is not a homomorphism as the following example shows.

Example 8.1.1. As we found above the series (8.1.1) is Cesàro summable to $\frac{1}{2}$. Consider the square of this series:

$$b = a^2 = \left(\sum_{m=0}^{\infty} (-1)^m\right)^2 = \sum_{m=0}^{\infty} \underbrace{((-1)^m + (-1)^m + \cdots + (-1)^m)}_{m+1} = \sum_{m=0}^{\infty} (-1)^m (m+1).$$

The partial sum of this series are

$$S_{2k} = \sum_{m=0}^{2k-1} (-1)^m (m+1) = -k, \qquad S_{2k-1} = \sum_{m=0}^{2k-2} (-1)^m (m+1) = k, \qquad k = 1, 2, \ldots.$$

The series b is not summable, since $S_{2k} \to +\infty$ and $S_{2k-1} \to -\infty$ as $k \to +\infty$. However, this series is also not Cesàro summable. Indeed, its Cesàro means have the form

$$\sigma_N = \frac{S_1 + S_1 + \cdots + S_N}{N} = \begin{cases} 0, & N = 2k, \\ \dfrac{k+1}{2k+1}, & N = 2k+1, \end{cases}$$

so the sequence σ_N has no limit, and b is not Cesàro summable.

Theorem 8.1.2. *If a series $\sum\limits_{n=0}^{\infty} a_n$ is Cesàro summable, then $a_n = o(n)$ as $n \to \infty$.*

Proof. Indeed, if $\sigma_N \to \sigma$ as $N \to \infty$, then $\dfrac{N+1}{N}\sigma_N \to \sigma$. Therefore, we have

$$\frac{(N+1)\sigma_{N+1} - N\sigma_N}{N} = \frac{S_{N+1}}{N} \underset{N\to\infty}{\longrightarrow} \sigma - \sigma = 0,$$

where $S_N = \sum\limits_{m=0}^{N-1} a_m$. Thus, we obtain

$$\frac{a_n}{n} = \frac{S_{n+1} - S_n}{n} = \frac{S_{n+1}}{n} - \frac{n-1}{n} \cdot \frac{S_n}{n-1} \underset{n\to\infty}{\longrightarrow} 0 - 1 \cdot 0 = 0.$$

□

It is clear that the series $\sum\limits_{m=0}^{\infty} (-1)^m (m+1)$ is non-Cesàro convergent according to this theorem.

The next summability method possesses all four properties of summability methods we mentioned above.

8.1.3 Abel summability method

For a given series

$$\sum_{m=0}^{\infty} a_m \tag{8.1.8}$$

consider the function

$$A(r) = \sum_{m=0}^{\infty} a_m r^m.$$

where the series $A(r)$ is supposed to be convergent for $0 \leqslant r < 1$. The series (8.1.8) is called *Abel summable* if there exists the final limit

$$S_A = \lim_{r\to 1-0} A(r).$$

In this case, the number S_A is called the *Abel sum* of the series (8.1.8).

First, note that the Abel summability method possesses all three properties of generalized summation operators.

1) *Abel summation method is regular.*

   This property can be proved directly. But it also follows from Theorem 8.2.2 below.

2) *Abel summation method is shift invariant.*

   Indeed,

$$\lim_{r\to 1-0} \sum_{m=0}^{\infty} a_m r^m = a_0 + \lim_{r\to 1-0} r \sum_{m=1}^{\infty} a_m r^{m-1} = a_0 + \lim_{r\to 1-0} \sum_{m=0}^{\infty} a_{m+1} r^m$$

3) *Abel summation method is linear.*

   This property is obvious.

Let us check whether the series b from Example 8.1.1 is Abel summable. For $0 \leqslant r < 1$, consider (formally) the series

$$A(r) = \sum_{m=0}^{\infty} (-1)^m (m+1) r^m = \sum_{m=0}^{\infty} (m+1)(-r)^m = \frac{d}{dz} \sum_{m=0}^{\infty} z^m \Bigg|_{z=-r} = \frac{d}{dz}\, \frac{1}{1-z}\Bigg|_{z=-r} = \frac{1}{(1+r)^2}.$$

Since we supposed that $r \in [0,1)$, the differentiation of the series is possible. Moreover, the series $A(r)$ converges for $r \in [0,1)$, and we have

$$S_A = \lim_{r \to 1-0} \frac{1}{(1+r)^2} = \frac{1}{4}.$$

The Abel summation is related to the analytic continuation of functions.

Example 8.1.3. Consider the Riemann zeta-function

$$\zeta(s) = \sum_{n=1}^{+\infty} \frac{1}{n^s}.$$

This series converges for $\operatorname{Re} s > 1$, but the function $\zeta(s)$ can be analytically continued to the domain $\mathbb{C} \setminus \{1\}$ where it is meromorphic. For example, by analytic continuation it is known that

$$\zeta(-1) = \sum_{m=1}^{\infty} m = -\frac{1}{12}.$$

We can get this result using the previous example. Indeed,

$$-3\zeta(-1) = \zeta(-1) - 4\zeta(-1) = \sum_{m=1}^{\infty} m - \sum_{m=1}^{\infty} 4m = 1 + (2-4) + 3 + (4-8) + 5 + (6-12) + \cdots$$

Therefore,

$$-3\zeta(-1) = \sum_{m=0}^{\infty} (-1)^m (m+1) \overset{A}{=} \frac{1}{4},$$

thus

$$\zeta(-1) \overset{A}{=} -\frac{1}{12}.$$

The following example will be useful later.

Example 8.1.4. Consider the series

$$\frac{1}{2} + \sum_{n=1}^{+\infty} \cos n\theta, \qquad \theta \neq 0. \tag{8.1.9}$$

The limit $\lim\limits_{n\to\infty} \cos n\theta$ does not exist, so the series (8.1.9) diverges. Let us find out whether this series Abel summable. To do this, consider the series

$$\frac{1}{2} + \sum_{n=1}^{+\infty} r^n \cos n\theta, \tag{8.1.10}$$

where $r \in [0,1)$. The series of absolute values of its terms

$$\frac{1}{2} + \sum_{n=1}^{+\infty} r^n |\cos n\theta|,$$

converges, for instance, by Cauchy test, since $\sqrt[n]{r^n |\cos n\theta|} = r \sqrt[n]{|\cos n\theta|} \leqslant r < 1$. Thus, the series (8.1.10) converges absolutely, and therefore converges. Let us find its sum. Note that

$$\frac{1}{2} + \sum_{n=1}^{+\infty} r^n \cos n\theta = \operatorname{Re}\left(\frac{1}{2} + \sum_{n=1}^{+\infty} r^n e^{in\theta}\right) = \operatorname{Re}\left(\frac{1}{2} + \frac{re^{i\theta}}{1 - re^{i\theta}}\right).$$

It is easy to see that

$$\frac{1}{2}+\sum_{n=1}^{+\infty} r^n e^{in\theta}=\frac{1}{2}+\frac{re^{i\theta}(1-re^{-i\theta})}{(1-re^{i\theta})(1-re^{-i\theta})}=\frac{1}{2}+\frac{re^{i\theta}-r^2}{1-2r\cos\theta+r^2}. \tag{8.1.11}$$

Taking the real part of this expression we have

$$\frac{1}{2}+\sum_{n=1}^{+\infty} r^n\cos n\theta=\frac{1}{2}+\frac{r\cos\theta-r^2}{1-2r\cos\theta+r^2}=\frac{1}{2}\cdot\frac{1-r^2}{1-2r\cos\theta+r^2}.$$

Thus, the Abel sum of the series (8.1.9) is the following.

$$\lim_{r\to 1-0}\operatorname{Re}\left(\frac{1}{2}+\sum_{n=1}^{+\infty} r^n e^{in\theta}\right)=\lim_{r\to 1-0}\frac{1}{2}\cdot\frac{1-r^2}{1-2r\cos\theta+r^2}=0, \qquad \theta\neq 0.$$

Furthermore, the series

$$\sum_{n=1}^{\infty}\sin n\theta, \qquad \theta\neq 0, \tag{8.1.12}$$

diverges. However, the function

$$\sum_{n=1}^{\infty} r^n\sin n\theta, \qquad 0\leqslant r<1,$$

is the imaginary part of the series (8.1.11), so

$$\sum_{n=1}^{\infty} r^n\sin n\theta=\operatorname{Im}\left(\sum_{n=1}^{\infty} r^n e^{in\theta}\right)=\frac{r\sin\theta}{1-2r\cos\theta+r^2},$$

and the Abel sum of the series (8.1.12) has the form

$$\lim_{r\to 1-0}\sum_{n=1}^{\infty} r^n\sin n\theta=\lim_{r\to 1-0}\frac{r\sin\theta}{1-2r\cos\theta+r^2}=\frac{1}{2}\cdot\frac{\sin\theta}{1-\cos\theta}=\frac{1}{2}\cdot\frac{2\sin\frac{\theta}{2}\sin\frac{\theta}{2}}{2\sin^2\frac{\theta}{2}}=\frac{1}{2}\cot\frac{\theta}{2}$$

Note that the series

$$\sum_{n=0}^{\infty} q^n, \qquad q>1,$$

is not Abel summable, since

$$\lim_{r\to 1-0}\sum_{n=0}^{\infty} r^n q^n=+\infty.$$

Thus, the series (8.1.5) and (8.1.6) are not Abel summable. Later we will introduce the Euler summability method that allow to "sum" the series (8.1.5).

8.1.4 Product of series

Historically, one of the first methods of "summing" divergent series was the Abel (Abel–Poisson) summation method which appeared to be useful for the product of series. Implicitly, Leibnitz and Euler used this method.

So, it is known that if two series $\sum\limits_{n=0}^{\infty} a_n$ and $\sum\limits_{n=0}^{\infty} b_n$ converge *absolutely* to some numbers A and B, respectively, then their (Cauchy) product

$$\sum_{n=0}^{\infty} a_n\cdot\sum_{n=0}^{\infty} b_n=\sum_{n=0}^{\infty}(a_0b_n+a_1b_{n-1}+\cdots+a_{n-1}b_1+a_nb_0)$$

converges to the number $C=A\cdot B$.

Example 8.1.5. Consider the series

$$\frac{1}{\sqrt{2}} = \sum_{m=0}^{\infty} (-1)^m \frac{(2m-1)!!}{(2m)!!} = 1 - \frac{1}{2} + \frac{1\cdot 3}{2\cdot 4} - \cdots\,, \tag{8.1.13}$$

that can be obtained from the functional series

$$\frac{1}{\sqrt{1-x}} = \sum_{m=0}^{\infty} \frac{(2m-1)!!}{(2m)!!} x^m, \tag{8.1.14}$$

whose convergence radius is 1. Indeed, from the Stirling formula

$$m! \approx \sqrt{2\pi m}\cdot\left(\frac{m}{e}\right)^m \qquad \text{as} \quad m \to +\infty$$

we have

$$a_m = \frac{(2m-1)!!}{(2m)!!} = \frac{1}{4^m}\cdot\frac{(2m)!}{m!\,m!} = \frac{1}{4^m}\binom{2m}{m} \approx \frac{1}{4^m}\cdot\frac{2\sqrt{\pi m}\cdot 4^m\cdot m^{2m}}{e^{2m}}\cdot\frac{e^{2m}}{2\pi m\cdot m^{2m}} = \frac{1}{\sqrt{\pi m}},$$

so

$$\frac{1}{R} = \overline{\lim_{m\to\infty}}\sqrt[m]{a_m} = \overline{\lim_{m\to\infty}}\sqrt[m]{\frac{1}{\sqrt{\pi m}}} = 1.$$

Thus, the series (8.1.14) converges absolutely for $|x| < 1$. It obviously diverges for $x = 1$, but it converges for $x = -1$. Indeed, since

$$a_m - a_{m+1} = \frac{1}{4^m}\cdot\frac{(2m)!}{m!\,m!} - \frac{1}{4^{m+1}}\cdot\frac{(2m+2)!}{(m+1)!(m+1)!} = a_m\cdot\frac{1}{2(m+1)} > 0,$$

and

$$a_m \approx \frac{1}{\sqrt{\pi m}} \underset{m\to\infty}{\longrightarrow} 0,$$

the series (8.1.13) converges by Leibnitz's test. But this series is not absolutely convergent! We prove now that its square is a divergent series. To do this let us recall the Chu-Vendermond identity

$$\sum_{m=0}^{n}\binom{\alpha}{k}\binom{\beta}{n-k} = \binom{\alpha+\beta}{n}.$$

If $\alpha = \beta = -\dfrac{1}{2}$, then we have

$$\binom{\alpha+\beta}{n} = \binom{-1}{n} = \frac{(-1)(-2)\cdots(-n)}{n!} = \frac{(-1)^n\cdot n!}{n!} = (-1)^n,$$

and

$$\binom{-\frac{1}{2}}{k} = \frac{\left(-\frac{1}{2}\right)\left(-\frac{1}{2}-1\right)\cdots\left(-\frac{1}{2}-k+1\right)}{k!} = \frac{(-1)^k(2k-1)!!}{2^k\cdot k!} = \frac{(-1)^k(2k)!}{2^k\cdot k!\,2^k\cdot k!} = \frac{(-1)^k}{4^k}\binom{2k}{k},$$

$$\binom{-\frac{1}{2}}{n-k} = \frac{(-1)^{n-k}}{4^{n-k}}\binom{2(n-k)}{n-k}.$$

So from the Chu-Vandermond identity it follows that

$$\sum_{m=0}^{n}\binom{-\frac{1}{2}}{k}\binom{-\frac{1}{2}}{n-k} = \binom{-1}{n},$$

or

$$\sum_{m=0}^{n} \binom{2k}{k}\binom{2n-2k}{n-k} = 4^n.$$

Consider now the square of the series (8.1.13):

$$\left(\sum_{m=0}^{\infty}(-1)^m a_m\right)^2 = \sum_{n=0}^{\infty}(-1)^n \sum_{k=0}^{n} a_k a_{n-k} = \sum_{n=0}^{\infty}(-1)^n \sum_{k=0}^{n} \frac{1}{4^k}\binom{2k}{k}\frac{1}{4^{n-k}}\binom{2n-2k}{n-k} = \sum_{n=0}^{\infty}(-1)^n.$$

However, the square of the left-hand side of (8.1.13) equals $\frac{1}{2}$. An we know that this is the Abel (and Cesàro) sum of the series $\sum\limits_{n=0}^{\infty}(-1)^n$.

The following theorem explain this phenomenon.

Theorem 8.1.6. *If series $\sum\limits_{n=0}^{\infty} a_n$ and $\sum\limits_{n=0}^{\infty} b_n$ are Abel summable with sums A and B, respectively, then their product is Abel summable with sum $A \cdot B$.*

Thus, the Abel summation operator is a homomorphism.

Proof. By assumption the series

$$\sum_{n=0}^{\infty} a_n r^n \qquad \text{and} \qquad \sum_{n=0}^{\infty} b_n r^n$$

converge for any $r \in [0,1)$. Therefore, the series

$$A(z) := \sum_{n=0}^{\infty} a_n z^n \qquad \text{and} \qquad B(z) := \sum_{n=0}^{\infty} b_n z^n$$

converge absolutely for $|z| < 1$ by Abel's theorem (from the power series theory). So we have

$$\lim_{r\to 1-0}\left(\sum_{n=0}^{\infty} a_n r^n\right)\left(\sum_{n=0}^{\infty} b_n r^n\right) = \lim_{r\to 1-0} A(r)\cdot B(r) = A\cdot B.$$

□

This theorem confirms that the square of the series (8.1.13) is Abel summable to $\frac{1}{2}$.

Example 8.1.7. Consider again the series

$$\sum_{m=0}^{\infty}(-1)^m.$$

It is easy to see that

$$\left(\sum_{m=0}^{\infty}(-1)^m\right)^2 = \sum_{m=0}^{\infty}\left(\underbrace{(-1)^m + (-1)^m + \cdots (-1)^m}_{m\text{times}}\right) = \sum_{m=0}^{\infty}(-1)^{m-1}m.$$

The initial series is Cesàro summable to $\frac{1}{2}$ but its square is not Cesàro summable. It is Abel summable to $\left(\frac{1}{2}\right)^2$.

Corollary 8.1.8. *If series $\sum\limits_{n=0}^{\infty} a_n$ and $\sum\limits_{n=0}^{\infty} b_n$ are summable with sums A and B, respectively, then their product is Abel summable with sum $A \cdot B$.*

8.2 Tauberian theorems. Higher methods of summation of divergent series.

8.2.1 Relation between Cesàro and Abel summability methods

To find out how Cesàro and Abel summability methods are related, we need the so-called *Abel summation formula.*

Theorem 8.2.1 (Abel summation formula)**.** *Let $(a_n)_{n=0}^N$ and $(b_n)_{n=0}^N$ be finite number sequences. Then*

$$\sum_{n=M}^{N} a_n b_n = a_N B_{N+1} - a_M B_M - \sum_{n=M}^{N-1} (a_{n+1} - a_n) B_{n+1}. \tag{8.2.1}$$

where

$$B_0 = 0, \quad B_m = \sum_{n=0}^{m-1} b_n$$

Proof. Since $b_m = B_{m+1} - B_m$, we have

$$\sum_{n=M}^{N} a_n b_n = \sum_{n=M}^{N} a_n (B_{n+1} - B_n) = \sum_{n=M}^{N} (a_n B_{n+1} - a_{n-1} B_n) - \sum_{n=M}^{N} (a_n B_n - a_{n-1} B_n) =$$

$$= a_N B_{N+1} - a_{M-1} B_M - \sum_{n=M-1}^{N-1} (a_{n+1} - a_n) B_{n+1} = a_N B_{N+1} - a_M B_M - \sum_{n=M}^{N-1} (a_{n+1} - a_n) B_{n+1},$$

as required. □

Theorem 8.2.2 (Frobenius)**.** *The Cesàro and Abel summation methods are consistent, and the Abel summation method is stronger.*

Proof. We have to prove that if the series (8.1.8) is Cesàro summable to a number σ, then it is Abel summable to the same number. Due to shift invariance of both methods, it is sufficient to prove the theorem in the case $\sigma = 0$.

So, by assumption the Cesàro means σ_N converge to zero as $N \to \infty$. This, in particular, means that the sequence $(\sigma_N)_{N=1}^{\infty}$ is bounded. Thus, there exists a positive number B such that $|\sigma_N| \leqslant B$ for any $N \in \mathbb{N}$, and for any $\varepsilon > 0$, there exists a number $N_1 \in \mathbb{N}$ such that for any $N \geqslant N_1$,

$$|\sigma_N| < \frac{\varepsilon}{3}. \tag{8.2.2}$$

Let $S_n = \sum\limits_{m=0}^{n-1} a_m$, $n \in \mathbb{N}$, and $S_0 := 0$. By Abel summation formula (8.2.1) we have

$$\sum_{m=0}^{N} a_m r^m = S_{N+1} r^N - S_0 - \sum_{m=0}^{N-1} (r^{m+1} - r^m) S_{m+1} = S_{N+1} r^N + (1 - r) \sum_{m=0}^{N-1} S_{m+1} r^m. \tag{8.2.3}$$

From the same formula and from (8.1.7) we get

$$\begin{aligned} \sum_{m=0}^{N-1} S_{m+1} r^m &= \sum_{m=0}^{N-1} [(m+1)\sigma_{m+1} - m\sigma_m] r^m = N\sigma_N r^{N-1} - \sum_{m=0}^{N-2} (m+1)\sigma_{m+1} (r^{m+1} - r^m) = \\ &= N\sigma_N r^{N-1} + (1 - r) \sum_{m=0}^{N-2} (m+1)\sigma_{m+1} r^m, \end{aligned} \tag{8.2.4}$$

where we supposed that $\sigma_0 = 0$. So, from (8.2.3)–(8.2.4) we obtain

$$\sum_{m=0}^{N} a_m r^m = S_{N+1} r^N + (1-r) N \sigma_N r^{N-1} + (1-r)^2 \sum_{m=0}^{N-2} (m+1)\sigma_{m+1} r^m. \tag{8.2.5}$$

Next, it is clear that

$$\frac{S_n}{n} = \sigma_n - \frac{n-1}{n}\sigma_{n-1} \underset{n\to\infty}{\longrightarrow} 0,$$

and that

$$|N\sigma_N| \leqslant BN \quad \forall N \in \mathbb{N},$$

so the sequence $N\sigma_N$ grows (if any) slower than N, that is $\sigma_N = O(N)$ and $S_N = o(N)$. Thus, for any $\varepsilon > 0$ there exists a number $N_2 \in \mathbb{N}$ such that for all $N \geqslant N_2$,

$$\left|S_{N+1} r^N\right| < \frac{\varepsilon}{3}. \qquad \text{and} \qquad |N\sigma_N r^{N-1}| < \frac{\varepsilon}{3}. \tag{8.2.6}$$

Therefore, (8.2.2), (8.2.5) and (8.2.6) imply that for any $N > \max\{N_1+2, N_2\}$,

$$\left|\sum_{m=0}^{N} a_m r^m\right| < \frac{\varepsilon}{3} + (1-r)\frac{\varepsilon}{3} + (1-r)^2 \sum_{m=0}^{N_1-1} (m+1)\sigma_{m+1} r^m + \frac{\varepsilon}{3}(1-r)^2 \sum_{m=N_1}^{N-2} (m+1) r^m \tag{8.2.7}$$

Let now $N \to \infty$. Then from (8.2.7) we get

$$|A(r)| = \left|\sum_{m=0}^{\infty} a_m r^m\right| < \frac{\varepsilon}{3} + (1-r)\frac{\varepsilon}{3} + (1-r)^2 B \sum_{m=0}^{N_1-1} (m+1) r^m + \frac{\varepsilon}{3}(1-r)^2 \frac{(1+N_1-N_1 r) r^{N_1}}{(1-r)^2} =$$

$$= (2-r)\frac{\varepsilon}{3} + (1-r)^2 B \sum_{m=0}^{N_1} (m+1) r^m + \frac{\varepsilon}{3}(1+N_1-N_1 r) r^{N_1}.$$

It is easy to see now that $|A(r)| < \varepsilon$ as $r \to 1-0$. Since ε is arbitrary, we have

$$\lim_{r\to 1-0} A(r) = 0,$$

as required. □

Example 8.2.3. Consider again the series (8.1.9). Its partial sums have the form

$$S_m(\theta) = \frac{1}{2} + \sum_{k=1}^{m-1} \cos k\theta = \frac{1}{2} + \frac{1}{\sin\frac{\theta}{2}} \sum_{k=1}^{m-1} \cos k\theta \sin\frac{\theta}{2} =$$

$$= \frac{1}{2} - \frac{1}{2\sin\frac{\theta}{2}} \sum_{k=1}^{m-1} \left[\sin\left(k-\frac{1}{2}\right)\theta - \sin\left(k+\frac{1}{2}\right)\theta\right] = \frac{\sin\left(m-\frac{1}{2}\right)\theta}{2\sin\frac{\theta}{2}},$$

so the Cesàro means are the following

$$\sigma_N(\theta) = \frac{1}{N}\sum_{m=1}^{N} S_m(\theta) = \frac{1}{2N\sin^2\frac{\theta}{2}} \sum_{m=1}^{N} \sin\left(m-\frac{1}{2}\right)\theta \sin\frac{\theta}{2} =$$

$$= \frac{1}{4N\sin^2\frac{\theta}{2}} \sum_{m=1}^{N} \left[\cos(m-1)\theta - \cos m\theta\right] = \frac{1-\cos N\theta}{4N\sin^2\frac{\theta}{2}} = \frac{1}{2N}\left(\frac{\sin\frac{N\theta}{2}}{\sin\frac{\theta}{2}}\right)^2.$$

It is easy to see that $\sigma_N(\theta) \to 0$ as $N \to \infty$ for any $\theta \neq 0$ that agrees with Example 8.1.4 and Theorem 8.2.2.

Consider again the series (8.1.12). Its partial sums have the form

$$S_m(\theta) = \sum_{k=1}^{m} \sin k\theta = \frac{1}{\sin\frac{\theta}{2}} \sum_{k=1}^{m} \sin k\theta \sin\frac{\theta}{2} =$$

$$= \frac{1}{2\sin\frac{\theta}{2}} \sum_{k=1}^{m} \left[\cos\left(k-\frac{1}{2}\right)\theta - \cos\left(k+\frac{1}{2}\right)\theta\right] = \frac{\cos\frac{\theta}{2} - \cos\left(m+\frac{1}{2}\right)\theta}{2\sin\frac{\theta}{2}},$$

so the Cesàro means are the following

$$\sigma_N(\theta) = \frac{1}{N}\sum_{m=1}^{N} S_m(\theta) = \frac{1}{2N\sin^2\frac{\theta}{2}} \sum_{m=1}^{N} \left[\cos\frac{\theta}{2} - \cos\left(m+\frac{1}{2}\right)\theta\right]\sin\frac{\theta}{2} =$$

$$= \frac{1}{2}\cot\frac{\theta}{2} + \frac{1}{4N\sin^2\frac{\theta}{2}} \sum_{m=1}^{N} \left[\sin m\theta - \sin\left(m+1\right)\theta\right] = \frac{1}{2}\cot\frac{\theta}{2} + \frac{\sin\theta - \sin(N+1)\theta}{4N\sin^2\frac{\theta}{2}}.$$

It is easy to see that $\sigma_N(\theta) \to \frac{1}{2}\cot\frac{\theta}{2}$ as $N \to \infty$ for any $\theta \neq 0$ that agrees with Example 8.1.4 and Theorem 8.2.2.

Remark 8.2.4. As we mentioned above, from this theorem it follows that the Abel summation method is regular. Indeed, if a series is summable, then it is Cesàro summable, since Cesàro summability method is regular. Now Theorem 8.2.2 guarantees that the series is Abel summable. This fact is usually known as Abel's theorem.

Corollary 8.2.5. *If series $\sum\limits_{n=0}^{\infty} a_n$ and $\sum\limits_{n=0}^{\infty} b_n$ are Cesàro summable with sums A and B, respectively, then their product is Abel summable with sum $A \cdot B$.*

Finally, let us prove a fact that is of use later.

Proposition 8.2.6. *Let the Abel partial sums $\sum\limits_{m=0}^{n-1} a_m r^m$ of a series $\sum\limits_{m=0}^{\infty} a_m$ be bounded as $r \to 1-0$. If, additionally, $a_m = O\left(\frac{1}{m}\right)$, then the partial sums $\sum\limits_{m=0}^{n-1} a_m$ of the given series are bounded.*

Proof. Let $r_N = 1 - \frac{1}{N}$. Then $N(1-r_N) = 1$, and $r_N \to 1-0$ whenever $N \to +\infty$. Moreover, since $a_m = O\left(\frac{1}{m}\right)$, there exists $M > 0$ such that $m|a_m| \leqslant M$ for any $m \in \mathbb{N}$ (for large m it follows from the assumption of the proposition, and for other, finitely many, m's we can always find such a bound).

Consider the difference

$$|S_N - A_{r_N}| = \left|\sum_{m=0}^{N-1}(a_m - a_m r_N^m) - \sum_{m=N}^{\infty} a_m r_N^m\right| \leqslant \sum_{m=0}^{N-1} |a_m| \cdot (1 - r_N^m) + \sum_{m=N}^{\infty} |a_m| \, r_N^m \leqslant$$

$$\leqslant (1 - r_N) \sum_{m=0}^{N-1} |a_m| \, (1 + r_N + \cdots + r_N^{m-1}) + \sum_{m=N}^{\infty} m \cdot |a_m| \cdot \frac{r_N^m}{m} \leqslant$$

$$\leqslant (1 - r_N) \sum_{m=0}^{N-1} m \cdot |a_m| + M \sum_{m=N}^{\infty} \frac{r_N^m}{m} \leqslant (1 - r_N) M \sum_{m=0}^{N-1} 1 + \frac{M}{N} \sum_{m=N}^{\infty} r_N^m \leqslant$$

$$\leqslant (1 - r_N) MN + \frac{M}{N} \sum_{m=0}^{\infty} r_N^m = M + \frac{M}{N} \cdot \frac{1}{1 - r_N} = 2M \quad \text{as} \quad N \to +\infty.$$

□

8.2.2 Tauberian theorems

From the properties of Abel and Cesàro summation methods we get that the Abel summation method is stronger than the Cesàro summation method, which, in its turn, is stronger than the standard summation method. This means that any convergent series is Abel and Cesàro convergent, and any Cesàro convergent series is Abel convergent. But are there some conditions on the given series such that the converse statements becomes true? In fact, such conditions exist, and the corresponding results on this topic are usually called *Tauberian* theorems after A. Tauber who proved the following theorem in 1897.

Theorem 8.2.7 (Tauber)**.** *Let the series* $\sum\limits_{m=0}^{\infty} a_m$ *be Abel summable to a number* S*, and let* $a_m = o\left(\frac{1}{m}\right)$. *Then* $\sum\limits_{m=0}^{\infty} a_m$ *is summable to* S.

Proof. Let $S_N = \sum\limits_{m=0}^{N-1} a_m$, and

$$\lim_{r\to 1-0} \sum_{m=0}^{\infty} a_m r^m = \lim_{r\to 1-0} A(r) = S. \tag{8.2.8}$$

Since $a_m = o\left(\frac{1}{m}\right)$, the sequence $(ma_m)_{m=1}^{\infty}$ is bounded, that is, there exists $B > 0$ such that $|ma_m| \leqslant B$ for any $m \in \mathbb{N}$. Moreover, for any $\varepsilon > 0$ there exists a number $N_1 \in \mathbb{N}$ such that for any $m \geqslant N_1$

$$|ma_m| \leqslant \frac{\varepsilon^2}{9B}.$$

Now from (8.2.8) it follows that for any $\varepsilon > 0$ there exists $\delta > 0$ such that

$$|A(r) - S| < \frac{\varepsilon}{3} \quad \text{whenever} \quad 0 \leqslant (1 - r) < \delta.$$

Choose a number $N_2 \in \mathbb{N}$ such that

$$\frac{\varepsilon}{3NB} < \delta \qquad \forall N \geqslant N_2.$$

Let $r := 1 - \frac{\varepsilon}{3NB}$. Then we have

$$1 - r = \frac{\varepsilon}{3NB} < \delta \quad \text{and} \quad (1 - r)N = \frac{\varepsilon}{3B}.$$

Thus, for any $N > \max\{N_1, N_2\}$, one obtains

$$|S_N - S| = \left|\sum_{m=0}^{N-1} a_m - S\right| = \left|\sum_{m=0}^{N-1} a_m(1-r^m) - \sum_{m=N}^{\infty} a_m r^m + \left(\sum_{m=0}^{\infty} a_m r^m - S\right)\right| \leqslant$$

$$\leqslant (1-r)\sum_{m=0}^{N-1} m|a_m| + \sum_{m=N}^{\infty} m|a_m|\frac{r^m}{m} + \left|\sum_{m=0}^{\infty} a_m r^m - S\right| \leqslant (1-r)NB + \frac{\varepsilon^2}{9B}\cdot\frac{1}{N(1-r)} + \frac{\varepsilon}{3} < \varepsilon.$$

□

From this theorem and from Theorem 8.2.2 we immediately obtain the following fact.

Corollary 8.2.8. *Let the series* $\sum\limits_{m=0}^{\infty} a_m$ *be Cesàro summable with the sum* σ*, and let* $a_m = o\left(\frac{1}{m}\right)$*. Then* $\sum\limits_{m=0}^{\infty} a_m$ *is summable with the sum* σ*.*

8.3 Some other summability methods

Here we briefly consider some additional regular summability methods.

8.3.1 Hölder summability method

This method is the iteration of Cesàro method. Namely, if we have a series

$$\sum_{m=0}^{\infty} a_m,$$

then its partial sums $\sigma_n^{(0)} = \sum\limits_{m=0}^{n-1} a_m$ are called initial or 0^{th} Hölder sums. Respectively, Cesàro means $\sigma_n^{(1)} = \frac{1}{n}\sum\limits_{k=1}^{n}\sigma_k^{(0)}$ are called 1^{st} Hölder sums. And in general, the l^{th} Hölder sums are

$$\sigma_n^{(l)} = \frac{1}{n}\sum_{k=1}^{n}\sigma_k^{(l-1)}.$$

So if the sequence $(\sigma_n^{(0)})_{n=1}^{\infty}$ diverges, we use the sequence $(\sigma_n^{(1)})_{n=1}^{\infty}$. If this diverges too, we use the next Hölder sums $(\sigma_n^{(2)})_{n=1}^{\infty}$, and so on.

The series is called l^{th} *Hölder summable* with the sum S if $\lim\limits_{n\to\infty}\sigma_n^{(l)} = S$.

Example 8.3.9. Consider the series

$$\sum_{m=0}^{\infty}(-1)^m(m+1).$$

As we established earlier, for $N \in \mathbb{N}$,

$$\sigma_N^{(0)} = \begin{cases} -k, & N = 2k, \\ k, & N = 2k-1, \end{cases} \qquad \sigma_N^{(1)} = \begin{cases} 0, & N = 2k, \\ \dfrac{k}{2k-1}, & N = 2k-1. \end{cases}$$

Then we have[1]

$$\sigma_N^{(2)} = \frac{1}{N}\sum_{k=1}^{\left[\frac{N+1}{2}\right]} \frac{k}{2k-1} = \frac{1}{2N}\sum_{k=1}^{\left[\frac{N+1}{2}\right]} \frac{2k-1+1}{2k-1} = \frac{1}{2N}\left(\left[\frac{N+1}{2}\right] + \sum_{k=1}^{\left[\frac{N+1}{2}\right]} \frac{1}{2k-1}\right) \leqslant$$

$$\leqslant \frac{1}{2N}\left[\frac{N+1}{2}\right] + \frac{1}{2N}\sum_{k=1}^{N}\frac{1}{k} \underset{N\to+\infty}{\longrightarrow} \frac{1}{4},$$

since the sequence $\frac{1}{k} \to 0$ as $k \to \infty$, so its Cesàro sums $\frac{1}{N}\sum\limits_{k=1}^{N}\frac{1}{k} \to 0$ as $N \to \infty$. On the other hand, from the formula above it follows that $\sigma_N^{(2)} \geqslant \frac{1}{4}$, thus $\sigma_N^{(2)} \to \frac{1}{4}$ as $N \to \infty$. Thus the series $\sum\limits_{m=0}^{\infty}(-1)^m(m+1)$ is second Hölder summable, and its Hölder sum coincides with the Abel sum of this series.

8.3.2 Voronoy summability method

Let $(p_n)_{n=1}^{\infty}$ be a sequence of some positive numbers. Construct the sequences

$$P_n = p_1 + p_2 + \cdots + p_n$$

and

$$\omega_n = \frac{p_n S_1 + p_{n-1}S_2 + \cdots + p_1 S_n}{P_n},$$

where $(S_n)_{n=1}^{\infty}$ is the sequence of partial sums of the given series.

The series is called *Voronoy summable* with the sum S if $\lim\limits_{n\to\infty}\omega_n = S$.

8.3.3 Higher Cesàro summability methods

These methods are particular cases of the Voronoy summability method and are denoted (C, l), $l = 1, 2, \ldots$. In Voronoy's method we set

$$p_n := \binom{n+l-2}{l-1}, \qquad n = 1, 2, \ldots.$$

Then we have by induction

$$P_n = \sum_{m=1}^{n}\binom{m+l-2}{l-1} = \binom{n+l-1}{l}.$$

Consider the numbers

$$\omega_n = \frac{\binom{n+l-2}{l-1}S_1 + \binom{n+l-3}{l-1}S_2 + \cdots + \binom{l-1}{l-1}S_n}{\binom{n+l-1}{l}},$$

where $(S_n)_{n=1}^{\infty}$ is the sequence of partial sums of the given series.

The series is called l^{th} *Cesàro summable* with the sum S if $\lim\limits_{n\to\infty}\omega_n = S$.

Note that the method $(C, 1)$ is the Cesàro summability method defined above.

[1] Here $[\alpha]$ denote the largest integer not exceeding α whenever $\alpha > 0$.

8.3.4 Borel summability method

For a given series $\sum_{m=0}^{\infty} a_m$, consider the function

$$F(x) = \frac{\sum_{m=0}^{\infty} S_{m+1} \frac{x^m}{m!}}{\sum_{m=0}^{\infty} \frac{x^m}{m!}} = e^{-x} \sum_{m=0}^{\infty} S_{m+1} \frac{x^m}{m!},$$

where $S_n = \sum_{m=0}^{n-1} a_m$.

The series $\sum_{m=0}^{\infty} a_m$ is called *Borel summable* with the sum S if $\lim_{x\to+\infty} F(x) = S$.

8.3.5 Euler summability method

For a given series $\sum_{m=0}^{\infty} a_m$, consider the following new series

$$\sum_{n=0}^{\infty} \frac{b_n}{2^{n+1}},$$

where

$$b_n = \binom{n}{0} a_0 + \binom{n}{1} a_1 + \cdots + \binom{n}{n} a_n = \Delta^n a_0. \tag{8.3.9}$$

Here $\Delta a_k = a_{k+1} - a_k$ is a finite difference.

The series $\sum_{m=0}^{\infty} a_m$ is called *Euler summable* with the sum S if the series $\sum_{n=0}^{\infty} b_n$ is summable (in the standard sense) with the sum S.

Note that if we formally set

$$\widehat{f}(x) = \sum_{m=0}^{\infty} a_m, \qquad \widehat{g}(y) = \sum_{n=0}^{\infty} \frac{b_n y^n}{2^{n+1}},$$

where b_n are defined in (8.3.9), then

$$\widehat{g}(y) = \widehat{f}\left(\frac{\frac{y}{2}}{1 - \frac{y}{2}}\right).$$

Example 8.3.10. Consider the series

$$\sum_{n=0}^{\infty} q^n. \tag{8.3.10}$$

Here $a_n = q^n$. The numbers b_n in the Euler summation method have the form

$$b_n = \binom{n}{0} a_0 + \binom{n}{1} a_1 + \cdots + \binom{n}{n} a_n = (1+q)^n,$$

so the Euler sum of the series (8.3.10) is the following

$$S = \sum_{n=0}^{\infty} \frac{b_n}{2^{n+1}} = \sum_{n=0}^{\infty} \frac{(1+q)^n}{2^{n+1}} = \frac{1}{1-q}.$$

Here the Euler series converges if $\left|\frac{1+q}{2}\right| < 1$, that is, if $-3 < q < 1$. Thus, the series (8.3.10) is Euler summable if $-3 < q < 1$, while it is ordinary summable only if $-1 < q < 1$. If $q = -2$ we have

$$\sum_{n=0}^{\infty}(-2)^n \overset{E}{=} \frac{1}{3}.$$

It is also easy to show that the series $\sum_{n=0}^{\infty} q^n$ is Borel summable with the (Borel) sum $\frac{1}{1-q}$ whenever $q < 1$.

8.4 Problems

Problem 8.1. Using Abel's summation formula, prove the Dirichlet test for convergence of a series: if the partial sums of the series $B_N = \sum_{m=0}^{N-1} b_m$ are bounded, and $(a_n)_{n=1}^{\infty}$ is a sequence of real numbers that decreases monotonically to 0, then the series $\sum_{n=0}^{\infty} a_n b_n$ converges.

From Dirichlet's test deduce Leibnitz theorem claiming the the series $\sum_{m=0}^{\infty} (-1)^m a_m$ converges if $(a_n)_{n=1}^{\infty}$ is a sequence of real numbers that decreases monotonically to 0.

Problem 8.2. Prove that any l^{th} Hölder summable series is $(l+1)^{\text{th}}$ Hölder summable to the same sum. Are the Hölder summability methods regular?

Problem 8.3. Prove that the Voronoy summability method is regular if, and only if, the following condition holds

$$\lim_{n\to\infty} \frac{p_n}{P_n} = 0.$$

Problem 8.4. Using the result of Problem 8.3 prove that the Cesàro summability methods (C, l), $l \in \mathbb{N}$, are regular.

Problem 8.5. Prove that the Euler summability method is regular.

Problem 8.6. Prove that the Borel summability method is regular.

Problem 8.7. Prove that if a series is k^{th} Cesàro summable, $l \in \mathbb{N}$, then it is $(k+1)^{\text{th}}$ Cesàro summable with the same sum. Show by example that converse is not true.

Hint: Consider the series $\sum_{n=1}^{\infty} (-1)^{n-1} n^k$, $k \in \mathbb{N}$.

Problem 8.8. Prove that if a series is l^{th} Cesàro summable, then it is Abel summable to the same sum.

Problem 8.9. A series $\displaystyle\sum_{n=0}^{+\infty} a_n$ is summable if, and only if, it is Cesàro summable with the same sum, and

$$u_n = a_1 + 2a_2 + \cdots + na_n = o(n) \quad \text{as} \quad n \to +\infty.$$

Problem 8.10. A series $\displaystyle\sum_{n=0}^{+\infty} a_n$ is summable if, and only if, it is Abel summable with the same sum, and

$$u_n = a_1 + 2a_2 + \cdots + na_n = o(n) \quad \text{as} \quad n \to +\infty.$$

Hint: For necessity, prove the formula $S_N - \sigma_N = \dfrac{u_n}{n}$. For sufficiency, prove that $\displaystyle\lim_{r\to 1-0} \sum_{n=1}^{+\infty} \frac{u_n r^{n+1}}{n(n+1)} = S - a_0$, where $\displaystyle S = \lim_{r\to 1-0} \sum_{n=0}^{+\infty} a_n r^n$, and use Theorem 8.1.2.

Problem 8.11. Prove that if a series is Cesàro summable (first Cesàro summable) and $ma_m > -C$, $m \in \mathbb{N}$, for some $C > 0$, then it is summable with the same sum.

Problem 8.12. Find Cesàro and Abel sums of the series

$$\sum_{n=1}^{\infty} (-1)^{n-1} n^{2k} \sin n\theta, \qquad k = 1, 2, \ldots,$$

and

$$\sum_{n=1}^{\infty} (-1)^{n-1} n^{2k} \cos n\theta, \qquad k = 0, 1, 2, \ldots$$

Problem 8.13. Find Cesàro and Abel sums of the series

$$\sum_{n=1}^{\infty}(-1)^{n-1}n^{2k+1}\sin n\theta,$$

and

$$\sum_{n=1}^{\infty}(-1)^{n-1}n^{2k+1}\cos n\theta,$$

for $k = 0, 1, 2, \ldots$.

Problem 8.14. Prove the identities

$$1^{2k} - 2^{2k} + 3^{2k} - \cdots = 0, \qquad k = 1, 2, \ldots,$$

$$1^{2k+1} - 3^{2k+1} + 5^{2k+1} - \cdots = 0, \qquad k = 0, 1, 2, \ldots,$$

$$1^{2k+1} - 2^{2k+1} + 3^{2k+1} - \cdots = \frac{2^{2k+2}-1}{2k+2}B_{2k+2}, \qquad k = 0, 1, 2, \ldots,$$

where B_k are Bernoulli numbers.

Here the "sums" of the given series are considered in Abel sense.

Hint: Consider the Taylor series of the function $\tan\theta$.

Problem 8.15. Find Cesàro and Abel sums of the series

$$\sum_{n=1}^{\infty}(-1)^{n-1}(2n-1)^{2k}\cos n\theta, \qquad k = 1, 2, \ldots,$$

and prove the identity

$$1^{2k} - 3^{2k} + 5^{2k} - \cdots = \frac{1}{2}E_{2k}, \qquad k = 1, 2, \ldots,$$

where E_k are Euler numbers, and the value in the right-hand side of the identity is supposed to be the Abel sum.

Hint: Consider the Taylor series of the function $\sec\theta$.

Chapter 9

Trigonometric Fourier series

I cover pages 2–15 from the book [10].

9.1 Trigonometric Fourier series

In Section 7.6.5 we studied properties of Fourier series w.r.t. an orthonormal basis $(e_n)_{n=1}^{\infty}$ in $L_2[a,b]$. In this section we will study mostly trigonometric Fourier series in $L_1[a,b]$, $C[a,b]$, and $C^{(k)}[a,b]$ as well as on the set of all piece-wise continuous or Lipschitz functions.

Definition 9.1.1. A function f is called periodic on $\mathbb{R}$ if there exists a number $A > 0$ such that

$$f(x) = f(x+A) \qquad \forall x \in \mathbb{R}. \tag{9.1.1}$$

The minimal number T satisfying (9.1.1) is called the *period* of the function f. In this case, the function f is called T-periodic.

Suppose now $f \in L_1[-\pi,\pi]$ and 2π-periodic. This, in fact, means that f is Lebesgue integrable over any interval of the real line (see Homework N13).

Definition 9.1.2. The series

$$\frac{a_0}{2} + \sum_{m=1}^{\infty} (a_m \cos mx + b_m \sin mx) \tag{9.1.2}$$

is called the *trigonometric Fourier series* of the function f if its coefficients are related to the function f as follows

$$a_m = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos mx dx, \qquad b_m = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin mx dx, \qquad m = 0, 1, 2, \ldots$$

The coefficients a_m and b_m exist, since f is integrable and $\cos mx$ and $\sin mx$ are bounded. We will also consider a complex form of the trigonometric Fourier series of the function f:

$$\sum_{n=-\infty}^{+\infty} c_n e^{inx}, \tag{9.1.3}$$

where the coefficients c_n have that form

$$c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-inx} dx, \qquad n \in \mathbb{Z}.$$

In what follows we will consider all functional spaces over the field of real numbers, $\mathbb{F} = \mathbb{R}$. In this case, the coefficients of the series (9.1.2) and (9.1.3) are related as follows

$$c_0 = \frac{a_0}{2}, \quad c_n = \frac{a_n - ib_n}{2}, \qquad c_{-n} = \frac{a_n + ib_n}{2}, \quad n = 1, 2, \ldots$$

One of the principal topic in the theory of Fourier series is the convergence which can be understanding in difference senses. We already studied convergence of Fourier series in $L_2[a,b]$. However, the theory of Fourier series in L_1 is more complicated. Also, if we try to study the pointwise (a.e.) convergence of Fourier series, we will see that from this point of view Fourier series might have rather weird behaviour even for continuous functions. Nevertheless, the uniform convergence[1] (the convergence in the space $C[-\pi,\pi]$) do not bring any troubles to researchers, as the following theorem shows.

Theorem 9.1.3. *Let a function $f(x)$ is defined on the interval $[-\pi,\pi]$ and is expanded into a trigonometric series of the form* (9.1.2) *which is uniformly convergent on $[-\pi,\pi]$. Then the coefficients a_m and b_m of the series are uniquely determined by the function f, and the series is the Fourier series of the function f.*

Proof. By assumption, we have $\forall x \in [-\pi,\pi]$,

$$f(x) = \frac{a_0}{2} + \sum_{m=1}^{\infty} (a_m \cos mx + b_m \sin mx), \tag{9.1.4}$$

where the series converges uniformly on $[-\pi,\pi]$. So we can integrate this series over the interval $[\pi,\pi]$:

$$\int_{-\pi}^{\pi} f(x)dx = \frac{a_0}{2}\int_{-\pi}^{\pi} dx + \sum_{m=1}^{\infty}\left(a_m \int_{-\pi}^{\pi} \cos mx dx + b_m \int_{-\pi}^{\pi} \sin mx dx\right) = a_0\pi,$$

so

$$a_0 = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)dx. \tag{9.1.5}$$

Now we multiply the series (9.1.5) by $\cos nx$. The resulting series is also uniformly convergent on $[-\pi,\pi]$, since $\cos nx$ is bounded. So we can integrate this series:

$$\int_{-\pi}^{\pi} f(x)\cos nx dx = \frac{a_0}{2}\int_{-\pi}^{\pi} \cos nx dx + \sum_{m=1}^{\infty}\left(a_m \int_{-\pi}^{\pi} \cos mx \cos nx dx + b_m \int_{-\pi}^{\pi} \sin mx \cos nx dx\right) = a_n\pi,$$

so

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos nx dx. \tag{9.1.6}$$

Analogously we obtain

$$b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin nx dx. \tag{9.1.7}$$

Thus, if the function is expanded into a uniformly convergent series on $[-\pi,\pi]$, then this series is its Fourier series which is uniquely determined by f. $\square$

[1] As we established in Section 7.4, the uniform convergence is the strongest convergence among all other types of convergence we considered.

For a given function $f \in L_1[-\pi, \pi]$, we will write

$$f \sim \frac{a_0}{2} + \sum_{m=1}^{\infty} (a_m \cos mx + b_m \sin mx),$$

or

$$f \sim \sum_{n=-\infty}^{+\infty} c_n e^{inx}$$

to emphasize that the corresponding trigonometric series (in real or complex form) is the Fourier series of the function f. We will use symbol "=" instead of "∼" if it is known that the Fourier series converges to f pointwise.

9.2 Basic theory of Fourier series

We start to study properties of trigonometric Fourier series in $L_1[-\pi\pi]$ from proving the Riemann-Lebesgue. For the space $L_2[-\pi, \pi]$, this lemma is a simple consequence of the convergence of the series of Fourier coefficients (see Corollaries 7.5.12 and 7.6.49).

First we prove a more general fact.

Theorem 9.2.4 (Lebesgue)**.** *Let $(\varphi_n)_{n=1}^{\infty}$ be a system of measurable functions defined on $[a, b]$. Suppose that there exists a constant $K > 0$ such that for any $n \in \mathbb{N}$ and $t \in [a, b]$,*

$$|\varphi_n(t)| \leqslant K. \tag{9.2.8}$$

If for any $c \in [a, b]$,

$$\lim_{n\to\infty} \int_a^c \varphi_n(t)dt = 0, \tag{9.2.9}$$

then for any $f \in L_1[a, b]$, the identity

$$\lim_{n\to\infty} \int_a^b f(t)\varphi_n(t)dt = 0 \tag{9.2.10}$$

holds.

Proof. Let $[\alpha, \beta] \subset [a, b]$. Then (9.2.9) implies that

$$\lim_{n\to\infty} \int_\alpha^\beta \varphi_n(t)dt = 0. \tag{9.2.11}$$

Consider now a continuous function f defined on $[a, b]$, and for a fixed $\varepsilon > 0$ split the interval $[a, b]$ into subintervals $[t_{k-1}, t_k]$, $k = 1, \dots, m$, $(t_0 = a, t_m = b)$, so that the oscillation[2] $\operatorname{osc}(f, t_k) < \varepsilon$, $k = 0, \dots, m$, that is possible because of continuity of $f(t)$. Then we have

$$\int_a^b f(t)\varphi_n(t)dt = \sum_{k=1}^{m-1} \int_{t_k}^{t_{k+1}} [f(t) - f(t_k)]\varphi_n(t)dt + \sum_{k=1}^{m-1} f(t_k) \int_{t_k}^{t_{k+1}} \varphi_n(t)dt. \tag{9.2.12}$$

But we have

$$\left| \int_{t_k}^{t_{k+1}} [f(t) - f(t_k)]\varphi_n(t)dt \right| \leqslant K\varepsilon(t_{k+1} - t_k), \tag{9.2.13}$$

[2]See Definition 5.4.3.

so the first sum in (9.2.12) does not exceed $K\varepsilon(b-a)$, while the second sum in (9.2.12) tends to zero as $n \to \infty$ according to (9.2.11). Thus, there exists $N \in \mathbb{N}$ such that for $n \geqslant N$,

$$\left| \int_a^b f(t)\varphi_n(t)dt \right| \leqslant \varepsilon[K(b-a)+1],$$

and (9.2.10) is established for continuous functions.

Now by Theorem 7.4.14, for any $f \in L_1[a,b]$ and for any $\varepsilon > 0$ there exists a continuous function g such that $\|f-g\|_{L_1} < \dfrac{\varepsilon}{2K(b-a)}$, so we have

$$\left| \int_a^b f(t)\varphi_n(t)dt \right| \leqslant \left| \int_a^b [f(t)-g(t)]\varphi_n(t)dt \right| + \left| \int_a^b g(t)\varphi_n(t)dt \right| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

for sufficiently large n. □

Example 9.2.5. Let $\varphi_n(t) = \cos nt$. Then

$$\int_a^c \varphi_n(t)dt = \frac{\sin nc - \sin na}{n} \underset{n\to\infty}{\longrightarrow} 0,$$

so the system $(\varphi_n(t)_{n=0}^{\infty}$ satisfy all the conditions of Theorem 9.2.4.

This implies the so-called Riemann–Lebesgue lemma.

Theorem 9.2.6 (Riemann-Lebesgue)**.** *For any function $f \in L_1[a,b]$, the following holds*

$$\lim_{n\to\infty} \int_a^b f(t)\cos nt dt = \lim_{n\to\infty} \int_a^b f(t)\sin nt dt = 0.$$

In particular, we have that for any function $f \in L_1[-\pi,\pi]$, its Fourier coefficients w.r.t. the system (7.6.13) vanish as $n \to \infty$:

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t)\cos nt dt \underset{n\to\infty}{\longrightarrow} 0,$$

and

$$b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t)\sin nt dt \underset{n\to\infty}{\longrightarrow} 0.$$

Remark 9.2.7. If, for a given system $(\varphi_n)_{n=1}^{\infty}$, the identity (9.2.10) holds for any $f \in L_1[a,b]$, then the system $(\varphi_n)_{n=1}^{\infty}$ is called *weakly convergent* to 0.

9.2.1 Series by sines or cosines with monotonically decreasing coefficients.

It is clear that if a function f is odd (resp. even), then $a_n = 0$ (resp. $b_n = 0$) for any n.

Theorem 9.2.8. *If $a_n \searrow 0$, then the series*

$$\frac{a_0}{2} + \sum_{n=1}^{+\infty} a_n \cos nx$$

converges at any point $x \in \mathbb{R}$ except possibly the points $2\pi m$, $m \in \mathbb{Z}$. Moreover, for any $\delta \in (0, \pi)$ this series converges uniformly on $[\delta, 2\pi - \delta]$.

If $b_n \searrow 0$, then the series

$$\sum_{n=1}^{+\infty} b_n \sin nx$$

converges at any point $x \in \mathbb{R}$ except possibly the points $2\pi m$, $m \in \mathbb{Z}$. Moreover, for any $\delta \in (0, \pi)$ this series converges uniformly on $[\delta, 2\pi - \delta]$.

Theorem 9.2.9. *Then the series*

$$\sum_{n=1}^{\infty} b_n \sin nx$$

converges uniformly on $[0, 2\pi -]$ if and only if $nb_n \to 0$ as $n \to +\infty$.

Remark 9.2.10. The previous theorem shows that there exist trigonometric series that converge uniformly but no absolutely. Indeed, by Theorem 9.2.9 one has that the series

$$\sum_{n=1}^{\infty} \frac{\sin nx}{n \ln n}$$

converges uniformly since $nb_n = \dfrac{1}{\ln n} \to 0$ as $n \to +\infty$. At the same time, it does not converge absolutely, since the series

$$\sum_{n=1}^{\infty} \frac{1}{n \ln n}$$

diverges.

Theorem 9.2.11. *If $b_n \searrow 0$, and there exists $C > 0$ such that $|nb_n| \leqslant C$ for any $n \in \mathbb{N}$. Then there exists a a constant $M > 0$ such that*

$$\left| \sum_{n=1}^{m} b_n \sin nx \right| \leqslant M \qquad \forall m \in \mathbb{N}.$$

Corollary 9.2.12. *There exists a constant $C > 0$ such that for any $m \in \mathbb{N}$ and any $x \in \mathbb{R}$,*

$$\left| \sum_{n=1}^{m} \frac{\sin nx}{n} \right| \leqslant C.$$

Theorem 9.2.13. *If $a_n \searrow 0$ and*

$$a_n - 2a_{n+1} + a_{n+2} \geqslant 0 \qquad \forall n \in \mathbb{N} \cup \{0\},$$

then the series

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos nx$$

converges at any point $x \in \mathbb{R}$, except possibly the points $2\pi m$, $m \in \mathbb{Z}$, to a nonnegative 2π-periodic even function $f \in L_1[-\pi, \pi]$. Moreover, this series is the Fourier series of f.

Corollary 9.2.14. *The series*

$$\sum_{n=2}^{\infty} \frac{\cos nx}{\ln n}.$$

is the Fourier series of a function $f \in L_1[-\pi, \pi]$.

9.3 Applications of Fourier series to ODE

I covered pages 17-22 of the book [10].

9.3.1 Differentiation of Fourier series

Theorem 9.3.1. *Let a function f be 2π-periodic and integrable on $[-\pi, \pi]$ (so on $\mathbb{R}$). Suppose that its Fourier coefficients $\widehat{f}(n)$ have the following estimate*

$$\widehat{f}(n) = \frac{\sigma_n}{n^k}, \quad \text{where} \quad \sum_{n=-\infty}^{+\infty} |\sigma_n|^2 < \infty,$$

then $f \in C^{(k-1)}[-\pi, \pi]$.

Proof. If we formally differentiate the Fourier series of the function f $k-1$ times, then the Fourier coefficients of its formal derivative have form

$$(in)^{k-1}\widehat{f}(n) = i^{k-1}\frac{\sigma_n}{n},$$

it converges, since

$$|(in)^{k-1}\widehat{f}(n)| \leqslant \left|\frac{\sigma_n}{n}\right| \leqslant \frac{|\sigma_n|^2}{2} + \frac{1}{2|n|^2}.$$

This justifies the differentiation, so f is $k-1$ times continuously differentiable. $\square$

9.3.2 Applications of Fourier series to finding periodic solutions of ODE

Let us consider a linear ordinary differential equation

$$p_0 y^{(n)}(x) + p_1 y^{(n-1)}(x) + \cdots + p_n y(x) = q(x) \tag{9.3.1}$$

with constant coefficients and a 2π-periodic function q. The problem is to find out whether there exists a periodic solution of the equation (9.3.1).

We will search the solution in the Fourier series form

$$y(x) = \sum_{k=-\infty}^{+\infty} y_k e^{ikx}.$$

Let us expand the right hand side of (9.3.1) in to the Fourier series:

$$q(x) = \sum_{k=-\infty}^{+\infty} q_k e^{ikx}.$$

then for all $k \in \mathbb{Z}$ one has

$$p_0(ik)^n y_k + p_1(ik)^{n-1} y_k + \cdots + p_n y_k = q_k$$

(this is the equality of k^{th} Fourier coefficients of the left hand side and the right hand side of (9.3.1)). Putting

$$P(z) = p_0 z^n + p_1 z^{n-1} + \cdots + p_n,$$

we obtain

$$P(ik)y_k = q_k \qquad (k \in \mathbb{Z}).$$

If

$$P(ik) \neq 0 \qquad (\forall k \in \mathbb{Z}), \tag{9.3.2}$$

then
$$y_k = \frac{q_k}{P(ik)},$$
that is, we found the Fourier coefficients of the expected periodic solution $y(x)$ of the equation (9.3.1), and therefore, the function itself is also found. But in order to the function $y(x)$ to be a periodic solution of (9.3.1) we must justify the possibility of n-times differentiating. To do this we require additionally that $q(x)$ is continuously differentiable.

From the asymptotics $P(ik) \sim p_0(ik)^n$ for $k \to \infty$ we obtain
$$|P(ik)| \geqslant c|k|^n$$
for some $c > 0$, so
$$|y_k| \leqslant \frac{|\sigma_k|}{c|k|^{n+1}},$$
where σ_k are the Fourier coefficients of the function $q'(x)$. Since q' is continuous by assumption, the series
$$\sum_{k=-\infty}^{+\infty} |\sigma_k|^2$$
converges, so we can apply Theorem 9.3.1 to infer that the sum of the series
$$\sum_{k=-\infty}^{+\infty} y_k e^{ikx}$$
converges uniformly to a n-times continuously differentiable function $y(x)$. So $y(x)$ is a 2π-periodic solution of (9.3.1).

Moreover, this solution is unique. In fact, if there exists on more 2π-periodic solution of (9.3.1), say, $\widetilde{y}(x)$, then the difference $f(x) = y(x) - \widetilde{y}(x)$ is also 2π-periodic and satisfies the equation
$$p_0 f^{(n)}(x) + p_1 f^{(n-1)}(x) + \cdots + p_n f(x) = 0. \tag{9.3.3}$$
The general solution of (9.3.3) can be expressed via the exponentials $e^{\mu x}$, where μ are the roots of the characteristic equation
$$P(z) = 0,$$
so the equation (9.3.3) has a periodic solution (except trivial $f(x) \equiv 0$) only if $\mu = ik$ for some $k \in \mathbb{Z}$, that is, $P(ik) = 0$ for some k. This contradicts with the assumption (9.3.2).

Thus we get that for 2π-periodic $q' \in C^1[-\pi, \pi]$ the condition (9.3.2) is the condition of the existence and uniqueness of 2π-periodic solution of the equation (9.3.1).

Note that if for some integer m
$$P(im) = 0,$$
and $q_m = 0$, then the equation (9.3.3) has a 2π-periodic solution
$$f(x) = Ae^{imx},$$
where A is an arbitrary constant. So in this case, the equation (9.3.1) has infinitely many periodic solutions.

Finally, if
$$P(im) = 0,$$
but $q_m \neq 0$, then the equation (9.3.3) has no periodic solutions.

9.3.3 Boundary problems of the theory of ODE

The main problem of the theory of ODE is the Cauchy problem (the initial value problem). But in applications to physics, more important problems are the so called boundary value problems. We consider now the boundary problems for ODE of the second order. In a sufficiently general form such a problem can be posed as follows.

On an interval $[a, b]$ we want find solutions $y = y(x)$ of the differential equation

$$p_2(x)y'' + p_1(x)y' + p_0(x)y = f(x), \tag{9.3.4}$$

satisfying the *boundary* conditions

$$\begin{cases} \alpha_0 y(a) + \alpha_1 y'(a) = c_1, \\ \beta_0 y(b) + \beta_1 y'(b) = c_2. \end{cases} \tag{9.3.5}$$

We assume that the functions p_0, p_1, p_2, and f are continuous on $[a, b]$. If $c_1 = c_2 = 0$, then the boundary conditions are called *uniform*. Here we consider this case. This restriction ($c_1 = c_2 = 0$) allows us to treat the set of all continuously differentiable functions on $[a, b]$ satisfying the uniform boundary conditions as a linear space. We denote this space as V_1. That is,

$$V_1 = \left\{ y(x) : y \in C^1[a, b] \quad \text{and} \quad \begin{array}{c} \alpha_0 y(a) + \alpha_1 y'(a) = 0, \\ \beta_0 y(b) + \beta_1 y'(b) = 0. \end{array} \right\}$$

We also denote by V_2 the subspace of V_1 consisting of two-times continuously differentiable functions. Then the boundary value problem tales the following form: to find $y \in V_2$ such that

$$L(y) = f, \tag{9.3.6}$$

where L is a linear differential operator defined on V_2 as follows

$$L(y) = p_2(x)y'' + p_1(x)y' + p_0(x)y.$$

The natural question is: for what kind of f the problem (9.3.6) has a solution? Is this solution unique? To answer this question, we need to study the *spectral properties* of the operator L. This means that we need to study the solvability of the eigenvalue problem of L, that is, to find all the pairs (L, y), where $\lambda \in \mathbb{C}$ and $y \in V_2$, $y \neq 0$ such that

$$L(y) = \lambda y.$$

We will see that the eigenvalue problem for the operators of boundary value problems is a source of many various orthonormal (in a certain sense) systems, the systems of eigenfunctions. If the eigenvalue problem for L is solved, and the complete orthonormal (in a certain sense) system of eigenfunctions φ_1, φ_1, ... is found, moreover, if there are no zero eigenvalues λ_n, then the solution of the problem (9.3.6) can be solved elementary. Indeed, let us expand the function f into the Fourier series with respect to the orthonormal system $\{\varphi_n\}_{n=0}^{\infty}$:

$$f = \sum_{n=1}^{+\infty} \widehat{f}(n)\varphi_n,$$

where $\widehat{f}(n)$ are the Fourier coefficients of the function f with respect to the

We will search the solution y of the boundary value problem in the form of Fourier series

$$y = \sum_{n=1}^{+\infty} \widehat{y}(n)\varphi_n.$$

Suppose that the rate of convergence of this series allows us to differentiate it sufficiently many times:

$$L\left(\sum_{n=1}^{+\infty} \widehat{y}(n)\varphi_n\right) = \sum_{n=1}^{+\infty} \widehat{y}(n)L(\varphi_n),$$

so the equation (9.3.6) has the form

$$\lambda_n \widehat{y}(n) = \widehat{f}(n), \qquad n = 1, 2, \dots$$

since $L(\varphi_n) = \lambda_n \varphi_n$. Thus, we have

$$y = \sum_{n=1}^{+\infty} \frac{\widehat{f}(n)}{\lambda_n} \varphi_n.$$

The question of the convergence of this series should be considered separately.

To study the boundary value problem, it is more convenient to rewrite the differential equation for eigenvalues in a symmetric form. Namely, we multiply the equation

$$p_2 y'' + p_1 y' + p_0 y = \lambda y$$

by a function ρ such that this equation turns into the following

$$-(py')' + qy = \lambda \rho y.$$

Obviously, the function ρ can be found from the equation

$$p_1 \rho = (p_2 \rho)',$$

where $p = -p_2 \rho$. Thus, if $p_2 \not\equiv 0$, then

$$p = Ce^{\int \frac{p_1}{p_2} dx}, \qquad \rho = -\frac{p}{p_2}.$$

The operator L defined formally by the equality

$$L(y) = \frac{-(py')' + qy}{\rho} \tag{9.3.7}$$

is called the *Sturm-Liouville operator*. The boundary value problem for eigenvalues and eigenfunctions (λ, y) of the operator L

$$-(py')' + qy = \lambda \rho y \tag{9.3.8}$$

$$\begin{cases} \alpha_0 y(a) + \alpha_1 y'(a) = 0, \\ \beta_0 y(b) + \beta_1 y'(b) = 0. \end{cases} \tag{9.3.9}$$

is called the *Sturm-liouville problem.* Here we assume that the functions p, q, and ρ are real continuous functions. Moreover, p is continuously differentiable, while q and ρ are **nonnegative**. The coefficients α_1, α_2, β_1, β_2 are assumed to be real and such that

$$\alpha_1^2 + \alpha_2^2 \neq 0, \tag{9.3.10}$$

$$\beta_1^2 + \beta_2^2 \neq 0.$$

Note that the boundary conditions of the form

$$y(a) = 0, \quad y(b) = 0$$

are called the *Dirichlet conditions*, while the boundary conditions of the form

$$y'(a) = 0, \quad y'(b) = 0$$

are called the *Neumann conditions.*

The differential equation of the Sturm-Liouville problem can be further transformed. Indeed, if we introduce the new variable t by the equality

$$t = \int \frac{dx}{p(x)},$$

where we assume that $p \neq 0$, then from the identity

$$\frac{d}{dt} = \frac{dx}{dt} \cdot \frac{d}{dx} = p\frac{d}{dx},$$

we obtain ta new form of our equation:

$$-\frac{d^2}{dx^2}y + pqy = \lambda p\rho y.$$

Let now

$$y = k(t)u(s), \qquad s = \int \frac{dt}{k^2},$$

where the function k is not defined yet, while s and u are the new variable and new sought-for function (instead of y), respectively. At the same time, we have

$$\frac{dy}{dt} = \frac{dk}{dt}u + k\frac{du}{ds} \cdot \frac{ds}{dt} = \frac{dk}{dt}u + \frac{1}{k} \cdot \frac{du}{ds},$$

$$\frac{dy^2}{dt^2} = \frac{d^2k}{dt^2}u + \frac{dk}{dt} \cdot \frac{du}{ds} \cdot \frac{ds}{dt} - \frac{1}{k^2} \cdot \frac{dk}{dt} \cdot \frac{du}{ds} + \frac{1}{k} \cdot \frac{d^2u}{ds^2} \cdot \frac{ds}{dt} = \frac{d^2k}{dt^2}u + \frac{1}{k^3} \cdot \frac{d^2u}{ds^2},$$

that implies

$$-\frac{1}{k^3} \cdot \frac{d^2u}{ds^2} + \left(pqk - \frac{d^2k}{dt^2}\right) = \lambda p\rho k u.$$

Now we choose the function k such that

$$p\rho k^4 = 1, \qquad r = pqk^4 - k^3\frac{dk^2}{dt^2}$$

to obtain the following new form of the considered equation:

$$-\frac{d^2u}{ds^2} + ru = \lambda u.$$

The transformations described above preserve the form of the uniform boundary conditions (with some new coefficients).

9.3.4 Regular Sturm-Liouville problem

The operator L defined in (9.3.7) is called *regular* Sturm-Liouville operator if $p, \rho > 0$ on $[a, b]$. the eigenvalue boundary problem (9.3.8)–(9.3.9) is called *regular* if L is regular.

Let now specify some properties of the regular Sturm-Liouville problem.

Proposition 9.3.2. *All the zeros of the eigenfunctions are simple (of multiplicity one).*

Proof. Indeed, if $y(x_0) = 0$ and $y'(x_0) = 0$ for some $x_0 \in [a, b]$, then $y(x) \equiv 0$ on $[a, b]$, so y cannot be an eigenfunction of L by definition. □

Proposition 9.3.3. *To every eigenvalue of L there corresponds a unique (up to a constant factor) eigenfunction, that is, all the eigenvalues of the regular Sturm-Liouville operator are simple.*

Proof. Let y_1 and y_2 be two eigenfunctions corresponding to an eigenvalue μ of the operator L. Since the uniform system (with respect to variables α_0 and α_1)

$$\begin{cases} \alpha_0 y_1(a) + \alpha_1 y_1'(a) = 0, \\ \alpha_0 y_2(a) + \alpha_1 y_2'(a) = 0 \end{cases} \tag{9.3.11}$$

must have a nontrivial solution by (9.3.10). Consequently, the determinant of this system must vanish. Note that the determinant of this system is the Wronskian $W[y_1, y_2]$ of the functions y_1 and y_2 calculated at the point a. The Wronskian of the functions y_1 and y_2 has the form

$$W[y_1, y_2](x) = \begin{vmatrix} y_1(x) & y_1'(x) \\ y_2(x) & y_2'(x) \end{vmatrix} = y_1(x)y_2'(x) - y_2(x)y_1'(x).$$

So the system (9.3.11) has a nontrivial solution if, and only if, $W[y_1, y_2](a) = 0$.

On the other hand, by (9.3.8) we have that

$$\frac{d}{dx}\{p(x)W[y_1, y_2](x)\} = y_1(x)(p(x)y_2'(x))' - y_2(x)(p(x)y_1'(x))' =$$

$$= y_2(x)[\lambda\rho(x) - q(x)]y_1(x) - y_1(x)[\lambda\rho(x) - q(x)]y_2(x) = 0,$$

so the function $\rho(x)W[y_1, y_2](x)$ is constant, so $W[y_1, y_2](x) \equiv 0$, since $W[y_1, y_2](a) = 0$. Consequently, $y_1(x) = cy_2(x)$, where c is a constant, as required. □

Proposition 9.3.4. *All the eigenvalues of the Sturm-Liouville problem are real. The corresponding eigenfunctions can be chosen real.*

Proof. Note that by integration by part one can obtain

$$\int_a^b [qf - (pf')']\overline{g}dx = pW[f, \overline{g}]\Big|_a^b + \int_a^b f[q\overline{g} - (p\overline{g}')']dx. \tag{9.3.12}$$

If the functions f and g satisfy the boundary conditions (9.3.9) (that is, $a, b \in V_2$), then $W[f, \overline{g}](a) = 0$ and $W[f, \overline{g}](b) = 0$[3]. Let us introduce the following weighted inner product

$$\langle f, g\rangle = \int_a^b f\overline{g}\rho dx, \qquad \|f\| = \sqrt{\langle f, f\rangle}. \tag{9.3.13}$$

Then the identity (9.3.12) takes the form

$$\langle L[f], g\rangle = \langle f, L[g]\rangle.$$

Such a property is called the symmetry of the operator L. If now y is an eigenfunction of L corresponding to an eigenvalue λ, then

$$\lambda\|y\|^2 = \langle L[y], y\rangle = \langle y, L[y]\rangle = \overline{\lambda}\|y\|^2.$$

Since $\|y\| \neq 0$ (by definition of eigenfunctions), we obtain $\lambda = \overline{\lambda}$, so $\lambda \in \mathbb{R}$, as required. □

Note that since the functions p, q and ρ are real, and since the equation $L[y] = \lambda y$ is real and linear, the eigenfunctions of the operator L can be chosen real (up to a constant complex factor). So in what follows we assume that the eigenfunctions of the operator L are *real*.

Proposition 9.3.5. *If $\lambda_1 \neq \lambda_2$, then the corresponding eigenfunctions y_1 and y_2 are orthogonal with respect to the inner product* (9.3.13).

[3] Here we use the fact that the boundary conditions have real coefficients.

Proof. Indeed,

$$(\lambda_1 - \lambda_2)\langle y_1, y_2\rangle = \langle L[y_1], y_2\rangle - \langle y_1, L[y_2]\rangle = 0.$$

So,

$$\langle y_1, y_2\rangle = 0,$$

since $\lambda_1 - \lambda_2 \neq 0$ by assumption. □

Proposition 9.3.6. *The sequence of eigenvalues of the operator L is infinite and monotone increasing*

$$\lambda_1 < \lambda_2 \cdots \lambda_n < \cdots, \qquad \lambda_n \xrightarrow[n\to\infty]{} \infty.$$

We will not prove this proposition.

Consider the space V_1 with the inner product (9.3.13). Suppose that the eigenfunctions y_n of the Sturm-Liouville problem are normed:

$$\|y_n\|^2 = \int_a^b |y_n(x)|^2 \rho(x)dx = 1.$$

Then they generate an orthonormal system. The Fourier coefficients of a function f (from the space V_1 with inner product (9.3.13)) with respect to this orthonormal system are defined by the following formula[4]

$$\widehat{f}(n) = \langle f, y_n\rangle = \int_a^b f(x)y_n(x)\rho(x)dx.$$

The expansion of the function f into the Fourier series with the eigenfunctions of the Sturm-Liouville problem has the form

$$f \sim \sum_{n=1}^{+\infty} \widehat{f}(n)y_n(x). \tag{9.3.14}$$

Let us establish that the system of the eigenfunctions of the Sturm-Liouville problem is closed (or complete), so the series (9.3.14) converges to the function f in the norm of the space V_1 with inner product (9.3.13):

$$\|f - \sum_{n=1}^{N} \widehat{f}(n)y_n(x)\| \xrightarrow[N\to\infty]{} 0. \tag{9.3.15}$$

We will give a scheme of the proof of (9.3.15) under additional conditions:

1. $\rho = 1$;
2. V_1 consists of **real** continuously differentiable functions satisfying Dirichlet's boundary conditions:

$$y(a) = y(b) = 0.$$

In this case, the inner product (9.3.13) has the form

$$\langle f, g\rangle = \int_a^b fg\,dx. \tag{9.3.16}$$

Let

$$I(y) \equiv \langle L(y), y\rangle = \int_a^b [-(py')' + qy]y\,dx = \int_a^b [p(y')^2 + qy^2]dx. \tag{9.3.17}$$

It can be proved (we skip this proof) that the minimal value of the functional $I(y)$ under conditions

$$y(a) = y(b) = 0, \quad \|y\| = 1,$$

[4]Recall that the eigenfunctions y_n are supposed to be real.

equals the following

$$\min I(y) = \lambda_1, \qquad \lambda_1 = I(y_1),$$

where λ_1 is the minimal eigenvalue of the considered Sturm-Liouville problem, and y_1 is the corresponding eigenfunction, $\|y\| = 1$.

Moreover, the following *variational principle* holds.

Theorem 9.3.7. *Let y_1, y_2, ..., y_{n-1} be the orthonormal system of eigenfunctions of the Sturm-Liouville problem satisfying the eigenvalues λ_1, λ_2, ..., λ_{n-1} such that*

$$\lambda_1, \lambda_2, \ldots, \lambda_{n-1}.$$

Then the minimal value of the functional $I(y)$ defined in (9.3.17) *under the conditions*

$$y(a) = y(b) = 0, \quad \|y\| = 1, \quad y \perp y_k, \forall k = 1, 2, \ldots, n-1,$$

can be achieved. Furthermore,

$$\min I(y) = \lambda_n, \quad \lambda_n = I(y_n),$$

where λ_n is the n-th eigenvalue of the Sturm-Liouville problem, and y_n is the corresponding eigenfunction.

We will use this theorem to prove the completeness of the system of eigenfunctions of the Sturm-Liouville operator L in the space V_1.

Let

$$r_n = f - \sum_{k=1}^{n-1} c_k y_k,$$

where $c_k = \langle f, y_k \rangle$ are the Fourier coefficients of the function f. So $y_k \perp r_n$ for $k < n$. Therefore,

$$\langle L(r_n), r_n \rangle = \langle L(f), r_n \rangle - \sum_{k=1}^{n-1} c_k \langle L(y_k), r_n \rangle = \langle L(f), r_n \rangle - \sum_{k=1}^{n-1} c_k \lambda_k \langle y_k, r_n \rangle,$$

and since $\langle y_k, r_n \rangle = 0$, we have

$$\langle L(r_n), r_n \rangle = \langle L(f), r_n \rangle.$$

Suppose that $\|r_n\| \neq 0$, so

$$r_n = \|r_n\| e_n,$$

where

$$\|e_n\| = 1, \qquad e_n \perp y_1, y_2, \ldots, y_{n-1}.$$

By Theorem 9.3.7, we have

$$\langle L(r_n), r_n \rangle = \|r_n\|^2 \langle L(e_n), e_n \rangle \geqslant \|r_n\|^2 \min_{\substack{\|y\|=1 \\ y \perp y_1, \ldots, y_{n-1}}} \langle L(y), y \rangle = \lambda_n \|r_n\|^2.$$

This implies, due to the Cauchy-Bunyakovsky-Schwarz inequality, that

$$\lambda_n \|r_n\|^2 \leqslant \langle L(f), r_n \rangle \leqslant \|L(f)\| \cdot \|r_n\|,$$

so since $\lambda_n \to +\infty$ as $n \to \infty$, we obtain

$$\lambda_n \|r_n\| \leqslant \frac{\|L(f)\|}{\lambda_n} \xrightarrow[n \to \infty]{} 0.$$

But this is equivalent to the closedness (completeness) of the system of the eigenvalues y_k of the Sturm-Liouville operator of the Dirichlet boundary problem with $\rho = 1$.

Let us return to the equation

$$L(y) = f,$$

and note that since L is symmetric, we have

$$\langle L(y), y_n\rangle = \langle y, L(y_n)\rangle = \langle y, \lambda_n y_n\rangle = \lambda_n\langle y, y_n\rangle,$$

so

$$\lambda_n\langle y, y_n\rangle = \langle f, y_n\rangle.$$

The question whether the function

$$y = \sum_{n=1}^{+\infty} \frac{\langle f, y_n\rangle}{\lambda_n} y_n$$

is, in fact, the solution of the Sturm-Liouville problem depends on the rate of convergence of this series.

9.4 Problems

9.4.1 Basic theory of Fourier series

Problem 9.1. Prove the formula

$$\sum_{m=1}^{n} \cos mx = -\frac{1}{2} + \frac{1}{2} \cdot \frac{\sin\left(n+\frac{1}{2}\right)x}{\sin\frac{x}{2}} \tag{9.1.1}$$

Hint: Use the formulæ $\cos y = \dfrac{e^{iy}+e^{-iy}}{2}$, $\sin y = \dfrac{e^{iy}-e^{-iy}}{2}$. An alternative proof is to exploit the trigonometric formula of the product of a sine and cosine.

Problem 9.2. Applying integration by parts, deduce from (9.1.1) that

$$\sum_{m=1}^{n} \frac{\sin mx}{m} = \frac{\pi - x}{2} - \frac{\cos\left(n+\frac{1}{2}\right)x}{(2n+1)\sin\frac{x}{2}} + \frac{1}{2n+1}\int_{x}^{\pi} \frac{\cos\left(n+\frac{1}{2}\right)t \cdot \cos\frac{t}{2}}{\sin^2\frac{t}{2}} dt \tag{9.2.2}$$

Problem 9.3. Prove that

$$\sum_{n=1}^{\infty} \frac{\sin nx}{n} = \frac{\pi - x}{2}$$

for any $x \in (0, \pi)$.

Analogously, prove that for any $x \in (-\pi, 0)$, the series converges to $-\dfrac{\pi + x}{2}$.

Problem 9.4. Prove the formula

$$\sum_{m=1}^{n} \cos(2m-1)x = \frac{1}{2} \cdot \frac{\sin 2nx}{\sin x}.$$

Using this formula prove that for any $x \in (0, \pi)$

$$\sum_{m=1}^{\infty} \frac{\sin(2n-1)x}{2n-1} = \frac{\pi}{4}. \tag{9.4.3}$$

Problem 9.5. Using the identity (9.4.3), prove Steklov's formula

$$f(x+\pi) - f(x) = -\frac{4}{\pi} \sum_{m=0}^{+\infty} \int_{x}^{x+\pi} f(u)\cos(2m+1)(u-x)du,$$

where $f(u)$ is differentiable on the interval $(x, x+\pi)$.

Problem 9.6. Expand into Fourier series the following periodic functions

1) $\{x\}$, the fractional part of x, $\{x\} = x - [x]$, where $[x]$ is the integer part of x;

2) (x), the distance from x to the closest integer;

Problem 9.7. Find the Fourier series of the a 3-periodic function defined on $[0, 3]$ by the following formulæ

$$\begin{cases} x & \text{if} \quad x \in [0,1], \\ 1 & \text{if} \quad x \in [1,2], \\ 3-x & \text{if} \quad x \in [2,3], \end{cases}$$

and study its convergence.

Problem 9.8. Verify that

$$\frac{1}{2i}\sum_{\substack{n=-\infty\\ n\neq 0}}^{+\infty}\frac{e^{inx}}{n}$$

is the Fourier series of the 2π-periodic **sawtooth** function defined by $f(0)=0$ and

$$f(x)=\begin{cases}-\frac{\pi}{2}-\frac{x}{2} & \text{if} \quad -\pi<x<0,\\[2ex] \frac{\pi}{2}-\frac{x}{2} & \text{if} \quad 0<x<\pi.\end{cases}$$

Note that this function is not continuous. Show that nevertheless, the series converges for every x (by which we mean, as usual, that the symmetric partial sums of the series converge). In particular, the value of the series at the origin, namely 0, is the average of the values of $f(x)$ as x approaches the origin from the left and the right.

Hint: Use Dirichlet's test for convergence of a series $\sum_n a_n b_n$.

Problem 9.9. Expand $f(x)=x^3$ into the cos-series in the interval $(0,\pi)$.

Problem 9.10. Expand $f(x)=x^3$ into the sin-series in the interval $(0,\pi)$.

Problem 9.11. Expand $f(x)=e^x$ into Fourier series in the interval $(-h,h)$.

Problem 9.12. Expand $f(x)=\cosh(ax)$ into Fourier series in the interval $(-\pi,\pi)$.

Problem 9.13. Expand $f(x)=\sinh(ax)$ into Fourier series in the interval $(-\pi,\pi)$.

Problem 9.14. Using the change of variables $t=e^{ix}$ prove that

1)

$$\frac{a\sin x}{1-2a\cos x+a^2}=\sum_{n=1}^{+\infty}a^n\sin(nx), \qquad |a|<1,$$

2)

$$\ln(1-2a\cos x+a^2)=-2\sum_{n=1}^{+\infty}\frac{a^n}{n}\cos(nx), \qquad |a|<1,$$

3)

$$\ln 2\cos\frac{x}{2}=\sum_{n=1}^{+\infty}\frac{(-1)^{n-1}\cos(nx)}{n},$$

where $x\in(-\pi,\pi)$.

Hint: For 2) use differentiation/integration, for 3) use the series 2).

Problem 9.15. Is it possible to find Fourier series of the 2π-periodic function

$$\tan\frac{x}{2}$$

on the interval $(-\pi,\pi)$? Explain the answer.

Problem 9.16. Prove that

1)
$$\frac{\cos 2x}{1\cdot 3}+\frac{\cos 4x}{3\cdot 5}+\frac{\cos 6x}{5\cdot 7}+\cdots=\frac{1}{2}-\frac{\pi}{4}\sin x, \qquad 0<x<\pi,$$

2)
$$\sum_{n=1}^{+\infty}\frac{\sin(2n+1)x}{(2n+1)^3}=\frac{\pi x}{8}(\pi-x), \qquad 0<x<\pi.$$

3)
$$\frac{\cos 3x}{1\cdot 3\cdot 5}-\frac{\cos 5x}{3\cdot 5\cdot 7}+\frac{\cos 7x}{5\cdot 7\cdot 9}-\cdots=\frac{\pi}{8}\cos^2 x-\frac{1}{3}\cos x, \qquad -\frac{\pi}{2}<x<\frac{\pi}{2}.$$

Problem 9.17. Prove that
$$\sum_{n=1}^{+\infty}\frac{\sin(2n+2)x}{n(n+1)}=\sin 2x-(\pi-2x)\sin^2 x-\sin x\cos x\ln(4\sin^2 x), \qquad 0<x<\pi.$$

Problem 9.18. Prove that
$$\sum_{n=1}^{+\infty}\frac{\cos(2n+1)x}{n^2(n+1)^2}=2(\pi-2x)\sin x+\left(\frac{\pi^2}{3}-2\pi x+2x^2-3\right)\cos x, \qquad 0<x<\pi.$$

Problem 9.19. Noting that $\frac{1}{n}=\int_0^1 t^{n-1}dt$ find the sum of the series
$$\sum_{n=1}^{+\infty}\frac{\sin nx}{n}, \qquad 0<x<\pi.$$

Problem 9.20. Find the sum of the following series

1)
$$\sum_{n=1}^{+\infty}\frac{\cos(nx)}{n^2}, \qquad 0<x<2\pi,$$

2)
$$\sum_{n=1}^{+\infty}\frac{\sin(nx)}{n^3}, \qquad 0<x<2\pi,$$

3)
$$\sum_{n=1}^{+\infty}\frac{\cos nx}{n}, \qquad 0<x<2\pi,$$

4)
$$\sum_{n=0}^{+\infty}\frac{\cos(2n+1)x}{2n+1}, \qquad 0<x<\pi,$$

5)
$$\sum_{n=2}^{+\infty}\frac{\cos(nx)}{n^2-1}, \qquad 0<x<2\pi,$$

6)
$$\sum_{n=2}^{+\infty}\frac{\sin(nx)}{n^2-1}, \qquad 0<x<2\pi,$$

7)
$$\sum_{n=1}^{+\infty} \frac{\sin nx}{n(n^2+a^2)}, \qquad 0 < x < \pi,$$

8)
$$\sum_{n=1}^{+\infty} \frac{(-1)^n \sin nx}{n(n^2+a^2)}, \qquad 0 < x < \pi,$$

9)
$$\sum_{n=0}^{+\infty} \frac{\sin(2n+1)x}{2n+1}, \qquad 0 < x < \pi,$$

10)
$$\sum_{n=0}^{+\infty} \frac{(-1)^n \cos nx}{n^2-4}, \qquad -\pi < x < \pi,$$

Hint: For 1) and 2) use the result of Problem 9.19 and differentiation.

Problem 9.21. Show that for α not an integer, the Fourier series of
$$\frac{\pi}{\sin \pi\alpha} e^{i(\pi-x)}\alpha$$
on $[0, 2\pi]$ is given by
$$\sum_{n=-\infty}^{+\infty} \frac{e^{inx}}{n+\alpha}$$
Apply Parseval's identity to show that
$$\sum_{n=-\infty}^{+\infty} \frac{1}{(n+\alpha)^2} = \frac{\pi^2}{(\sin \pi\alpha)^2}.$$

Problem 9.22. Suppose f is a periodic function of period 2π which belongs to the class C^k. Show that if $\sum\limits_n c_n e^{inx}$ is the Fourier series of $f(x)$, then
$$c_n = o\left(\frac{1}{|n|^k}\right) \qquad \text{as} \quad |n| \to \infty.$$

Hint: Use Riemann-Lebesgue lemma.

Problem 9.23. Let $f \in C[-\pi, \pi]$. Prove that
$$\left|\int_{-\pi}^{\pi} f(x) \cos nx dx\right| \leqslant C\omega\left(\frac{\pi}{n}, f\right), \qquad \left|\int_{-\pi}^{\pi} f(x) \sin nx dx\right| \leqslant C\omega\left(\frac{\pi}{n}, f\right)$$
where $C > 0$ is a constant and
$$\omega(\delta, f) = \sup_{|x-y|<\delta} |f(x) - f(y)|, \qquad x, y \in [-\pi, \pi].$$

Problem 9.24. Let f is continuous and 2π-periodic. Prove that

$$f_n(x) = \frac{1}{2\pi} \cdot \frac{4^n (n!)^2}{(2n)!} \int_{-\pi}^{\pi} f(x+u) \left(\cos \frac{u}{2}\right)^{2n} du$$

is a singular integral, that is, $f_n(x)$ tends to $f(x)$ uniformly on $\mathbb{R}$. Show also that $f_n(x)$ is a trigonometric polynomial.

Problem 9.25. Improve the rate of convergence of the following series

1)
$$\sum_{n=1}^{+\infty} \frac{n^2 \sin nx}{n^3 + 1}$$
from order n^{-1} to order n^{-4} using the result of Problem 6,

2)
$$\sum_{n=1}^{+\infty} \frac{n^2 + 1}{n^4 + 1} \cos nx$$
from order n^{-2} to order n^{-4} using Problem 6.

Problem 9.26. Find the sum of the series

$$\sum_{n=2}^{+\infty} \frac{\cos nx}{n^2 - 1}, \qquad 0 < x < 2\pi.$$

and use it to improve the rate of convergence of the series

$$\sum_{n=1}^{+\infty} \frac{n^2 + 1}{n^4 + 1} \cos nx$$

from order n^{-2} to order n^{-6}.

Problem 9.27. Let $f \in L_1[a, b]$. Prove that

$$\lim_{R \to \infty} \int_a^b f(x) e^{iRx} dx = 0.$$

9.4.2 Applications of Fourier series

Problem 9.28. Use the method of eigenfunctions expansion to solve the following boundary value problem:

$$\begin{cases} y'' + 4y = x^3 \\ y(0) = 0, \quad y'(1) = 0. \end{cases}$$

Problem 9.29. Find the eigenvalues and eigenfunctions of the following boundary value problem:

$$\begin{cases} (xy')' + \dfrac{2y}{x} = -\lambda \dfrac{y}{x} \\ y'(1) = 0, \quad y'(2) = 0. \end{cases}$$

Problem 9.30. Use the method of eigenfunctions expansion to solve the following boundary value problem:

$$\begin{cases} (xy')' + \dfrac{y}{x} = \dfrac{1}{x} \\ y(1) = 0, \quad y(e) = 0. \end{cases}$$

Problem 9.31. The eigenvalue problem $x^2y'' - \lambda xy' + \lambda y = 0$ with $y(1) = y(2) = 0$ is not a Sturm-Liouville eigenvalue problem. Show that none of the eigenvalues are real by solving this eigenvalue problem.

Problem 9.32. Find 2π-periodic solutions of the following ordinary differential equation:

$$y''' + y = \sin 2x$$

Problem 9.33. Find the solution (tending to zero as $|x| \to \infty$) of the following difference equations:

$$f(x+h) - 2f(x) = \frac{1}{1+x^2},$$

Hint: Represent the solution as a **functional** (not power or Fourier!) series depending on parameter h. When $h = 0$, the series must degenerate to the function $f(x) = -\dfrac{1}{1+x^2}$.

Problem 9.34. Find 2π-periodic solutions of the following difference equation:

$$f(x+h) + 3f(x) + f(x-h) = g(x), \quad h \neq 2\pi,$$

where $g(x)$ is the 2π-periodic function defined by the equality

$$g(x) = x, \quad \text{for} \quad x \in (0, 2\pi).$$

Hint: Compare the Fourier coefficients of the left and right hand sides of the equation.

Problem 9.35. Find the eigenvalues and eigenfunctions of the following problem

$$\begin{cases} y'' + y = \lambda y, \\ y(1) = y'(l) = 0. \end{cases}$$

and solve the problem

$$\begin{cases} y'' + y = x^2, \\ y(1) = y'(l) = 0. \end{cases}$$

Problem 9.36. Find the eigenvalues and eigenfunctions of the following problem

$$\begin{cases} x^2y'' + \dfrac{1}{4}y = \lambda y, \\ y(1) = y'(e) = 0. \end{cases}$$

and solve the problem

$$\begin{cases} x^2y'' + \dfrac{1}{4}y = x, \\ y(1) = y'(e) = 0. \end{cases}$$

Bibliography

[1] R. Bass, *Real analysis for graduate students: measure and integration theory*, 2011.

[2] P. Cohen, *Set theory and the continuum hypothesis*, Dover Publications, Mineola, New York, 1966.

[3] G.B. Folland, *Real Analysis*, 2nd ed., Wiley, New York, 1999.

[4] J.K. Hunter, *Measure theory*, Lecture notes: `https://www.math.ucdavis.edu/~hunter/measure_theory/measure_theory.html`.

[5] A. Kolmogorov and S. Fomin, *Introduction to Real Analysis*, Dover Publication Inc., New York, 1970.

[6] A. Kolmogorov and S. Fomin, *Measure, Lebesgue integrals, and Hilbert space*, Academic Press, New York and London, 1960.

[7] T. Körner, *Excersices for Fourier Analysis*, Cambridge University Press, 1993.

[8] I. Natanson, *Theory of Functions of a Real Variable*, Vol. I and Vol. II, Frederick Ungar Publishing Co, New York, 1955, 1961, 1964.

[9] W. Rudin, *Principles of Mathematical Analysis*, McGraw-Hill Inc., 1976.

[10] E. Stein and R. Shakarchi, *Fourier Analysis*, Princeton University Press, Princeton and London, 2003.

[11] E. Stein and R. Shakarchi, *Real Analysis*, Princeton University Press, Princeton and London, 2005.