Core Module in Analysis
Honours MAM

Measure Theory and Integration

Associate Professor Elena Berdysheva

Version: 5 June 2023

Contents

Literature

- P.R. Halmos. *Measure Theory*. Graduate Texts in Mathematics. New York, NY: Springer, 1950
- D.L. Cohn. *Measure Theory*. Birkhäuser Advanced Texts Basler Lehrbücher. New York, NY: Birkhäuser, 2013
- M. Loève. *Probability Theory I*. Graduate Texts in Mathematics. New York, NY: Springer, 1977
- V.I. Bogachev. *Measure Theory*. Berlin, Heidelberg: Springer, 2007

Chapter 0
Motivation

Let us consider the following question. We want to assign an area A to sets $F \subseteq \mathbb{R}^2$ in the plane. What is a sensible way to do this? Is it possible to assign an area to all subsets of $\mathbb{R}^2$?

I believe we all will easily agree on a formula for an area of a rectangle. Let R be a rectangle whose sides have lenghts x and y. Per definition, its area is $A(R) = xy$ (Figure 0.1).

This definition is a starting point for our consideration. Now we must find a way to extend this definition to more complicated figures in the plane.

We require the following properties.

(1) If we put two disjoint figures together, the area of the resulting figure must be the sum of the areas of the original figures.

(2) Congruent figures must have the same areas.

(3) Lines mist have zero areas.

Applying this rules, we are able to deduce a formula for the area of a triangle. Let T be a triangle with the base g and the height h. We extend the triangle to

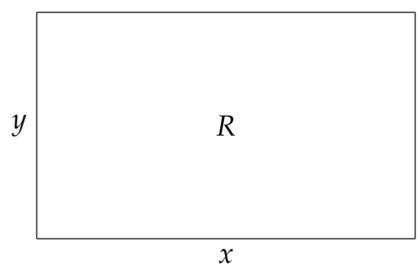

Figure 0.1: A rectangle with the area $A(R) = xy$.

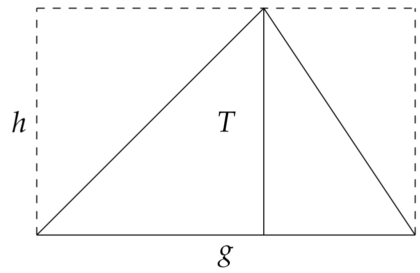

Figure 0.2: A triangle with the area $A(T) = \frac{1}{2}hg$.

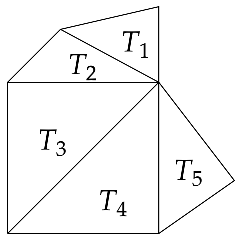

Figure 0.3: A triangulation of a polygon.

the rectangle with sides g and h as shown in Figure 0.2. We see that the area of the triangle T must be the half of the area of the rectangle. Thus, $A(T) = \frac{1}{2}hg$.

Areas of polygons can be now calculated using triangulation. For example, for the area of the 7-gon P on Figure 0.3 we have $A(P) = A(T_1) + \cdots + A(T_5)$.

For figures whose boundaries are parts of curves, we need some kind of a limit process. For example, to determine the area of a disc D we "fill" the disc with triangles like shown in the Figure 0.4. In the limit, we "fill" the disc with countably many pairwise disjoint triangles $\{T_i\}_{i\in\mathbb{N}}$: $D = \bigcup_{i=1}^{\infty} T_i$.

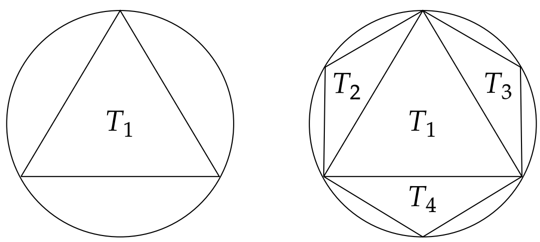

Figure 0.4: Approximating the area of a disc.

A union of pairwise disjoint sets is called a *disjoint union*. For disjoint unions, we will use the notation $\sum_{i=1}^{\infty} T_i$. (Note that the is a not a common notation and other notations are also in use in the literature, for example, $\sqcup_{i=1}^{\infty} T_i$, or $\dot{\bigcup}_{i=1}^{\infty} T_i$.) We will write $A + B$ for the disjoint union of two sets A, B, $A \cap B = \emptyset$.

To proceed with the calculation of the area of the disc, we require that property (1) from the above list holds also for disjoint unions of countably many figures. We obtain $A(D) = \sum_{i=1}^{\infty} A(T_i)$.

Let us summarize. As a first attempt to arrive at a mathematical definition of an area for sets in $\mathbb{R}^2$, we made the following steps.

A. We gave a definition for the area of a rectangle in a natural way. At this point we have got a function A that maps the set of all rectangles in the plain to the set of non-negative real numbers,

$$A : \{R : R \subseteq \mathbb{R}^2, R \text{ is a rectangle}\} \to [0, \infty).$$

B. Areas of other sets in $\mathbb{R}^2$ should be obtained via an extension of the function A to a function $\widetilde{A}$ with the following properties:

(1) $\widetilde{A}(F) = \sum_{i=1}^{\infty} \widetilde{A}(F_i)$, if $F = \sum_{i=1}^{\infty} F_i$, a disjoint union.

(2) $\widetilde{A}(F_1) = \widetilde{A}(F_2)$, if F_1 and F_2 are congruent.

It is a nontrivial task to figure out, how and for which sets $F \subseteq \mathbb{R}^2$ we can define $\widetilde{A}$.

The crucial property is property (1), called *σ-additivity*.

Chapter 1

Families of sets

Let $\Omega \neq \varnothing$ be a nonempty set. The *power set* of Ω is the set of all subsets of Ω, i.e. $P(\Omega) = 2^{\Omega} = \{A : A \subseteq \Omega\}$.

In this chapter we will study several types of families of sets (or set systems) $\mathscr{A} \subseteq P(\Omega)$. For example, in the previous chapter we worked with $\Omega = \mathbb{R}^2$ and $\mathscr{A}$ being the set of all rectangles in the plane.

In this lecture notes we will use the abbreviation “p.d.” for “pairwise disjoint”.

Given a family of sets $\mathscr{A} \subseteq P(\Omega)$, we denote by $\mathscr{A}^+$ the family of all disjoint finite unions of elements from $\mathscr{A}$, i.e.

$$\mathscr{A}^+ := \left\{ \sum_{i=1}^{n} A_i : A_i \in \mathscr{A}, A_1, \dots, A_n \text{ p.d.}, n \in \mathbb{N} \cup \{0\} \right\},$$

where $\sum_{i=1}^{0} A_i = \varnothing$ is understood as an empty union.

Definition 1.1 A family of sets $\mathscr{A} \subseteq P(\Omega)$ is called a *semiring*, if

(1) $\varnothing \in \mathscr{A}$,

(2) $A, B \in \mathscr{A} \implies A \cap B \in \mathscr{A}$ (i.e., $\mathscr{A}$ is closed under intersections),

(3) $A, B \in \mathscr{A} \implies A \setminus B \in \mathscr{A}^+$.

$\triangle$

Example A. Let Ω be a arbitrary nonempty set, then $\mathscr{A} = \{\varnothing\}$ is a semiring.

B. Let $\Omega = \mathbb{R}$. The family $\mathscr{A} = \{(x, y] : x, y \in \mathbb{R}\} =: \mathscr{J}$ of left open, right closed finite intervals is a semiring. Let us look at the properties from the definition a semiring:

(1) $\varnothing \in \mathscr{J}$ (indeed, for example, $(0, 0] = \varnothing \in \mathscr{J}$).

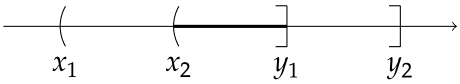

Figure 1.1: $A \cap B = (x_2, y_1] \in \mathscr{J}$.

Figure 1.2: $A \setminus B = (x_1, x_2] \in \mathscr{J} \subseteq \mathscr{J}^+$.

(2) $A, B \in \mathscr{J} \implies A \cap B \in \mathscr{J}$.

If $A, B \in \mathscr{J}$, then the form of $A \cap B$ depends on the relative position of the intervals A and B. For example, if $A = (x_1, y_1]$ and $B = (x_2, y_2]$ with $x_1 < x_2 < y_1 < y_2$, then $A \cap B = (x_2, y_1] \in \mathscr{J}$ (see Figure 1.1).

In the general case we write

$$\begin{aligned} A \cap B &= \{t \in \mathbb{R} : x_1 < t \leq y_1\} \cap \{t \in \mathbb{R} : x_2 < t \leq y_2\} \\ &= (\max\{x_1, x_2\}, \min\{y_1, y_2\}] \in \mathscr{J}. \end{aligned}$$

(3) $A, B \in \mathscr{J} \implies A \setminus B \in \mathscr{J}^+$.

The case when $A = (x_1, y_1]$ and $B = (x_2, y_2]$ with $x_1 < x_2 < y_1 < y_2$ is illustrated in Figure 1.2, in this case we have $A \setminus B = (x_1, x_2] \in \mathscr{J} \subseteq \mathscr{J}^+$. The general situation will be discussed in the Homework.

△

The semiring $\mathscr{J}$ is a starting point for determining lengths of subsets of $\mathbb{R}$ (in the same manner how we used rectangles as a starting point for determining areas of sets in the plane in Chapter 0).

Definition 1.2 A family of sets $\mathscr{A} \subseteq P(\Omega)$ is called a *ring*, if

(1) $\varnothing \in \mathscr{A}$,

(2) $A, B \in \mathscr{A} \implies A \setminus B \in \mathscr{A}$,

(3) $A, B \in \mathscr{A} \implies A \cup B \in \mathscr{A}$,

(i.e., $\mathscr{A}$ is closed under unions and differences).

△

Remark Each ring is a semiring. A semiring does not need to be a ring.

△

Proof. Homework. □

Definition 1.3 Let $\mathscr{E} \subseteq P(\Omega)$ be an arbitrary family of sets. The ring

$$r(\mathscr{E}) := \bigcap_{\substack{\mathscr{A} \supseteq \mathscr{E} \\ \mathscr{A} \text{ is a ring}}} \mathscr{A}$$

is called the *ring generated by $\mathscr{E}$*.

△

Remark The intersection of arbitrarily many rings is a ring.

△

Proof. Homework. □

In the next two lemmas we study the structure of a semiring.

Lemma 1.4 *Let $\mathscr{A}$ be a semiring, and let $A, A_1, \dots, A_n \in \mathscr{A}$ be such that $A_1, \dots, A_n$ are p.d. and $\sum_{i=1}^{n} A_i \subseteq A$. Then there exist the sets $A_{n+1}, \dots, A_N \in \mathscr{A}$ such that $A = \sum_{i=1}^{N} A_i$ (a disjoint union).*

△

Proof. We prove the statement using induction in n.

Let first $n = 1$. Given is the set $A_1 \subseteq A$. Since $\mathscr{A}$ is a semiring, we have $A \setminus A_1 \in \mathscr{A}^+$, i.e., there exist p.d. sets $A_2, \dots, A_N \in \mathscr{A}$ such that $A \setminus A_1 = \sum_{j=2}^{N} A_j$. It follows that

$$A = \sum_{j=1}^{N} A_j.$$

For the induction step, assume that the statement of the lemma holds for a certain $n \in \mathbb{N}$. We will show that it also holds for $n + 1$.

Let $A_1, \dots, A_{n+1} \in \mathscr{A}$ be p.d. sets with $\sum_{i=1}^{n+1} A_i \subseteq A$. Then, in particular, $\sum_{i=1}^{n} A_i \subseteq A$, and by the induction hypothesis there exist sets $B_1, \dots, B_m \in \mathscr{A}$ such that

$$A = \sum_{i=1}^{n} A_i + \sum_{j=1}^{m} B_j,$$

a disjoint union.

We have $A_{n+1} \subseteq \sum_{j=1}^m B_j$. It follows that

$$A_{n+1} = A_{n+1} \cap \left(\sum_{j=1}^m B_j \right) = \sum_{j=1}^m (A_{n+1} \cap B_j),$$

and $A_{n+1} \cap B_j \in \mathscr{A}$, because $\mathscr{A}$ is a semiring. Then, in particular, $B_j \setminus (A_{n+1} \cap B_j) \in \mathscr{A}^+$, and for each $j = 1, \dots, m$ we can find a number $r_j \in \mathbb{N}$ and p.d. sets $B_{jk} \in \mathscr{A}$, $k = 1, \dots, r_j$, such that

$$B_j = (A_{n+1} \cap B_j) + \sum_{k=1}^{r_j} B_{jk}.$$

Altogether we obtain

$$\begin{aligned} A &= \sum_{i=1}^n A_i + \sum_{j=1}^m \left[(A_{n+1} \cap B_j) + \sum_{k=1}^{r_j} B_{jk} \right] \\ &= \sum_{i=1}^n A_i + \underbrace{\sum_{j=1}^m (A_{n+1} \cap B_j)}_{=A_{n+1}} + \sum_{j=1}^m \sum_{k=1}^{r_j} B_{jk} \\ &= \sum_{i=1}^{n+1} A_i + \sum_{j=1}^m \sum_{k=1}^{r_j} B_{jk}. \end{aligned}$$

This is the desired representation.

□

Lemma 1.5 *Let $\mathscr{A}$ be a semiring, and let $A_1, \dots, A_n \in \mathscr{A}$. Then there exists a collection of p.d. sets $B_1, \dots, B_N \in \mathscr{A}$ such that for each $k = 1, \dots, n$ we have*

$$A_k = \sum_{j \in M_k} B_j, \quad M_k \subseteq \{1, \dots, N\}.$$

△

Proof. Again, we prove the statement by induction in n.

For $n = 1$, the statement is obvious: just take $B_1 = A_1$.

For the induction step, assume that the statement of the lemma holds for a certain $n \in \mathbb{N}$. We will show that it also holds for $n + 1$.

Let $A_1, \dots, A_n, A_{n+1} \in \mathscr{A}$. According to the induction hypothesis, applied to the sets $A_1, \dots, A_n$, there exist p.d. sets $B_1, \dots, B_N \in \mathscr{A}$ such that $A_k = \sum_{j \in M_k} B_j$, $M_k \subseteq \{1, \dots, N\}$, $k = 1, \dots, n$.

Since $A_{n+1}, B_j \in \mathscr{A}$ and $\mathscr{A}$ is a semiring, $A_{n+1} \cap B_j \in \mathscr{A}$. We now apply Lemma 1.4 to the p.d. sets $\{A_{n+1} \cap B_j\}_{j=1}^N$ and the set A_{n+1}. Thus, there exist p.d. sets $\{D_\ell\}_{\ell=1}^L \subseteq \mathscr{A}$ such that

$$A_{n+1} = \sum_{j=1}^{N}(A_{n+1} \cap B_j) + \sum_{\ell=1}^{L} D_\ell,$$

a disjoint union.

For each j we have $B_j \setminus (A_{n+1} \cap B_j) \in \mathscr{A}^+$. It follows that there exist p.d. sets $B_{j2}, \ldots B_{jr_j} \in \mathscr{A}$ such that

$$B_j = \underbrace{(A_{n+1} \cap B_j)}_{:=B_{j1} \in \mathscr{A}} + \sum_{k=2}^{r_j} B_{jk} = \sum_{k=1}^{r_j} B_{jk}.$$

The collections of the sets $\{D_\ell\}_{\ell=1}^L \cup \{B_{jk}\}_{\substack{j=1,\ldots,N \\ k=1,\ldots,r_j}}$ has the desired properties.

□

Theorem 1.6 *(The structure of a ring generated by a semiring)* *Let $\mathscr{A}$ be a semiring. Then*

$$r(\mathscr{A}) = \mathscr{A}^+.$$

△

Proof. Recall that

$$\mathscr{A}^+ = \left\{ \sum_{i=1}^{n} A_i : A_i \in \mathscr{A}, \text{ p.d.}, n \in \mathbb{N} \cup \{0\} \right\},$$

the collection of all finite disjoint unions of elements from $\mathscr{A}$.

A. First we show that $\mathscr{A}^+$ is a ring. We check the properties from Definition 1.2.

(1) $\varnothing \in \mathscr{A}^+$ per definition.

(2) Let $A, B \in \mathscr{A}^+$. We have to show that $A \setminus B \in \mathscr{A}^+$. Let $A = \sum_{k=1}^n A_k$, $A_k \in \mathscr{A}$, p.d., and $B = \sum_{j=1}^m B_j$, $B_j \in \mathscr{A}$, p.d. We have $\sum_{j=1}^m (A_k \cap B_j) \subseteq A_k$, and consequently by Lemma 1.4 there are p.d. sets $\{C_{ki}\}_{i=1}^{m_k} \subseteq \mathscr{A}$ such that

$$A_k = \sum_{j=1}^{m}(A_k \cap B_j) + \sum_{i=1}^{m_k} C_{ki}, \quad k = 1, \ldots, n.$$

Moreover, the sets $\{C_{ki} : i = 1, \dots, m_k;\ k = 1, \dots, n\}$ are p.d., since $C_{ki} \subseteq A_k$ and A_k, $k = 1, \dots, n$, are p.d. We obtain

$$A = \sum_{k=1}^{n} A_k = \underbrace{\sum_{k=1}^{n} \sum_{j=1}^{m} (A_k \cap B_j)}_{=A\cap B} + \sum_{k=1}^{n} \sum_{i=1}^{m_k} C_{ki},$$

and finally

$$A \setminus B = A \setminus (A \cap B) = \sum_{k=1}^{n} \sum_{i=1}^{m_k} C_{ki} \in \mathscr{A}^+.$$

(3) Let, again, $A, B \in \mathscr{A}^+$. We have to show that $A \cup B \in \mathscr{A}^+$.

Let $A = \sum_{k=1}^{n} A_k$ with $A_k \in \mathscr{A}$, p.d., and $B = \sum_{j=1}^{m} B_j$ with $B_j \in \mathscr{A}$, p.d. By Lemma 1.5 applied to the sets $A_1, \dots, A_n, B_1, \dots, B_m$, there exist p.d. sets $\{E_j\}_{j=1}^{N} \subseteq \mathscr{A}$ such that

$$A_k = \sum_{\ell \in M_k} E_\ell, \quad k = 1, \dots, n,$$
$$B_j = \sum_{\ell \in N_j} E_\ell, \quad j = 1, \dots, m,$$

where $M_k, N_j \subseteq \{1, \dots, N\}$. We may assume that all sets E_j, $j = 1, \dots, N$, are used (otherwise we remove those sets E_j that are not used). But then we have $A \cup B = \sum_{\ell=1}^{N} E_\ell \in \mathscr{A}^+$.

Therefore, $\mathscr{A}^+$ is a ring. By Definition 1.3 $r(\mathscr{A}) \subseteq \mathscr{A}^+$.

B. Let now $A \in \mathscr{A}^+$. By the definition of $\mathscr{A}^+$ we have $A = \sum_{i=1}^{n} A_i$ with some p.d. sets $A_1, \dots, A_n \in \mathscr{A}$. Since $\mathscr{A} \subseteq r(\mathscr{A})$, we also have $A_1, \dots, A_n \in r(\mathscr{A})$. But $r(\mathscr{A})$ is a ring and in particular closed under unions. This implies $A = \sum_{i=1}^{n} A_i \in r(\mathscr{A})$.

We therefore have $\mathscr{A}^+ \subseteq r(\mathscr{A})$.

To finish the proof, note that the inclusions $r(\mathscr{A}) \subseteq \mathscr{A}^+$ and $\mathscr{A}^+ \subseteq r(\mathscr{A})$ yield $\mathscr{A}^+ = r(\mathscr{A})$.

□

With this theorem at hand, we can now present an example of a ring: $\mathscr{J}^+$, the set of finite disjoint unions of intervals of the form $(x, y]$, $x, y \in \mathbb{R}$, is a ring.

Definition 1.7 A ring $\mathscr{A}$ is called an *algebra*, if $\Omega \in \mathscr{A}$.

△

Now we are at position to define the main object of this chapter.

Definition 1.8 A family of sets $\mathscr{A} \subset P(\Omega)$ is called a *σ-algebra*, if

(1) $\Omega \in \mathscr{A}$,

(2) $A \in \mathscr{A} \implies \overline{A} := \Omega \setminus A \in \mathscr{A}$,

(3) $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathscr{A} \implies \bigcup\limits_{n=1}^{\infty} A_n \in \mathscr{A}$.

△

Remark It follows from (1) und (2) that $\varnothing = \overline{\Omega} \in \mathscr{A}$. Indeed,

$$\Omega \in A \text{ and } \varnothing = \overline{\Omega} \implies \varnothing \in A.$$

△

Remark Each σ-algebra is a ring. A ring is not necessarily a σ-algebra.

△

Proof. Homework. □

Example The following families of sets are examples of σ-algebras.

A. $\{\varnothing, \Omega\}$, the so-called trivial σ-algebra.

B. $P(\Omega)$, the so-called full σ-algebra.

C. Fix a set $A \subseteq \Omega$, then $\{\varnothing, A, \overline{A}, \Omega\}$ is a σ-algebra.

D. The family

$$\mathscr{A} = \{A \subseteq \Omega : A \text{ is (at most) countable or } \overline{A} \text{ is (at most) countable}\}$$

is a σ-algebra. The proof of this statement is homework.

△

Definition 1.9 Let $\mathscr{E} \subseteq P(\Omega)$ be an arbitrary family of sets. The σ-algebra

$$\sigma(\mathscr{E}) := \bigcap_{\substack{\mathscr{A} \supseteq \mathscr{E} \\ \mathscr{A} \text{ is a } \sigma\text{-algebra}}} \mathscr{A}$$

is called the *σ-algebra generated by* $\mathscr{E}$. The family $\mathscr{E}$ is called a *generator* of $\mathscr{A}$.

△

Remark The intersection of arbitrarily many σ-algebras is a σ-algebra. The proof of this fact is very similar to the proof of the corresponding statement for rings. Prove it as an exercise!

△

Proposition 1.10 (*Properties of the operator σ*) *For any $\mathscr{E}, \mathscr{E}_1, \mathscr{E}_2 \subseteq P(\Omega)$ we have*

(1) *$\mathscr{E}$ is a σ-algebra $\implies \sigma(\mathscr{E}) = \mathscr{E}$.*

(2) *$\sigma(\sigma(\mathscr{E})) = \sigma(\mathscr{E})$ (i.e., σ is idempotent).*

(3) *$\mathscr{E}_1 \subseteq \mathscr{E}_2 \implies \sigma(\mathscr{E}_1) \subseteq \sigma(\mathscr{E}_2)$ (i.e., σ is monotone).*

(4) *$\mathscr{E}_1 \subseteq \mathscr{E}_2 \subseteq \sigma(\mathscr{E}_1) \implies \sigma(\mathscr{E}_2) = \sigma(\mathscr{E}_1)$.*

△

Remark The same properties hold for generated rings.

△

Proof. (1) We have

$$\mathscr{E} \underbrace{\subseteq}_{\text{by definition}} \sigma(\mathscr{E}) = \bigcap_{\substack{\mathscr{A} \supseteq \mathscr{E} \\ \mathscr{A} \text{ is a } \sigma\text{-algebra}}} \mathscr{A} \subseteq \mathscr{E},$$

the last inclusion holds because $\mathscr{E}$ is one of the sets in the intersection. It follows that $\mathscr{E} = \sigma(\mathscr{E})$.

(2) Since $\sigma(\mathscr{E})$ is a σ-algebra, the previous statement for $\sigma(\mathscr{E})$ gives immediately that $\sigma(\sigma(\mathscr{E})) = \sigma(\mathscr{E})$.

(3) We have

$$\sigma(\mathscr{E}_1) = \bigcap_{\substack{\mathscr{A} \supseteq \mathscr{E}_1 \\ \mathscr{A} \text{ is a } \sigma\text{-algebra}}} \mathscr{A} \subseteq \bigcap_{\substack{\mathscr{A} \supseteq \mathscr{E}_2 \\ \mathscr{A} \text{ is a } \sigma\text{-algebra}}} \mathscr{A} = \sigma(\mathscr{E}_2).$$

Because $\mathscr{E}_1 \subseteq \mathscr{E}_2$, for each $\mathscr{A} \supseteq \mathscr{E}_2$ on the right-hand side we also have $\mathscr{A} \supseteq \mathscr{E}_1$. In other words, each $\mathscr{A}$ in the intersection on the right-hand side is also included in the intersection on the left-hand side. Thus, the intersection on the left-hand side is taken over more sets, and it is smaller.

(4) From $\mathscr{E}_1 \subseteq \mathscr{E}_2 \subseteq \sigma(\mathscr{E}_1)$ it follows that

$$\sigma(\mathscr{E}_1) \subseteq \sigma(\mathscr{E}_2) \subseteq \sigma(\sigma(\mathscr{E}_1)) = \sigma(\mathscr{E}_1),$$

where the first and the second inclusions are justified by (3) and the last equality is justified by (2). This implies $\sigma(\mathscr{E}_2) = \sigma(\mathscr{E}_1)$.

□

Remark The σ-algebra $\sigma(\mathscr{E})$ is the smallest σ-algebra (with respect to inclusion) that contains $\mathscr{E}$. Indeed, for each σ-algebra $\widetilde{\mathscr{A}}$ with $\mathscr{E} \subseteq \widetilde{\mathscr{A}}$ we have

$$\sigma(\mathscr{E}) = \bigcap_{\substack{\mathscr{A} \supseteq \mathscr{E} \\ \mathscr{A} \text{ is a } \sigma\text{-algebra}}} \mathscr{A} \subseteq \widetilde{\mathscr{A}}.$$

△

Definition 1.11 Let $\Omega = \mathbb{R}$. The σ-algebra $\mathscr{B} := \sigma(\mathscr{J})$ is called the *Borel σ-algebra*. Its elements are called *Borel sets*.

△

We have $\mathscr{J} \subseteq \mathscr{B} \subset P(\mathbb{R})$. On the other hand, it can be shown that $\mathscr{B} \neq P(\mathbb{R})$. We will not prove this fact.

Proposition 1.12 *All intervals, all open sets and all closed sets in $\mathbb{R}$ are Borel sets.*

△

The proof of this statement for intervals is homework. It can be shown that every open set in $\mathbb{R}$ can be represented as a union of (at most) countably many intervals. We will not prove this fact. It follows from the properties of a σ-algebra that all open sets are Borel sets. Since each closed set is a complement of an open set, we immediately obtain that also closed sets are Borel sets.

Remark Instead of σ-algebras, one can consider δ-algebras. A *δ-algebra* is a family of sets $\mathscr{A} \subseteq P(\Omega)$ such that

(1) $\Omega \in \mathscr{A}$,

(2) $A \in \mathscr{A} \implies \overline{A} := \Omega \setminus A \in \mathscr{A}$,

(3) $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathscr{A} \implies \bigcap_{n=1}^{\infty} A_n \in \mathscr{A}$.

Each σ-algebra is a δ-algebra and vice versa.

△

Proof. Homework. □

Chapter 2
Measures

Let $\mathscr{A} \subseteq P(\Omega)$ be a family of subsets of a nonempty set Ω. A function μ from $\mathscr{A}$ to a set of numbers (like $\mathbb{R}$, $\overline{\mathbb{R}}$, $\mathbb{C}$) is called a *set function*.

The set $\overline{\mathbb{R}}$ above is the set of *extended real numbers*:

$$\overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, \infty\} = [-\infty, \infty].$$

Here, we add to the set $\mathbb{R}$ of real numbers two further numbers $-\infty$, ∞ with the properties $-\infty \neq \infty$ and $-\infty < x < \infty$ for any $x \in \mathbb{R}$.

The rules for calculation with extended real numbers are the following.

A. $x + \infty = \infty + x = \infty \quad \forall\, x \in (-\infty, \infty]$.

B. $x - \infty = -\infty + x = -\infty \quad \forall\, x \in [-\infty, \infty)$.

C. The expressions $\infty - \infty$ and $-\infty + \infty$ are not defined.

D. $x \cdot (\pm\infty) = (\pm\infty) \cdot x = \begin{cases} \pm\infty, & x > 0, \\ 0, & x = 0, \\ \mp\infty, & x < 0. \end{cases}$

E. $\frac{x}{\pm\infty} = 0 \quad \forall\, x \in \mathbb{R}$.

F. $\frac{\pm\infty}{x} = \begin{cases} \pm\infty, & 0 < x < \infty, \\ \mp\infty, & -\infty < x < 0. \end{cases}$

G. The expressions $\frac{\infty}{\infty}$ and $\frac{x}{0}$ for any $x \in \overline{\mathbb{R}}$ are not defined.

In what follows, we will frequently work with set functions, functions, etc., with values in $\overline{\mathbb{R}}$.

Definition 2.1 *(Properties of set functions)* Let $\mathscr{A} \subseteq P(\Omega)$ be a family of sets such that $\varnothing \in \mathscr{A}$, and let $\mu : \mathscr{A} \to \overline{\mathbb{R}}$ be a set function.

(1) μ is called *finitely additive*, if for any finite family $\{A_n\}_{n=1}^N \subseteq \mathscr{A}$ of p.d. sets with $A = \sum_{n=1}^N A_n \in \mathscr{A}$ we have

$$\mu(A) = \sum_{n=1}^{N} \mu(A_n).$$

(2) μ is called *σ-additive*, if for any countable family $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathscr{A}$ of p.d. sets with $A = \sum_{n=1}^\infty A_n \in \mathscr{A}$ we have

$$\mu(A) = \sum_{n=1}^{\infty} \mu(A_n).$$

(3) μ is called *finitely subadditive*, if for any finite family of sets $\{A_n\}_{n=1}^N \subseteq \mathscr{A}$ and any $A \in \mathscr{A}$ such that $A = \bigcup_{n=1}^N A_n$ we have

$$\mu(A) \leq \sum_{n=1}^{N} \mu(A_n).$$

(4) μ is called *σ-subadditive*, if for any countable family $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathscr{A}$ and $A \in \mathscr{A}$ such that $A = \bigcup_{n=1}^\infty A_n$ we have

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

$\triangle$

Remark Each σ-additive set function is also finitely additive. Indeed, to deal with a finite family $\{A_n\}_{n=1}^N \subseteq \mathscr{A}$ of p.d. sets, take $A_{N+1} = A_{N+2} = \cdots = \varnothing$. The same is true for σ-subadditivity.

$\triangle$

Definition 2.2 Let $\mathscr{A}$ be a semiring and $\mu : \mathscr{A} \to \overline{\mathbb{R}}$ be a set function. μ is called a *finitely additive measure* if

(1) $\mu \geq 0$, $\mu(\varnothing) = 0$,

(2) μ is finitely additive.

$\triangle$

Remark If there is $A \in \mathscr{A}$ such that $\mu(A) < \infty$, then then property $\mu(\emptyset) = 0$ follows from the other two properties. Indeed, since $A = A + \emptyset$, we have $\mu(A) = \mu(A) + \mu(\emptyset)$. If $\mu(A) < \infty$, we have

$$\mu(\emptyset) = \mu(A) - \mu(A) = 0.$$

Of course, this is the usual situation, the case when $\mu(A) = \infty$ for each $A \in \mathscr{A}$ is quite exotic.

△

Example Consider the semring $\mathscr{J} = \{(x, y] : x, y \in \mathbb{R}\}$. The set function $\mu : \mathscr{J} \to \mathbb{R}$, $\mu((x, y]) = y - x$ if $y > x$, $\mu(\emptyset) = 0$ is a finitely additive measure.

△

Theorem 2.3 *Let μ be a finitely additive measure on a semiring $\mathscr{A}$. Let $A_1, A_2, \ldots, A_n \in \mathscr{A}$ be p.d. sets, $A \in \mathscr{A}$, and assume that $\sum_{i=1}^{n} A_n \subseteq A$. Then*

$$\sum_{i=1}^{n} \mu(A_i) \leq \mu(A).$$

△

Proof. By Lemma 1.4 there are sets $A_{n+1}, \ldots, A_N \in \mathscr{A}$ such that $A = \sum_{i=1}^{N} A_i$. With these sets we have

$$\mu(A) = \sum_{i=1}^{N} \mu(A_i) \geq \sum_{i=1}^{n} \mu(A_i).$$

□

Definition 2.4 A set function μ defined on a family of sets $\mathscr{A}$ is called *monotone*, if

$$A, B \in \mathscr{A}, A \subseteq B \implies \mu(A) \leq \mu(B).$$

△

Corollary 2.5 *Let μ be a finitely additive measure on a semiring $\mathscr{A}$. Then μ is monotone.*

△

Proof. The statement follows immediately from Theorem 2.3 applied to $A_1, A \in \mathscr{A}$, $A_1 \subseteq A$.

□

Theorem 2.6 *Let μ be a finitely additive measure on a semiring $\mathscr{A}$. If $A, A_1, A_2, \dots, A_n \in \mathscr{A}$ and $A \subseteq \bigcup_{i=1}^n A_i$, then*

$$\mu(A) \leq \sum_{i=1}^n \mu(A_i).$$

In particular, μ is finitely subadditive.

△

Proof. Let $A, A_1, A_2, \dots, A_n \in \mathscr{A}$ and $A \subseteq \bigcup_{i=1}^n A_i$. We want to show that $\mu(A) \leq \sum_{i=1}^n \mu(A_i)$.

By Lemma 1.5 applied to the sets $A, A_1, \dots, A_n$ there exist p.d. sets $\{B_j\}_{j=1}^m \subseteq \mathscr{A}$ such that

$$\begin{aligned} A_k &= \sum_{j \in M_k} B_j, \quad M_k \subseteq \{1, \dots, m\}, \quad k = 1, \dots, n, \\ A &= \sum_{j \in M} B_j, \quad M \subseteq \{1, \dots, m\}. \end{aligned}$$

Let us assume that each of the sets B_j, $j = 1, \dots, m$, is used at least once (otherwise we just remove those sets B_j that are not used). We have

$$\mu(A) = \sum_{j \in M} \mu(B_j) \leq \sum_{j=1}^m \mu(B_j).$$

On the other hand,

$$\sum_{i=1}^n \mu(A_i) = \sum_{i=1}^n \sum_{j \in M_i} \mu(B_j) = \sum_{j=1}^m \mu(B_j).$$

It follows

$$\mu(A) \leq \sum_{j=1}^m \mu(B_j) = \sum_{i=1}^n \mu(A_i).$$

□

Definition 2.7 Let $\mathscr{A}_1, \mathscr{A}_2$ be two semirings, $\mathscr{A}_1 \subseteq \mathscr{A}_2$, and let μ_1, μ_2 by finitely additive measures on $\mathscr{A}_1, \mathscr{A}_2$, respectively. Assume that $\mu_1(A) = \mu_2(A)$ for all $A \in \mathscr{A}_1$. Then

(1) μ_2 is called an *extension* of μ_1 from $\mathscr{A}_1$ to $\mathscr{A}_2$,

(2) μ_1 is called a *restriction* of μ_2 from $\mathscr{A}_2$ to $\mathscr{A}_1$.

△

Theorem 2.8 *Let $\mathscr{A}$ be a semiring and μ a finitely additive measure on $\mathscr{A}$. Then μ can be extended to a finitely additive measure $\tilde{\mu}$ on $r(\mathscr{A})$, and the extension is unique.*

△

Proof. Recall that

$$r(\mathscr{A}) = \mathscr{A}^+ = \left\{ \sum_{i=1}^{n} A_i : A_i \in \mathscr{A},\ \text{p.d.},\ n \in \mathbb{N} \cup \{0\} \right\}.$$

We define the set function $\tilde{\mu}$ on $r(\mathscr{A})$ as follows: For $A = \sum_{i=1}^{n} A_i \in r(\mathscr{A})$ we set

$$\tilde{\mu}(A) := \sum_{i=1}^{n} \mu(A_i).$$

A. First we show that $\tilde{\mu}$ is well-defined, i.e., $\tilde{\mu}(A)$ does not depend on a concrete choice of a representation of $A \in r(\mathscr{A})$ as a finite disjoint union of elements of $\mathscr{A}$. Let

$$A = \sum_{i=1}^{n} A_i = \sum_{j=1}^{m} B_j, \quad A_i, B_j \in \mathscr{A}.$$

We have to show that

$$\sum_{i=1}^{n} \mu(A_i) = \sum_{j=1}^{m} \mu(B_j).$$

For each $k = 1, \dots, n$ we have

$$A_k = A_k \cap A = A_k \cap \left(\sum_{j=1}^{m} B_j \right) = \sum_{j=1}^{m} (A_k \cap B_j),$$

where $A_k \cap B_j \in \mathscr{A}$. It follows that

$$\mu(A_k) = \sum_{j=1}^{m} \mu(A_k \cap B_j).$$

Similarly, for each $j = 1, \dots, m$

$$\mu(B_j) = \sum_{k=1}^{n} \mu(A_k \cap B_j).$$

Now we have

$$\tilde{\mu}(A) = \sum_{k=1}^{n} \mu(A_k) = \sum_{k=1}^{n} \sum_{j=1}^{m} \mu(A_k \cap B_j) = \sum_{j=1}^{m} \sum_{k=1}^{n} \mu(A_k \cap B_j) = \sum_{j=1}^{m} \mu(B_j)$$

as desired.

B. It is clear that by construction $\tilde{\mu}$ is a finitely additive measure on $r(\mathscr{A})$ and that $\tilde{\mu}$ is an extension of μ from $\mathscr{A}$ to $r(\mathscr{A})$.

C. It remains to show that there is only one such extension.

Let $\bar{\mu}$ be another finitely additive extension of μ from $\mathscr{A}$ to $r(\mathscr{A})$. Let $A \in r(\mathscr{A})$, $A = \sum_{i=1}^{n} A_i$, where $A_i \in \mathscr{A}$. Since $\bar{\mu}$ is finitely additive and $\bar{\mu}(A_i) = \mu(A_i)$, $i = 1, \dots, n$, we have

$$\bar{\mu}(A) = \sum_{i=1}^{n} \bar{\mu}(A_i) = \sum_{i=1}^{n} \mu(A_i) = \tilde{\mu}(A),$$

i.e. $\bar{\mu}(A) = \tilde{\mu}(A)$ for all $A \in r(\mathscr{A})$ which means that $\bar{\mu}$ and $\tilde{\mu}$ coincide.

□

Theorem 2.9 *Let μ be a finitely additive measure on a semiring $\mathscr{A}$, and assume that μ is σ-additive. Then its extension to $r(\mathscr{A})$ is also σ-additive.*

△

Proof. Let $\{A_n\}_{n\in\mathbb{N}} \subseteq r(\mathscr{A})$ be p.d. sets, and let $A \in r(\mathscr{A})$ such that $A = \sum_{n=1}^{\infty} A_n$. As above, we denote the extension of μ to $r(\mathscr{A})$ by $\tilde{\mu}$. We have to show that

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

Since $A \in r(\mathscr{A}) = \mathscr{A}^+$, we have $A = \sum_{k=1}^{N} B_k$ with some $B_k \in \mathscr{A}$. Similarly, $A_n = \sum_{k=1}^{N_n} C_{nk}$ with $C_{nk} \in \mathscr{A}$, $n \in \mathbb{N}$. By the construction of $\tilde{\mu}$ we have

$$\tilde{\mu}(A) = \sum_{k=1}^{N} \mu(B_k),$$

$$\tilde{\mu}(A_n) = \sum_{k=1}^{N_n} \mu(C_{nk}), \quad n \in \mathbb{N}.$$

For each $k = 1, \dots, N$ we have

$$B_k = B_k \cap A = B_k \cap \left(\sum_{n=1}^{\infty} A_n \right) = B_k \cap \left(\sum_{n=1}^{\infty} \sum_{j=1}^{N_n} C_{nj} \right) = \sum_{n=1}^{\infty} \sum_{j=1}^{N_n} \underbrace{(B_k \cap C_{nj})}_{\in \mathscr{A}}.$$

Since μ is σ-additive on $\mathscr{A}$, we have

$$\mu(B_k) = \sum_{n=1}^{\infty} \sum_{j=1}^{Nn} \mu(B_k \cap C_{nj}).$$

It follows

$$\tilde{\mu}(A) = \sum_{k=1}^{N} \mu(B_k) = \sum_{k=1}^{N} \sum_{n=1}^{\infty} \sum_{j=1}^{Nn} \mu(B_k \cap C_{nj}).$$

All summands here are non-negative and thus the series is absolutely convergent (possibly to the value $\infty \in \overline{\mathbb{R}}$). This means that we may change the order of summation. We obtain

$$\begin{aligned} \tilde{\mu}(A) &= \sum_{n=1}^{\infty} \sum_{j=1}^{N_n} \sum_{k=1}^{N} \mu(B_k \cap C_{nj}) = \sum_{n=1}^{\infty} \sum_{j=1}^{N_n} \mu(\underbrace{A \cap C_{nj}}_{=C_{nj}}) \\ &= \sum_{n=1}^{\infty} \sum_{j=1}^{N_n} \mu(C_{nj}) = \sum_{n=1}^{\infty} \tilde{\mu}(A_n) \end{aligned}$$

as desired.

□

Remark If the extension $\tilde{\mu}$ on $r(\mathscr{A})$ is σ-additive, then, trivially, also μ on $\mathscr{A}$ is σ-additive.

△

Example Consider the semiring $\mathscr{J} = \{(x, y] : x, y \in \mathbb{R}\}$. The length $\mu((x, y]) = y - x$ for $y > x$, $\mu(\emptyset) = 0$ is σ-additive on $\mathscr{J}$.

△

Definition 2.10 A set function $\mu : \mathscr{A} \to \overline{\mathbb{R}}_+ := [0, +\infty]$ on a family of sets $\mathscr{A} \ni \emptyset$ is called *pre-measure*, if

(1) $\mu(\emptyset) = 0$,

(2) μ is σ-additive.

△

Usually we take in this definition $\mathscr{A}$ to be a semiring or a ring. In this case, a pre-measure is a finitely additive measure which is σ-additive.

Definition 2.11 A pre-measure μ is called

(1) *finite*, if $\mu(A) < \infty$ for all $A \in \mathscr{A}$,

(2) *σ-finite*, if there exists a sequence $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathscr{A}$ such that $\Omega = \bigcup_{n\in\mathbb{N}} A_n$ and $\mu(A_n) < \infty$, $n \in \mathbb{N}$.

△

Theorem 2.12 *Let μ be a finitely additive measure on a ring $\mathscr{A}$. The following properties are equivalent:*

(1) *μ is σ-additive,*

(2) *μ is σ-subadditve,*

(3) *μ is continuous from below in the following sense: If $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathscr{A}$, $A_n \subseteq A_{n+1}$, $n \in \mathbb{N}$, and $\bigcup_{n=1}^{\infty} A_n \in \mathscr{A}$, then*

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

△

Proof. We will prove the implications (1) $\Longrightarrow$ (3) $\Longrightarrow$ (2) $\Longrightarrow$ (1).

A. (1) $\Longrightarrow$ (3)

Assume that μ is σ-additive. Consider a sequence of sets $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathscr{A}$, $A_n \subseteq A_{n+1}$, $n \in \mathbb{N}$. Set $A := \bigcup_{n=1}^{\infty} A_n$, and assume that $A \in \mathscr{A}$. We have to show that

$$\mu(A) = \lim_{n\to\infty} \mu(A_n).$$

We have $A = A_1 + \sum_{n=2}^{\infty}(A_n \setminus A_{n-1})$. Since μ is σ-additive, we have

$$\begin{aligned}
\mu(A) &= \mu(A_1) + \sum_{n=2}^{\infty} \mu(A_n \setminus A_{n-1}) \\
&= \lim_{N\to\infty} \left(\mu(A_1) + \sum_{n=2}^{N} \mu(A_n \setminus A_{n-1})\right) \\
&= \lim_{N\to\infty} \mu\left(A_1 + \sum_{n=2}^{N}(A_n \setminus A_{n-1})\right) \\
&= \lim_{N\to\infty} \mu(A_N)
\end{aligned}$$

as desired.

B. (3) $\implies$ (2)

Assume that μ is continuous from below. Let $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathscr{A}$, $A \in \mathscr{A}$, $A = \bigcup_{n\in\mathbb{N}} A_n$. We have to show that

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

Put $B_N := \bigcup_{n=1}^{N} A_n$, $N \in \mathbb{N}$. Then $B_N \subseteq B_{N+1}$, $N \in \mathbb{N}$, and $A = \bigcup_{n=1}^{\infty} A_n = \bigcup_{N=1}^{\infty} B_N$. Recall that μ is finitely subadditive by Theorem 2.6. Using (3), we obtain

$$\begin{aligned}\mu(A) &= \mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \mu\left(\bigcup_{N=1}^{\infty} B_N\right) = \lim_{N\to\infty} \mu(B_N) = \lim_{N\to\infty} \mu\left(\bigcup_{n=1}^{N} A_n\right)\\ &\leq \lim_{N\to\infty} \sum_{n=1}^{N} \mu(A_n) = \sum_{n=1}^{\infty} \mu(A_n).\end{aligned}$$

C. (2) $\implies$ (1)

Assume that μ is finitely additive and σ-subadditive. Let $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathscr{A}$, p.d., such that $A := \sum_{n=1}^{\infty} A_n \in \mathscr{A}$. We have to show that

$$\mu(A) = \sum_{n=1}^{\infty} \mu(A_n).$$

By (2) we have $\mu(A) \leq \sum_{n=1}^{\infty} \mu(A_n)$. Thus, it remains to show that $\mu(A) \geq \sum_{n=1}^{\infty} \mu(A_n)$. For each $N \in \mathbb{N}$ we have $\sum_{n=1}^{N} A_n \subseteq A$. It follows by Theorem 2.3 that

$$\sum_{n=1}^{N} \mu(A_n) \leq \mu(A).$$

Since this inequality holds for each $N \in \mathbb{N}$, we conclude that the series $\sum_{n=1}^{\infty} \mu(A_n)$ is convergent and

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

The two inequalities together imply

$$\sum_{n=1}^{\infty} \mu(A_n) = \mu(A).$$

□

Theorem 2.13 *Let μ be like in Theorem 2.12. The equivalent properties (1)-(3) imply the following properties:*

(4) *μ is continuous from above in the following sense: If $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathscr{A}$ is such that $A_n \supseteq A_{n+1}$, $n \in \mathbb{N}$, $\bigcap_{n=1}^{\infty} A_n \in \mathscr{A}$ and $\mu(A_1) < \infty$, then*

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

(5) *μ is continuous at zero in the following sense: If $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathscr{A}$ such that $A_n \supseteq A_{n+1}$, $n \in \mathbb{N}$, $\bigcap_{n=1}^{\infty} A_n = \varnothing$ and $\mu(A_1) < \infty$, then*

$$\lim_{n\to\infty} \mu(A_n) = 0.$$

△

Proof. A. (3) $\implies$ (4)

Put $A := \bigcap_{n=1}^{\infty} A_n$. It follows from $\mu(A_1) < \infty$ and the monotonicity property that $\mu(A_n) < \infty$, $n \in \mathbb{N}$, and $\mu(A) < \infty$. Consider the sets $B_n := A_1 \setminus A_n$, $n \in \mathbb{N}$. We have $B_n \subseteq B_{n+1}$, $n \in \mathbb{N}$,

$$\bigcup_{n=1}^{\infty} B_n = \bigcup_{n=1}^{\infty} (A_1 \setminus A_n) = A_1 \setminus \left(\bigcap_{n=1}^{\infty} A_n\right) = A_1 \setminus A.$$

By (3) we have

$$\lim_{n\to\infty} \mu(A_1 \setminus A_n) = \mu(A_1 \setminus A).$$

Since $A_1 = (A_1 \setminus A_n) + A_n$ and μ is finitely additive (even σ-additive), we have

$$\mu(A_1) = \mu(A_1 \setminus A_n) + \mu(A_n), \quad n \in \mathbb{N}.$$

Similarly,

$$\mu(A_1) = \mu(A_1 \setminus A) + \mu(A).$$

Since all the values here are finite, we may subtract. We obtain

$$\begin{aligned}\lim_{n\to\infty} \mu(A_n) &= \lim_{n\to\infty} (\mu(A_1) - \mu(A_1 \setminus A_n)) = \mu(A_1) - \lim_{n\to\infty} \mu(A_1 \setminus A_n) \\ &= \mu(A_1) - \mu(A_1 \setminus A) = \mu(A)\end{aligned}$$

as desired.

B. (4) $\implies$ (5)

This implication is obvious, because (5) is the special case of (4) with $A = \bigcap_{n=1}^{\infty} A_n = \varnothing$, so that

$$\lim_{n\to\infty} \mu(A_n) = \mu(A) = \mu(\varnothing) = 0.$$

$\square$

Remark The statement of Theorem 2.13 does not hold it $\mu(A_1) = \infty$.

$\triangle$

Example Let $A_n = [n, +\infty)$, $n \in \mathbb{N}$. Clearly, $\mu(A_n) = \infty$, $n \in \mathbb{N}$. But $A = \bigcap_{n=1}^{\infty} A_n = \varnothing$ and $\mu(A) = 0$.

$\triangle$

Theorem 2.14 *Let μ be like in Theorem 2.12, and assume that μ is finite. Then the properties (1)-(5) are equivalent.*

$\triangle$

Proof. We already know that (1) $\iff$ (2) $\iff$ (3) $\implies$ (4) $\implies$ (5). We will now prove that (5) $\implies$ (3), and the equivalence of all properties will follow.

Let $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathscr{A}$ such that $A_n \subseteq A_{n+1}$, $n \in \mathbb{N}$, $A = \bigcup_{n=1}^{\infty} A_n \in \mathscr{A}$. We have to show that

$$\mu(A) = \lim_{n\to\infty} \mu(A_n).$$

Put $B_n := A \setminus A_n$, $n \in \mathbb{N}$. We have $B_n \supseteq B_{n+1}$, $n \in \mathbb{N}$, and $\bigcap_{n=1}^{\infty} B_n = \varnothing$. Since μ is finite, $\mu(B_1) < \infty$. By (5), $\lim_{n\to\infty} \mu(B_n) = 0$.

On the other hand, $A = A_n + B_n$, and thus $\mu(A) = \mu(A_n) + \mu(B_n)$, $n \in \mathbb{N}$. It follows that

$$\mu(A_n) = \mu(A) - \underbrace{\mu(B_n)}_{\to 0} \to \mu(A), \quad n \to \infty.$$

$\square$

Definition 2.15 A set function $\mu : \mathscr{A} \to \overline{\mathbb{R}}_+$ is called a *measure*, if

(1) $\mathscr{A}$ is a σ-algebra,

(2) $\mu(\emptyset) = 0$,

(3) μ is σ-additive, i.e., for all families $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathscr{A}$ of p.d. sets

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

△

Remark Note that in this case we automatically have $\sum_{n=1}^{\infty} A_n \in \mathscr{A}$, because $\mathscr{A}$ is a σ-algebra. △

Remark Each σ-algebra is a ring and a semiring, and each measure is a pre-measure and a finitely additive measure. In particular, all properties that we proved in Chapter 2 are valid for measures.

△

Definition 2.16 A triple $(\Omega, \mathscr{A}, \mu)$, where Ω is a nonempty set, $\mathscr{A}$ is a σ-algebra on Ω, and μ is a measure on $\mathscr{A}$, is called a *measure space*.

△

Example Let $\Omega \neq \emptyset$, $\mathscr{A} = P(\Omega)$. The following set functions are measures.

A. $\mu(A) = 0$ for all $A \in \mathscr{A}$, the so-called zero measure.

B. $\mu(A) = \begin{cases} 0, & \text{if } A = \emptyset, \\ \infty, & \text{if } A \neq \emptyset, \end{cases}$ the so-called trivial measure.

C. $\mu(A) = \begin{cases} |A|, & \text{if } A \text{ is finite}, \\ \infty, & \text{if } A \text{ is infinite}, \end{cases}$
where $|A|$ denotes the number of elements in A. This measure is frequently used if A is countable. It is called the counting measure.

D. The *Dirac measure* at the point $\omega \in \Omega$:

$$\delta_\omega(A) = \begin{cases} 1, & \text{if } \omega \in A, \\ 0, & \text{if } \omega \notin A. \end{cases}$$

△

Definition 2.17 If $\mathscr{A}_1$ and $\mathscr{A}_2$ are two σ-algebras on Ω and $\mathscr{A}_2 \subseteq \mathscr{A}_1$, then $\mathscr{A}_2$ is called a *sub-σ-algebra* of $\mathscr{A}_1$.

$\triangle$

Remark If $\mu : \mathscr{A}_1 \to \overline{\mathbb{R}}_+$ is a measure on $\mathscr{A}_1$ and $\mathscr{A}_2$ is a sub-σ-algebra of $\mathscr{A}_1$, then the restriction $\mu|_{\mathscr{A}_2}$ is a measure on $\mathscr{A}_2$.

In particular, all measures defined in the above examples are also measures on any σ-algebra on Ω.

$\triangle$

Definition 2.18 A measure μ is called a *probability measure*, if $\mu(\Omega) = 1$. The measure space $(\Omega, \mathscr{A}, \mu)$ is called in this case a *probability space*.

$\triangle$

Example The Dirac measures are probability measures. $\triangle$

Chapter 3

Extension of measures

Let $\mathscr{A}$ be a ring and $\mu : \mathscr{A} \to \overline{\mathbb{R}}_+$ a pre-measure, i.e., a finitely additive measure on $\mathscr{A}$ which is also σ-additive. Our aim in this section is to extend μ to a measure on a suitable σ-algebra. There are different approaches to this task. We will consider the so-called *Carathéodory extension*.

We start with a quite general construction.

Definition 3.1 A set function $\mu^* : P(\Omega) \to \overline{\mathbb{R}}_+$ is called an *outer measure*, if

(1) $\mu^*(\emptyset) = 0$,

(2) $A \subseteq B \implies \mu^*(A) \leq \mu^*(B)$,

(3) μ^* is σ-subadditive.

△

We will see examples of outer measures later in the lecture and in the tutorial. We will see that an outer measure does not need to be additive.

Definition 3.2 Let $\mu^* : P(\Omega) \to \overline{\mathbb{R}}_+$ be an outer measure. A set $A \subseteq \Omega$ is called *Carathéodory-measurable* (or just *measurable*), if

$$\mu^*(A \cap E) + \mu^*(\overline{A} \cap E) = \mu^*(E) \quad \forall\, E \subseteq \Omega.$$

The class of all Carathéodory-measurable sets will be denoted by $\mathscr{M}(\mu^*)$.

△

It is not intuitively clear what this property means. The set A intersects any set $E \subseteq \Omega$ in such a way that the additivity property of μ^* is fulfilled for the parts $A \cap E$ and $\overline{A} \cap E$.

Remark In Definition 3.2, it is enough to require that

$$\mu^*(A \cap E) + \mu^*(\overline{A} \cap E) \leq \mu^*(E) \quad \forall\, E \subseteq \Omega.$$

Indeed, the reverse inequality follows from the fact that $E = (A \cap E) \cup (\overline{A} \cap E)$ and the subadditivity of μ^*:

$$\mu^*(A \cap E) + \mu^*(\overline{A} \cap E) \geq \mu^*(E).$$

△

Theorem 3.3 *Let $\mu^* : P(\Omega) \to \overline{\mathbb{R}}_+$ be an outer measure. We have:*

(1) *$\mathscr{M}(\mu^*)$ is a σ-algebra,*

(2) *the restriction $\mu^*|_{\mathscr{M}(\mu^*)}$ is a measure.*

△

Proof. A. We first prove that $\mathscr{M}(\mu^*)$ fulfills properties (1) and (2) from the definition of a σ-algebra. We have $\Omega \in \mathscr{M}(\mu^*)$, because

$$\mu^*(\underbrace{\Omega \cap E}_{=E}) + \mu^*(\underbrace{\overline{\Omega} \cap E}_{=\varnothing}) = \mu^*(E) + \mu^*(\varnothing) = \mu^*(E) \quad \forall\, E \subseteq \Omega.$$

The equivalence $A \in \mathscr{M}(\mu^*) \iff \overline{A} \in \mathscr{M}(\mu^*)$ is obvious.

It remains to prove that $\mathscr{M}(\mu^*)$ is closed under countable unions and that $\mu^*|_{\mathscr{M}(\mu^*)}$ is σ-additive.

B. We first consider a union of two sets. Let $A_1, A_2 \in \mathscr{M}(\mu^*)$, and let $E \subseteq \Omega$. We have

$$\begin{aligned}
\mu^*(E) &= \mu^*(E \cap A_1) + \mu^*(E \cap \overline{A_1}) \\
&= \mu^*(E \cap A_1) + \mu^*(E \cap \overline{A_1} \cap A_2) + \mu^*(E \cap \overline{A_1} \cap \overline{A_2}) \\
&= \mu^*(E \cap (A_1 \cup A_2) \cap A_1) + \mu^*(E \cap (A_1 \cup A_2) \cap \overline{A_1}) + \mu^*(E \cap \overline{(A_1 \cup A_2)}) \\
&= \mu^*(E \cap (A_1 \cup A_2)) + \mu^*(E \cap \overline{(A_1 \cup A_2)}),
\end{aligned}$$

and thus $A_1 \cup A_2$ is measurable, i.e. $A_1 \cup A_2 \in \mathscr{M}(\mu^*)$.

Now assume that $A_1, A_2 \in \mathscr{M}(\mu^*)$ and $A_1 \cap A_2 = \varnothing$. Let $E \subseteq \Omega$. We have

$$\begin{aligned}
\mu^*(E \cap (A_1 + A_2)) &= \mu^*(E \cap (A_1 + A_2) \cap A_1) + \mu^*(E \cap (A_1 + A_2) \cap \overline{A_1}) \\
&= \mu^*(E \cap A_1) + \mu^*(E \cap A_2).
\end{aligned}$$

If we repeat this rule $N-1$ times, we obtain

$$\mu^*\left(E \cap \sum_{n=1}^{N} A_n\right) = \sum_{n=1}^{N} \mu^*(E \cap A_n)$$

for p.d. sets $A_1, \dots, A_N \in \mathscr{M}(\mu^*)$. In particular, if we take $E = \Omega$, we get

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

We have shown so far that $\mathscr{M}(\mu^*)$ is closed under finite unions and that $\mu^*|_{\mathscr{M}(\mu^*)}$ is finitely additive.

c. Now consider $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathscr{M}(\mu^*)$, p.d. It follows from the previous step that for each $N \in \mathbb{N}$ we have $\sum_{n=1}^N A_n \in \mathscr{M}(\mu^*)$ and

$$\mu^*\left(E \cap \sum_{n=1}^N A_n\right) = \sum_{n=1}^N \mu^*(E \cap A_n) \quad \forall\, E \subseteq \Omega.$$

Using this, we write

$$\begin{aligned}
\mu^*(E) &= \mu^*\left(E \cap \sum_{n=1}^N A_n\right) + \mu^*\left(E \cap \overline{\sum_{n=1}^N A_n}\right) \\
&= \sum_{n=1}^N \mu^*(E \cap A_n) + \mu^*\left(E \cap \overline{\sum_{n=1}^N A_n}\right) \\
&\geq \sum_{n=1}^N \mu^*(E \cap A_n) + \mu^*\left(E \cap \overline{\sum_{n=1}^\infty A_n}\right),
\end{aligned}$$

where in the last step we used the monotonicity property. Since this is true for all $N \in \mathbb{N}$, we have

$$\mu^*(E) \geq \sum_{n=1}^\infty \mu^*(E \cap A_n) + \mu^*\left(E \cap \overline{\sum_{n=1}^\infty A_n}\right).$$

Using the σ-additivity property twice, we further estimate

$$\begin{aligned}
\mu^*(E) &\geq \sum_{n=1}^\infty \mu^*(E \cap A_n) + \mu^*\left(E \cap \overline{\sum_{n=1}^\infty A_n}\right) \\
&\geq \mu^*\left(E \cap \sum_{n=1}^\infty A_n\right) + \mu^*\left(E \cap \overline{\sum_{n=1}^\infty A_n}\right) \geq \mu^*(E).
\end{aligned}$$

This proves that

$$\mu^*(E) = \mu^*\left(E \cap \sum_{n=1}^\infty A_n\right) + \mu^*\left(E \cap \overline{\sum_{n=1}^\infty A_n}\right),$$

and thus $\sum_{n=1}^{\infty} A_n \in \mathscr{M}(\mu^*)$. It also follows that

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

Taking $E = \Omega$ in the above formula, we obtain $\sum_{n=1}^{\infty} \mu^*(A_n) = \mu^* \left(\sum_{n=1}^{\infty} A_n \right)$.

We have shown so far that $\mathscr{M}(\mu^*)$ is closed under finite unions as well under countable unions of p.d. sets and that $\mu^*|_{\mathscr{M}(\mu^*)}$ is σ-additive.

D. For $A_1, A_2 \in \mathscr{M}(\mu^*)$ we have by above

$$A_1 \cap A_2 = \overline{\overline{A_1 \cap A_2}} = \overline{\overline{A_1} \cup \overline{A_2}} \in \mathscr{M}(\mu^*),$$

and thus $\mathscr{M}(\mu^*)$ is also closed under finite intersections.

E. In the last step we will show that $\mathscr{M}(\mu^*)$ is closed under countable unions (of not necessarily pairwise disjoint sets). Let $\{A_n\}_{n \in \mathbb{N}} \subseteq \mathscr{M}(\mu^*)$. We have

$$\bigcup_{n=1}^{\infty} A_n = \sum_{n=1}^{\infty} \left(A_n \setminus \bigcup_{m=1}^{n-1} A_m \right) = \sum_{n=1}^{\infty} \left(A_n \cap \overline{\bigcup_{m=1}^{n-1} A_m} \right).$$

The facts we proved so far imply that $\bigcup_{m=1}^{n-1} A_m \in \mathscr{M}(\mu^*)$, and thus also $\overline{\bigcup_{m=1}^{n-1} A_m} \in \mathscr{M}(\mu^*)$ and $A_n \cap \overline{\bigcup_{m=1}^{n-1} A_m} \in \mathscr{M}(\mu^*)$. It follows that

$$\bigcup_{n=1}^{\infty} A_n = \sum_{n=1}^{\infty} \left(A_n \cap \overline{\bigcup_{m=1}^{n-1} A_m} \right) \in \mathscr{M}(\mu^*)$$

as a countable union of p.d. sets from $\mathscr{M}(\mu^*)$.

□

Now we introduce an important example of an outer measure.

Definition 3.4 Let $\mathscr{E} \subseteq P(\Omega)$ be a family of sets with $\varnothing \in \mathscr{E}$, and let $\mu : \mathscr{E} \to \overline{\mathbb{R}}_+$ be a set function such that $\mu(\varnothing) = 0$. The set function $\mu^* : P(\Omega) \to \overline{\mathbb{R}}_+$ defined as

$$\mu^*(A) := \inf \left\{ \sum_{n=1}^{\infty} \mu(E_n) : \{E_n\}_{n \in \mathbb{N}} \subseteq \mathscr{E},\ A \subseteq \bigcup_{n=1}^{\infty} E_n \right\}$$

is called the *outer measure induced by μ*. A family $\{E_n\}_{n \in \mathbb{N}}$ as above is called a *covering* of A. If A does not possess any covering, we set $\mu^*(A) := \infty$.

△

Theorem 3.5 *μ^* from Definition 3.4 is an outer measure.*

△

Proof. We will show that the properties (1)-(3) from Definition 3.1 are fulfilled for μ^*.

(1) Since $\varnothing \in \mathscr{E}$, the family $\{\varnothing\}$ is a covering of $\varnothing$. It follows that

$$0 \leq \mu^*(\varnothing) \leq \mu(\varnothing) = 0,$$

and thus $\mu^*(\varnothing) = 0$.

(2) Let $A \subseteq B$. If $\{E_n\}_{n\in\mathbb{N}}$ is a covering of B, then $\{E_n\}_{n\in\mathbb{N}}$ is also a covering A. This implies the inequality $\mu^*(A) \leq \mu^*(B)$.

(3) It remains to show that μ^* is σ-subadditive. Let $\{A_n\}_{n\in\mathbb{N}} \subseteq P(\Omega)$. We have to show that

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

If there is $n \in \mathbb{N}$ such that $\mu^*(A_n) = \infty$, then by monotonicity also $\mu^*\left(\bigcup_{n=1}^{\infty} A_n\right) = \infty$, and the inequality is fulfilled as both sides in it are ∞. Therefore, let us now assume that $\mu^*(A_n) < \infty$ for all $n \in \mathbb{N}$. Let $\varepsilon > 0$. By definition, for each $n \in \mathbb{N}$ there exists a covering $\{E_{nm}\}_{m\in\mathbb{N}} \subseteq \mathscr{E}$ of A_n such that $A_n \subseteq \bigcup_{m=1}^{\infty} E_{nm}$ and

$$\sum_{m=1}^{\infty} \mu(E_{nm}) \leq \mu^*(A_n) + \frac{\varepsilon}{2^n}.$$

It follows that $\bigcup_{n=1}^{\infty} A_n \subseteq \bigcup_{n=1}^{\infty}\bigcup_{m=1}^{\infty} E_{mn}$ so that $\{E_{nm}\}_{n,m\in\mathbb{N}}$ is a covering of $\bigcup_{n=1}^{\infty} A_n$. We have

$$\begin{aligned}\mu^*\left(\bigcup_{n=1}^{\infty} A_n\right) &\leq \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} \mu(E_{nm}) \leq \sum_{n=1}^{\infty}\left(\mu^*(A_n) + \frac{\varepsilon}{2^n}\right) \\ &= \sum_{n=1}^{\infty} \mu^*(A_n) + \varepsilon \sum_{n=1}^{\infty} \frac{1}{2^n} = \sum_{n=1}^{\infty} \mu^*(A_n) + \varepsilon.\end{aligned}$$

Since this is true for any $\varepsilon > 0$, it holds $\mu^*\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} \mu^*(A_n)$.

□

Now we are in the position to extend a pre-measure defined on a ring to a measure on the σ-algebra generated by this ring.

Theorem 3.6 *(Carathéodory Extension Theorem)* *Let $\mathscr{A}$ be a ring and let $\mu : \mathscr{A} \to \overline{\mathbb{R}}_+$ be a pre-measure on $\mathscr{A}$. Let μ^* be the outer measure generated by μ. Then*

(1) $\mu^*|_{\mathscr{A}} = \mu$,

(2) $\mathscr{A} \subseteq \mathscr{M}(\mu^*)$.

In particular, $\mu^*|_{\sigma(\mathscr{A})}$ is an extension of μ to a measure on $\sigma(\mathscr{A})$.

△

Proof. We already know from Theorem 3.3 that $\mathscr{M}(\mu^*)$ is a σ-algebra and $\mu^*|_{\mathscr{M}(\mu^*)}$ is a measure. (2) implies $\sigma(\mathscr{A}) \subseteq \mathscr{M}(\mu^*)$, and (1) and (2) together imply the last statement.

Let us now prove (1) and (2).

(1) Let $A \in \mathscr{A}$ and $\{E_n\}_{n\in\mathbb{N}} \subseteq \mathscr{A}$ be a covering of A, i.e. $A \subseteq \bigcup_{n=1}^{\infty} E_n$. We have

$$\mu(A) = \mu\left(\bigcup_{n=1}^{\infty} \underbrace{(A \cap E_n)}_{\in\mathscr{A}}\right) \le \sum_{n=1}^{\infty} \mu(A \cap E_n) \le \sum_{n=1}^{\infty} \mu(E_n),$$

where we used σ-subadditivity in the penultimate estimate, and monotonicity in the last estimate. Since this inequality is true for any covering, we have $\mu(A) \le \mu^*(A)$.

On the other hand, $\{A\}$ is a covering of A. This gives the estimate $\mu^*(A) \le \mu(A)$.

Altogether we have $\mu(A) = \mu^*(A)$, $A \in \mathscr{A}$. This proves (1).

(2) Let $A \in \mathscr{A}$. We want to show that $A \in \mathscr{M}(\mu^*)$, i.e.

$$\mu^*(A \cap E) + \mu^*(\overline{A} \cap E) = \mu^*(E) \quad \forall\, E \subseteq \Omega.$$

If E does not possess a covering, then at least one of the sets $A \cap E$, $\overline{A} \cap E$ does not possess a covering (otherwise the union of coverings of $A \cap E$ and $\overline{A} \cap E$ would be a covering of E). Thus, in this case both sides of the equation are ∞.

Now assume that $\{E_n\}_{n\in\mathbb{N}} \subseteq \mathscr{A}$ is a covering of E, i.e $E \subseteq \bigcup_{n=1}^{\infty} E_n$. We have

$$A \cap E \subseteq \bigcup_{n=1}^{\infty} \underbrace{(A \cap E_n)}_{\in \mathscr{A}}$$

and

$$\overline{A} \cap E \subseteq \bigcup_{n=1}^{\infty} \underbrace{(\overline{A} \cap E_n)}_{=E_n \setminus A \in \mathscr{A}}.$$

Thus, $\{A \cap E_n\}_{n\in\mathbb{N}}$ is a covering of $A \cap E$, and $\{\overline{A} \cap E_n\}_{n\in\mathbb{N}}$ is a covering of $\overline{A} \cap E$. It follows

$$\begin{aligned} \mu^*(A \cap E) + \mu^*(\overline{A} \cap E) &\leq \sum_{n=1}^{\infty} \mu(A \cap E_n) + \sum_{n=1}^{\infty} \mu(\overline{A} \cap E_n) \\ = \sum_{n=1}^{\infty} \Big(\mu(A \cap E_n) + \mu(\overline{A} \cap E_n)\Big) &= \sum_{n=1}^{\infty} \mu(E_n), \end{aligned}$$

where in the last step we used the additivity of μ.

Since this inequality is true for any covering $\{E_n\}_{n\in\mathbb{N}}$ of E, we obtain

$$\mu^*(A \cap E) + \mu^*(\overline{A} \cap E) \leq \mu^*(E).$$

The reverse inequality follows from the σ-subadditivity (see Remark after Definition 3.2). This proves that $A \in \mathscr{M}(\mu^*)$.

□

Remark It can be shown that the extension of a pre-measure μ from a ring $\mathscr{A}$ to a measure on $\sigma(\mathscr{A})$ is unique if μ is σ-finite.

△

Example We have already constructed a pre-measure μ on the ring

$$r(\mathscr{J}) = \mathscr{J}^+ = \left\{ \sum_{i=1}^{n} (x_i, y_i] : x_i, y_i \in \mathbb{R} \text{ for } i = 1, \dots, n,\ n \in \mathbb{N} \cup \{0\} \right\},$$

namely,

$$\mu\left(\sum_{i=1}^{n} (x_i, y_i]\right) = \sum_{i=1}^{n} \max(y_i - x_i, 0).$$

By Theorem 3.6, μ can be extended to a measure λ on the Borel σ-algebra $\mathscr{B}$. The resulting measure λ is called the *Lebesgue measure* (or the *Lebesgue-Borel measure*) on $\mathbb{R}$.

It can be shown that this extension is unique (this is so because the measure is σ-finite).

$\triangle$

Completeness of a measure

Definition 3.7 Let $(\Omega, \mathscr{A}, \mu)$ be a measure space. A set $N \in \mathscr{A}$ is called *μ-null set*, if $\mu(N) = 0$.

$\triangle$

Example Each point on the real line $\mathbb{R}$ is a λ-null set.

Indeed, let us consider the set $\{x\}$, where $x \in \mathbb{R}$. For each $n \in \mathbb{N}$ we have $\{x\} \subseteq \left(x - \frac{1}{n}, x + \frac{1}{n}\right]$, and by the monotonicity property

$$0 \leq \lambda(\{x\}) \leq \lambda\left(\left(x - \frac{1}{n}, x + \frac{1}{n}\right]\right) = \frac{2}{n} \to 0, \quad n \to \infty,$$

which implies $\lambda(\{x\}) = 0$.

$\triangle$

Definition 3.8 A measure space $(\Omega, \mathscr{A}, \mu)$ is called *complete*, if each subset of any null set is measurable (with the measure zero), i.e.

$$A \subseteq N \in \mathscr{A},\ \mu(N) = 0 \implies A \in \mathscr{A},\ \mu(A) = 0.$$

$\triangle$

Not every measure space is complete.

Theorem 3.9 *Let $\mu^* : P(\Omega) \to \overline{\mathbb{R}}_+$ be an outer measure. Then its restriction $\mu^*|_{\mathscr{M}(\mu^*)} : \mathscr{M}(\mu^*) \to \overline{\mathbb{R}}_+$ is a complete measure.*

$\triangle$

Proof. Let $N \in \mathscr{M}(\mu^*)$, $\mu^*(N) = 0$. Take a set $A \subseteq N$. Because of the monotonicity property of the outer measure μ^* on $P(\Omega)$, we have $\mu^*(A) = 0$. We will now show that all sets $A \subseteq \Omega$ with $\mu^*(A) = 0$ are Carathéodory-measurable, i.e. $A \in \mathscr{M}(\mu^*)$.

Let $E \subseteq \Omega$ be an arbitrary set. We have

$$0 \leq \mu^*(A \cap E) \leq \mu^*(A) = 0,$$

and thus $\mu^*(A \cap E) = 0$. On the other hand,

$$\mu^*(E) \leq \underbrace{\mu^*(E \cap A)}_{=0} + \mu^*(E \cap \overline{A}) = \mu^*(E \cap \overline{A}) \leq \mu^*(E)$$

which implies $\mu^*(E \cap \overline{A}) = \mu^*(E)$. Together this gives

$$\mu^*(E \cap A) + \mu^*(E \cap \overline{A}) = 0 + \mu^*(E) = \mu^*(E),$$

and it follows that A is measurable.

□

If a measure space $(\Omega, \mathscr{A}, \mu)$ is not complete, it can be completed. In this process μ is extended to a somewhat larger σ-algebra. See homework for details.

A short remark on Lebesgue's extension of measures

Lebesgue's extension is a different approach to extending a pre-measure on a ring to a measure on the generated σ-algebra. For simplicity, we describe the construction of Lebesgue for the Lebesgue measure on $\mathbb{R}$.

We have already constructed a pre-measure μ on the ring $r(\mathscr{J}) = \mathscr{J}^+$,

$$\mu\left(\sum_{i=1}^{n}(x_i, y_i]\right) = \sum_{i=1}^{n} \max(y_i - x_i, 0).$$

Let μ^* be the outer measure induced by μ. A bounded set $A \subseteq \mathbb{R}$ is called *Lebesgue-measurable* if for any $\varepsilon > 0$ there is a set $A_\varepsilon \in \mathscr{J}^+$ such that

$$\mu^*(A \triangle A_\varepsilon) < \varepsilon.$$

Here, $\triangle$ is the symmetric difference:

$$A \triangle A_\varepsilon = (A \setminus A_\varepsilon) \cup (A_\varepsilon \setminus A).$$

For a bounded measurable set, we set $\lambda(A) := \mu^*(A)$. For an unbounded set $A \subseteq \mathbb{R}$, we say that A is measurable if there exists

$$\lim_{n \to \infty} \lambda(A \cap [-n, n]) =: \lambda(A).$$

One can similarly define Lebesgue's extension of a pre-measure from a ring to a measure on the generated σ-algebra in the general case, for σ-finite measures. Lebesgue's extension and Carathéodory's extension coincide under reasonable assumptions (for example, for the Lebesgue measure on $\mathbb{R}$), but Carathéodory's extension is more general.

On Lebesgue-Stieltjes measures on $\mathbb{R}$

Let $\mu : \mathscr{B} \to \overline{\mathbb{R}}_+$ be a measure on the Borel σ-algebra $\mathscr{B} \subseteq P(\mathbb{R})$ such that the measure of each compact interval is finite. Such measures are called *Borel measures*.

Consider the function $g : \mathbb{R} \to \mathbb{R}$,

$$g(x) := \begin{cases} \mu((0,x]), & x > 0, \\ -\mu((x,0]), & x \leq 0. \end{cases}$$

It follows from the monotonicity of μ that g is monotonically non-decreasing. One can also show hat g is right continuous, i.e. $\lim_{t \to x, t \geq x} g(t) = g(x)$.

One can show that every monotonically non-decreasing, right continuous function $g : \mathbb{R} \to \mathbb{R}$ defines a Borel measure μ with the property

$$\mu((x,y]) = g(y) - g(x), \quad y > x.$$

Such measures are called *Lebesgue-Stieltjes measures*.

Example A. Take $g(x) = x$, then the Lebesgue-Stieltjes measure is the usual Lebesgue measure.

B. Take $g(x) = \arctan x$. The corresponding Lebesgue-Stieltjes measure satisfies

$$\mu((x,y]) = \arctan y - \arctan x, \quad y > x.$$

In particular,

$$\mu(\mathbb{R}) = \lim_{t \to \infty} \arctan t - \lim_{t \to -\infty} \arctan t = \frac{\pi}{2} - \left(-\frac{\pi}{2}\right) = \pi.$$

△

On the existence of non-measurable sets

The measure space $(\mathbb{R}, \mathscr{B}, \lambda)$ can be extended by completion to a larger measure space that contains sets which are measurable but not Borel. A question arises: Do there exist sets in $\mathbb{R}$ that are not measurable?

The answer depends on the *Axiom of Choice* that says: For any family $\mathscr{A}$ of nonempty sets, there exists a function defined on $\mathscr{A}$ which maps each set $A \in \mathscr{A}$ to an element $a \in A$.

If one does not accept the Axiom of Choice, a model of set theory can be constructed in which every subset of $\mathbb{R}$ is measurable.

In this course, I will accept the Axiom of Choice. Below we will present an example of a non-measurable subset of $\mathbb{R}$.

Example (*Vitali*) We consider the interval $[0,1]$ and the following equivalence relation:

$$x \sim y \quad \text{if} \quad x - y \in \mathbb{Q}.$$

Let $[x]$ denote the equivalence class of $x \in [0,1]$. Using the Axiom of Choice, we choose one representative from each equivalence class. Let E be the set of these representatives. We claim that E is not Lebesgue-measurable.

The Lebesgue measure λ is *translation invariant*, i.e. $\lambda(A) = \lambda(x + A)$ for every measurable set $A \subseteq \mathbb{R}$ and every $x \in \mathbb{R}$. We will not prove this fact in detail; intuitively this is quite clear because

$$\lambda((a,b]) = \lambda((a+x, b+x]) = b - a \quad \forall\, x \in \mathbb{R}.$$

Let us show that E is not measurable. The case that $\lambda(E) = 0$ is impossible because $[0,1] \subseteq \bigcup_{q\in\mathbb{Q}}(E+q)$, and $\lambda(E) = 0$ would imply that

$$\lambda([0,1]) \leq \sum_{q\in\mathbb{Q}} \lambda(E+q) = 0.$$

But also $\lambda(E) = c > 0$ is not possible. Indeed, it follows from the definition of the equivalence classes that $(E+p) \cap (E+q) = \emptyset$ for $p, q \in \mathbb{Q}$, $p \neq q$. We have $\bigcup_{p\in\mathbb{Q}\cap[0,1]}(E+p) \subseteq [0,2]$, and this is a disjoint union. Now $\lambda(E) = c > 0$ would imply

$$\lambda([0,2]) \geq \sum_{p\in\mathbb{Q}\cap[0,1]} \lambda(E+p) = \infty.$$

These contradictions show that E is not measurable.

$\triangle$

Chapter 4

Measurable functions

Definition 4.1 A pair $(\Omega, \mathscr{A})$, where $\Omega \neq \varnothing$ and $\mathscr{A}$ is a σ-algebra on Ω, is called a *measurable space*. Sets $A \in \mathscr{A}$ are called *$\mathscr{A}$-measurable sets*.

$\triangle$

In the probability theory, measurable sets are called *events*.

Definition 4.2 Let $(\Omega, \mathscr{A})$, $(\Omega', \mathscr{A}')$ be two measurable spaces. A mapping $T : \Omega \to \Omega'$ is called *$(\mathscr{A}, \mathscr{A}')$-measurable*, if

$$\forall A' \in \mathscr{A}' \ : \ T^{-1}(A') \in \mathscr{A}.$$

$\triangle$

Here, $T^{-1}(A') = \{\omega \in \Omega : T(\omega) \in A'\}$ is the pre-image of A'. Thus, measurability of T means that pre-images of measurable sets are measurable.

For a family of sets $\mathscr{E}' \subseteq P(\Omega')$ we use the notation

$$T^{-1}(\mathscr{E}') := \{T^{-1}(E') : E' \in \mathscr{E}'\}.$$

Proposition 4.3 *We have*

$$T \text{ is } (\mathscr{A}, \mathscr{A}')\text{-measurable} \iff T^{-1}(\mathscr{A}') \subseteq \mathscr{A}.$$

$\triangle$

This is just a reformulation of the definition. We will also frequently write just "T is measurable", when it is clear from the context which σ-algebras are considered.

If $(\Omega, \mathscr{A}, P)$ is a probability space and $(E, \tilde{\mathscr{A}})$ is a measurable space, a $(\mathscr{A}, \tilde{\mathscr{A}})$-measurable function $X : \Omega \to E$ is called a *random variable*.

Example A. A constant mapping is always measurable (for all possible σ-algebras $\mathscr{A}$, $\mathscr{A}'$).

B. Each mapping $T : \Omega \to \Omega'$ is $(P(\Omega), \mathscr{A}')$-measurable.

C. Fix a set $A \subseteq \Omega$. The function $\chi_A : \Omega \to \mathbb{R}$,

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

is called the *characteristic function* of the set A.

We claim that χ_A is measurable if and only if $A \in \mathscr{A}$.

Indeed, for a set $B \in \mathscr{B}$ we have

$$\chi_A^{-1}(B) = \begin{cases} \Omega, & \text{if } 0 \in B, 1 \in B, \\ \varnothing, & \text{if } 0 \notin B, 1 \notin B, \\ A, & \text{if } 1 \in B, 0 \notin B, \\ \overline{A}, & \text{if } 0 \in B, 1 \notin B. \end{cases}$$

We see that χ_A is measurable if and only if $A, \overline{A} \in \mathscr{A}$. It remains to note that $A \in \mathscr{A} \iff \overline{A} \in \mathscr{A}$.

△

Proposition 4.4 *Let $(\Omega_1, \mathscr{A}_1)$, $(\Omega_2, \mathscr{A}_2)$, $(\Omega_3, \mathscr{A}_3)$ be three measurable spaces, $T_1 : \Omega_1 \to \Omega_2$ be $(\mathscr{A}_1, \mathscr{A}_2)$-measurable, $T_2 : \Omega_2 \to \Omega_3$ be $(\mathscr{A}_2, \mathscr{A}_3)$-measurable. Then $T_2 \circ T_1 : \Omega_1 \to \Omega_3$ is $(\mathscr{A}_1, \mathscr{A}_3)$-measurable.*

△

Proof. Let $A_3 \in \mathscr{A}_3$, then

$$(T_2 \circ T_1)^{-1}(A_3) = \underbrace{T_1^{-1}(\underbrace{T_2^{-1}(A_3)}_{\in \mathscr{A}_2})}_{\in \mathscr{A}_1}.$$

□

Recall the following rules for pre-images: Let $T : \Omega \to \Omega'$, $A_i' \subseteq \Omega'$, $i \in I$, where I is an arbitrary index set. We have

$$T^{-1}\left(\bigcup_{i \in I} A_i'\right) = \bigcup_{i \in I} T^{-1}(A_i'),$$

$$T^{-1}\left(\bigcap_{i \in I} A_i'\right) = \bigcap_{i \in I} T^{-1}(A_i'),$$

$$T^{-1}(\overline{A_i'}) = \overline{T^{-1}(A_i')}.$$

These rules imply the following statements.

Proposition 4.5 (1) *Let $\Omega \neq 0$, $(\Omega', \mathscr{A}')$ be a measurable space, $T : \Omega \to \Omega'$. Then the family $T^{-1}(\mathscr{A}')$ is a σ-algebra on Ω.*

(2) *Let $(\Omega, \mathscr{A})$ be a measurable space, $\Omega' \neq \emptyset$, $T : \Omega \to \Omega'$. Then the family*

$$\{A' \subseteq \Omega' : T^{-1}(A') \in \mathscr{A}\}$$

is a σ-algebra on Ω'.

△

Proof. Homework. □

Proposition 4.6 *Let $\Omega, \Omega' \neq \emptyset$, $T : \Omega \to \Omega'$, $\mathscr{E}' \subseteq P(\Omega')$ be an arbitrary family of sets. Then*

$$T^{-1}(\sigma(\mathscr{E}')) = \sigma(T^{-1}(\mathscr{E}')).$$

△

Note that here $\sigma(\mathscr{E}')$ is a σ-algebra on Ω', and $\sigma(T^{-1}(\mathscr{E}'))$ is a σ-algebra on Ω.

Proof. Obviously, $T^{-1}(\mathscr{E}') \subseteq T^{-1}(\sigma(\mathscr{E}'))$. On the other hand, $T^{-1}(\sigma(\mathscr{E}'))$ is a σ-algebra by Proposition 4.5. It follows that $\sigma(T^{-1}(\mathscr{E}')) \subseteq T^{-1}(\sigma(\mathscr{E}'))$.

Let us show the reverse inclusion. The family

$$\mathscr{A}' := \{A' \subseteq \Omega' : T^{-1}(A') \in \sigma(T^{-1}(\mathscr{E}'))\}$$

is a σ-algebra on Ω' by Proposition 4.5. For any $E' \in \mathscr{E}'$ we have $T^{-1}(E') \in T^{-1}(\mathscr{E}') \subseteq \sigma(T^{-1}(\mathscr{E}'))$ which implies that $E' \in \mathscr{A}'$. Consequently, $\mathscr{E}' \subseteq \mathscr{A}'$.

Since $\mathscr{A}'$ is a σ-algebra, it follows that $\sigma(\mathscr{E}') \subseteq \mathscr{A}'$. This implies

$$T^{-1}(\sigma(\mathscr{E}')) \subseteq T^{-1}(\mathscr{A}') = \sigma(T^{-1}(\mathscr{E}')).$$

Altogether we obtain $T^{-1}(\sigma(\mathscr{E}')) = \sigma(T^{-1}(\mathscr{E}'))$ as desired.

□

Theorem 4.7 *Let $(\Omega, \mathscr{A})$, $(\Omega', \mathscr{A}')$ be measurable spaces, $T : \Omega \to \Omega'$, and let $\mathscr{E}' \subseteq P(\Omega')$ be a generator of $\mathscr{A}'$, i.e. $\mathscr{A}' = \sigma(\mathscr{E}')$. Then*

$$T \text{ is } (\mathscr{A}, \mathscr{A}')\text{-measurable} \iff T^{-1}(\mathscr{E}') \subseteq \mathscr{A}.$$

△

Remark The statement of the theorem means the following: To show mesuarability of a mapping, it is enough to consider a generator of $\mathscr{A}'$.

△

Proof. The implication $\Longrightarrow$ holds trivially. Let us show the implication $\Longleftarrow$. With the use of Proposition 4.6 we write

$$T^{-1}(\mathscr{A}') = T^{-1}(\sigma(\mathscr{E}')) = \sigma(T^{-1}(\mathscr{E}')) \subseteq \mathscr{A}.$$

The last inclusion follows from the fact that $T^{-1}(\mathscr{E}') \subseteq \mathscr{A}$ and $\mathscr{A}$ is a σ-algebra. The formula above means that T is measurable.

□

In the remaining part of this chapter we will consider real-valued functions $f : \mathbb{R} \to \mathbb{R}$ or extended real-valued functions $f : \mathbb{R} \to \overline{\mathbb{R}}$. We will always consider $\mathbb{R}$ with the Borel σ-algebra $\mathscr{B}$.

Example Let $f : \mathbb{R} \to \mathbb{R}$.

(1) f is continuous $\Longrightarrow$ f is measurable.

(2) f is monotone $\Longrightarrow$ f is measurable.

Indeed, by Theorem 4.7 it is enough to consider a generator of $\mathscr{B}$. We take the generator $\{(x,y) : x,y \in \mathbb{R}\}$. If f is continuous, then $f^{-1}((x,y))$ is an open set and thus $f^{-1}((x,y)) \in \mathscr{B}$. If f is monotone, then $f^{-1}((x,y))$ is an interval and thus, again, $f^{-1}((x,y)) \in \mathscr{B}$. In both cases we see that f is measurable.

△

The Borel σ-algebra on $\overline{\mathbb{R}}$ is

$$\begin{aligned}\overline{\mathscr{B}} &:= \{B \in \overline{\mathbb{R}} : B \cap \mathbb{R} \in \mathscr{B}\}\\ &= \mathscr{B} \cup \{B \cup \{\infty\}, B \cup \{-\infty\}, B \cup \{\infty, -\infty\} : B \in \mathscr{B}\}.\end{aligned}$$

One can easily check directly that $\overline{\mathscr{B}}$ is a σ-algebra. Moreover,

$$\overline{\mathscr{B}} = \sigma(\{[-\infty, x] : x \in \mathbb{R}\}).$$

(Homework.)

In what follows we will use the following notation: For $f, g : \Omega \to \overline{\mathbb{R}}$, $B \subseteq \overline{\mathbb{R}}$, $x \in \overline{\mathbb{R}}$ we write

$$
\begin{aligned}
\{f \in B\} &:= \{\omega \in \Omega : f(\omega) \in B\} = f^{-1}(B), \\
\{f \le x\} &:= \{\omega \in \Omega : f(\omega) \le x\}, \\
\{f \le g\} &:= \{\omega \in \Omega : f(\omega) \le g(\omega)\}, \\
\{f + g = x\} &:= \{\omega \in \Omega : f(\omega) + g(\omega) = x\},
\end{aligned}
$$

and so on.

In all following statements we assume that $(\Omega, \mathscr{A})$ is a measurable space.

Proposition 4.8 *Let $f : \Omega \to \mathbb{R}$ (or $f : \Omega \to \overline{\mathbb{R}}$). The following properties are equivalent:*

(1) *f is $(\mathscr{A}, \mathscr{B})$-measurable (or f is $(\mathscr{A}, \overline{\mathscr{B}})$-measurable),*

(2) $\forall\, x \in \mathbb{R} : \{f \le x\} \in \mathscr{A}$,

(3) $\forall\, x \in \mathbb{R} : \{f < x\} \in \mathscr{A}$,

(4) $\forall\, x \in \mathbb{R} : \{f \ge x\} \in \mathscr{A}$,

(5) $\forall\, x \in \mathbb{R} : \{f > x\} \in \mathscr{A}$.

△

Proof. We only show the equivalence (1) $\iff$ (2), all other equivalences are homework.

Let first $f : \Omega \to \mathbb{R}$, and $\mathbb{R}$ is considered with the σ-algebra $\mathscr{B}$. To show the implication (1) $\implies$ (2), recall that $\{f \le x\} = f^{-1}((-\infty, x])$. Since $(-\infty, x] \in \mathscr{B}$ and f is measurable, we have $f^{-1}((-\infty, x]) \in \mathscr{A}$. The implication (2) $\implies$ (1) follows directly from Theorem 4.7, since $\mathscr{B} = \sigma(\{(-\infty, x] : x \in \mathbb{R}\})$.

To proof the equivalence (1) $\iff$ (2) for $f : \Omega \to \overline{\mathbb{R}}$, we repeat the above proof almost verbatim. To show the implication (1) $\implies$ (2), recall that $\{f \le x\} = f^{-1}([-\infty, x])$. Since $[-\infty, x] \in \overline{\mathscr{B}}$ and f is measurable, we have $f^{-1}([-\infty, x]) \in \mathscr{A}$. The implication (2) $\implies$ (1) follows directly from Theorem 4.7, since $\overline{\mathscr{B}} = \sigma(\{[-\infty, x] : x \in \mathbb{R}\})$.

□

Notice that if $f : \Omega \to \mathbb{R}$, then $f^{-1}((-\infty, x]) = f^{-1}([-\infty, x])$, $x \in \mathbb{R}$. It follows that in this case f is $(\mathscr{A}, \mathscr{B})$-measurable if and only if f is $(\mathscr{A}, \overline{\mathscr{B}})$-measurable.

Proposition 4.9 *Let $f, g : \Omega \to \overline{\mathbb{R}}$ be measurable functions. Then the sets $\{f < g\}, \{f \leq g\}, \{f = g\}, \{f \neq g\} \in \mathscr{A}$.*

△

Proof. Observe that

$$\begin{aligned}
\{f < g\} &= \bigcup_{r \in \mathbb{Q}} (\underbrace{\{f < r\}}_{\in \mathscr{A}} \cap \underbrace{\{r < g\}}_{\in \mathscr{A}}) \in \mathscr{A}, \\
\{f \leq g\} &= \overline{\{g < f\}} \in \mathscr{A}, \\
\{f = g\} &= \{f \leq g\} \cap \{g \leq f\} \in \mathscr{A}, \\
\{f \neq g\} &= \overline{\{f = g\}} \in \mathscr{A}.
\end{aligned}$$

□

In the next statement we will consider sequences $\{f_n\}_{n \in \mathbb{N}}$ of measurable functions. All operations will be always understood pointwisely, for example $f_n \to f$, $n \to \infty$ is the pointwise convergence, i.e.

$$\forall\, \omega \in \Omega : f_n(\omega) \to f(\omega), \quad n \to \infty.$$

We will also consider the upper and the lower limits

$$\begin{aligned}
\limsup_{n \to \infty} f_n &:= \inf_{n \in \mathbb{N}} \sup_{m \geq n} f_m, \\
\liminf_{n \to \infty} f_n &:= \sup_{n \in \mathbb{N}} \inf_{m \geq n} f_m.
\end{aligned}$$

Recall that a sequence $\{f_n\}_{n \in \mathbb{N}}$ is convergent if and only if $\limsup_{n \to \infty} f_n = \liminf_{n \to \infty} f_n$, the common value in the case is $\lim_{n \to \infty} f_n$.

Proposition 4.10 *Let $\{f_n\}_{n \in \mathbb{N}}$ be a sequence of measurable functions. Then*

(1) $\sup_{n \in \mathbb{N}} f_n$, $\inf_{n \in \mathbb{N}} f_n$, $\limsup_{n \to \infty} f_n$, $\liminf_{n \to \infty} f_n$ *are measurable. In particular,* $\max_{1 \leq i \leq n} f_i$, $\min_{1 \leq i \leq n} f_i$ *are measurable.*

(2) *If $f_n \to f$, $n \to \infty$, then f is measurable.*

△

Proof. (1) For $x \in \mathbb{R}$ we have $\left\{\sup_{n \in \mathbb{N}} f_n \leq x\right\} = \bigcap_{n=1}^{\infty} \underbrace{\{f_n \leq x\}}_{\in \mathscr{A}} \in \mathscr{A}$. This implies that $\sup_{n \in \mathbb{N}} f_n$ is measurable. The proof for $\inf_{n \in \mathbb{N}} f_n$ is similar. Applying these rules twice, we directly see that also $\limsup_{n \to \infty} f_n = \inf_{n \in \mathbb{N}} \sup_{m \geq n} f_n$ and $\liminf_{n \to \infty} f_n = \sup_{n \in \mathbb{N}} \inf_{m \geq n} f_m$ are measurable.

(2) If $\{f_n\}_{n\in\mathbb{N}}$ converges to f, then

$$f = \lim_{n\to\infty} f_n = \limsup_{n\to\infty} f_n = \liminf_{n\to\infty} f_n,$$

and it follows by (1) that f is measurable.

□

We will now show that arithmetic operations preserve the measurability. For simplicity, we restrict ourselves to real-valued functions.

Theorem 4.11 *Let $f, g : \Omega \to \mathbb{R}$ be measurable. Then $f+g$, $f-g$, fg and $\frac{f}{g}$ if $g \neq 0$ are measurable.*

△

Proof, Part 1. We first consider $f+g$. For any $x \in \mathbb{R}$ we have $\{f+g \leq x\} = \{f \leq x-g\}$. Since for any $y \in \mathbb{R}$ we have $\{x-g \leq y\} = \{x-y \leq g\} \in \mathscr{A}$, the function $x-g$ is measurable. It follows that $\{f+g \leq x\} = \{f \leq x-g\} \in \mathscr{A}$, i.e. $f+g$ is measurable.

The proof for $f-g$ is similar.

□

To prove the statement for fg and $\frac{f}{g}$, we need some theory.

Definition 4.12 (1) For $x \in \overline{\mathbb{R}}$ we call

$x^+ := \max(x, 0)$ the *positive part* of x,
$x^- := \max(-x, 0)$ the *negative part* of x.

(2) For $f : \Omega \to \overline{\mathbb{R}}$ we define the functions $f^\pm : \Omega \to \overline{\mathbb{R}}_+$,

$f^+ := \max(f, 0)$ the *positive part* of f,
$f^- := \max(-f, 0)$ the *negative part* of f.

△

We have $x^\pm \geq 0$, $x = x^+ - x^-$, $|x| = x^+ + x^-$ as well as $f^\pm \geq 0$, $f = f^+ - f^-$, $|f| = f^+ + f^-$.

Lemma 4.13 (1) *f is measurable $\iff$ $f^\pm$ are measurable.*

(2) *f is measurable $\implies$ $|f|$ is measurable.*

△

Proof. If f is measurable, then $-f = 0 - f$ is measurable as the difference of two measurable functions. Furthermore, $f^{\pm} = \max(\pm f, 0)$ are measurable by Proposition 4.10. Finally, $|f| = f^{+} + f^{-}$ is measurable as the sum of two measurable functions.

Now assume that $f^{\pm}$ are measurable, then $f = f^{+} - f^{-}$ is measurable as the difference of two measurable functions.

$\square$

Remark If $|f|$ is measurable, it does not follow that f is measurable. Take, for example, $A \subseteq \Omega$, $A \notin \mathscr{A}$, and consider the function

$$f(\omega) := \begin{cases} 1, & \omega \in \mathscr{A}, \\ -1, & \omega \notin \mathscr{A}. \end{cases}$$

It is easy to see that f is not measurable, but $|f| \equiv 1$ is measurable.

$\triangle$

Proof of Theorem 4.11, Part 2. We have $(fg)^{+} = f^{+}g^{+} + f^{-}g^{-}$. Indeed,

$$\begin{aligned}(fg)^{+} &= \max(fg, 0) \\ &= \underbrace{fg \cdot \chi_{\{f\geq 0\}}\chi_{\{g\geq 0\}}}_{=f^{+}g^{+}} + 0 \cdot \chi_{\{f\geq 0\}}\chi_{\{g<0\}} \\ &\quad + 0 \cdot \chi_{\{f<0\}}\chi_{\{g\geq 0\}} + \underbrace{fg\chi_{\{f<0\}}\chi_{\{g<0\}}}_{=f^{-}g^{-}} \\ &= f^{+}g^{+} + f^{-}g^{-}.\end{aligned}$$

Analogously, $(fg)^{-} = f^{+}g^{-} + f^{-}g^{+}$.

It follows that

$$fg = (fg)^{+} - (fg)^{-} = f^{+}g^{+} + f^{-}g^{-} - f^{+}g^{-} - f^{-}g^{+},$$

and the functions $f^{\pm}, g^{\pm}$ are measurable and non-negative. Thus, it is enough to prove that a product of two measurable, non-negative functions is measurable.

Now assume that f, g are measurable and non-negative. Take $x \in \mathbb{R}$. We have

$$\{fg < x\} = \emptyset \in \mathscr{A}, \quad \text{if } x \leq 0,$$

and

$$\{fg < x\} = \bigcup_{\substack{0 \leq r,s \in \mathbb{Q} \\ rs < x}} (\{f \leq r\} \cap \{g \leq s\}) \in \mathscr{A}, \quad \text{if } x > 0.$$

This proves that fg is measurable.

The proof for $\frac{f}{g}$, $g \neq 0$, is similar.

□

The previous statements can be extended to functions $f : \Omega \to \overline{\mathbb{R}}$, however, when considering arithmetic operations, one should add assumptions that guarantee that the operations are defined (for example, to avoid $\infty - \infty$ which is not defined). In the proofs, the cases when $f = \pm\infty$ or $g = \pm\infty$ should be considered separately. We will not give details.

Approximation of measurable functions by simple functions

Definition 4.14 A function $f : \Omega \to \mathbb{R}$ is called a *simple function* if f is measurable and takes only finitely many values.

△

Lemma 4.15 *A function f is a simple function if and only if there are sets $A_1, \dots, A_n \in \mathscr{A}$ and numbers $a_1, \dots, a_n \in \mathbb{R}$ such that*

$$f = \sum_{i=1}^{n} a_i \chi_{A_i}.$$

△

Proof. A function defined by the above formula is clearly measurable and simple.

To show the opposite direction, assume that f is a simple function with values $a_1, \dots, a_n \in \mathbb{R}$. Then $A_i := f^{-1}(\{a_i\}) \in \mathscr{A}$, $i = 1, \dots, n$, and $f = \sum_{i=1}^{n} a_i \chi_{A_i}$.

□

Theorem 4.16 *Let $f : \Omega \to \overline{\mathbb{R}}_+$ be a measurable function. Then there exists a sequence $\{f_n\}_{n \in \mathbb{N}}$ of simple functions $f_n : \Omega \to \mathbb{R}_+$ such that*

(1) $0 \leq f_n(\omega) \leq f_{n+1}(\omega) \leq f(\omega)$ *for all* $\omega \in \Omega$, $n \in \mathbb{N}$,

(2) $f_n(\omega) \to f(\omega)$, $n \to \infty$, *for all* $\omega \in \Omega$.

△

Proof. We denote by

$$\lfloor t \rfloor := \max\{z \in \mathbb{Z} : z \leq t\}$$

the integer part of $t \in \mathbb{R}$.

Define the auxiliary functions $g_n : \mathbb{R} \to \mathbb{R}$, $g_n(x) := \frac{\lfloor 2^n x \rfloor}{2^n}$, $n \in \mathbb{N}$. Each function g_n is monotonically non-decreasing and therefore measurable.

Moreover, $g_n(x) \leq g_{n+1}(x)$ for all $x \in \mathbb{R}$. Indeed, this inequality is equivalent to $\frac{\lfloor 2^n x \rfloor}{2^n} \leq \frac{\lfloor 2^{n+1} x \rfloor}{2^{n+1}}$, and the latter inequality is equivalent to $2\lfloor 2^n x \rfloor \leq \lfloor 2^{n+1} x \rfloor$. But $2\lfloor 2^n x \rfloor \leq 2 \cdot 2^n x = 2^{n+1} x$, and since $2\lfloor 2^n x \rfloor \in \mathbb{Z}$, it follows from the definition of the integer part of the number $2^{n+1} x$ that $2\lfloor 2^n x \rfloor \leq \lfloor 2^{n+1} x \rfloor$.

Next we prove that $0 \leq x - g_n(x) \leq \frac{1}{2^n}$. Indeed,

$$g_n(x) = \frac{\lfloor 2^n x \rfloor}{2^n} \leq \frac{2^n x}{2^n} = x,$$

so that $x - g_n(x) \geq 0$. On the other hand, $2^n x \leq \lfloor 2^n x \rfloor + 1$, and thus $x \leq \frac{\lfloor 2^n x \rfloor}{2^n} + \frac{1}{2^n}$, i.e. $x - g_n(x) \leq \frac{1}{2^n}$.

The inequality $0 \leq x - g_n(x) \leq \frac{1}{2^n}$ implies that $g_n(x) \to x$, $n \to \infty$. Moreover, $0 \leq g_n(x) \leq g_{n+1}(x) \leq x$.

Now define $f_n(\omega) := g_n(\min(f(\omega), n))$, $n \in \mathbb{N}$. The function f_n is measurable as the composition of two measurable functions g_n and $\min(f, n)$. Since $0 \leq \min(f(\omega), n) \leq n$, the function f_n takes only finitely many values (possible values are $0, \frac{1}{2^n}, \frac{2}{2^n}, \ldots, n$). In other words, f_n is a simple function satisfying $f_n \geq 0$.

Finally, it follows from the above properties that $0 \leq f_n(\omega) \leq f_{n+1}(\omega) \leq f(\omega)$, $n \in \mathbb{N}$, and $f_n(\omega) \to f(\omega)$, $n \to \infty$.

□

Chapter 5
Lebesgue integral

Let $(\Omega, \mathscr{A}, \mu)$ be a measure space, and let $f : \Omega \to \overline{\mathbb{R}}$ be a measurable function. Our aim is to define

$$\int_\Omega f d\mu.$$

We will use the following notation: For $f, g : \Omega \to \overline{\mathbb{R}}$, $B \in \overline{\mathscr{B}}$, $x \in \overline{\mathbb{R}}$ we write

$$\begin{aligned}
&\mu(f \in B) = \mu(\{f \in B\}) = \mu(\{\omega \in \Omega : f(\omega) \in B\}),\\
&\mu(f \leq g) = \mu(\{f \leq g\}) = \mu(\{\omega \in \Omega : f(\omega) \leq g(\omega)\}),\\
&\mu(f \leq x) = \mu(\{f \leq x\}) = \mu(\{\omega \in \Omega : f(\omega) \leq x\})
\end{aligned}$$

and so on.

Lebesgue integral of non-negative simple functions

Definition 5.1 Let $f : \Omega \to \mathbb{R}_+$ be a simple function.

(1) The *Lebesgue integral* of f is defined by the formula

$$\int f d\mu := \sum_{x \in f(\Omega)} x\mu(f = x) \in \overline{\mathbb{R}}_+.$$

(2) Let $A \in \mathscr{A}$. The Lebesgue integral of f over A is defined by

$$\int_A f d\mu := \int f \chi_A d\mu,$$

where χ_A is the characteristic function of the set A.

△

Remark A. The sum in the definition of the Lebesgue integral is finite since f takes finitely many values.

B. For a simple function f, also $f\chi_A$ is a simple function.

△

The following terminology is used for the Lebesgue integral in the probability theory: $(\Omega, \mathscr{A}, P)$ is a probability space, a measurable function $X : \Omega \to \mathbb{R}$ is a random variable, and the integral $E(X) = \int XdP$ is the *expectation* of X.

Proposition 5.2 (1) *Let $f : \Omega \to \mathbb{R}_+$ be a simple function, $a \in \mathbb{R}_+$. Then the function $af : \Omega \to \mathbb{R}_+$ is a simple function and*

$$\int af d\mu = a \int f d\mu.$$

(2) *Let $f, g : \Omega \to \mathbb{R}_+$ be simple functions. Then the function $f + g : \Omega \to \mathbb{R}_+$ is a simple function and*

$$\int (f + g) d\mu = \int f d\mu + \int g d\mu.$$

(3) *Let $f, g : \Omega \to \mathbb{R}_+$ be simple functions and $f \leq g$. Then*

$$\int f d\mu \leq \int g d\mu.$$

(4) *If $f = \sum_{i=1}^{n} a_i \chi_{A_i}$, where $a_1, \dots, a_n \in \mathbb{R}_+$, $A_1, \dots, A_n \in \mathscr{A}$, then $f : \Omega \to \mathbb{R}_+$ is a simple function and*

$$\int f d\mu = \sum_{i=1}^{n} a_i \mu(A_i).$$

In particular, for $A \in \mathscr{A}$ we have

$$\int \chi_A d_\mu = \int_A 1 d\mu = \mu(A)$$

and for $0 = \chi_\varnothing$ we have $\int 0 d\mu = 0$.

△

Proof. (1) The function $af : \Omega \to \mathbb{R}_+$ is obviously a simple function. For $x \in \mathbb{R}$ we have $x \in f(\Omega) \iff ax \in (af)(\Omega)$ and $\{f = x\} = \{af = ax\}$.

Consequently,

$$\begin{aligned}\int af d\mu &= \sum_{y\in(af)(\Omega)} y\mu(af = y) = \sum_{x\in f(\Omega)} ax\mu(af = ax)\\ &= a\sum_{x\in f(\Omega)} x\mu(f = x) = a\int f d\mu.\end{aligned}$$

(2) Obviously, $f + g : \Omega \to \mathbb{R}_+$ is a simple function. By definition,

$$\int (f+g)d\mu = \sum_{x\in(f+g)(\Omega)} x\mu(f+g = x).$$

We have

$$\{f+g = x\} = \sum_{\substack{u\in f(\Omega)\\ v\in g(\Omega)\\ u+v=x}} (\{f = u\}\cap\{g = v\}).$$

Thus,

$$\begin{aligned}\int (f+g)d\mu &= \sum_{x\in(f+g)(\Omega)} \sum_{\substack{u\in f(\Omega)\\ v\in g(\Omega)\\ u+v=x}} (u+v)\mu(\{f = u\}\cap\{g = v\})\\ &= \sum_{\substack{u\in f(\Omega)\\ v\in g(\Omega)}} (u+v)\mu(\{f = u\}\cap\{g = v\}) \sum_{\substack{x\in(f+g)(\Omega)\\ x=u+v}} 1.\end{aligned}$$

If $\{f = u\}\cap\{g = v\} \neq \emptyset$, then $\sum_{\substack{x\in(f+g)(\Omega)\\ x=u+v}} 1 = 1$, because for each pair $(u, v) \in f(\Omega)\times g(\Omega)$ with this property there is exactly one x with $x = u+v$. If $\{f = u\}\cap\{g = v\} = \emptyset$, then $\mu(\{f = u\}\cap\{g = v\}) = 0$ and the entire term is zero. It follows that

$$\begin{aligned}\int (f+g)d\mu &= \sum_{\substack{u\in f(\Omega)\\ v\in g(\Omega)}} (u+v)\mu(\{f = u\}\cap\{g = v\})\\ &= \sum_{u\in f(\Omega)} u \sum_{v\in g(\Omega)} \mu(\{f = u\}\cap\{g = v\})\\ &+ \sum_{v\in g(\Omega)} v \sum_{u\in f(\Omega)} \mu(\{f = u\}\cap\{g = v\}).\end{aligned}$$

Since $\sum_{v\in g(\Omega)}\{g = v\} = \Omega$ and $\sum_{u\in f(\Omega)}\{f = u\} = \Omega$, we have

$$\begin{aligned}\int (f+g)d\mu &= \sum_{u\in f(\Omega)} u\mu(f=u) + \sum_{v\in g(\Omega)} v\mu(g=v)\\ &= \int f d\mu + \int g d\mu.\end{aligned}$$

(3) We have

$$\int f d\mu = \sum_{x\in f(\Omega)} x\mu(f=x) = \sum_{\substack{x\in f(\Omega)\\ y\in g(\Omega)}} x\mu(\{f=x\}\cap\{g=y\}).$$

If $\omega \in \{f = x\} \cap \{g = y\}$, then $x = f(\omega) \le g(\omega) = y$. On the other hand, for $x > y$ we have $\{f = x\} \cap \{g = y\} = \emptyset$, and thus $\mu(\{f = u\} \cap \{g = v\}) = 0$. This implies

$$\begin{aligned}\int f d\mu &\le \sum_{\substack{x\in f(\Omega)\\ y\in g(\Omega)}} y\mu(\{f=x\}\cap\{g=y\})\\ &= \sum_{y\in g(\Omega)} y\mu(\{g=y\}) = \int g d\mu.\end{aligned}$$

(4) Obviously, $f : \Omega \to \mathbb{R}_+$, and f is a simple function by Lemma 4.15. We have

$$\int \chi_A d\mu = 1\cdot\mu(\chi_A = 1) + 0\cdot\mu(\chi_A = 0) = \mu(A).$$

Using the properties (1) and (2), we obtain

$$\int f d\mu = \sum_{i=1}^{n} a_i\mu(A_i).$$

□

Lebesgue integral of non-negative measurable functions

Definition 5.3 Let $f : \Omega \to \overline{\mathbb{R}}_+$ be a measurable function.

(1) The *Lebesgue integral* of f is defined by

$$\int f d\mu := \sup\left\{\int g d\mu : g : \Omega \to \mathbb{R}_+ \text{ is a simple function, } g \le f\right\} \in \overline{\mathbb{R}}_+.$$

(2) Let $A \in \mathscr{A}$. The Lebesgue integral of f over A is defined by

$$\int_A f d\mu := \int f \chi_A d\mu.$$

$\triangle$

Remark A. Since $0 \le f$, and the function $0 : \Omega \to \mathbb{R}_+$ is a simple function, the set in the definition of the integral is not empty.

B. It follows immediately from the definition that for two measurable functions f, g we have $0 \le f \le g \implies 0 \le \int f d\mu \le \int g d\mu$.

$\triangle$

Remark Definition 5.1 and Definition 5.3 are consistent in the sense that for a simple non-negative function the integrals coincide. This follows from the monotonicity property (Proposition 5.2 (3)).

$\triangle$

Theorem 5.4 (*Monotone Convergence Theorem*) *Let $\{f_n\}_{n\in\mathbb{N}}$ be a sequence of measurable functions $f_n : \Omega \to \overline{\mathbb{R}}_+$ such that $f_n(\omega) \le f_{n+1}(\omega)$ for all $\omega \in \Omega$, $n \in \mathbb{N}$. Then the function $f := \sup_{n\in\mathbb{N}} f_n = \lim_{n\to\infty} f_n : \Omega \to \overline{\mathbb{R}}_+$ is measurable and*

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

$\triangle$

Proof. The function f is measurable by Proposition 4.10. Since $f_n \le f_{n+1} \le f$, we have

$$\int f_n d\mu \le \int f_{n+1} d\mu \le \int f d\mu.$$

Thus, the sequence $\left\{\int f_n d\mu\right\}_{n\in\mathbb{N}}$ is monotone non-decreasing in $\overline{\mathbb{R}}_+$ and therefore it converges in $\overline{\mathbb{R}}_+$ and

$$\lim_{n\to\infty} \int f_n d\mu \le \int f d\mu.$$

We will now show the reverse inequality. It is enough to show the following: For any simple function $g : \Omega \to \mathbb{R}_+$ such that $g \le f$ we have

$$\int g d\mu \le \lim_{n\to\infty} \int f_n d\mu.$$

Suppose that $g : \Omega \to \mathbb{R}_+$ is a simple function such that $g \leq f$. Let $a \in (0,1)$. For $n \in \mathbb{N}$ put

$$B_n := \{f_n \geq ag\} \in \mathscr{A}.$$

It holds $B_n \subseteq B_{n+1}$, since $f_{n+1}(\omega) \geq f_n(\omega) \geq ag(\omega)$ for $\omega \in B_n$.

We show that $\Omega = \bigcup_{n\in\mathbb{N}} B_n$. We consider two cases.

1st case: $f(\omega) > 0$. Then $\lim_{n\to\infty} f_n(\omega) = f(\omega) > ag(\omega)$, because $g(\omega) \leq f(\omega)$ $g(\omega) < \infty$, $a < 1$. It follows that $\omega \in B_n$ for n large enough.

2nd case: $f(\omega) = 0$. In this case also $f_n(\omega) = 0$ for all $n \in \mathbb{N}$ and $g(\omega) = 0$. Then $\omega \in B_n$ for all $n \in \mathbb{N}$.

Altogether this shows that $\Omega = \bigcup_{n\in\mathbb{N}} B_n$.

According to Lemma 4.15 applied to the simple function g, there exist $a_1, \dots, a_m \in \mathbb{R}_+$ and $A_1, \dots, A_m \in \mathscr{A}$ such that $g = \sum_{i=1}^m a_i\chi_{A_i}$. Then

$$\begin{aligned}
a\int g d\mu &= a\sum_{i=1}^m a_i\mu(A_i)\\
&= a\sum_{i=1}^m a_i \lim_{n\to\infty}\mu(A_i\cap B_n) \quad \text{(continuity from below)}\\
&= \lim_{n\to\infty}\sum_{i=1}^m aa_i\mu(A_i\cap B_n)\\
&= \lim_{n\to\infty}\int\sum_{i=1}^m aa_i \underbrace{\chi_{A_i\cap B_n}}_{\chi_{A_i}\cdot\chi_{B_n}} d\mu\\
&= \lim_{n\to\infty}\int a\underbrace{\left(\sum_{i=1}^m a_i\chi_{A_i}\right)}_{=g}\chi_{B_n} d\mu\\
&= \lim_{n\to\infty}\int \underbrace{ag\chi_{\{f_n\geq ag\}}}_{\leq f_n} d\mu\\
&\leq \lim_{n\to\infty}\int f_n d\mu.
\end{aligned}$$

Since this is true for any $a \in (0,1)$, we have $\int g d\mu \leq \lim_{n\to\infty}\int f_n d\mu$ as desired. □

Remark For a sequence of sets $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathscr{A}$, $A_n \subseteq A_{n+1}$, $A := \bigcup_{n=1}^\infty A_n$ we have $0 \leq \chi_{A_n} \leq \chi_{A_{n+1}} \leq \chi_A$ and $\chi_{A_n} \to \chi_A$ $n \to \infty$. Thus,

$$\lim_{n\to\infty}\int\chi_{A_n} d\mu = \int\chi_A d\mu$$

is the same as

$$\lim_{n\to\infty} \mu(A_n) = \mu(A).$$

Theorem 5.4 carries over the continuity of μ from below from characteristic functions of sets to arbitrary functions.

△

Proposition 5.5 (1) *Let $f : \Omega \to \overline{\mathbb{R}}_+$ be a measurable function and $a \in \overline{\mathbb{R}}_+$. Then*

$$\int af d\mu = a \int f d\mu.$$

(2) *Let $f, g : \Omega \to \overline{\mathbb{R}}_+$ be measurable functions, then*

$$\int (f+g) d\mu = \int f d\mu + \int g d\mu.$$

△

Proof. We will only prove (1), the prove of (2) is similar.

Let $f : \Omega \to \overline{\mathbb{R}}_+$. By Theorem 4.16 there exists a sequence $\{f_n\}_{n\in\mathbb{N}}$ of non-negative simple functions that monotonically converges to f. Then $\{af_n\}_{n\in\mathbb{N}}$ is a sequence of non-negative simple functions that monotonically converges to af. By Theorem 5.4 and Proposition 5.2 we have

$$\int af d\mu = \lim_{n\to\infty} \int af_n d\mu = a \lim_{n\to\infty} \int f_n d\mu = a \int f d\mu.$$

□

Lebesgue integral of measurable functions

Now we will define the Lebesgue integral in the general case of measurable functions of arbitrary sign. We will see, however, that not all measurable functions are integrable.

Definition 5.6 Let $f : \Omega \to \overline{\mathbb{R}}$ be a measurable function.

(1) f is called *integrable*, if

$$\int f^+ d\mu < \infty \quad \text{and} \quad \int f^- d\mu < \infty.$$

If f is integrable, then the *Lebesgue integral* of f is defined as

$$\int f d\mu := \int f^+ d\mu - \int f^- d\mu \in \mathbb{R}.$$

(2) Let $A \in \mathscr{A}$. A function f is integrable over A, when $f\chi_A$ is integrable. In this case we put

$$\int_A f d\mu := \int f\chi_A d\mu.$$

△

Remark If $f \geq 0$, then $f \equiv f^+$ and $f^- \equiv 0$, and in particular $\int f^- d\mu = 0$. It follows that Definition 5.3 and Definition 5.6 are consistent.

△

In the next statement we present very useful criteria for integrability of a function.

Theorem 5.7 *Let $f : \Omega \to \overline{\mathbb{R}}$ be a measurable function. The following properties are equivalent:*

(1) *f is integrable,*

(2) *$|f|$ is integrable,*

(3) *there exist integrable functions $g, h : \Omega \to \overline{\mathbb{R}}$ such that $g \leq f \leq h$.*

△

Proof. We will show the implications (1) $\implies$ (2) $\implies$ (3) $\implies$ (1).

A. (1) $\implies$ (2)

Since $|f| \geq 0$, we have $|f|^- \equiv 0$ and $\int |f|^- d\mu = 0 < \infty$. On the other hand, $|f|^+ = |f| = f^+ + f^-$, so that

$$\int |f|^+ d\mu = \underbrace{\int f^+ d\mu}_{<\infty} + \underbrace{\int f^- d\mu}_{<\infty} < \infty,$$

which implies that $|f|$ is integrable.

B. (2) $\implies$ (3)

We have $-|f| \leq f \leq |f|$. The function $|f|$ is integrable by assumption. It suffices to show that $-|f|$ is integrable.

We have $(-|f|)^+ \equiv 0$, $(-|f|)^- \equiv |f|$. It follows that $\int (-|f|)^+ d\mu = 0 < \infty$, $\int (-|f|)^- d\mu = \int |f| d\mu < \infty$. Thus, $-|f|$ is integrable.

c. (3) $\implies$ (1)

It follows from $g \leq f \leq h$ that $f^+ \leq h^+$ and $f^- \leq g^-$. Consequently,

$$\int f^+ d\mu \leq \int h^+ d\mu < \infty, \quad \int f^- d\mu \leq \int g^- d\mu < \infty.$$

This means that f is integrable.

□

Theorem 5.8 *Let $f, g : \Omega \to \overline{\mathbb{R}}$ be integrable functions.*

(1) *For any $a \in \mathbb{R}$, the function af is integrable and*

$$\int (af) d\mu = a \int f d\mu.$$

(2) *If $f + g$ is defined on the whole of Ω, then $f + g$ is integrable and*

$$\int (f + g) d\mu = \int f d\mu + \int g d\mu.$$

(3) $f \leq g \implies \int f d\mu \leq \int g d\mu.$

(4) $\left|\int f d\mu\right| \leq \int |f| d\mu.$

△

Proof. (1) For $a = 0$ the statement is trivial.

We consider in detail the case when $a < 0$. Then

$$(af)^+ = \max(0, af) = \max(0, (-a)(-f)) = (-a)\max(0, -f) = (-a) \cdot f^-.$$

Similarly, $(af)^- = (-a)f^+$. Consequently, with the use of Proposition 5.5 we obtain

$$\int (af)^{\pm} d\mu = \int (-a) f^{\mp} d\mu = (-a) \int f^{\mp} d\mu < \infty,$$

which implies that af is integrable. Moreover,

$$\begin{aligned}
\int (af)d\mu &= \int (af)^+ d\mu - \int (af)^- d\mu \\
&= \int (-a)f^- d\mu - \int (-a)f^+ d\mu \\
&= (-a)\int f^- d\mu - (-a)\int f^+ d\mu \\
&= (-a)\left(\int f^- d\mu - \int f^+ d\mu\right) \\
&= a\left(\int f^+ d\mu - \int f^- d\mu\right) \\
&= a\int f d\mu.
\end{aligned}$$

The proof for $a > 0$ is similar.

(2) We have $0 \le |f+g| \le |f|+|g|$, where $|f|$ and $|g|$ are integrable by Theorem 5.7, which implies by Proposition 5.5 that $|f| + |g|$ is integrable. Now it follows by Theorem 5.7 again that $|f + g|$ is integrable, and finally that $f + g$ is integrable.

To consider the integral, we first observe that

$$f + g = (f + g)^+ - (f + g)^- = f^+ - f^- + g^+ - g^-,$$

which is equivalent to

$$(f + g)^+ + f^- + g^- = (f + g)^- + f^+ + g^+.$$

Integrating this equation and applying Proposition 5.5, we obtain

$$\int (f + g)^+ d\mu + \int f^- d\mu + \int g^- d\mu = \int (f + g)^- d\mu + \int f^+ d\mu + \int g^+ d\mu.$$

This implies

$$\begin{aligned}
\int (f + g)d\mu &= \int (f + g)^+ d\mu - \int (f + g)^- d\mu \\
&= \int f^+ d\mu - \int f^- d\mu + \int g^+ d\mu - \int g^- d\mu \\
&= \int f d\mu + \int g d\mu.
\end{aligned}$$

(3) The inequality $f \le g$ implies $f^+ \le g^+, g^- \le f^-$, and consequently $\int f^+ d\mu \le \int g^+ d\mu$, $\int g^- d\mu \le \int f^- d\mu$. It follows

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

(4) We have

$$\begin{aligned}\left|\int f d\mu\right| &= \left|\int f^+ d\mu - \int f^- d\mu\right| \\ &\le \int f^+ d\mu + \int f^- d\mu = \int (f^+ + f^-) d\mu = \int |f| d\mu.\end{aligned}$$

□

Corollary 5.9 (1) *If f is integrable over a set $A \in \mathscr{A}$, then f is also integrable over all $A' \in \mathscr{A}$, $A' \subseteq A$.*

(2) *If $A, B \in \mathscr{A}$, $A \cap B = \emptyset$, f is integrable over A and over B, then f is integrable over $A + B$. Moreover,*

$$\int_{A+B} f d\mu = \int_A f d\mu + \int_B f d\mu.$$

△

Proof. (1) follows from $0 \le |f\chi_{A'}| \le |f\chi_A|$.

(2) follows from $f\chi_{A+B} = f\chi_A + f\chi_B$.

□

Theorem 5.10 *Let $f : \Omega \to \overline{\mathbb{R}}_+$ be a measurable function. Then the set function $\nu_f : \mathscr{A} \to \overline{\mathbb{R}}_+$,*

$$\nu_f(A) := \int_A f d\mu, \quad A \in \mathscr{A},$$

is a measure on $(\Omega, \mathscr{A})$.

△

Proof. Clearly, $\nu_f \ge 0$ and

$$\nu_f(\emptyset) = \int \chi_\emptyset f d\mu = \int 0 d\mu = 0.$$

It remains to show that ν_f is σ-additive. Let $\{A_i\}_{i\in\mathbb{N}} \subseteq \mathscr{A}$ be p.d. sets and $A := \sum_{i=1}^{\infty} A_i$. We have

$$\begin{aligned}
\nu_f(A) &= \int_A f d\mu = \int \chi_A f d\mu = \int \sum_{i=1}^{\infty} \chi_{A_i} f d\mu \\
&= \int \lim_{n\to\infty} \sum_{i=1}^{n} \chi_{A_i} f d\mu \quad \text{(monotone convergence)} \\
&= \lim_{n\to\infty} \int \sum_{i=1}^{n} \chi_{A_i} f d\mu \\
&= \lim_{n\to\infty} \sum_{i=1}^{n} \int \chi_{A_i} f d\mu \\
&= \sum_{i=1}^{\infty} \int_{A_i} f d\mu = \sum_{i=1}^{\infty} \nu_f(A_i)
\end{aligned}$$

as desired.

□

Theorem 5.11 (*Chebyshev's Inequality*) *Let $f : \Omega \to \overline{\mathbb{R}}_+$ be a measurable function and $\varepsilon > 0$. Then*

$$\mu(f \geq \varepsilon) \leq \frac{1}{\varepsilon} \int_{\{f\geq\varepsilon\}} f d\mu \leq \frac{1}{\varepsilon} \int_\Omega f d\mu.$$

△

Proof. We have

$$\chi_{\{f\geq\varepsilon\}} \leq \frac{f}{\varepsilon} \chi_{\{f\geq\varepsilon\}} \leq \frac{f}{\varepsilon}.$$

Integrating this inequality, we obtain

$$\int_\Omega \chi_{\{f\geq\varepsilon\}} d\mu \leq \int_\Omega \frac{f}{\varepsilon} \chi_{\{f\geq\varepsilon\}} d\mu \leq \int_\Omega \frac{f}{\varepsilon} d\mu$$

which is the same as

$$\mu(f \geq \varepsilon) \leq \frac{1}{\varepsilon} \int_{\{f\geq\varepsilon\}} f d\mu \leq \frac{1}{\varepsilon} \int_\Omega f d\mu.$$

□

Corollary 5.12 *For an integrable function $f : \Omega \to \overline{\mathbb{R}}$ and $\varepsilon > 0$ we have*

$$\mu(|f| \geq \varepsilon) \leq \frac{1}{\varepsilon} \int_{\{|f|\geq\varepsilon\}} |f| d\mu \leq \frac{1}{\varepsilon} \int_\Omega |f| d\mu.$$

△

In the probability theory, this inequality means that for a random variable X we have

$$P(|X| \geq \varepsilon) \leq \frac{1}{\varepsilon} E(|X|).$$

This inequality is also known as Markov's inequality.

Definition 5.13 We say that a property $P(\omega)$ holds *μ-almost everywhere (a.e.)* in Ω, if there exists a set $N \in \mathscr{A}$, $\mu(N) = 0$ such that $P(\omega)$ holds true for all $\omega \in \Omega \setminus N$.

△

Remark It is not required that the set $\{\omega \in \Omega : \neg P(\Omega)\}$ is measurable.

△

We frequently say just almost everywhere (a.e.), when it is clear from the context which measure μ is considered.

Example A. For $f, g : \Omega \to \overline{\mathbb{R}}$, we say that $f = g$ μ-a.e. if $f(\omega) = g(\omega)$ for all $\omega \in \Omega \setminus N$ with $\mu(N) = 0$. If f, g are measurable, then $\{f \neq g\} \in \mathscr{A}$ and $\mu(f \neq g) = 0$.

B. For $f : \Omega \to \overline{\mathbb{R}}$, we say that f is μ-a.e finite, if $|f| < \infty$ μ-a.e., i.e. if $|f(\omega)| < \infty$ for all $\omega \in \Omega \setminus N$ with $\mu(N) = 0$. If f is measurable, then $\{|f| = \infty\} \in \mathscr{A}$ and $\mu(|f| = \infty) = 0$.

△

Theorem 5.14 *Let $f : \Omega \to \overline{\mathbb{R}}$ be μ–integrable. Then f is μ-a.e. finite.*

△

Proof. Note that f is measurable by assumption. Thus we have to show that $\mu(|f| = \infty) = 0$. For each $n \in \mathbb{N}$ we have by Chebyshev's inequality

$$0 \leq \mu(|f| = \infty) \leq \mu(|f| \geq n) \leq \frac{1}{n} \underbrace{\int_\Omega |f| d\mu}_{<\infty} \to 0, \quad n \to \infty,$$

which implies $\mu(|f| = \infty) = 0$.

□

Theorem 5.15 *Let $f : \Omega \to \overline{\mathbb{R}}_+$ be a measurable function, then*

$$\int_\Omega f d\mu = 0 \iff f = 0 \ \mu\text{-a.e.}$$

△

Proof. A. $\Longrightarrow$

Suppose that $\int_\Omega f d\mu = 0$. We have to show that $\mu(f > 0) = 0$. We have

$$\{f > 0\} = \bigcup_{n=1}^{\infty} \left\{ f \geq \frac{1}{n} \right\},$$

and $\left\{ f \geq \frac{1}{n} \right\} \subseteq \left\{ f \geq \frac{1}{n+1} \right\}$. It follows that

$$0 \leq \mu(f > 0) = \lim_{n\to\infty} \mu\left(f \geq \frac{1}{n}\right) \leq \lim_{n\to\infty} n \underbrace{\int_\Omega f d\mu}_{=0} = 0.$$

Thus, $\mu(f > 0) = 0$.

B. $\Longleftarrow$

Suppose that $f = 0$ μ-a.e. We have

$$0 \leq f \leq \infty \cdot \chi_{\{f>0\}} = \lim_{n\to\infty} n\chi_{\{f>0\}}$$

and therefore, using monotone convergence,

$$\begin{aligned} 0 \leq \int_\Omega f d\mu &\leq \int_\Omega \lim_{n\to\infty} n\chi_{\{f>0\}} d\mu \\ &= \lim_{n\to\infty} \int_\Omega n\chi_{\{f>0\}} d\mu = \lim_{n\to\infty} n \underbrace{\mu(f > 0)}_{=0} = 0. \end{aligned}$$

This implies that $\int_\Omega f d\mu = 0$.

□

Corollary 5.16 *If $N \in \mathscr{A}$, $\mu(N) = 0$, then any measurable function $f : \Omega \to \overline{\mathbb{R}}$ is integrable over N and*

$$\int_N f d\mu = 0.$$

△

Proof. Since $\{|f|\chi_N \neq 0\} \subseteq N$ we see that $|f|\chi_N = 0$ μ-a.e. By Theorem 5.15,

$$\int_\Omega |f|\chi_N d\mu = 0.$$

It follows that $f\chi_N$ is integrable (and thus by definition f is integrable over N) and $0 \le \left|\int_\Omega f\chi_N d\mu\right| \le \int_\Omega |f|\chi_N d\mu = 0$ which implies

$$\int_N f d\mu = \int_\Omega f\chi_n d\mu = 0.$$

□

Theorem 5.17 *Let $f, g : \Omega \to \overline{\mathbb{R}}$ be measurable functions and $f = g$ μ-a.e. If f is integrable, then also g is integrable and*

$$\int_\Omega f d\mu = \int_\Omega g d\mu.$$

△

Proof. The fact that $f = g$ μ-a.e. implies that also $f^\pm = g^\pm$ μ-a.e. We have

$$\begin{aligned}\int_\Omega f^\pm d\mu &= \int_{\{f^\pm = g^\pm\}} f^\pm d\mu + \underbrace{\int_{\{f^\pm \neq g^\pm\}} f^\pm d\mu}_{=0,\text{ since } \mu(f^\pm \neq g^\pm)=0} \\ &= \int_{\{f^\pm = g^\pm\}} g^\pm d\mu + \underbrace{\int_{\{f^\pm \neq g^\pm\}} g^\pm d\mu}_{=0,\text{ since } \mu(f^\pm \neq g^\pm)=0} = \int_\Omega g^\pm d\mu.\end{aligned}$$

It follows that $\int_\Omega g^\pm d\mu < \infty$, so that g is integrable, and

$$\int_\Omega g d\mu = \int_\Omega g^+ d\mu - \int_\Omega g^- d\mu = \int_\Omega f^+ d\mu - \int_\Omega f^- d\mu = \int_\Omega f d\mu.$$

□

Chapter 6

Convergence theorems for Lebesgue integral

In the previous chapter we already discussed a very important convergence theorem: The Monotone Convergence Theorem (Theorem 5.4). In this chapter we will prove two further very important statements.

Theorem 6.1 (*Fatou's Lemma*) *Let $\{f_n\}_{n\in\mathbb{N}}$ be a sequence of measurable functions $f_n : \Omega \to \overline{\mathbb{R}}_+$. Then*

$$\int_\Omega \liminf_{n\to\infty} f_n d\mu \leq \liminf_{n\to\infty} \int_\Omega f_n d\mu.$$

△

Proof. By Proposition 4.10, the function $\liminf_{n\to\infty} f_n : \Omega \to \overline{\mathbb{R}}_+$ is measurable and non-negative, and thus the integral $\int_\Omega \liminf_{n\to\infty} f_n d\mu$ is well defined. We have

$$\begin{aligned}
\int_\Omega \liminf_{n\to\infty} f_n d\mu &= \int_\Omega \lim_{n\to\infty} \underbrace{\inf_{m\geq n} f_m}_{\substack{\geq 0 \text{ and} \\ \text{non-decreasing in } n}} d\mu \quad \text{(monotone convergence)} \\
&= \lim_{n\to\infty} \underbrace{\int_\Omega \inf_{m\geq n} f_m d\mu}_{\leq \int_\Omega f_k d\mu \ \forall\ k\geq n} \\
&\leq \lim_{n\to\infty} \inf_{k\geq n} \int_\Omega f_k d\mu = \liminf_{n\to\infty} \int_\Omega f_n d\mu.
\end{aligned}$$

□

Theorem 6.2 (*Lebesgue's Dominated Convergence Theorem*) *Let $\{f_n\}_{n\in\mathbb{N}}$ be a sequence of measurable functions $f_n : \Omega \to \overline{\mathbb{R}}$ such that*

(1) $f_n \to f$, $n \to \infty$, μ-a.e., where $f : \Omega \to \overline{\mathbb{R}}$ *is a measurable function,*

(2) *there exists an integrable function* $g : \Omega \to \overline{\mathbb{R}}$ *("an integrable majorant") such that*

$$|f_n| \leq g \quad \mu\text{-a.e.} \quad \text{for all } n \in \mathbb{N}.$$

Then:

(1) f *is integrable,*

(2) $\int_\Omega f d\mu = \lim_{n\to\infty} \int_\Omega f_n d\mu.$

△

Proof. We consider the exceptional set

$$N := \{f_n \not\to f\} \cup \left(\bigcup_{n=1}^{\infty} \{|f_n| > g\} \right) \cup \{g = \infty\} \in \mathscr{A}.$$

The sets $\{f_n \not\to f\}$, $\{|f_n| > g\}$ for $n \in \mathbb{N}$ are null sets by the assumptions, and the set $\{g = \infty\}$ is a null set by Theorem 5.14. Thus, $\mu(N) = 0$.

Put $A := \Omega \setminus N$, $\tilde{f}_n := f_n \chi_A$, $n \in \mathbb{N}$, $\tilde{f} := f\chi_A$, $\tilde{g} := g\chi_A$. Then $\tilde{g}$ is integrable (this follows by Theorem 5.17 from the facts that $\tilde{g} = g$ μ-a.e. and g is integrable),

(1) $\tilde{f}_n \to \tilde{f}$, $n \to \infty$, for all $\omega \in \Omega$,

(2) $|\tilde{f}_n| \leq \tilde{g}$, $n \in \mathbb{N}$, for all $\omega \in \Omega$,

and all functions $\tilde{g}$, $\tilde{f}_n$, $\tilde{f}$ are finite ($\tilde{g}$ by its definition, $\tilde{f}_n$ and $\tilde{f}$ since $|\tilde{f}_n| \leq \tilde{g}$, $|\tilde{f}| \leq \tilde{g}$).

Since $|\tilde{f}_n| \leq \tilde{g}$, $|\tilde{f}| \leq \tilde{g}$ and $\tilde{g}$ is integrable, the functions $\tilde{f}_n$, $n \in \mathbb{N}$, and $\tilde{f}$ are integrable.

From $-\tilde{g} \leq \tilde{f}_n \leq \tilde{g}$ it follows that $\tilde{f}_n + \tilde{g} \geq 0$, $\tilde{g} - \tilde{f}_n \geq 0$. We have

$$\begin{aligned}
\int_\Omega \tilde{f} d\mu + \int_\Omega \tilde{g} d\mu &= \int_\Omega (\tilde{f} + \tilde{g}) d\mu \\
&= \int_\Omega \liminf_{n\to\infty} (\tilde{f}_n + \tilde{g}) d\mu \quad \text{(Fatou's Lemma)} \\
&\leq \liminf_{n\to\infty} \int_\Omega (\tilde{f}_n + \tilde{g}) d\mu \\
&= \liminf_{n\to\infty} \left(\int_\Omega \tilde{f}_n d\mu + \int_\Omega \tilde{g} d\mu \right) \\
&= \liminf_{n\to\infty} \int_\Omega \tilde{f}_n d\mu + \int_\Omega \tilde{g} d\mu,
\end{aligned}$$

and this implies

$$\int_\Omega \tilde{f} d\mu \le \liminf_{n\to\infty} \int_\Omega \tilde{f}_n d\mu.$$

Similarly,

$$\begin{aligned}
\int_\Omega \tilde{g} d\mu - \int_\Omega \tilde{f} d\mu &= \int_\Omega (\tilde{g} - \tilde{f}) d\mu \\
&= \int_\Omega \liminf_{n\to\infty} (\tilde{g} - \tilde{f}_n) d\mu \quad \text{(Fatou's Lemma)} \\
&\le \liminf_{n\to\infty} \int_\Omega (\tilde{g} - \tilde{f}_n) d\mu \\
&= \liminf_{n\to\infty} \left(\int_\Omega \tilde{g} d\mu - \int_\Omega \tilde{f}_n d\mu \right) \\
&= \int_\Omega \tilde{g} d\mu - \limsup_{n\to\infty} \int_\Omega \tilde{f}_n d\mu,
\end{aligned}$$

and this implies

$$\limsup_{n\to\infty} \int_\Omega \tilde{f}_n d\mu \le \int_\Omega \tilde{f} d\mu.$$

Altogether we have

$$\int_\Omega \tilde{f} d\mu \le \liminf_{n\to\infty} \int_\Omega \tilde{f}_n d\mu \le \limsup_{n\to\infty} \int_\Omega \tilde{f}_n d\mu \le \int_\Omega \tilde{f} d\mu.$$

It follows that

$$\liminf_{n\to\infty} \int_\Omega \tilde{f}_n d\mu = \limsup_{n\to\infty} \int_\Omega \tilde{f}_n d\mu = \int_\Omega \tilde{f} d\mu.$$

This means that $\lim_{n\to\infty} \int_\Omega \tilde{f}_n d\mu$ exists and

$$\lim_{n\to\infty} \int_\Omega \tilde{f}_n d\mu = \int_\Omega \tilde{f} d\mu.$$

Recalling that $\tilde{f} = f$, $\tilde{f}_n = f_n$ μ-a.e., we conclude that f is integrable. Moreover, $\int_\Omega f d\mu = \int_\Omega \tilde{f} d\mu$, $\int_\Omega f_n d\mu = \int_\Omega \tilde{f}_n d\mu$, so that

$$\lim_{n\to\infty} \int_\Omega f_n d\mu = \int_\Omega f d\mu.$$

□

Chapter 7

Product measures and Fubini's theorem

Let $(\Omega_1, \mathscr{A}_1, \mu_1)$ and $(\Omega_2, \mathscr{A}_2, \mu_2)$ be two σ-finite measure spaces. Our aim is to construct a measure μ on the Cartesian product $\Omega_1 \times \Omega_2$ such that

$$\mu(A_1 \times A_2) = \mu_1(A_1)\mu_2(A_2), \quad \text{for all } A_1 \in \mathscr{A}_1,\ A_2 \in \mathscr{A}_2.$$

For $A_1 \in \mathscr{A}_1$, $A_2 \in \mathscr{A}_2$, we call the set $A_1 \times A_2 \subseteq \Omega_1 \times \Omega_2$ a *measurable rectangle*. We refer to A_1 and A_2 as its sides.

Proposition 7.1

The family of all measurable rectangles

$$\mathscr{R} := \{A_1 \times A_2 : A_1 \in \mathscr{A}_1, A_2 \in \mathscr{A}_2\} \subseteq P(\Omega_1 \times \Omega_2)$$

is a semiring on $\Omega_1 \times \Omega_2$, and $\Omega_1 \times \Omega_2 \in \mathscr{R}$.

△

Proof. This can be checked directly (Homework). □

Definition 7.2 The σ-algebra $\sigma(\mathscr{R})$ on $\Omega_1 \times \Omega_2$ is called the *product σ-algebra* of $\mathscr{A}_1$ and $\mathscr{A}_2$. We use the notation $\sigma(\mathscr{R}) =: \mathscr{A}_1 \otimes \mathscr{A}_2$.

△

To proceed with the construction of the measure, we need a concept of another set system.

Definition 7.3 Let $\Omega \neq \emptyset$ and $\mathscr{A} \subseteq P(\Omega)$. $\mathscr{A}$ is called a *Dynkin system*, if

(1) $\Omega \in \mathscr{A}$,

(2) $A \in \mathscr{A} \implies \overline{A} \in \mathscr{A}$,

(3) $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathscr{A}$, p.d. $\implies \sum_{n=1}^{\infty} A_n \in \mathscr{A}$.

△

Remark Every σ-algebra is a Dynkin system. △

Lemma 7.4 *Let $\mathscr{A}$ be a Dynkin system and $A, B \in \mathscr{A}$, $B \subseteq A$. Then $A \setminus B \in \mathscr{A}$.*

△

Proof. We have

$$\overline{A \setminus B} = \overline{A \cap \overline{B}} = \overline{A} \cup B = \overline{A} + B \in \mathscr{A},$$

and it follows that $A \setminus B \in \mathscr{A}$.

□

Theorem 7.5 *Let $\mathscr{A} \subseteq P(\Omega)$. The following statements are equivalent:*

(1) *$\mathscr{A}$ is a σ-algebra,*

(2) *$\mathscr{A}$ is a Dynkin system and $\mathscr{A}$ is closed under finite intersections.*

△

Proof. The implication (1) $\implies$ (2) is obvious.

We will prove that (2) $\implies$ (1). We only have to prove that $\mathscr{A}$ fulfills the property (3) from the definition of a σ-algebra (Definition 1.8).

We first consider two sets $A, B \in \mathscr{A}$. We have

$$A \cup B = \overline{\overline{A \cup B}} = \overline{\overline{A} \cap \overline{B}} \in \mathscr{A},$$

since $\mathscr{A}$ is closed under complements and intersections. Thus, $\mathscr{A}$ is closed under finite unions.

Now let $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathscr{A}$. We have

$$\bigcup_{n=1}^{\infty} A_n = \sum_{n=1}^{\infty}\left(A_n \setminus \bigcup_{m=1}^{n-1} A_m\right) = \sum_{n=1}^{\infty}\left(A_n \setminus \bigcup_{m=1}^{n-1}(A_m \cap A_n)\right).$$

Since $\mathscr{A}$ is closed under intersections, we have $A_m \cap A_n \in \mathscr{A}$, and by the above $\bigcup_{m=1}^{n-1}(A_m \cap A_n) \in \mathscr{A}$. Since $\bigcup_{m=1}^{n-1}(A_m \cap A_n) \subseteq A_n$, Lemma 7.4 implies that $A_n \setminus \bigcup_{m=1}^{n-1}(A_m \cap A_n) \in \mathscr{A}$. Finally,

$$\bigcup_{n=1}^{\infty} A_n = \sum_{n=1}^{\infty}\left(A_n \setminus \bigcup_{m=1}^{n-1}(A_m \cap A_n)\right) \in \mathscr{A}$$

as a disjoint union of sets from $\mathscr{A}$.

□

Definition 7.6 For $\mathscr{E} \subseteq P(\Omega)$, the family

$$\delta(\mathscr{E}) := \bigcap_{\substack{\mathscr{A} \supseteq \mathscr{E} \\ \mathscr{A} \text{ is a Dynkin system}}} \mathscr{A}$$

is called the *Dynkin system generated by $\mathscr{E}$*.

△

Theorem 7.7 *Assume that $\mathscr{E} \subseteq P(\Omega)$ is closed under finite intersections. Then*

$$\sigma(\mathscr{E}) = \delta(\mathscr{E}).$$

△

Proof. Since every σ-algebra is a Dynkin system, the inclusion $\delta(\mathscr{E}) \subseteq \sigma(\mathscr{E})$ is obvious. We will show that $\sigma(\mathscr{E}) \subseteq \delta(\mathscr{E})$. It is enough to show that $\delta(\mathscr{E})$ is closed under finite intersection. Indeed, in this case Theorem 7.5 implies that $\delta(\mathscr{E})$ is a σ-algebra and thus $\sigma(\mathscr{E}) \subseteq \delta(\mathscr{E})$.

Let us show that $\delta(\mathscr{E})$ is closed under finite intersection. We will need several steps.

A. Fix a set $D \in \delta(\mathscr{E})$ and consider the family

$$\mathscr{A}_D := \{A \in \delta(\mathscr{E}) : A \cap D \in \delta(\mathscr{E})\}.$$

We claim that $\mathscr{A}_D$ is a Dynkin system. Let us check the conditions from the definition of a Dynkin system.

(1) $\Omega \in \mathscr{A}_D$, since $\Omega \cap D = D \in \delta(\mathscr{E})$.

(2) Let $A \in \mathscr{A}_D$. We have to show that $\overline{A} \in \mathscr{A}_D$. We have

$$\overline{A} \cap D = D \setminus (A \cap D).$$

Note that $A \cap D \in \delta(\mathscr{E})$, since $A \in \mathscr{A}_D$. On the other hand, $A \cap D \subseteq D$, and it follows by Lemma 7.4 that $\overline{A} \cap D = D \setminus (A \cap D) \in \delta(\mathscr{E})$. This means that $\overline{A} \in \mathscr{A}_D$.

(3) Let $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathscr{A}_D$ be a family of p.d. sets. Then $\{A_n \cap D\}_{n\in\mathbb{N}} \subseteq \delta(\mathscr{E})$ and are p.d. It follow that

$$\left(\sum_{n=1}^{\infty} A_n\right) \cap D = \sum_{n=1}^{\infty} (A_n \cap D) \in \delta(\mathscr{E})$$

and thus $\sum_{n=1}^{\infty} A_n \in \mathscr{A}_D$.

We have shown that $\mathscr{A}_D$ is a Dynkin system.

B. Consider a set $D \in \mathscr{E}$. Since $\mathscr{E}$ is closed under finite intersections, for each $A \in \mathscr{E}$ we have $A \cap D \in \mathscr{E} \subseteq \delta(\mathscr{E})$ and thus $A \in \mathscr{A}_D$. This show that $\mathscr{E} \subseteq \mathscr{A}_D$.

Since $\mathscr{A}_D$ is a Dynkin system, it follows that $\delta(\mathscr{E}) \subseteq \mathscr{A}_D$ for any $D \in \mathscr{E}$. In particular, if $D \in \mathscr{E}$, $A \in \delta(\mathscr{E})$, then $A \cap D \in \delta(\mathscr{E})$.

C. Finally, consider an arbitrary $D \in \delta(\mathscr{E})$. By above, for any $A \in \mathscr{E}$ we have $A \cap D \in \delta(\mathscr{E})$, and thus $A \in \mathscr{A}_D$. Consequently, $\mathscr{E} \subseteq \mathscr{A}_D$. Since $\mathscr{A}_D$ is a Dynkin system, this implies $\delta(\mathscr{E}) \subseteq \mathscr{A}_D$.

In other words, for all $D \in \delta(\mathscr{E})$, $A \in \delta(\mathscr{E})$ we have $A \cap D \in \delta(\mathscr{E})$. Thus, $\delta(\mathscr{E})$ is closed under intersections, as required.

□

Definition 7.8 For $A \subseteq \Omega_1 \times \Omega_2$, $\omega_1 \in \Omega_1$, $\omega_2 \in \Omega_2$ we call

(1) $A_{(\omega_1,\cdot)} := \{\omega_2 \in \Omega_2 : (\omega_1, \omega_2) \in A\}$ the *Ω_1-section of A determined by ω_1*,

(2) $A_{(\cdot,\omega_2)} := \{\omega_1 \in \Omega_1 : (\omega_1, \omega_2) \in A\}$ the *Ω_2-section of A determined by ω_2*.

△

Theorem 7.9 *For any $A \in \mathscr{A}_1 \otimes \mathscr{A}_2$ we have*

(1) (I) $A_{(\omega_1,\cdot)} \in \mathscr{A}_2$ *for any* $\omega_1 \in \Omega_1$,
   (II) $A_{(\cdot,\omega_2)} \in \mathscr{A}_1$ *for any* $\omega_2 \in \Omega_2$.

(2) (I) *The function* $\Omega_1 \ni \omega_1 \mapsto \mu_2(A_{(\omega_1,\cdot)}) \in \overline{\mathbb{R}}_+$ *is $\mathscr{A}_1$-measurable.*
   (II) *The function* $\Omega_2 \ni \omega_2 \mapsto \mu_1(A_{(\cdot,\omega_2)}) \in \overline{\mathbb{R}}_+$ *is $\mathscr{A}_2$-measurable.*

△

Proof. We will prove only (I); (II) can be considered similarly.

(1) Fix $\omega_1 \in \Omega_1$ and consider

$$\mathscr{A} := \{A \in \mathscr{A}_1 \otimes \mathscr{A}_2 : A_{(\omega_1,\cdot)} \in \mathscr{A}_2\}.$$

It is easy to check that $\mathscr{A}$ is a σ-algebra. (Prove this!) We have $\mathscr{R} \subseteq \mathscr{A}$, because for $A_1 \times A_2 \in \mathscr{R}$

$$(A_1 \times A_2)_{(\omega_1,\cdot)} = \{\omega_2 \in \Omega_2 : (\omega_1, \omega_2) \in A_1 \times A_2\} = \begin{cases} A_2, & \omega_1 \in A_1 \\ \varnothing, & \omega_1 \notin A_1 \end{cases} \in \mathscr{A}_2.$$

It follows that $\mathscr{A}_1 \otimes \mathscr{A}_2 = \sigma(\mathscr{R}) \subseteq \mathscr{A} \subseteq \mathscr{A}_1 \otimes \mathscr{A}_2$, and thus $\mathscr{A}_1 \otimes \mathscr{A}_2 = \mathscr{A}$.

(2) A. We first consider the case when μ_2 is finite, i.e. $\mu_2(\Omega_2) < \infty$.

For $A \in \mathscr{A}_1 \otimes \mathscr{A}_2$ we consider the function

$$f_A(\omega_1) := \mu_2(A_{(\omega_1,\cdot)}), \quad \omega_1 \in \Omega_1.$$

Set

$$\mathscr{A} := \{A \in \mathscr{A}_1 \otimes \mathscr{A}_2 : f_A : \Omega_1 \to \mathbb{R}_+ \text{ is } \mathscr{A}_1\text{-measurable}\}.$$

Clearly, $\mathscr{R} \subseteq \mathscr{A}$, because for $A_1 \times A_2 \in \mathscr{R}$ we have

$$f_{A_1 \times A_2}(\omega_1) = \mu_2((A_1 \times A_2)_{(\omega_1,\cdot)}) = \mu_2(A_2)\chi_{A_1}(\omega_1),$$

and this function is $\mathscr{A}_1$-measurable, since $A_1 \in \mathscr{A}_1$.

We will now show that $\mathscr{A}$ is a Dynkin system. Let us check the conditions in the definition of a Dynkin system

(1) $\Omega_1 \times \Omega_2 \in \mathscr{R}$ and we know that $\mathscr{R} \subseteq \mathscr{A}$. Thus, $\Omega_1 \times \Omega_2 \in \mathscr{A}$.

(2) Let $A \in \mathscr{A}$. Then

$$f_{\overline{A}}(\omega_1) = \mu_2(\Omega_2) - f_A(\omega_1)$$

is $\mathscr{A}_1$-measurable as the difference of two $\mathscr{A}_1$-measurable functions. It follows that $\overline{A} \in \mathscr{A}$.

(3) Let $\{A_n\}_{n \in \mathbb{N}}$ be a family if p.d. sets in $\mathscr{A}$. We have

$$\begin{aligned} f_{\sum_{n=1}^{\infty} A_n}(\omega_1) &= \mu_2\left(\left(\sum_{n=1}^{\infty} A_n\right)_{(\omega_1,\cdot)}\right) = \mu_2\left(\sum_{n=1}^{\infty} (A_n)_{(\omega_1,\cdot)}\right) \\ &= \sum_{n=1}^{\infty} \mu_2((A_n)_{(\omega_1,\cdot)}) = \sum_{n=1}^{\infty} f_{A_n}(\omega_1) = \lim_{N \to \infty} \sum_{n=1}^{N} f_{A_n}(\omega_1) \end{aligned}$$

is $\mathscr{A}_1$-measurable by known statements about a sum and a limit of measurable functions.

Thus, $\mathscr{A}$ is a Dynkin system.

The family $\mathscr{R}$ of measurable rectangles is a semiring and, in particular, closed under intersections. By Theorem 7.7 we obtain that $\delta(\mathscr{R}) = \sigma(\mathscr{R})$. This implies

$$\mathscr{A}_1 \otimes \mathscr{A}_2 = \sigma(\mathscr{R}) = \delta(\mathscr{R}) \subseteq \mathscr{A} \subseteq \mathscr{A}_1 \otimes \mathscr{A}_2,$$

and thus $\mathscr{A}_1 \otimes \mathscr{A}_2 = \mathscr{A}$. This proves the statement in the case when μ_2 is finite.

B. Now assume that μ_2 is σ-finite. Then there are sets $\{C_{2,n}\}_{n\in\mathbb{N}} \subseteq \mathscr{A}_2$ such that $C_{2,n} \subseteq C_{2,n+1}$, $\mu(C_{2,n}) < \infty$, $n \in \mathbb{N}$, and $\Omega_2 = \bigcup_{n=1}^{\infty} C_{2,n}$. For each $n \in \mathbb{N}$ we define

$$\mu_2^{(n)}(A_2) := \mu_2(A_2 \cap C_{2,n}), \quad A_2 \in \mathscr{A}_2.$$

The set functions $\mu_2^{(n)}$ are finite measures on $\mathscr{A}_2$. In particular, the functions

$$f_A^{(n)}(\omega_1) := \mu_2^{(n)}(A_{(\omega_1,\cdot)})$$

are $\mathscr{A}_1$-measurable by the first part of the proof. But it follows from the continuity of μ_2 from below that

$$\mu_2(A_{(\omega_1,\cdot)}) = \lim_{n\to\infty} \mu_2(A_{(\omega_1,\cdot)} \cap C_{2,n}) = \lim_{n\to\infty} \mu_2^{(n)}(A_{(\omega_1,\cdot)}).$$

Thus, $f_A = \lim_{n\to\infty} f_A^{(n)}$ is $\mathscr{A}_1$-measurable as the limit of a sequence of measurable functions. This observation finishes the proof of the theorem.

□

Theorem 7.10 *For any $A \in \mathscr{A}_1 \otimes \mathscr{A}_2$ we have*

$$\int_{\Omega_1} \mu_2(A_{(\omega_1,\cdot)}) d\mu_1(\omega_1) = \int_{\Omega_2} \mu_1(A_{(\cdot,\omega_2)}) d\mu_2(\omega_2).$$

The formula

$$\mu(A) := \int_{\Omega_1} \mu_2(A_{(\omega_1,\cdot)}) d\mu_1(\omega_1), \quad A \in \mathscr{A}_1 \otimes \mathscr{A}_2,$$

defines a σ-finite measure in $\mathscr{A}_1 \otimes \mathscr{A}_2$ such that

$$\mu(A_1 \times A_2) = \mu_1(A_1)\mu_2(A_2), \quad A_1 \in \mathscr{A}_1,\ A_2 \in \mathscr{A}_2. \tag{$*$}$$

The condition $(*)$ determines μ uniquely.

△

We first prove a lemma which is useful on its own.

Lemma 7.11 *Let $\mathscr{R} \subseteq P(\Omega)$ be closed under intersections. Let μ, ν be two finite measures on $\sigma(\mathscr{R})$ such that*

(1) $\mu(A) = \nu(A)$ *for all* $A \in \mathscr{R}$,

(2) $\mu(\Omega) = \nu(\Omega) < \infty$.

Then $\mu = \nu$, *i.e.* $\mu(A) = \nu(A)$ *for all* $A \in \sigma(\mathscr{R})$.

△

Proof. Set

$$\mathscr{A} := \{A \in \sigma(\mathscr{R}) : \mu(A) = \nu(A)\}.$$

Clearly, $\mathscr{R} \subseteq \mathscr{A}$. Moreover, $\mathscr{A}$ is a Dynkin system. Indeed:

(1) $\Omega \in \mathscr{A}$, because $\mu(\Omega) = \nu(\Omega)$ by assumption.

(2) If $A \in \mathscr{A}$, then

$$\mu(\overline{A}) = \mu(\Omega) - \mu(A) = \nu(\Omega) - \nu(A) = \nu(\overline{A}),$$

so that $\overline{A} \in \mathscr{A}$. Here we used the fact that $\mu(\Omega) < \infty$.

(3) If $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathscr{A}$ is a sequence of p.d. sets, then

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

so that $\sum_{n=1}^{\infty} A_n \in \mathscr{A}$.

Thus, $\mathscr{A}$ is a Dynkin system.

Since $\mathscr{R}$ is closed under intersections, it follows by Theorem 7.7 that $\sigma(\mathscr{R}) = \delta(\mathscr{R})$. We have

$$\sigma(\mathscr{R}) = \delta(\mathscr{R}) \subseteq \mathscr{A} \subseteq \sigma(\mathscr{R}).$$

This implies

$$\sigma(\mathscr{R}) = \mathscr{A}.$$

□

Proof of Theorem 7.10. A. Recall that the functions $\omega_1 \mapsto \mu_2(A_{(\omega_1,\cdot)}) \in \overline{\mathbb{R}}_+$, $\omega_2 \mapsto \mu_1(A_{(\cdot,\omega_2)}) \in \overline{\mathbb{R}}_+$ are measurable by Theorem 7.9. We first show that

$$\mu(A) := \int_{\Omega_1} \mu_2(A_{(\omega_1,\cdot)}) d\mu_1(\omega_1)$$

is a measure on $\mathscr{A}_1 \otimes \mathscr{A}_2$. Indeed, $\mathscr{A}_1 \otimes \mathscr{A}_2$ is a σ-algebra, and obviously $\mu(\varnothing) = 0$. It remains to show that μ is σ-additive.

Consider a family of p.d. sets $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathscr{A}_1 \otimes \mathscr{A}_2$. We have

$$\begin{aligned}
\mu\left(\sum_{n=1}^{\infty} A_n\right) &= \int_{\Omega_1} \mu_2\left(\left(\sum_{n=1}^{\infty} A_n\right)_{(\omega_1,\cdot)}\right) d\mu_1(\omega_1) \\
&= \int_{\Omega_1} \mu_2\left(\sum_{n=1}^{\infty} (A_n)_{(\omega_1,\cdot)}\right) d\mu_1(\omega_1) \\
&= \int_{\Omega_1} \sum_{n=1}^{\infty} \mu_2((A_n)_{(\omega_1,\cdot)}) d\mu_1(\omega_1) \\
&= \int_{\Omega_1} \lim_{N\to\infty} \sum_{n=1}^{N} \mu_2((A_n)_{(\omega_1,\cdot)}) d\mu_1(\omega_1) \\
&= \lim_{N\to\infty} \sum_{n=1}^{N} \int_{\Omega_1} \mu_2((A_n)_{(\omega_1,\cdot)}) d\mu_1(\omega_1) \\
&= \sum_{n=1}^{\infty} \int_{\Omega_1} \mu_2((A_n)_{(\omega_1,\cdot)}) d\mu_1(\omega_1) \\
&= \sum_{n=1}^{\infty} \mu(A_n),
\end{aligned}$$

where we used the Monotone Convergence Theorem to interchange the limit and the integration. Thus, μ is σ-additive, and therefore μ is a measure.

B. Now we show that μ satisfies $(*)$. We have

$$\begin{aligned}
\mu(A_1 \times A_2) &= \int_{\Omega_1} \mu_2((A_1 \times A_2)_{(\omega_1,\cdot)}) d\mu_1(\omega_1) \\
&= \int_{\Omega_1} \mu_2(A_2) \chi_{A_1}(\omega_1) d\mu_1(\omega_1) \\
&= \mu_2(A_2) \int_{\Omega_1} \chi_{A_1}(\omega_1) d\mu_1(\omega_1) = \mu_2(A_2)\mu_1(A_1).
\end{aligned}$$

C. Since μ_1 and μ_2 are σ-finite, there are sequences $\{C_n\}_{n\in\mathbb{N}} \subseteq \mathscr{A}_1, \Omega_1 = \bigcup_{n=1}^{\infty} C_n$, $\mu_1(C_n) < \infty$, $C_n \subseteq C_{n+1}$, and $\{D_n\}_{n\in\mathbb{N}} \subseteq \mathscr{A}_2$, $\Omega_2 = \bigcup_{n=1}^{\infty} D_n$, $\mu_2(D_n) < \infty$, $D_n \subseteq D_{n+1}$, $n \in \mathbb{N}$. But then $\Omega_1 \times \Omega_2 = \bigcup_{n=1}^{\infty}(C_n \times D_n)$, $\mu(C_n \times D_n) = \mu_1(C_n)\mu_2(D_n) < \infty$, $C_n \times D_n \subseteq C_{n+1} \times D_{n+1}$, $n \in \mathbb{N}$. It follows that μ is σ-finite.

D. Now suppose that ν is another σ-finite measure on $\mathscr{A}_1 \otimes \mathscr{A}_2$ that satisfies $(*)$. It follows from $(*)$ that $\mu(A) = \nu(A)$ for all $A \in \mathscr{R}$.

If μ is finite, then also ν is finite, and $\mu = \nu$ by Lemma 7.11.

We now consider the case when μ and ν are σ-finite. Define for $n \in \mathbb{N}$, $A \in \mathscr{A}_1 \otimes \mathscr{A}_2$

$$\mu_n(A) := \mu(A \cap (C_n \times D_n)), \quad \nu_n(A) := \nu(A \cap (C_n \times D_n)).$$

The set functions μ_n and ν_n are finite measures on $\mathscr{A}_1 \otimes \mathscr{A}_2$,

$$\mu_n(\Omega_1 \times \Omega_2) = \mu(C_n \times D_n) = \nu(C_n \times D_n) = \nu_n(\Omega_1 \times \Omega_2).$$

Moreover, if $A \in \mathscr{R}$, then $A \cap (C_n \times D_n) \in \mathscr{R}$ and thus $\mu_n(A) = \nu_n(A)$. We see that the assumptions of Lemma 7.11 are fulfilled for μ_n and ν_n. It follows that $\mu_n = \nu_n$.

The continuity of μ and ν from below yields that $\mu = \nu$ (argue like in the proof on Theorem 7.9(2)).

This shows that there is only one σ-finite measure on $\mathscr{A}_1 \otimes \mathscr{A}_2$ that satisfies $(*)$.

E. Finally,

$$\tilde{\mu}(A) := \int_{\Omega_2} \mu_1(A_{(\cdot,\omega_2)}) d\mu_2(\omega_2)$$

is also a σ-finite measure on $\mathscr{A}_1 \otimes \mathscr{A}_2$ that satisfies $(*)$ (the proof is the same as above with interchanging Ω_1 and Ω_2). It follows that $\mu = \tilde{\mu}$.

□

Definition 7.12 The measure μ from Theorem 7.10 is called the *product measure* of the measures μ_1 and μ_2. We denote this measure by $\mu = \mu_1 \times \mu_2 = \mu_1 \otimes \mu_2$.

The measure space $(\Omega_1 \times \Omega_2, \mathscr{A}_1 \otimes \mathscr{A}_2, \mu_1 \times \mu_2)$ is called the *product of measure spaces* $(\Omega_1, \mathscr{A}_1, \mu_1)$ and $(\Omega_2, \mathscr{A}_2, \mu_2)$.

△

Definition 7.13 Let $f : \Omega_1 \times \Omega_2 \to \overline{\mathbb{R}}$, $\omega_1 \in \Omega_1$, $\omega_2 \in \Omega_2$.

(1) The function $f_{(\omega_1,\cdot)} : \Omega_2 \to \overline{\mathbb{R}}$, $f_{(\omega_1,\cdot)}(\omega_2) = f(\omega_1,\omega_2)$ is called a *Ω_1-section of f determined by ω_1*.

(2) The function $f_{(\cdot,\omega_2)} : \Omega_1 \to \overline{\mathbb{R}}$, $f_{(\cdot,\omega_2)}(\omega_1) = f(\omega_1,\omega_2)$ is called a *Ω_2-section of f determined by ω_2*.

△

Proposition 7.14 *If $f : \Omega_1 \times \Omega_2 \to \overline{\mathbb{R}}$ is $\mathscr{A}_1 \otimes \mathscr{A}_2$-measurable, then $f_{(\omega_1,\cdot)}$ is $\mathscr{A}_2$-measurable for all $\omega_1 \in \Omega_1$, and $f_{(\cdot,\omega_2)}$ is $\mathscr{A}_1$-measurable for all $\omega_2 \in \Omega_2$.*

△

Proof. We consider only $f_{(\omega_1,\cdot)}$. For $x \in \mathbb{R}$ and $\omega_1 \in \Omega_1$ we have

$$\begin{aligned}\{f_{(\omega_1,\cdot)} \le x\} &= \{\omega_2 \in \Omega_2 : f(\omega_1,\omega_2) \le x\} \\ &= \{\omega_2 \in \Omega_2 : (\omega_1,\omega_2) \in \{f \le x\}\} = \{f \le x\}_{(\omega_1,\cdot)}.\end{aligned}$$

Observe that $\{f \le x\} \in \mathscr{A}_1 \otimes \mathscr{A}_2$, because f is $\mathscr{A}_1 \otimes \mathscr{A}_2$-measurable. It follows by Theorem 7.9(1) that $\{f \le x\}_{(\omega_1,\cdot)} \in \mathscr{A}_2$, and this means that $f_{(\omega_1,\cdot)}$ is $\mathscr{A}_2$-measurable.

□

Theorem 7.15 (*Tonelli's Theorem*) *Let $f : \Omega_1 \times \Omega_2 \to \overline{\mathbb{R}}_+$ be a $\mathscr{A}_1 \otimes \mathscr{A}_2$-measurable function. Then*

(1) *The function $\Omega_1 \ni \omega_1 \mapsto \int_{\Omega_2} f(\omega_1,\omega_2)d\mu_2(\omega_2) \in \overline{\mathbb{R}}_+$ is $\mathscr{A}_1$-measurable, and the function $\Omega_2 \ni \omega_2 \mapsto \int_{\Omega_1} f(\omega_1,\omega_2)d\mu_1(\omega_1) \in \overline{\mathbb{R}}_+$ is $\mathscr{A}_2$-measurable.*

(2) *We have*

$$\begin{aligned}\int_{\Omega_1\times\Omega_2} f d(\mu_1 \times \mu_2) &= \int_{\Omega_1}\int_{\Omega_2} f(\omega_1,\omega_2)d\mu_2(\omega_2)d\mu_1(\omega_1) \\ &= \int_{\Omega_2}\int_{\Omega_1} f(\omega_1,\omega_2)d\mu_1(\omega_1)d\mu_2(\omega_2).\end{aligned}$$

△

Proof. We prove the statement in two steps.

A. Let $f : \Omega_1 \times \Omega_2 \to \mathbb{R}_+$ be a simple function. Then $f = \sum_{i=1}^m a_i \chi_{A_i}$, where $a_1, \dots, a_m \in \mathbb{R}_+$, $A_1, \dots, A_m \in \mathscr{A}_1 \otimes \mathscr{A}_2$.

For $\omega_1 \in \Omega_1$, $\omega_2 \in \Omega_2$ we have

$$\chi_{A_i}(\omega_1, \omega_2) = 1 \iff (\omega_1, \omega_2) \in A_i \iff \omega_2 \in (A_i)_{(\omega_1, \cdot)} \iff \chi_{(A_i)_{(\omega_1,\cdot)}}(\omega_2) = 1.$$

It follows that

$$\begin{aligned}\int_{\Omega_2} f(\omega_1, \omega_2) d\mu_2(\omega_2) &= \sum_{i=1}^m a_i \int_{\Omega_2} \chi_{A_i}(\omega_1, \omega_2) d\mu_2(\omega_2) \\ &= \sum_{i=1}^m a_i \int_{\Omega_2} \chi_{(A_i)_{(\omega_1,\cdot)}}(\omega_2) d\mu(\omega_2) \\ &= \sum_{i=1}^m a_i \mu_2((A_i)_{(\omega_1,\cdot)}),\end{aligned}$$

and thus this function is $\mathscr{A}_1$-measurable by Theorem 7.9. It follows

$$\begin{aligned}&\int_{\Omega_1} \int_{\Omega_2} f(\omega_1, \omega_2) d\mu_2(\omega_2) d\mu_1(\omega_1) \\ &= \sum_{i=1}^m a_i \int_{\Omega_1} \mu_2((A_i)_{(\omega_1,\cdot)}) d\mu_1(\omega_1) \quad \text{(Theorem 7.10)} \\ &= \sum_{i=1}^m a_i (\mu_1 \times \mu_2)(A_i) \\ &= \int_{\Omega_1 \times \Omega_2} \sum_{i=1}^m a_i \chi_{A_i} d(\mu_1 \times \mu_2) \\ &= \int_{\Omega_1 \times \Omega_2} f d(\mu_1 \times \mu_2).\end{aligned}$$

The other integral can be considered similarly.

B. Now we consider the case when $f : \Omega_1 \times \Omega_2 \to \overline{\mathbb{R}}_+$ is an arbitrary $\mathscr{A}_1 \otimes \mathscr{A}_2$-measurable function.

Let $\{f_n\}_{n \in \mathbb{N}}$ be a sequence of non-negative simple functions $f_n : \Omega_1 \times \Omega_2 \to \overline{\mathbb{R}}_+$ such that $0 \le f_n \le f_{n+1}$, $f_n \to f$, $n \to \infty$. Then also $0 \le f_n(\omega_1, \cdot) \le f_{n+1}(\omega_1, \cdot)$ and $f_n(\omega_1, \cdot) \to f(\omega_1, \cdot)$, $n \to \infty$. By the Monotone Convergence Theorem,

$$\int_{\Omega_2} f_n(\omega_1, \cdot) d\mu_2 \to \int_{\Omega_2} f(\omega_1, \cdot) d\mu_2, \quad n \to \infty,$$

for all $\omega_1 \in \Omega_1$. Since the functions $\omega_1 \mapsto \int_{\Omega_2} f_n(\omega_1, \cdot) d\mu_2$ are $\mathscr{A}_1$-measurable for all $n \in \mathbb{N}$, their limit function $\omega_1 \mapsto \int_{\Omega_2} f(\omega_1, \cdot) d\mu_2$ is also $\mathscr{A}_1$-measurable. Using the Monotone Convergence Theorem again, we write

$$\begin{aligned}
\int_{\Omega_1} \int_{\Omega_2} f(\omega_1, \cdot) d\mu_2 d\mu_1(\omega_1) &= \lim_{n\to\infty} \int_{\Omega_1} \int_{\Omega_2} f_n(\omega_1, \cdot) d\mu_2 d\mu_1(\omega_1) \\
&= \lim_{n\to\infty} \int_{\Omega_1 \times \Omega_2} f_n d(\mu_1 \times \mu_2) \\
&= \int_{\Omega_1 \times \Omega_2} f d(\mu_1 \times \mu_2).
\end{aligned}$$

The other integral can be considered similarly.

□

Theorem 7.16 (*Fubini's Theorem*) *Let $f : \Omega_1 \times \Omega_2 \to \overline{\mathbb{R}}$ is a $\mu_1 \times \mu_2$ -integrable function. Then*

(1) *The function $f(\omega_1, \cdot) : \Omega_2 \to \overline{\mathbb{R}}$ is μ_2-integrable for μ_1-almost all $\omega_1 \in \Omega_1$. The function $f(\cdot, \omega_2) : \Omega_1 \to \overline{\mathbb{R}}$ is μ_1-integrable for μ_2-almost all $\omega_2 \in \Omega_2$.*

(2) *The μ_1-a.e. defined function on Ω_1,*

$$\omega_1 \mapsto \int_{\Omega_2} f(\omega_1, \cdot) d\mu_2$$

is μ_1-integrable. The μ_2-a.e. defined function on Ω_2,

$$\omega_2 \mapsto \int_{\Omega_1} f(\cdot, \omega_2) d\mu_1$$

is μ_2-integrable.

(3) *We have*

$$\begin{aligned}
\int_{\Omega_1 \times \Omega_2} f d(\mu_1 \times \mu_2) &= \int_{\Omega_1} \int_{\Omega_2} f(\omega_1, \omega_2) d\mu_2(\omega_2) d\mu_1(\omega_1) \\
&= \int_{\Omega_2} \int_{\Omega_1} f(\omega_1, \omega_2) d\mu_1(\omega_1) d\mu_2(\omega_2).
\end{aligned}$$

△

Proof. Recall that f is $\mu_1 \times \mu_2$ -integrable if and only if $|f|$ is $\mu_1 \times \mu_2$ -integrable.

(1) Set

$$N_1 := \left\{ \omega_1 \in \Omega_1 : \int_{\Omega_2} |f(\omega_1, \cdot)| d\mu_2 = \infty \right\}.$$

We have

$$N_1 = \left\{ \omega_1 \in \Omega_1 : f(\omega_1, \cdot) \text{ is not } \mu_2\text{-integrable} \right\}.$$

By Theorem 7.15(1), $N_1 \in \mathscr{A}_1$. By Theorem 7.15(2) we have

$$\int_{\Omega_1} \int_{\Omega_2} |f(\omega_1, \omega_2)| d\mu_2(\omega_2) d\mu_1(\omega_1) = \int_{\Omega_1 \times \Omega_2} |f| d(\mu_1 \times \mu_2) < \infty.$$

It follows that the function $\omega_1 \mapsto \int_{\Omega_2} |f(\omega_1, \cdot)| d\mu_2$ is μ_1-integrable, and consequently μ_1-a.e. finite. This proves that $\mu_1(N_1) = 0$.

The function $f(\cdot, \omega_2)$ can be considered similarly.

(2) We have

$$\begin{aligned} &\int_{\Omega_1 \setminus N_1} \left| \int_{\Omega_2} f(\omega_1, \omega_2) d\mu_2(\omega_2) \right| d\mu_1(\omega_1) \\ &\leq \int_{\Omega_1} \int_{\Omega_2} |f(\omega_1, \omega_2)| d\mu_2(\omega_2) d\mu_1(\omega_1) < \infty. \end{aligned}$$

Thus, the function $\omega_1 \mapsto \int_{\Omega_2} f(\omega_1, \omega_2) d\mu_2(\omega_2)$ is μ_1-integrable (with any continuation on the null set N_1).

The other case can be considered similarly.

(3) We have

$$\begin{aligned}
&\int_{\Omega_1\times\Omega_2} f d(\mu_1\times\mu_2)\\
&= \int_{\Omega_1\times\Omega_2} f^+ d(\mu_1\times\mu_2) - \int_{\Omega_1\times\Omega_2} f^- d(\mu_1\times\mu_2) \text{ (Tonelli's Theorem)}\\
&= \int_{\Omega_1}\int_{\Omega_2} f^+ d\mu_2 d\mu_1 - \int_{\Omega_1}\int_{\Omega_2} f^- d\mu_2 d\mu_1\\
&= \int_{\Omega_1\setminus N_1} \underbrace{\int_{\Omega_2} f^+(\omega_1,\omega_2) d\mu_2\, d\mu_1}_{<\infty} - \int_{\Omega_1\setminus N_1} \underbrace{\int_{\Omega_2} f^-(\omega_1,\omega_2) d\mu_2\, d\mu_1}_{<\infty}\\
&= \int_{\Omega_1\setminus N_1}\int_{\Omega_2} (f^+ - f^-) d\mu_2 d\mu_1\\
&= \int_{\Omega_1}\int_{\Omega_2} f d\mu_1 d\mu_2.
\end{aligned}$$

The other case can be considered similarly.

□

Product of finitely many measure spaces

The theory can be extended to the case of finitely many measure spaces. Let $(\Omega_1,\mathscr{A}_1,\mu_1),\dots,(\Omega_n,\mathscr{A}_n,\mu_n)$ be σ-finite measure spaces. The Cartesian product of the sets $\Omega_1,\dots,\Omega_n$ is

$$\bigtimes_{i=1}^n \Omega_i = \{(\omega_1,\dots,\omega_n) : \omega_i \in \Omega_i,\ i=1,\dots,n\}.$$

The product σ-algebra of $\mathscr{A}_1,\dots,\mathscr{A}_n$ on $\bigtimes_{i=1}^n \Omega_i$ is

$$\bigotimes_{i=1}^n \mathscr{A}_i := \sigma(\{A_1\times\cdots\times A_n : A_i\in\mathscr{A}_i,\ i=1,\dots,n\}).$$

The *product of measures* $\mu_1,\dots,\mu_n$ is the unique σ-finite measure $\mu = \bigtimes_{i=1}^n \mu_i = \bigotimes_{i=1}^n \mu_i$ on $\bigotimes_{i=1}^n \mathscr{A}_i$ such that

$$\mu\left(\bigtimes_{i=1}^n A_i\right) = \prod_{i=1}^n \mu_i(A_i), \quad A_i\in\mathscr{A}_i,\ i=1,\dots,n.$$

This measure can be constructed inductively. First construct the measure $\mu_{12} = \mu_1 \times \mu_2$ on $\mathscr{A}_{12} = \mathscr{A}_1 \otimes \mathscr{A}_2$. Then construct $\mu_{123} = \mu_{12} \times \mu_3$ on $\mathscr{A}_{123} = \mathscr{A}_{12} \otimes \mathscr{A}_3$, and so on.

Tonelli's and Fubini's Theorems remain true for products of n measure spaces (for the proof apply them inductively as above).

The d-dimensional Lebesgue measure

On $\mathbb{R}^d$, we consider the semiring

$$\mathscr{J}_d := \left\{ \bigtimes_{i=1}^{d} (a_i, b_i] : a_i, b_i \in \mathbb{R} \right\}.$$

The generated σ-algebra $\mathscr{B}_d := \sigma(\mathscr{J}_d)$ is the Borel σ-algebra on $\mathbb{R}^d$. As in the one-dimensional case, it can be shown that $\mathscr{B}_d$ is exactly the σ-algebra generated by open sets in $\mathbb{R}^d$.

According to the above theory, there exists a unique σ-finite measure λ_d in $\mathscr{B}_d$ such that

$$\lambda_d \left(\bigtimes_{i=1}^{d} (a_i, b_i] \right) = \prod_{i=1}^{d} \max(b_i - a_i, 0).$$

This measure λ_d is called the *d-dimensional Lebesgue measure*.

As in the one-dimensional case, it can be completed.

Tonelli's and Fubini's Theorems hold true for λ_d. In particular, for an integrable function $f : \mathbb{R}^d \to \overline{\mathbb{R}}$ we have

$$\int_{\mathbb{R}^d} f d\lambda_d = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, \dots, x_d) dx_d \dots dx_1.$$

Example A. For any $x \in (0, \infty)$ we have

$$\int_0^\infty y e^{-(1+x^2)y^2} dy = -\frac{1}{2} \frac{1}{1+x^2} e^{-(1+x^2)y^2} \Big|_{y=0}^{y=\infty} = \frac{1}{2(1+x^2)}$$

and

$$\int_0^\infty \int_0^\infty y e^{-(1+x^2)y^2} dy dx = \frac{1}{2} \int_0^\infty \frac{1}{1+x^2} dx = \frac{1}{2} \arctan x \Big|_{x=0}^{x=\infty} = \frac{1}{2} \frac{\pi}{2} = \frac{\pi}{4}.$$

Since $ye^{-(1+x^2)y^2} \geq 0$ for $x, y \in (0, \infty)$, we can apply Tonelli's Theorem. We have

$$\begin{aligned}
&\int_0^\infty \int_0^\infty ye^{-(1+x^2)y^2} dydx = \int_0^\infty \int_0^\infty ye^{-(1+x^2)y^2} dxdy \\
&= \int_0^\infty ye^{-y^2} \left(\int_0^\infty e^{-x^2y^2} dx \right) dy \quad (\text{with } xy = t) \\
&= \int_0^\infty ye^{-y^2} \left(\int_0^\infty e^{-t^2} \frac{1}{y} dt \right) dy \\
&= \int_0^\infty e^{-y^2} dy \int_0^\infty e^{-x^2} dx = \left(\int_0^\infty e^{-x^2} dx \right)^2 .
\end{aligned}$$

It follows that

$$\int_0^\infty e^{-x^2} dx = \frac{\sqrt{\pi}}{2} \quad \text{and} \quad \int_{-\infty}^\infty e^{-x^2} dx = \sqrt{\pi}.$$

B. This example shows that the order of integration cannot be interchanged for an arbitrary function.

Consider $f(x, y) = \frac{x^2 - y^2}{(x^2+y^2)^2}$. We have

$$\frac{d}{dy} \left[\frac{y}{x^2 + y^2} \right] = \frac{(x^2 + y^2) - y \cdot 2y}{(x^2 + y^2)^2} = f(x, y),$$

and thus

$$\begin{aligned}
\int_0^1 \int_0^1 f(x, y) dydx &= \int_0^1 \int_0^1 \frac{d}{dy} \left[\frac{y}{x^2 + y^2} \right] dydx \\
&= \int_0^1 \left[\frac{y}{x^2 + y^2} \right]_{y=0}^{y=1} = \int_0^1 \frac{1}{1 + x^2} dx \\
&= \arctan x |_{x=0}^{x=1} = \frac{\pi}{4}.
\end{aligned}$$

On the other hand,

$$\int_0^1 \int_0^1 f(x, y) dxdy = - \int_0^1 \int_0^1 f(y, x) dxdy = -\frac{\pi}{4}.$$

This does not contradict Fubini's Theorem because $f(x,y)$ is not integrable on $(0,1)^2$. Indeed,

$$\int_0^1 \int_0^1 \left| \frac{x^2 - y^2}{(x^2+y^2)^2} \right| dydx \geq \int_0^1 \int_0^x \frac{x^2-y^2}{(x^2+y^2)^2} dydx$$
$$= \int_0^1 \left[\frac{y}{x^2+y^2} \right]_{y=0}^{y=x} dx = \int_0^1 \frac{x}{x^2+x^2} dx = \int_0^1 \frac{1}{2x} dx = \infty.$$

c. The diagonal $D = \{(x,y) \in \mathbb{R}^2 : x = y\}$ is closed and thus a Borel set. We have

$$\lambda_2(D) = \int_{\mathbb{R}^2} \chi_D d\lambda_2 = \int_{\mathbb{R}} \int_{\mathbb{R}} \chi_D(x,y) dxdy$$
$$= \int_{\mathbb{R}} \int_{\mathbb{R}} \chi_{\{y\}}(x) dxdy = \int_{\mathbb{R}} \underbrace{\lambda(\{y\})}_{=0} dy = 0.$$

△

Chapter 8

L^p-spaces

Definition 8.1 Let X be a vector space over the field $\mathbb{R}$ or $\mathbb{C}$. A *norm* on X is a mapping $\|\cdot\| : X \to \mathbb{R}$ with the following properties: for all $x, y \in X$, $\alpha \in \mathbb{R}$ or $\mathbb{C}$

(1) $\|x\| \geq 0$ and ($\|x\| = 0 \iff x = 0$),

(2) $\|\alpha x\| = |\alpha|\|x\|$,

(3) $\|x + y\| \leq \|x\| + \|y\|$ (the triangle inequality).

The pair $(X, \|\cdot\|)$ is called a *normed space*.

△

We will now define the L^p-spaces, $1 \leq p \leq \infty$.
Let $(\Omega, \mathscr{A}, \mu)$ be a measure space. We first consider $1 \leq p < \infty$. Let

$$\mathscr{L}^p(\Omega, \mathscr{A}, \mu) := \left\{ f : \Omega \to \mathbb{R} \text{ (or } \mathbb{C}) : f \text{ is measurable}, \int_\Omega |f|^p d\mu < \infty \right\}.$$

Consider the mapping

$$|f|_p := \left(\int_\Omega |f|^p d\mu \right)^{\frac{1}{p}}.$$

This mapping is almost a norm. We have $|f|_p \geq 0$, $|\alpha f|_p = |\alpha||f|_p$. The triangle inequality is not obvious, but we will show it. However, $|f|_p = 0$ does not imply that $f = 0$, it only implies that $f = 0$ μ-a.e. By Theorem 5.15 we have

$$|f|_p = 0 \iff f = 0 \ \mu\text{-a.e.}$$

and by Theorem 5.17 $f = g$ μ-a.e. $\implies |f|_p = |g|_p$.
We introduce an equivalence relation on $\mathscr{L}^p(\Omega, \mathscr{A}, \mu)$:

$$f \sim g \iff f = g \ \mu\text{-a.e.}$$

The *space* $L^p(\Omega, \mathscr{A}, \mu)$*,* $1 \le p < \infty$, is defined as the quotient space

$$L^p(\Omega, \mathscr{A}, \mu) := \mathscr{L}^p(\Omega, \mathscr{A}, \mu)/\sim,$$

i.e.

$$L^p(\Omega, \mathscr{A}, \mu) = \{[f] : f \in \mathscr{L}^p(\Omega, \mathscr{A}, \mu)\},$$

where $[f]$ denotes the equivalence class of f.

We usually still call elements of $L^p(\Omega, \mathscr{A}, \mu)$ *L^p-functions*. We sometimes write simply $L^p(\mu)$ or even just L^p.

For $f \in L^p(\Omega, \mathscr{A}, \mu)$ we define the *L^p-norm*

$$\|f\|_p := \|f\|_{L^p(\Omega, \mathscr{A}, \mu)} := \left(\int_\Omega |f|^p d\mu\right)^{\frac{1}{p}}.$$

The case when $p = \infty$ needs a special consideration. We say that a function $f : \Omega \to \mathbb{R}$ (or $\mathbb{C}$) is *essentially bounded*, if

$$\exists\, M > 0 : \mu(\{x \in \Omega : |f(x)| > M\}) = 0.$$

A number M with this property is called an *essential bound* for f. The *essential supremum* of $|f|$ on Ω is

$$\operatorname*{ess\,sup}_{x \in \Omega} |f(x)| := \inf\{M > 0 : M \text{ is an essential bound for } f\}.$$

Notice that $\operatorname{ess\,sup}_{x \in \Omega} |f(x)|$ is itself an essential bound for f, i.e. the infimum in the definition of the essential supremum is in fact a minimum. In particular,

$$|f| \le \operatorname*{ess\,sup}_{x \in \Omega} |f(x)| \quad \mu\text{-a.e.}$$

Indeed, put $\widetilde{M} := \operatorname{ess\,sup}_{x \in \Omega} |f(x)|$. We have

$$\left\{x \in \Omega : |f(x)| > \widetilde{M}\right\} = \bigcup_{n \in \mathbb{N}} \left\{x \in \Omega : |f(x)| > \widetilde{M} + \frac{1}{n}\right\},$$

where the right-hand side is a union of null sets. It follows that $\mu(\{x \in \Omega : |f(x)| > \widetilde{M}\}) = 0$.

We now put

$$\mathscr{L}^\infty(\Omega, \mathscr{A}, \mu) := \{f : \Omega \to \mathbb{R} \text{ (or } \mathbb{C}) : f \text{ is measurable and essentially bounded}\}$$

and

$$|f|_\infty := \operatorname*{ess\,sup}_{x\in\Omega} |f(x)|.$$

It is easy to see that $|f|_\infty = 0 \iff f = 0$ μ-a.e. and $f \sim g \implies |f|_\infty = |g|_\infty$.

We define the space $L^\infty(\Omega, \mathscr{A}, \mu)$ as the quotient space

$$L^\infty(\Omega, \mathscr{A}, \mu) := \mathscr{L}^\infty(\Omega, \mathscr{A}, \mu)/\sim,$$

i.e.

$$L^\infty(\Omega, \mathscr{A}, \mu) = \{[f] : f \in \mathscr{L}^\infty(\Omega, \mathscr{A}, \mu)\}.$$

The *L^∞-norm* of $f \in L^\infty(\Omega, \mathscr{A}, \mu)$ is defined by

$$\|f\|_\infty := \|f\|_{L^\infty(\Omega,\mathscr{A},\mu)} := \operatorname*{ess\,sup}_{x\in\Omega} |f(x)|.$$

Now we will justify that L^p-spaces are indeed vector spaces, and that $\|\cdot\|_p$ is indeed a norm. Property (1) from the definition of the norm holds by construction. Let us check the second property. Let $f \in L^p$ and $\alpha \in \mathbb{R}$ (or $\mathbb{C}$). For $1 \le p < \infty$ we have

$$\left(\int_\Omega |\alpha f|^p d\mu\right)^{\frac{1}{p}} = \left(|\alpha|^p \int_\Omega |f|^p d\mu\right)^{\frac{1}{p}} = (|\alpha|^p)^{\frac{1}{p}} \left(\int_\Omega |f|^p d\mu\right)^{\frac{1}{p}} = |\alpha| \|f\|_p,$$

and thus $\alpha f \in L^p$ and $\|\alpha f\|_p = |\alpha| \|f\|_p$. For $p = \infty$, one can see directly that $\operatorname{ess\,sup}_{x\in\Omega} |\alpha f(x)| = |\alpha| \operatorname{ess\,sup}_{x\in\Omega} |f(x)|$, and the same follows.

The proof of the fact that the sum of two L^p-functions is an L^p-function is more involved. Let $f, g \in L^p$. Consider first $1 \le p < \infty$. We have

$$\begin{aligned}\int_\Omega |f+g|^p d\mu &\le \int_\Omega \left(|f| + |g|\right)^p d\mu \le \int_\Omega \left(2\max(|f|, |g|)\right)^p d\mu \\ &= 2^p \int_\Omega \max(|f|^p, |g|^p) d\mu \le 2^p \int_\Omega \left(|f|^p + |g|^p\right) d\mu \\ &= 2^p \left(\int_\Omega |f|^p d\mu + \int_\Omega |g|^p d\mu\right) < \infty,\end{aligned}$$

and thus $f + g \in L^p$. The triangle inequality will be shown below in Theorem 8.5.

Now let $p = \infty$. Then

$$\left\{|f+g| > \|f\|_\infty + \|g\|_\infty\right\} \subseteq \left\{|f| > \|f\|_\infty\right\} \cup \left\{|g| > \|g\|_\infty\right\},$$

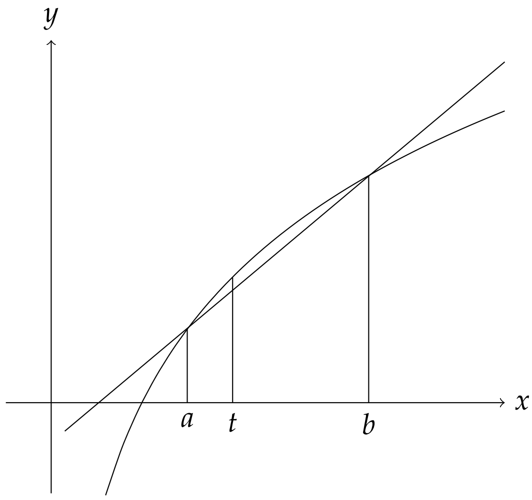

Figure 8.1: The graph of the logarithm function and a secant, with $t = ra+(1-r)b$.

so that

$$\mu\left(|f+g| > \|f\|_\infty + \|g\|_\infty\right) \le \mu\left(|f| > \|f\|_\infty\right) + \mu\left(|g| > \|g\|_\infty\right) = 0.$$

It follows that $f+g$ is essentially bounded, i.e., $f+g \in L^\infty$, and

$$\|f+g\|_\infty \le \|f\|_\infty + \|g\|_\infty.$$

We will now prove several inequalities that are interesting and useful on its own.

Lemma 8.2 (*Young's Inequality*) *Let $a, b \ge 0$, $0 < r < 1$. Then*

$$a^r b^{1-r} \le ra + (1-r)b.$$

△

Proof. If $a = 0$ or $b = 0$, then the inequality holds trivially.

Therefore assume that $a, b > 0$. The function $z \mapsto \ln z$, $z > 0$, is concave down; indeed,

$$(\ln z)'' = \left(\frac{1}{z}\right)' = -\frac{1}{z^2} < 0.$$

This means that the segment connecting the points $(a, \ln a)$ and $(b, \ln b)$ lies below the graph of the logarithm function (see Figure 8.1). Thus we have

$$\ln\left(a^r b^{1-r}\right) = r\ln a + (1-r)\ln b \le \ln(ra + (1-r)b),$$

and the desired inequality follows.

□

Corollary 8.3 (*Young's Inequality*) *Let $a, b \geq 0$, and let $1 < p, q < \infty$ be such that $\frac{1}{p} + \frac{1}{q} = 1$. Then*

$$a^{\frac{1}{p}} b^{\frac{1}{q}} \leq \frac{a}{p} + \frac{b}{q}.$$

△

Proof. Take $r = \frac{1}{p}, 1 - r = 1 - \frac{1}{p} = \frac{1}{q}$ in Lemma 8.2.

□

Remark In the particular case $p = q = 2$ we obtain the well-known inequality of arithmetic and geometric means

$$\sqrt{ab} \leq \frac{a+b}{2}.$$

△

Theorem 8.4 (*Hölder's Inequality*) *Let $1 \leq p, q \leq \infty$, $\frac{1}{p} + \frac{1}{q} = 1$. Let $f \in L^p(\Omega, \mathscr{A}, \mu)$ and $g \in L^q(\Omega, \mathscr{A}, \mu)$. Then $fg \in L^1(\Omega, \mathscr{A}, \mu)$ and*

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

△

Proof. A. We first consider the case when $p = 1$, $q = \infty$. (Note that the case when $p = \infty$, $q = 1$ can be directly deduced by interchanging f and g.)

Let

$$N := \left\{x \in \Omega : |g(x)| > \|g\|_\infty\right\}, \quad \mu(N) = 0.$$

We have

$$\int_\Omega |fg| d\mu = \int_{\Omega \setminus N} |f||g| d\mu \leq \|g\|_\infty \int_{\Omega \setminus N} |f| d\mu = \|g\|_\infty \int_\Omega |f| d\mu = \|g\|_\infty \|f\|_1 < \infty.$$

We conclude that $fg \in L^1(\Omega, \mathscr{A}, \mu)$ and $\|fg\|_1 \leq \|f\|_1 \|g\|_\infty$.

B. Now let us consider the case when $1 < p, q < \infty$, $\frac{1}{p} + \frac{1}{q} = 1$.

If $\|f\|_p = 0$, then $f = 0$ μ-a.e. and thus also $fg = 0$ μ-a.e. It follows that $fg \in L^1(\Omega, \mathscr{A}, \mu)$ and $\|fg\|_1 = 0 = \|f\|_p \|g\|_q$. The same argument applies when $\|g\|_q = 0$.

Now assume that $\|f\|_p > 0$, $\|g\|_q > 0$. Let $x \in \Omega$. Corollary 8.3 with $a = \left(\frac{|f(x)|}{\|f\|_p}\right)^p$, $b = \left(\frac{|g(x)|}{\|g\|_q}\right)^q$ yields

$$\frac{|f(x)|}{\|f\|_p} \frac{|g(x)|}{\|g\|_q} \leq \frac{1}{p}\left(\frac{|f(x)|}{\|f\|_p}\right)^p + \frac{1}{q}\left(\frac{|g(x)|}{\|g\|_q}\right)^q.$$

It follows

$$|f(x)g(x)| \leq \|f\|_p \|g\|_q \left(\frac{1}{p}\frac{|f(x)|^p}{\|f\|_p^p} + \frac{1}{q}\frac{|g(x)|^q}{\|g\|_q^q}\right).$$

We integrate this inequality over Ω and obtain

$$\begin{aligned}
\|fg\|_1 = \int_\Omega |f(x)g(x)| d\mu &\leq \|f\|_p \|g\|_q \left(\frac{1}{p}\frac{\int_\Omega |f(x)|^p d\mu}{\|f\|_p^p} + \frac{1}{q}\frac{\int_\Omega |g(x)|^q d\mu}{\|g\|_q^q}\right) \\
&= \|f\|_p \|g\|_q \left(\frac{1}{p}\frac{\|f\|_p^p}{\|f\|_p^p} + \frac{1}{q}\frac{\|g\|_q^q}{\|g\|_q^q}\right) = \|f\|_p \|g\|_q \left(\frac{1}{p} + \frac{1}{q}\right) = \|f\|_p \|g\|_q
\end{aligned}$$

as desired. It follows in particular that $fg \in L^1(\Omega, \mathscr{A}, \mu)$.

□

Theorem 8.5 (*Minkowski's Inequality*) *Let $1 \leq p \leq \infty$, and let $f, g \in L^p(\Omega, \mathscr{A}, \mu)$. Then*

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p.$$

△

Proof. We have shown already that $f + g \in L^p(\Omega, \mathscr{A}, \mu)$, if $f, g \in L^p(\Omega, \mathscr{A}, \mu)$. Moreover, we have shown the inequality in the case when $p = \infty$. It remains to consider $1 \leq p < \infty$.

A. First we consider $p = 1$. We have

$$\|f + g\|_1 = \int_\Omega |f + g| d\mu \leq \int_\Omega (|f| + |g|) d\mu = \int_\Omega |f| d\mu + \int_\Omega |g| d\mu = \|f\|_1 + \|g\|_1.$$

B. Now assume that $1 < p < \infty$. Take $1 < q < \infty$ such that $\frac{1}{p} + \frac{1}{q} = 1$. We have

$$\begin{aligned}
\|f+g\|_p^p &= \int_\Omega |f+g|^p d\mu \leq \int_\Omega (|f|+|g|)|f+g|^{p-1} d\mu \\
&= \int_\Omega |f||f+g|^{p-1} d\mu + \int_\Omega |g||f+g|^{p-1} d\mu \quad \text{(Hölder's inequality)} \\
&\leq \|f\|_p \left\| |f+g|^{p-1} \right\|_q + \|g\|_p \left\| |f+g|^{p-1} \right\|_q \\
&= (\|f\|_p + \|g\|_p) \left(\int_\Omega |f+g|^{(p-1)q} d\mu \right)^{\frac{1}{q}}.
\end{aligned}$$

Since $\frac{1}{p} + \frac{1}{q} = 1$, we have $\frac{1}{q} = 1 - \frac{1}{p} = \frac{p-1}{p}$ and $q = \frac{p}{p-1}$. We continue

$$\begin{aligned}
\|f+g\|_p^p &\leq (\|f\|_p + \|g\|_p) \left(\int_\Omega |f+g|^{(p-1)\frac{p}{p-1}} d\mu \right)^{1-\frac{1}{p}} \\
&= (\|f\|_p + \|g\|_p) \left(\int_\Omega |f+g|^p d\mu \right)^{1-\frac{1}{p}} \\
&= (\|f\|_p + \|g\|_p) \|f+g\|_p^{p\left(1-\frac{1}{p}\right)} = (\|f\|_p + \|g\|_p) \|f+g\|_p^{p-1}.
\end{aligned}$$

Summarizing, we have shown that

$$\|f+g\|_p^p \leq (\|f\|_p + \|g\|_p) \|f+g\|_p^{p-1}.$$

This implies

$$\|f+g\|_p \leq \|f\|_p + \|g\|_p.$$

$\square$

Completeness of L^p-spaces

Definition 8.6 Let $(X, \|\cdot\|)$ be a normed space. A sequence $\{x_n\}_{n\in\mathbb{N}} \subseteq X$ is called a *Cauchy sequence* if

$$\forall\, \varepsilon > 0\ \exists\, N \in \mathbb{N}\ \forall\, n, m \geq N : \|x_n - x_m\| < \varepsilon.$$

$\triangle$

Each convergent sequence is a Cauchy sequence. Indeed, suppose that $x_n \to x \in X$, $n \to \infty$. Take $\varepsilon > 0$. By the definition of the limit, there exists $N \in \mathbb{N}$ such that $\forall\, n \geq N : \|x_n - x\| < \frac{\varepsilon}{2}$. But then $\forall\, n, m \geq N$

$$\|x_n - x_m\| \leq \|x_n - x\| + \|x - x_m\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

However, a Cauchy sequence is not necessarily convergent. Consider, for example, the sequence in $\mathbb{Q}$

$$1.4, 1.41, 1.414, 1.4142, 1.41421, \ldots$$

that converges to $\sqrt{2} \in \mathbb{R} \setminus \mathbb{Q}$. If we consider $X = \mathbb{Q}$, then this sequence does not have a limit, i.e. it is not convergent.

Definition 8.7 (1) A normed space X is called *complete* if each Cauchy sequence in X converges in X.

(2) A complete normed space is called *Banach space*.

△

Example $\mathbb{R}$ is complete, $\mathbb{Q}$ is not complete. △

We will prove that the spaces $L^p(\Omega, \mathscr{A}, \mu)$, $1 \leq p \leq \infty$, are complete, i.e. Banach spaces.

Lemma 8.8 *Let $1 \leq p < \infty$. Let $\{u_j\}_{j\in\mathbb{N}}$ be a sequence in $L^p(\Omega, \mathscr{A}, \mu)$ such that $\sum_{j=1}^{\infty} \|u_j\|_p < \infty$. Put $f_k := \sum_{j=1}^{k} u_j$. Then*

(1) *$f := \lim_{k\to\infty} f_k$ exists μ-a.e.,*

(2) *$f \in L^p(\Omega, \mathscr{A}, \mu)$,*

(3) *$\|f - f_k\|_p \to 0$, $k \to \infty$.*

△

Proof. Consider $g_k := \sum_{j=1}^{k} |u_j|$, $k \in \mathbb{N}$, $g := \sum_{j=1}^{\infty} |u_j|$ with values in $\overline{\mathbb{R}}_+$. We have $0 \leq g_k^p(x) \leq g_{k+1}^p(x)$, $k \in \mathbb{N}$, and $g_k^p(x) \to g^p(x)$, $k \to \infty$, for all $x \in \Omega$. Using the Monotone Convergence Theorem and Minkowski's inequality, we estimate

$$\begin{aligned}\|g\|_p &= \left(\int_\Omega |g|^p d\mu\right)^{\frac{1}{p}} = \left(\int_\Omega \lim_{k\to\infty} |g_k|^p d\mu\right)^{\frac{1}{p}} = \left(\lim_{k\to\infty} \int_\Omega |g_k|^p d\mu\right)^{\frac{1}{p}} \\ &= \lim_{k\to\infty} \|g_k\|_p = \lim_{k\to\infty} \left\| \sum_{j=1}^{k} |u_j| \right\|_p \leq \lim_{k\to\infty} \sum_{j=1}^{k} \|u_j\|_p = \sum_{j=1}^{\infty} \|u_j\|_p < \infty,\end{aligned}$$

and thus the function g^p is integrable. It follows that $N := \{g^p = \infty\} = \{g = \infty\}$ is a null set.

Let $x \in \Omega \setminus N$. The real series $\sum_{j=1}^{\infty} u_j(x)$ is absolutely convergent in $\mathbb{R}$, i.e.

$$\forall\, \varepsilon > 0\ \exists\, K(\varepsilon) \in \mathbb{N}\ \forall\, k \geq K(\varepsilon) : \sum_{j=k+1}^{\infty} |u_j(x)| < \varepsilon.$$

For the sequence $\{f_k(x)\}_{k\in\mathbb{N}}$ we then have the following: For all $\ell > k \geq K(\varepsilon)$

$$\begin{aligned} |f_\ell(x) - f_k(x)| &= \left|\sum_{j=1}^{\ell} u_j(x) - \sum_{j=1}^{k} u_j(x)\right| = \left|\sum_{j=k+1}^{\ell} u_j(x)\right| \\ &\leq \sum_{j=k+1}^{\ell} |u_j(x)| \leq \sum_{j=k+1}^{\infty} |u_j(x)| < \varepsilon. \end{aligned}$$

Thus, $\{f_k(x)\}_{k\in\mathbb{N}}$ for $x \in \Omega \setminus N$ is a Cauchy sequence in $\mathbb{R}$. But $\mathbb{R}$ is complete, and consequently there exists

$$f(x) := \lim_{k\to\infty} f_k(x), \quad x \in \Omega \setminus N.$$

The function f is hereby defined μ-a.e.

Since $|f_k|^p \to |f|^p$, $k \to \infty$, μ-a.e., and $|f_k|^p \leq g^p$ μ-a.e. for all $k \in \mathbb{N}$ (so that g^p is an integrable majorant), by Lebesgue's Dominated Convergence Theorem we obtain that $|f|^p$ is integrable. In other words, $f \in L^p(\Omega, \mathscr{A}, \mu)$.

Moreover, we have that $f - f_k \to 0$, $k \to \infty$, μ-a.e., and for all $k \in \mathbb{N}$

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

μ- a.e., and thus $2^p g^p$ is an integrable majorant. Using Lebesgue's Dominated Convergence Theorem again, we conclude that

$$\lim_{k\to\infty} \int_\Omega |f - f_k|^p d\mu = \int_\Omega 0 d\mu = 0,$$

i.e., $\|f - f_k\|_p \to 0$, $k \to \infty$.

□

Theorem 8.9 (*Riesz-Fischer Theorem*) *Let $1 \leq p \leq \infty$. The space $L^p(\Omega, \mathscr{A}, \mu)$ is complete, i.e. a Banach space.*

△

Proof. A. First let $1 \le p < \infty$. Let $\{f_k\}_{k\in\mathbb{N}} \subseteq L^p(\Omega, \mathscr{A}, \mu)$ be a Cauchy sequence w.r.t. the norm $\|\cdot\|_p$. It is enough to construct a subsequence of $\{f_k\}_{k\in\mathbb{N}}$ that converges in $L^p(\Omega, \mathscr{A}, \mu)$ to a certain function $f \in L^p(\Omega, \mathscr{A}, \mu)$. If this is true, then also the entire sequence converges to f in $L^p(\Omega, \mathscr{A}, \mu)$.

Since $\{f_k\}_{k\in\mathbb{N}}$ is a Cauchy sequence, we may w.l.o.g. assume that

$$\|f_{k+1} - f_k\|_p \le \frac{1}{2^{k+1}}, \quad k \in \mathbb{N}.$$

(If this is not the case for $\{f_k\}_{k\in\mathbb{N}}$, select a suitable subsequence.)

Set $u_1 := f_1$, $u_j := f_j - f_{j-1}$ for $j = 2, 3, \ldots$. Then $f_k = \sum_{j=1}^{k} u_j$, $k \in \mathbb{N}$. Since $\|u_1\|_p = \|f_1\|_p$ and for $j = 2, 3, \ldots$

$$\|u_j\|_p = \|f_j - f_{j-1}\|_p \le \frac{1}{2^j},$$

we see that $u_j \in L^p(\Omega, \mathscr{A}, \mu)$, $j \in \mathbb{N}$, and

$$\sum_{j=1}^{\infty} \|u_j\|_p \le \|f_1\|_p + \sum_{j=2}^{\infty} \frac{1}{2^j} < \infty.$$

Therefore the assumptions of Lemma 8.8 are fulfilled in our case, and we conclude that $\{f_k\}_{k\in\mathbb{N}}$ converges in $L^p(\Omega, \mathscr{A}, \mu)$ as well as pointwisely μ-a.e. to a function $f \in L^p(\Omega, \mathscr{A}, \mu)$. This proves the statement for $1 \le p < \infty$.

B. Now let us consider the case $p = \infty$. Let $\{f_k\}_{k\in\mathbb{N}}$ be a Cauchy sequence in $L^\infty(\Omega, \mathscr{A}, \mu)$. This means that

$$\forall\, \varepsilon > 0\ \exists\, K(\varepsilon) \in \mathbb{N}\ \forall\, k, \ell \ge K(\varepsilon) : \|f_k - f_\ell\|_\infty < \varepsilon.$$

We have

$$\|f_k\|_\infty = \|f_k - f_\ell + f_\ell\|_\infty \le \|f_k - f_\ell\|_\infty + \|f_\ell\|_\infty,$$

and consequently

$$\|f_k\|_\infty - \|f_\ell\|_\infty \le \|f_k - f_\ell\|_\infty.$$

Interchanging k and ℓ gives us the inequality $\|f_\ell\|_\infty - \|f_k\|_\infty \le \|f_k - f_\ell\|_\infty$, and altogether we obtain

$$\big|\|f_k\|_\infty - \|f_\ell\|_\infty\big| \le \|f_k - f_\ell\|_\infty < \varepsilon, \quad k, \ell \ge K(\varepsilon).$$

This shows that $\{\|f_k\|_\infty\}_{k\in\mathbb{N}}$ is a Cauchy sequence in $\mathbb{R}$. Since $\mathbb{R}$ is complete, there exists a finite limit $\lim_{k\to\infty} \|f_k\|_\infty \in \mathbb{R}$.

By the definition of the norm $\|\cdot\|_\infty$, the sets

$$N_k := \left\{|f_k| > \|f_k\|_\infty\right\}, \quad k \in \mathbb{N},$$

and

$$\widetilde{N}_{k,\ell} := \left\{|f_k - f_\ell| > \|f_k - f_\ell\|_\infty\right\}, \quad k, \ell \in \mathbb{N},$$

are null sets. This implies that also $N := \left(\bigcup_{k=1}^\infty N_k\right) \cup \left(\bigcup_{k=1}^\infty \bigcup_{\ell=1}^\infty \widetilde{N}_{k,\ell}\right)$ is a null set.

For $x \in \Omega \setminus N$ we have

$$|f_k(x) - f_\ell(x)| \le \|f_k - f_\ell\|_\infty < \varepsilon, \quad k, \ell \ge K(\varepsilon),$$

and thus $\{f_k(x)\}_{k\in\mathbb{N}}$ is a Cauchy sequence in $\mathbb{R}$. It follows that there exists

$$f(x) := \lim_{k\to\infty} f_k(x), \quad x \in \Omega \setminus N.$$

In particular, the function f is defined μ-a.e. For $x \in \Omega \setminus N$ we have

$$|f(x)| \le \lim_{k\to\infty} |f_k(x)| \le \lim_{k\to\infty} \|f_k\|_\infty \in \mathbb{R},$$

and thus $f \in L^\infty(\Omega, \mathscr{A}, \mu)$. Moreover, for all $x \in \Omega \setminus N$ we have

$$|f_k(x) - f(x)| \le \lim_{\ell\to\infty} |f_k(x) - f_\ell(x)| \le \lim_{\ell\to\infty} \|f_k - f_\ell\|_\infty \le \varepsilon, \quad k \ge K(\varepsilon),$$

and thus $\|f_k - f\|_\infty \to 0$, $k \to \infty$.

□

Notice that, for $p = \infty$, the sequence $\{f_k\}_{k\in\mathbb{N}}$ converges pointwisely μ-a.e. to f. This is, in general, not the case if $1 \le p < \infty$. However, the following statement is a part of the proof of the theorem in the case $1 \le p < \infty$. It is an important fact on its own.

Corollary 8.10 *Let $1 \le p < \infty$, and let $\{f_k\}_{k\in\mathbb{N}}$ be a sequence in $L^p(\Omega, \mathscr{A}, \mu)$ that converges in $L^p(\Omega, \mathscr{A}, \mu)$ to a function $f \in L^p(\Omega, \mathscr{A}, \mu)$. Then there exists a subsequence $\{f_{k_m}\}_{m\in\mathbb{N}}$ that converges to f pointwisely μ-a.e.*

△

Chapter 9

Signed measures

Let $(\Omega, \mathscr{A}, \mu)$ be a measure space. Let $f : \Omega \to \overline{\mathbb{R}}_+$ be a measurable function. We proved in Theorem 5.10 that in this case the set function

$$\nu_f(A) := \int_A f d\mu, \quad A \in \mathscr{A},$$

is a measure on $(\Omega, \mathscr{A})$.

Now we want to consider the following generalization. Take an integrable function $f : \Omega \to \overline{\mathbb{R}}$ and consider the set function $\nu_f : \mathscr{A} \to \overline{\mathbb{R}}$,

$$\nu_f(A) := \int_A f d\mu, \quad A \in \mathscr{A}.$$

Clearly, $\nu_f(\varnothing) = 0$. If $\{A_i\}_{i \in \mathbb{N}} \subseteq \mathscr{A}$ is a sequence of p.d. sets, then

$$\int_{\sum_{i=1}^{\infty} A_i} f d\mu = \sum_{i=1}^{\infty} \int_{A_i} f d\mu, \tag{$*$}$$

and the series on the right-hand side is absolutely convergent. Indeed, in Theorem 5.10 we proved that $(*)$ holds for $f \geq 0$. This implies that for an integrable $f : \Omega \to \overline{\mathbb{R}}$ we have

$$\sum_{i=1}^{\infty} \left| \int_{A_i} f d\mu \right| \leq \sum_{i=1}^{\infty} \int_{A_i} |f| d\mu = \int_{\sum_{i=1}^{\infty} A_i} |f| d\mu \leq \int_{\Omega} |f| d\mu < \infty,$$

and thus the series $\sum_{i=1}^{\infty} \int_{A_i} f d\mu$ converges absolutely. To prove $(*)$, apply it to f^+ and f^-.

Summarizing, $\nu_f : \mathscr{A} \to \overline{\mathbb{R}}$ is a σ-additive set function with $\nu_f(\varnothing) = 0$. However, it is not necessarily monotone: It can happen that $A \subseteq B$, but $\nu_f(A) > \nu_f(B)$. We will term such set functions signed measures.

Definition 9.1 Let $(\Omega, \mathscr{A})$ be a measurable space. A *finite signed measure* (*real measure*) is a set function $\mu : \mathscr{A} \to \mathbb{R}$ such that $\mu(\varnothing) = 0$ and μ is σ-additive.

△

Note that a finite signed measure has the property

$$-\infty < \mu(A) < \infty \quad \forall\, A \in \mathscr{A}.$$

Since there is no monotonicity, it is not enough to assume that $-\infty < \mu(\Omega) < \infty$.

Example If μ_1, μ_2 are two finite measures on $(\Omega, \mathscr{A})$, then $\mu = \mu_1 - \mu_2$ is a finite signed measure on $(\Omega, \mathscr{A})$.

△

Recall that for (non-negative) measures we also allowed the value $+\infty$. The situation becomes more complicated for signed measures because the expression $\infty - \infty$ is not defined. Thus we cannot work with set functions that can take both values ∞ and $-\infty$.

Definition 9.2 Let $(\Omega, \mathscr{A})$ be a measurable space. An *extended (extended real) signed measure* is a set function $\mu : \mathscr{A} \to \mathbb{R} \cup \{\infty\}$ or $\mu : \mathscr{A} \to \mathbb{R} \cup \{-\infty\}$ such that $\mu(\varnothing) = 0$ and μ is σ-additive.

△

Of course, a finite signed measure is a particular case of an extended signed measure. In this course, the expression "signed measure" always means an extended signed measure.

Definition 9.3 Let μ be a signed measure on a measurable space $(\Omega, \mathscr{A})$.

(1) A set $E \in \mathscr{A}$ is called a *positive set* for μ if $\mu(A) \geq 0$ for all $A \in \mathscr{A}$, $A \subseteq E$.

(2) A set $E \in \mathscr{A}$ is called a *negative set* for μ if $\mu(A) \leq 0$ for all $A \in \mathscr{A}$, $A \subseteq E$.

(3) A set $E \in \mathscr{A}$ is called a *μ-null set* if E is both positive and negative.

△

If E is a positive (negative) set for μ, then $\mu(E) \geq 0$ ($\mu(E) \leq 0$). If E is a μ-null set, then $\mu(E) = 0$. However, the inverse implications do not hold.

If μ is a (non-negative) measure on Ω, then any set $E \in \mathscr{A}$ is a positive set for μ and only null sets are negative sets.

The following proposition lists simple properties of positive and negative sets.

Proposition 9.4 (1) *If E is a positive set for μ, then μ is monotone on E, i.e.*

$$A \subseteq B \subseteq E \implies 0 \le \mu(A) \le \mu(B) \le \mu(E)$$

for $A, B \in \mathscr{A}$. Similarly, if E is a negative set for μ, then for $A, B \in \mathscr{A}$

$$A \subseteq B \subseteq E \implies \mu(E) \le \mu(B) \le \mu(A) \le 0.$$

(2) *Any measurable subset of a positive (negative) set is a positive (negative) set.*

(3) *Countable unions of positive (negative) sets are positive (negative) sets.*

△

Proof. Homework. □

Theorem 9.5 *Let μ be a signed measure on a measurable space $(\Omega, \mathscr{A})$. Then:*

(1) *μ is continuous from below, i.e. if $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathscr{A}$ is such that $A_n \subseteq A_{n+1}$, $n \in \mathbb{N}$, then*

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

(2) *μ is continuous from above, i.e. if $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathscr{A}$ is such that $A_n \supseteq A_{n+1}$, $n \in \mathbb{N}$, and $-\infty < \mu(A_1) < \infty$, then*

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

△

Proof. (1) Since $\bigcup_{n=1}^{\infty} A_n = A_1 + \sum_{n=2}^{\infty}(A_n \setminus A_{n-1})$ and μ is σ-additive, we have

$$\begin{aligned}
\mu\left(\bigcup_{n=1}^{\infty} A_n\right) &= \mu(A_1) + \sum_{n=2}^{\infty} \mu(A_n \setminus A_{n-1}) \\
&= \lim_{N\to\infty}\left(\mu(A_1) + \sum_{n=2}^{N} \mu(A_n \setminus A_{n-1})\right) \\
&= \lim_{N\to\infty} \mu\left(A_1 + \sum_{n=2}^{N}(A_n \setminus A_{n-1})\right) \\
&= \lim_{N\to\infty} \mu(A_N).
\end{aligned}$$

(2) Put $A := \bigcap_{n=1}^{\infty} A_n$, $B_n := A_1 \setminus A_n$, $n \in \mathbb{N}$. We have $B_n \subseteq B_{n+1}$, $n \in \mathbb{N}$, and

$$\bigcup_{n=1}^{\infty} B_n = \bigcup_{n=1}^{\infty} (A_1 \setminus A_n) = A_1 \setminus \left(\bigcap_{n=1}^{\infty} A_n \right) = A_1 \setminus A.$$

By the first part we have

$$\lim_{n\to\infty} \mu(A_1 \setminus A_n) = \mu(A_1 \setminus A).$$

Now $A_1 = (A_1 \setminus A_n) + A_n$ implies

$$\mu(A_1) = \mu(A_1 \setminus A_n) + \mu(A_n).$$

By assumption, $\mu(A_1) \in (-\infty, \infty)$. This implies that also $\mu(A_1 \setminus A_n), \mu(A_n) \in (-\infty, \infty)$. Indeed, if $\mu(A_1 \setminus A_n) = \infty$, then

$$\mu(A_n) = \underbrace{\mu(A_1)}_{\in\mathbb{R}} - \underbrace{\mu(A_1 \setminus A_n)}_{=\infty} = -\infty,$$

and this contradicts the assumption that μ cannot attain both values ∞ and $-\infty$. Similarly, if $\mu(A_1 \setminus A_n) = -\infty$, then

$$\mu(A_n) = \underbrace{\mu(A_1)}_{\in\mathbb{R}} - \underbrace{\mu(A_1 \setminus A_n)}_{=-\infty} = \infty,$$

which is again impossible by the same reason. The same consideration shows that $\mu(A_n)$ cannot be ∞ or $-\infty$.

We also have $A_1 = (A_1 \setminus A) + A$ and thus $\mu(A_1) = \mu(A_1 \setminus A) + \mu(A)$. The same arguments as above show that $\mu(A_1 \setminus A), \mu(A) \in (-\infty, \infty)$.

Since all the values here are finite, we may finish the proof as follows:

$$\begin{aligned}\lim_{n\to\infty} \mu(A_n) &= \lim_{n\to\infty} (\mu(A_1) - \mu(A_1 \setminus A_n)) = \mu(A_1) - \lim_{n\to\infty} \mu(A_1 \setminus A_n) \\ &= \mu(A_1) - \mu(A_1 \setminus A) = \mu(A).\end{aligned}$$

□

Lemma 9.6 *Let μ be a signed measure on a measurable space $(\Omega, \mathscr{A})$. Suppose that there exists a set $B \in \mathscr{A}$ such that $-\infty < \mu(B) < 0$. Then there is a set $E \in \mathscr{A}$, $E \subseteq B$ such that E is a negative set for μ and $-\infty < \mu(E) \leq \mu(B) < 0$.*

△

Proof. Put
$$r_1 := \sup\{\mu(A) : A \in \mathscr{A},\ A \subseteq B\}.$$
If $r_1 \leq 0$, then B itself is a negative set for μ and we are done.

Therefore assume that $r_1 > 0$ (note that it is possible that $r_1 = \infty$). By the construction of r_1, there is a set $C_1 \in \mathscr{A}$, $C_1 \subseteq B$ such that

$$\mu(C_1) \geq \min\left\{\frac{1}{2}r_1, 1\right\} > 0.$$

Put $B_1 := B \setminus C_1$. We have $\mu(B) = \mu(B_1) + \mu(C_1)$. Since $-\infty < \mu(B) < 0$, $\mu(C_1) > 0$, we can subtract and obtain

$$\mu(B_1) = \underbrace{\mu(B)}_{\in\mathbb{R}} - \underbrace{\mu(C_1)}_{>0} < \mu(B).$$

On the other hand, $\mu(C_1) = \infty$ would imply

$$\mu(B_1) = \underbrace{\mu(B)}_{\in\mathbb{R}} - \underbrace{\mu(C_1)}_{=\infty} = -\infty,$$

which is impossible because μ cannot attain both values ∞ and $-\infty$. We conclude that
$$-\infty < \mu(B_1) < \mu(B) < 0.$$

Now put
$$r_2 := \sup\{\mu(A) : A \in \mathscr{A},\ A \subseteq B_1\}.$$
If $r_2 \leq 0$, then B_1 is a negative set for μ and we can take $E = B_1$.

Otherwise we have $r_2 > 0$. As above, take $C_2 \in \mathscr{A}$, $C_2 \subseteq B_1$ such that

$$\mu(C_2) \geq \min\left\{\frac{1}{2}r_2, 1\right\} > 0.$$

and put $B_2 := B_1 \setminus C_2$. We conclude by the same argument that

$$-\infty < \mu(B_2) < \mu(B_1) < 0.$$

Continue in the same manner. If at the k-th step we have $r_k \leq 0$, then $E = B_{k-1}$ is the desired set. If $r_k > 0$ for all $k \in \mathbb{N}$, we obtain a sequence of sets $\{B_k\}_{k\in\mathbb{N}} \subseteq \mathscr{A}$, $B_{k+1} \subseteq B_k$, $-\infty < \mu(B_{k+1}) < \mu(B_k) < 0$, $k \in \mathbb{N}$. Put $E := \bigcap_{k=1}^{\infty} B_k$.

It follows immediately that $E \in \mathscr{A}$ and $E \subseteq B$. From the continuity of μ from above it follows that

$$\mu(E) = \lim_{k\to\infty} \mu(B_k) \leq \mu(B) < 0.$$

Since $B = E + (B \setminus E)$ and thus $\mu(B) = \mu(E) + \mu(B \setminus E)$, the assumption $\mu(E) = -\infty$ would lead to

$$\mu(B \setminus E) = \underbrace{\mu(B)}_{\in\mathbb{R}} - \underbrace{\mu(E)}_{=-\infty} = \infty,$$

which is impossible. This proves

$$-\infty < \mu(E) \leq \mu(B) < 0.$$

It remains to show that E is a negative set for μ. The sets $C_k = B_{k-1} \setminus B_k$, $k = 2, 3, \dots$, $C_1 = B \setminus B_1$ are pairwise disjoint by construction. We have

$$\sum_{k=1}^{n} \mu(C_k) = \mu\left(\sum_{k=1}^{n} C_k\right) = \mu(B \setminus B_n) = \mu(B) - \mu(B_n) \to \mu(B) - \mu(E) \in \mathbb{R}, \quad n \to \infty,$$

i.e. the series $\sum_{k=1}^{n} \mu(C_k)$ is convergent in $\mathbb{R}$. It follows that $\mu(C_k) \to 0$, $k \to \infty$. Since $\mu(C_k) \geq \min\left\{\frac{1}{2} r_k, 1\right\} > 0$, this implies that $r_k \to 0$, $k \to \infty$.

Now let $A \in \mathscr{A}$, $A \subseteq E$. Since $E = \bigcap_{k=1}^{\infty} B_k$, we have $A \subseteq B_k$ for all $k \in \mathbb{N}$, and thus $\mu(A) \leq r_k \to 0$, $k \to \infty$. This implies $\mu(A) \leq 0$. Since this is true for any measurable $A \subseteq E$, we conclude that E is a negative set for μ.

□

Theorem 9.7 (*Hahn Decomposition Theorem*) *Let μ be a signed measure on a measurable space $(\Omega, \mathscr{A})$. Then there exist two sets $P, N \in \mathscr{A}$ such that*

(1) *P is a positive set for μ,*

(2) *N is a negative set for μ,*

(3) *$\Omega = P \cup N$, $P \cap N = \varnothing$.*

Moreover, this decomposition is essentially unique, i.e. if (P', N') is another pair of such sets, then the symmetric differences $P \triangle P'$, $N \triangle N'$ are μ-null sets.

△

Definition 9.8 A pair (P, N) from Theorem 9.7 is called a *Hahn decomposition* of the signed measure μ.

$\triangle$

Proof of Theorem 9.7. A. First we show that a Hahn decomposition is essentially unique. Let (P, N) and (P', N') be two Hahn decompositions of μ. Since $P \setminus P' \subseteq P$, it is a positive set for μ. On the other hand, since $P' \cap N' = \varnothing$, $P \setminus P' \subseteq N'$ and thus it is a negative set for μ. It follows that $P \setminus P'$ is a μ-null set. Similarly, $P' \setminus P$ is a μ-null set, and thus also $P \triangle P = (P \setminus P') \cup (P' \setminus P)$ is a μ-null set.

The same argument shows that $N \triangle N'$ is a μ-null set.

B. Now let us show that a Hahn decomposition of μ exists.

Without loss of generality we may assume that $-\infty < \mu(A) \leq \infty$ for all $A \in \mathscr{A}$ (otherwise we consider $-\mu$ instead of μ). Put

$$\alpha := \inf\{\mu(E) : E \in \mathscr{A},\ E \text{ is a negative set for } \mu\}.$$

Since $\varnothing$ is a negative set for μ, we have $\alpha \leq \mu(\varnothing) = 0$.

Let $\{N_k\}_{k\in\mathbb{N}}$ be a sequence of negative sets such that $\alpha = \lim_{k\to\infty} \mu(N_k)$. Put $N := \bigcup_{k=1}^{\infty} N_k$. N is a negative set for μ as a countable union of negative sets. Then, by the definition of α, we have $\alpha \leq \mu(N)$. For each $k \in \mathbb{N}$ we have

$$\mu(N) = \mu(N_k) + \underbrace{\mu(N \setminus N_k)}_{\leq 0, \text{ since } N \text{ is a negative set}} \leq \mu(N_k).$$

Altogether we have

$$\alpha \leq \mu(N) \leq \mu(N_k) \to \alpha, \quad k \to \infty,$$

and it follows that $\mu(N) = \alpha$. Moreover, $\alpha = \mu(N) > -\infty$, as μ does not attain the value $-\infty$.

Now put $P := \Omega \setminus N$. If we show that P is a positive set for μ, we are done.

Assume the opposite. Then there is a set $B \in \mathscr{A}$, $B \subseteq P$ such that $\mu(B) < 0$. Since μ does not take the value $-\infty$, we have $-\infty < \mu(B) < 0$. By Lemma 9.6 there exists a set $E \in \mathscr{A}$, $E \subseteq B$ with $-\infty < \mu(E) \leq \mu(B) < 0$ that is a negative set for μ.

Consider now that set $N \cup E$. Since N and E are negative sets for μ, also $N \cup E$ is a negative set. This implies $\alpha \leq \mu(N \cup E)$.

On the other hand, $E \subseteq B \subseteq P$ and $N \cap P = \emptyset$, so that $N \cap E = \emptyset$. This implies

$$\mu(N \cup E) = \mu(N) + \mu(E) = \alpha + \underbrace{\mu(E)}_{<0} < \alpha,$$

and this contradicts the definition of α. The contradiction shows that our assumption was wrong and P is a positive set for μ.

□

Theorem 9.9 (*Jordan Decomposition Theorem*) *Let μ be a signed measure on a measurable space $(\Omega, \mathscr{A})$. Then μ can be represented as*

$$\mu = \mu_1 - \mu_2,$$

where μ_1, μ_2 are (non-negative) measures on $(\Omega, \mathscr{A})$, and at least one of these measures is finite. We can take

$$\mu_1(A) = \mu^+(A) := \mu(A \cap P), \quad \mu_2(A) = \mu^-(A) := -\mu(A \cap N), \quad A \in \mathscr{A},$$

where (P, N) is a Hahn decomposition of μ.

△

Note that $\mu^+(A)$, $\mu^-(A)$ as defined above do not depend on the choice of the particular Hahn decomposition of μ. Indeed, if (P', N') is another Hahn decomposition of μ, then

$$A \cap P' = (A \cap P' \cap P) + (A \cap P' \cap \overline{P}),$$

where $A \cap P' \cap \overline{P} \subseteq P' \cap \overline{P} = P' \setminus P$ which is a μ-null set. Thus,

$$\mu(A \cap P') = \mu(A \cap P' \cap P) + \underbrace{\mu(A \cap P' \cap \overline{P})}_{=0} = \mu(A \cap P' \cap P).$$

By the same argument, $\mu(A \cap P) = \mu(A \cap P' \cap P)$, and it follows that $\mu(A \cap P) = \mu(A \cap P')$. In the same way one can prove that $\mu(A \cap N) = \mu(A \cap N')$.

Definition 9.10 The measure

$$\mu^+(A) = \mu(A \cap P), \quad A \in \mathscr{A},$$

is called the *positive part* of μ. The measure

$$\mu^-(A) = -\mu(A \cap N), \quad A \in \mathscr{A},$$

is called the *negative part* if μ. The representation

$$\mu = \mu^+ - \mu^-$$

is called the *Jordan decomposition* of μ. The measure

$$|\mu| = \mu^+ + \mu^-$$

is called the *total variation* of μ.

△

Proof of Theorem 9.9. It is easy to see that μ^+, μ^- as defined above are (non-negative) measures. Indeed, $\mathscr{A}$ is a σ-algebra, $\mu^+(\varnothing) = \mu^-(\varnothing) = 0$, $\mu^+, \mu^- \geq 0$, and μ^+, μ^- are σ-additive.

Since for any $A \in \mathscr{A}$ we have

$$\mu(A) = \mu(A \cap P) + \mu(A \cap N) = \mu^+(A) - \mu^-(A),$$

the representation $\mu = \mu^+ - \mu^-$ holds.

It remains to show that at least one of the measures μ^+, μ^- is finite. Suppose, by contrary, that there are sets $A, B \in \mathscr{A}$ such that $\mu^+(A) = \infty$, $\mu^-(B) = \infty$. But then $\mu(A \cap P) = \mu^+(A) = \infty$, $\mu(B \cap N) = -\mu^-(B) = -\infty$, and thus μ takes both values ∞, $-\infty$, which is impossible. This shows that at least one of the measures μ^+, μ^- is finite.

□

Absolute continuity and the Radon-Nikodym Theorem

Definition 9.11 Let ν be a signed measure on a measure space $(\Omega, \mathscr{A}, \mu)$. We say that ν is *absolutely continuous* w.r.t. μ, if

$$\forall A \in \mathscr{A} : \mu(A) = 0 \implies \nu(A) = 0.$$

We write $\nu \ll \mu$.

△

Example Consider the measure space $(\mathbb{R}, \mathscr{B}, \lambda)$.

A. Let f be a λ-integrable function on $\mathbb{R}$. Then

$$\nu_f(A) = \int_A f \, d\lambda, \quad A \in \mathscr{B},$$

is absolutely continuous w.r.t. λ. Indeed, if $\lambda(A) = 0$, the integral is taken over a null set, and so $\nu_f(A) = 0$.

B. The Dirac measure δ_x, $x \in \mathbb{R}$, is not absolutely continuous w.r.t. λ. Indeed, $\lambda(\{x\}) = 0$, but $\delta_x(\{x\}) = 1$.

△

We will show that under reasonable conditions all absolutely continuous signed measures have the form $\int_A f d\mu$ with a certain function f.

Proposition 9.12 *Let ν be a signed measure on a measure space $(\Omega, \mathscr{A}, \mu)$. Then*

$$\nu \ll \mu \iff (\nu^+ \ll \mu \text{ and } \nu^- \ll \mu).$$

△

Proof. Homework. □

Theorem 9.13 (*The Radon-Nikodym Theorem*) *Let $(\Omega, \mathscr{A}, \mu)$ be a σ-finite measure space, and let ν be a finite signed measure on $(\Omega, \mathscr{A})$. The following properties are equivalent:*

(1) $\nu \ll \mu$,

(2) *there exists a μ-integrable function $f : \Omega \to \overline{\mathbb{R}}$ such that*

$$\nu(A) = \int_A f d\mu, \quad A \in \mathscr{A}.$$

The function f is essentially unique in the sense that if $\nu(A) = \int_A g d\mu$ for all $A \in \mathscr{A}$, then $f = g$ μ-a.e.

If ν is a measure, then f can be taken non-negative.

△

Proof. We already have discussed the implication (2) $\implies$ (1).

Let us prove that (1) $\implies$ (2). Suppose that $\nu \ll \mu$. Since $\nu = \nu^+ - \nu^-$ and $\nu^+ \ll \mu$, $\nu^- \ll \mu$, we can consider ν^+, ν^- separately. This means that it is enough to prove the existence of f in the case when ν is a (non-negative) finite measure.

A. First we consider the case when μ is finite.

Put

$$G := \left\{ g : \Omega \to \overline{\mathbb{R}}_+ : g \text{ is } \mu\text{-integrable}, \int_A g d\mu \le \nu(A) \ \forall\, A \in \mathscr{A} \right\}.$$

We have $G \neq \emptyset$ because $g \equiv 0 \in G$.

The set G has the following property: If $g_1, g_2 \in G$, then also $\max(g_1, g_2) \in G$. Indeed, for any $A \in \mathscr{A}$ we have

$$\begin{aligned}\int_A \max(g_1, g_2) d\mu &= \int_{A\cap\{g_1 \le g_2\}} g_2 d\mu + \int_{A\cap\{g_1 > g_2\}} g_1 d\mu \\ &\le \nu(A \cap \{g_1 \le g_2\}) + \nu(A \cap \{g_1 > g_2\}) = \nu(A).\end{aligned}$$

Set

$$\gamma := \sup\left\{\int_\Omega g d\mu : g \in G\right\} \le \nu(\Omega) < \infty.$$

By the definition of γ, there exists a sequence $\{\tilde{g}_n\}_{n\in\mathbb{N}} \subseteq G$ such that

$$\gamma = \lim_{n\to\infty} \int_\Omega \tilde{g}_n d\mu.$$

Set $g_n := \max(\tilde{g}_1, \dots, \tilde{g}_n)$. Then the functions g_n are measurable, $\int_\Omega g_n d\mu \to \gamma$, $n \to \infty$, $g_n \ge 0$, $g_n \in G$, $g_{n+1} \ge g_n$, $n \in \mathbb{N}$.

Now put $f := \lim_{n\to\infty} g_n$. It follows from above that f is measurable and $f \ge 0$. Moreover, by the Monotone Convergence Theorem we have for any $A \in \mathscr{A}$

$$\int_A f d\mu = \lim_{n\to\infty} \int_A g_n d\mu \le \nu(A),$$

and in particular $\int_\Omega f d\mu \le \nu(\Omega) < \infty$, so that f is μ-integrable and $f \in G$. Furthermore, we have

$$\int_\Omega f d\mu = \lim_{n\to\infty} \int_\Omega g_n d\mu = \gamma.$$

We will show that f is the desired function.

Consider the set function

$$\tau(A) := \nu(A) - \int_A f d\mu \ge 0, \quad A \in \mathscr{A}.$$

Clearly, τ is σ-additive and $\tau(\emptyset) = 0$, thus τ is a finite measure on $\mathscr{A}$.

We will show that $\tau(\Omega) = 0$. Note that this implies that $\tau(A) = 0$ for any $A \in \mathscr{A}$, which means that $\nu(A) = \int_A f d\mu$ and, finally, that f is the desired function.

Assume by contrary that $\tau(\Omega) > 0$. Since $\mu(\Omega) < \infty$, we can find $\varepsilon > 0$ such that

$$\tau(\Omega) > \varepsilon\mu(\Omega).$$

Consider the finite signed measure $\tau - \varepsilon\mu$. Let (P, N) be a Hahn decomposition for $\tau - \varepsilon\mu$. Then for each $A \in \mathscr{A}$

$$(\tau - \varepsilon\mu)(A \cap P) \geq 0,$$

i.e. $\tau(A \cap P) \geq \varepsilon\mu(A \cap P)$.

It holds $\mu(P) > 0$. Indeed, if $\mu(P) = 0$, then also $\nu(P) = 0$ (this is because $\nu \ll \mu$), so that $\tau(P) = \nu(P) - \int_P f d\mu = 0$, and finally $(\tau - \varepsilon\mu)(P) = 0$. But this would imply

$$(\tau - \varepsilon\mu)(\Omega) = \underbrace{(\tau - \varepsilon\mu)(P)}_{=0} + \underbrace{(\tau - \varepsilon\mu)(N)}_{\leq 0} \leq 0,$$

which contradicts $\tau(\Omega) > \varepsilon\mu(\Omega)$. Thus, $\mu(P) > 0$.

For each $A \in \mathscr{A}$ we have

$$\int_A \varepsilon\chi_P d\mu = \varepsilon\mu(A \cap P) \leq \tau(A \cap P) \leq \tau(A) = \nu(A) - \int_A f d\mu,$$

so that

$$\int_A \left(f + \varepsilon\chi_P\right) d\mu \leq \nu(A).$$

It follows that $f + \varepsilon\chi_P \in G$. But then

$$\int_\Omega \left(f + \varepsilon\chi_P\right) d\mu = \int_\Omega f d\mu + \varepsilon\mu(P) = \gamma + \underbrace{\varepsilon\mu(P)}_{>0} > \gamma,$$

and this is a contradiction to the definition of γ. The contradiction shows that $\tau(\Omega) = 0$ and thus f is the desired function.

B. Now assume that μ is σ-finite.

In this case there exist sets $\{\Omega_n\}_{n\in\mathbb{N}} \subseteq \mathscr{A}$, p.d., such that $\Omega = \sum_{n=1}^\infty \Omega_n$ and $\mu(\Omega_n) < \infty$, $n \in \mathbb{N}$. For $n \in \mathbb{N}$ and $A \in \mathscr{A}$ set

$$\mu_n(A) := \mu(A \cap \Omega_n), \quad \nu_n(A) := \nu(A \cap \Omega_n).$$

Clearly, μ_n and ν_n are finite measures. If $\mu_n(A) = \mu(A \cap \Omega_n) = 0$, then, since $\nu \ll \mu$, we have $\nu_n(A) = \nu(A \cap \Omega_n) = 0$. This shows that $\nu_n \ll \mu_n$.

By the first part of the proof there exist functions $f_n : \Omega \to \overline{\mathbb{R}}_+$ such that f_n is μ_n-integrable, $f_n \geq 0$ and $\nu_n(A) = \int_A f_n d\mu_n$ for all $A \in \mathscr{A}$. Put $f := \sum_{n=1}^{\infty} f_n \chi_{\Omega_n} : \Omega \to \overline{\mathbb{R}}_+$. This is a non-negative measurable function and for any $A \in \mathscr{A}$ we have, using the Monotone Convergence Theorem to justify the first step,

$$\begin{aligned}\int_A f d\mu &= \sum_{n=1}^{\infty} \int_A f_n \chi_{\Omega_n} d\mu = \sum_{n=1}^{\infty} \int_{A\cap\Omega_n} f_n d\mu_n \\ &= \sum_{n=1}^{\infty} \nu_n(A \cap \Omega_n) = \sum_{n=1}^{\infty} \nu(A \cap \Omega_n) = \nu(A).\end{aligned}$$

In particular, $\int_\Omega f d\mu = \nu(\Omega) < \infty$, so that f is μ-integrable. We see that f is the desired function.

c. Now let ν be a finite signed measure, $\nu \ll \mu$. Let f and g be two functions such that

$$\nu(A) = \int_A f d\mu = \int_A g d\mu \quad \text{for all } A \in \mathscr{A}.$$

For the set $\{f > g\}$ we obtain

$$\int_{\{f>g\}} \underbrace{(f-g)}_{>0} d\mu = \nu(\{f > g\}) - \nu(\{f > g\}) = 0,$$

which implies that $\mu(\{f > g\}) = 0$. Similarly, $\mu(\{g > f\}) = 0$. It follows that $f = g$ μ-a.e.

□

Remark The Radon-Nikodym Theorem can be extended to the case when ν is not necessarily finite. For example, if μ is a σ-finite (non-negative) measure, ν is any (non-negative) measure and $\nu \ll \mu$, then there exists a measurable function $f : \Omega \to \overline{\mathbb{R}}_+$ such that

$$\nu(A) = \int_A f d\mu, \quad A \in \mathscr{A}.$$

The function f is called the *Radon-Nikodym derivative* of ν with respect to μ. The notation is $f = \frac{d\nu}{d\mu}$.

△

The dual space of $L^p(\Omega, \mathscr{A}, \mu)$

Definition 9.14 Let X be a vector space over the field $\mathbb{F}$. A *linear functional* on X is a mapping $\varphi : X \to \mathbb{F}$ such that

$$\varphi(\alpha x + \beta y) = \alpha\varphi(x) + \beta\varphi(y) \quad \forall\, x, y \in X,\ \alpha, \beta \in \mathbb{F}.$$

△

Example A. Let $X = \mathbb{R}^d$. For a fixed $a = (a_1, \dots, a_d)^T \in \mathbb{R}^d$ the mapping

$$\varphi(x) = a \cdot x = \sum_{i=1}^{d} a_i x_i, \quad x = (x_1, \dots x_d)^T \in \mathbb{R}^d,$$

is a linear functional on $\mathbb{R}^d$.

B. Let $C([a,b])$ denote the vector space of continuous functions on $[a,b]$, $a, b \in \mathbb{R}$. Then

$$\varphi(f) = \int_a^b f\,dx$$

is a linear functional on $C([a,b])$.

△

In what follows we will assume that X is a normed space over $\mathbb{R}$ with the norm $\|\cdot\|$.

Definition 9.15 Let $(X, \|\cdot\|)$ be a normed space over $\mathbb{R}$. A linear functional $\varphi : X \to \mathbb{R}$ is called *continuous at $x_0 \in X$*, if

$$\forall\, \varepsilon > 0\, \exists\, \delta > 0\, \forall\, x \in X : \|x - x_0\| < \delta \implies |\varphi(x) - \varphi(x_0)| < \varepsilon.$$

A linear functional φ is called *continuous*, if it is continuous at each point $x_0 \in X$.

△

Definition 9.16 A linear functional $\varphi : X \to \mathbb{R}$ is called *bounded*, if there exists a number $M \geq 0$ such that

$$|\varphi(x)| \leq M\|x\| \quad \forall\, x \in X.$$

△

It is not difficult to prove the following statement (we will not discuss the details).

Proposition 9.17 *Let $\varphi : X \to \mathbb{R}$ be a linear functional. The following properties are equivalent:*

(1) *φ is continuous,*

(2) *φ is continuous at the point $0 \in X$,*

(3) *φ is bounded.*

△

Definition 9.18 Let φ be a bounded linear functional. The number

$$\|\varphi\| := \sup_{\substack{x\in X \\ x\neq 0}} \frac{|\varphi(x)|}{\|x\|} = \sup_{\|x\|=1} |\varphi(x)|$$

is called the *norm* of the functional φ.

△

It follows from the definition of the norm that

$$|\varphi(x)| \leq \|\varphi\| \, \|x\| \quad \text{for all} \quad x \in X.$$

Example A. Consider $C([a, b])$ with the norm

$$\|f\|_\infty := \max_{x\in[a,b]} |f(x)|.$$

The linear functional $\varphi(f) = \int_a^b f dx$ is bounded. Indeed,

$$|\varphi(f)| = \left|\int_a^b f(x)dx\right| \leq \int_a^b |f(x)|dx \leq \int_a^b \|f\|_\infty dx = \|f\|_\infty(b-a).$$

It follows that φ is bounded and $\|\varphi\| \leq b - a$.

In fact, $\|\varphi\| = b - a$. To see this, take the function $\tilde{f} \equiv 1$. For this function we have

$$\varphi(\tilde{f}) = \int_a^b 1dx = b - a = (b - a)\|\tilde{f}\|_\infty,$$

and thus

$$\|\varphi\| = \sup_{f\neq 0} \frac{|\varphi(f)|}{\|f\|_\infty} \geq \frac{|\varphi(\tilde{f})|}{\|\tilde{f}\|_\infty} = b - a.$$

Altogether we obtain $\|\varphi\| = b - a$.

B. Now consider the vector space of continuously differentiable functions on $[-\pi, \pi]$ with the same norm $\|\cdot\|_\infty$. Consider the linear functional $D(f) = f'(0)$. We claim that D is not bounded.

Indeed, take $f_n(x) = \sin nx$, $n \in \mathbb{N}$. Then $\|f_n\|_\infty = \max_{x\in[-\pi,\pi]} |\sin nx| = 1$. But

$$D(f_n) = f_n'(0) = n \cos nx|_{x=0} = n$$

and so

$$\frac{|D(f_n)|}{\|f_n\|_\infty} = \frac{n}{1} = n \to \infty, \quad n \to \infty.$$

△

Definition 9.19 Let $(X, \|\cdot\|)$ be a normed space. The space X^* of all linear bounded functionals on X is called the *dual space* of X.

△

The dual space X^* is a normed space with the norm $\|\varphi\|$, $\varphi \in X^*$, as described above.

Our aim for the remaining part of the chapter is to study the dual space of $L^p(\Omega, \mathscr{A}, \mu)$. We will use the the theory developed above and restrict ourselves to the case $1 < p < \infty$. For simplicity, we consider the real case.

Fix a function $g \in L^q(\Omega, \mathscr{A}, \mu)$ with $\frac{1}{p} + \frac{1}{q} = 1$. Then the functional

$$\varphi_g(f) := \int_\Omega f g d\mu, \quad f \in L^p(\Omega, \mathscr{A}, \mu),$$

is a bounded linear functional on $L^p(\Omega, \mathscr{A}, \mu)$ with the norm $\|\varphi_g\| = \|g\|_q$. Indeed, it follows from Hölder's inequality that

$$|\varphi_g(f)| = \left| \int_\Omega f g d\mu \right| \le \int_\Omega |fg| d\mu \le \|f\|_p \|g\|_q,$$

and thus φ_g is bounded with $\|\varphi_g\| \le \|g\|_q$. To show the equality, take

$$\tilde{f} := \frac{g}{|g|} \left(\frac{|g|}{\|g\|_q} \right)^{q/p}.$$

We have

$$\|\tilde{f}\|_p^p = \int_\Omega \frac{|g|^p}{|g|^p} \frac{|g|^q}{\|g\|_q^q} d\mu = \frac{\|g\|_q^q}{\|g\|_q^q} = 1,$$

i.e., $\|\tilde{f}\|_p = 1$, and

$$\varphi_g(\tilde{f}) = \int_\Omega \frac{g}{|g|} \frac{|g|^{q/p}}{\|g\|_q^{q/p}} g d\mu = \frac{1}{\|g\|_q^{q/p}} \int_\Omega |g|^{2-1+\frac{q}{p}} d\mu = \frac{1}{\|g\|_q^{q/p}} \int_\Omega |g|^{1+\frac{q}{p}} d\mu$$

$$= \frac{1}{\|g\|_q^{q/p}} \int_\Omega |g|^{q\overbrace{\left(\frac{1}{q}+\frac{1}{p}\right)}^{=1}} d\mu = \frac{1}{\|g\|_q^{q/p}} \|g\|_q^q = \|g\|_q^{q\left(1-\frac{1}{p}\right)} = \|g\|_q^{q\cdot\frac{1}{q}} = \|g\|_q.$$

It follows that $\|\varphi_g\| = \|g\|_q$.

In fact, all bounded linear functionals on $L^p(\Omega, \mathscr{A}, \mu)$, $1 < p < \infty$, have this form. For simplicity, we will show this only in the case when μ is a finite measure.

Theorem 9.20 *Let $(\Omega, \mathscr{A}, \mu)$ be a finite measure space and $1 < p < \infty$. Let φ be a linear bounded functional on $L^p(\Omega, \mathscr{A}, \mu)$. Then there exists a function $g \in L^q(\Omega, \mathscr{A}, \mu)$, $\frac{1}{p} + \frac{1}{q} = 1$, such that*

$$\varphi(f) = \int_\Omega f g d\mu, \quad f \in L^p(\Omega, \mathscr{A}, \mu).$$

Moreover, $\|\varphi\| = \|g\|_q$.

△

Proof. Let φ be a linear bounded functional on $L^p(\Omega, \mathscr{A}, \mu)$.

Consider the set function $\nu : \mathscr{A} \to \mathbb{R}$,

$$\nu(A) := \varphi(\chi_A), \quad A \in \mathscr{A}.$$

Since μ is finite, we always have $\chi_A \in L^p(\Omega, \mathscr{A}, \mu)$, so that $\nu(A)$ is well defined. In fact, ν is a signed measure on $\mathscr{A}$. It is obvious that $\nu(\emptyset) = 0$ and that ν is additive. We will show that ν is σ-additive.

Let $\{A_n\}_{n\in\mathbb{N}}$ be a sequence of p.d. sets in $\mathscr{A}$, and let $A := \sum_{n=1}^\infty A_n$. Then

$$\left\| \chi_A - \sum_{k=1}^n \chi_{A_k} \right\|_p = \left\| \chi_{\sum_{k=n+1}^\infty A_k} \right\|_p = \left(\mu\left(\sum_{k=n+1}^\infty A_k \right) \right)^{\frac{1}{p}} = \left(\sum_{k=n+1}^\infty \mu(A_k) \right)^{\frac{1}{p}}.$$

Since $\mu(A) = \sum_{k=1}^\infty \mu(A_k) < \infty$, the series $\sum_{k=1}^\infty \mu(A_k)$ is convergent in $\mathbb{R}$ and thus $\sum_{k=n+1}^\infty \mu(A_k) \to 0$, $n \to \infty$. It follows that

$$\left\| \chi_A - \sum_{k=1}^n \chi_{A_k} \right\|_p \to 0, \quad n \to \infty.$$

Since φ is continuous, we have

$$\left|\nu(A) - \sum_{k=1}^{n} \nu(A_k)\right| = \left|\varphi(\chi_A) - \sum_{k=1}^{n} \varphi(\chi_{A_k})\right| = \left|\varphi\left(\chi_A - \sum_{k=1}^{n} \chi_{A_k}\right)\right|$$
$$\leq \|\varphi\| \left\|\chi_A - \sum_{k=1}^{n} \chi_{A_k}\right\|_p \to 0, \quad n \to \infty,$$

and it follows that $\nu(A) = \sum_{k=1}^{\infty} \nu(A_k)$. This shows that ν is σ-additive.

Now take $A \in \mathscr{A}$ with $\mu(A) = 0$. In this case $\chi_A = 0$ μ-a.e., i.e. $\chi_A = 0$ as the element of the space $L^p(\Omega, \mathscr{A}, \mu)$. This implies

$$\nu(A) = \varphi(\chi_A) = \varphi(0) = 0.$$

Consequently, $\nu \ll \mu$.

By the Radon-Nikodym Theorem there exists a μ-integrable function g such that

$$\nu(A) = \int_A g d\mu, \quad A \in \mathscr{A}.$$

In particular,

$$\varphi(\chi_A) = \nu(A) = \int_A g d\mu = \int_\Omega \chi_A g d\mu.$$

By linearity, for any simple function $f = \sum_{i=1}^{n} a_i \chi_{A_i}$ we have

$$\varphi(f) = \int_\Omega f g d\mu.$$

We need to show that this holds also for an arbitrary function $f \in L^p(\Omega, \mathscr{A}, \mu)$.

Consider first $f \in L^\infty(\Omega, \mathscr{A}, \mu)$. One can show[1] that there exists a sequence of simple functions $\{f_k\}_{k\in\mathbb{N}}$ such that $\|f_k - f\|_\infty \to 0$, $k \to \infty$. Since $\mu(\Omega) < \infty$, we have

$$\|f_k - f\|_p^p = \int_\Omega |f_k - f|^p d\mu \leq \|f_k - f\|_\infty^p \int_\Omega d\mu = \|f_k - f\|_\infty^p \mu(\Omega),$$

[1]Indeed, denote $M := \|f\|_\infty$, take $k \in \mathbb{N}$ and set $I_j^k := \left[-M + 2M\frac{j-1}{k}, -M + 2M\frac{j}{k}\right)$ for $j = 1, \dots, k-1$, $I_k^k := \left[M - \frac{2M}{k}, M\right]$. Let $A_j^k := f_0^{-1}(I_j^k) \in \mathscr{A}$, $j = 1, \dots, k$, where f_0 is a representative of the equivalence class f. Notice that $\mu\left(\Omega \setminus \sum_{j=1}^{k} A_j^k\right) = 0$. Denote by a_j^k the midpoints of the intervals I_j^k. Then $f_k := \sum_{j=1}^{k} a_j^k \chi_{A_j^k}$ is a simple function and $\|f - f_k\|_\infty \leq \frac{M}{k}$.

and thus

$$\|f_k - f\|_p \le (\mu(\Omega))^{\frac{1}{p}} \|f_k - f\|_\infty \to 0, \quad k \to \infty.$$

In other words, $\{f_k\}_{k\in\mathbb{N}}$ converges to f also in $L^p(\Omega, \mathscr{A}, \mu)$. From the continuity of φ it follows that

$$|\varphi(f_k) - \varphi(f)| \le \|\varphi\| \, \|f_k - f\|_p \to 0, \quad k \to \infty.$$

On the other hand,

$$\left|\int_\Omega f_k g d\mu - \int_\Omega f g d\mu\right| \le \int_\Omega |f_k - f||g| d\mu \le \|f_k - f\|_\infty \underbrace{\int_\Omega |g| d\mu}_{<\infty} \to 0, \quad k \to \infty.$$

Thus,

$$\varphi(f) = \lim_{k\to\infty} \varphi(f_k) = \lim_{k\to\infty} \int_\Omega f_k g d\mu = \int_\Omega f g d\mu.$$

In the next step we show that $g \in L^q(\Omega, \mathscr{A}, \mu)$. Consider

$$\tilde{f}(t) := \begin{cases} \frac{|g(t)|^q}{g(t)}, & g(t) \neq 0, \\ 0, & g(t) = 0. \end{cases}$$

The function $\tilde{f}$ is measurable, $\tilde{f}(t)g(t) = |g(t)|^q$ and

$$|\tilde{f}(t)|^p = |g(t)|^{(q-1)p} = |g(t)|^{(q-1)\frac{q}{q-1}} = |g(t)|^q.$$

For $n \in \mathbb{N}$ set $E_n := \{t \in \Omega : |g(t)| \le n\}$. Then $\chi_{E_n}\tilde{f} \in L^\infty(\Omega, \mathscr{A}, \mu)$, and thus

$$\begin{aligned} \int_\Omega \chi_{E_n} |g|^q d\mu &= \int_\Omega \chi_{E_n} \tilde{f} g d\mu = \varphi(\chi_{E_n}\tilde{f}) \le \|\varphi\| \, \|\chi_{E_n}\tilde{f}\|_p \\ &= \|\varphi\| \left(\int_\Omega \chi_{E_n} |\tilde{f}|^p d\mu\right)^{\frac{1}{p}} = \|\varphi\| \left(\int_\Omega \chi_{E_n} |g|^q d\mu\right)^{\frac{1}{p}}. \end{aligned}$$

It follows that

$$\left(\int_\Omega \chi_{E_n} |g|^q d\mu\right)^{1-\frac{1}{p}} = \left(\int_\Omega \chi_{E_n} |g|^q d\mu\right)^{\frac{1}{q}} \le \|\varphi\|.$$

The sequence $\{\chi_{E_n}|g|^q\}_{n\in\mathbb{N}}$ is monotone non-decreasing, $\chi_{E_n}|g|^q \to |g|^q$ pointwisely as $n \to \infty$, and $\int_\Omega \chi_{E_n}|g|^q d\mu \le \|\varphi\|^q$ for all $n \in \mathbb{N}$. It follows by the Monotone Convergence Theorem that

$$\int_\Omega |g|^q d\mu = \lim_{n\to\infty} \int_\Omega \chi_{E_n} |g|^q d\mu \le \|\varphi\|^q.$$

We conclude that $|g|^q$ is μ-integrable, i.e. $g \in L^q(\Omega, \mathscr{A}, \mu)$, and $\|g\|_q \leq \|\varphi\|$.

Finally, it can be shown[2] that for any $f \in L^p(\Omega, \mathscr{A}, \mu)$ there is a sequence $\{f_k\}_{k \in \mathbb{N}}$ of L^∞-functions such that $\|f_k - f\|_p \to 0$, $k \to \infty$. Then

$$|\varphi(f_k) - \varphi(f)| \leq \|\varphi\| \, \|f_k - f\|_p \to 0, \quad k \to \infty,$$

and

$$\left| \int_\Omega f_k g d\mu - \int_\Omega f g d\mu \right| \leq \int_\Omega |f_k - f||g| d\mu \leq \|f_k - f\|_p \|g\|_q \to 0, \quad k \to \infty.$$

It follows that

$$\varphi(f) = \lim_{k \to \infty} \varphi(f_k) = \lim_{k \to \infty} \int_\Omega f_k g d\mu = \int_\Omega f g d\mu$$

as desired.

□

Remark According to Theorem 9.20, the dual space of $L^p(\Omega, \mathscr{A}, \mu)$, $1 < p < \infty$, can be identified with $L^q(\Omega, \mathscr{A}, \mu)$, $\frac{1}{p} + \frac{1}{q} = 1$. We proved this for finite measure spaces, but this can be also shown for arbitrary measure spaces $(\Omega, \mathscr{A}, \mu)$.

△

Remark If μ is σ-finite, it can be shown that $L^1(\Omega, \mathscr{A}, \mu)$ can be identified with $L^\infty(\Omega, \mathscr{A}, \mu)$ in the same way.

Regarding $L^\infty(\Omega, \mathscr{A}, \mu)$, it can be shown that any function $g \in L^1(\Omega, \mathscr{A}, \mu)$ defines a linear bounded functional

$$\varphi_g(f) = \int_\Omega f g d\mu$$

on $L^\infty(\Omega, \mathscr{A}, \mu)$ with $\|\varphi_g\| = \|g\|_1$. However, (unless Ω is degenerate, e.g., consisting of finitely many points), the space $(L^\infty(\Omega, \mathscr{A}, \mu))^*$ is larger than $L^1(\Omega, \mathscr{A}, \mu)$, i.e., there exist linear bounded functionals on $L^\infty(\Omega, \mathscr{A}, \mu)$ that cannot be represented in the form $\int_\Omega f g d\mu$.

△

[2] Take a representative f_0 of the equivalence class $f \in L^p(\Omega, \mathscr{A}, \mu)$ and set $E_k := \{t \in \Omega : |f_0(t)| \leq k\}$. Put $f_k := f_0 \chi_{E_k}$. Clearly, $f_k \in L^\infty(\Omega, \mathscr{A}, \mu)$. We have $|f - f_k|^p \to 0$ μ-a.e. as $k \to \infty$ and $|f - f_k|^p \leq 2^p |f|^p$, the latter function being an integrable majorant. By the Lebesgue Dominated Convergence Theorem, $\|f - f_k\|_p \to 0$, $k \to \infty$.

Index