The Language of Probability

Michael Hediger

University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland

April 23, 2024

References: The given script is inspired by the books of Patrick Billingsley [1] and Jean-François Le Gall [2].

Contents

1 Introduction: Part I

The following gives a list of greek letters:

α	alpha	ν	nu
β	beta	ξ,Ξ	xi
γ,Γ	gamma	o	o, omicron
δ,Δ	delta	π,Π	pi
ϵ,ε	epsilon	ρ	rho
ζ	zeta	σ,Σ	sigma
η	eta	τ	tau
θ,Θ	theta	υ,Υ	upsilon
ι	iota	ϕ,φ,Φ	phi
κ	kappa	χ	chi
λ,Λ	lambda	ψ,Ψ	psi
μ	mu	ω,Ω	omega

We clarify some logical operations on statements. If S_1 and S_2 are two statements, we write $S_1 \Rightarrow S_2$ if S_1 implies S_2. S_1 if and only if S_2 ($S_1 \Leftrightarrow S_2$) means that S_1 implies S_2 and S_2 implies S_1. Further, the following logical operators are helpful:

$$
\begin{aligned}
&: \text{ means } \textit{such that}\\
&\forall \text{ means } \textit{for all}\\
&\exists \text{ means } \textit{there exists}.
\end{aligned}
$$

1.1 Sets

Sets can be defined by their elements $A = \{\omega_1, \omega_2, \dots, \omega_n\}$ or upon a certain property

$$A = \{\omega \colon \omega \text{ has property } P\}.$$

Let A be a set, then

$$
\begin{aligned}
&\omega \in A \text{ means that } \omega \text{ is an element of } A,\\
&\omega \notin A \text{ means that } \omega \text{ is not an element of } A.
\end{aligned}
$$

Example 1.1. *The set which contains the strictly positive integers* $1, 2, 3, \dots$ *is denoted with* $\mathbb{N}$. *If* $n \in \mathbb{N}$, *then so is* $n + 1$. *It is a matter of convention whether* $0 \in \mathbb{N}$ *or not. For us,* $0 \notin \mathbb{N}$.

Example 1.2. *The set of integers is denoted with* $\mathbb{Z}$, *it contains* 0, $\mathbb{N}$, *and the set of points* $\{-n \colon n \in \mathbb{N}\}$.

Example 1.3. $\mathbb{Q}$ *is the set of rational numbers:*

$$\mathbb{Q} = \big\{q \colon q = \frac{n}{m},\ n, m \in \mathbb{Z},\ m \neq 0\big\}.$$

Example 1.4. *It can be shown that there does not exists* $q \in \mathbb{Q}$ *s.t.* $q^2 = 2$. *This shows that* $\sqrt{2} \notin \mathbb{Q}$. *The same is true for* π *and Euler's number* e. *The latter numbers belong to the set of real numbers, denoted with* $\mathbb{R}$. *In particular,* $\mathbb{R}$ *contains all the integers and rational numbers.*

Example 1.5. *Let* $A_1, \dots, A_n$, $n \in \mathbb{N}$, *be a family of sets. The Cartesian product of* $A_1, \dots, A_n$ *is given by*

$$\prod_{i=1}^{n} A_i = A_1 \times \cdots \times A_n = \{\omega \colon \omega = (\omega_1, \dots, \omega_n) \colon \omega_i \in A_i,\ i = 1, \dots, n\}.$$

An element ω of $A_1 \times \cdots \times A_n$ is referred to as a vector with coordinates $\omega_i \in A_i$, $i = 1, \ldots, n$. If $A_i = A$, $i = 1, \ldots, n$, we write $A_1 \times \cdots \times A_n = A^n$. The space $\mathbb{R}^k$ is referred to as the real coordinate space of dimension k.

Definition 1.1. *Let A and B be two sets. Then, we define the following set operations for A and B.*

Equality of sets: *$A = B$ if and only if A and B contain the same elements. That is, any element of A is also an element of B and any element if B is also an element of A.*

Inclusion: *$A \subset B$ if and only if $\omega \in A$ implies that $\omega \in B$. If $A \subset B$, we say that A is a subset of B.*

Intersection: *The intersection of A and B is the set*

$$A \cap B = \{\omega \colon \omega \in A \text{ and } \omega \in B\}.$$

Union: *The union of A and B is the set*

$$A \cup B = \{\omega \colon \omega \in A \text{ or } \omega \in B\}.$$

Set difference: *The difference between A and B is the set*

$$A \setminus B = \{\omega \colon \omega \in A \text{ and } \omega \notin B\}.$$

It is often of interest to consider the intersection and union of more than just two sets. Let $\{A_i \colon i \in I\}$ be a family of sets where I is some set. Then, the intersection of A_i, $i \in I$, is defined as the set

$$\bigcap_{i \in I} A_i = \{\omega \colon \omega \in A_i \ \forall i \in I\}. \tag{1}$$

The union of A_i, $i \in I$, is defined as

$$\bigcup_{i \in I} A_i = \{\omega \colon \exists i \in I \text{ s.t. } \omega \in A_i\}. \tag{2}$$

As an example one could think of $I = \mathbb{N}$ or $I = \{1, \ldots, N\}$, where $N \in \mathbb{N}$.

We can apply Definition 1.1 to derive several elementary properties of set operations.

Proposition 1.1. *Let A, B and C be some sets.*

Properties of the inclusion:

(1.1) $A \subset A$, *i.e., each set is a subset of itself;*

(1.2) $A \subset B$ *and* $B \subset A$ *if and only if* $A = B$*;*

(1.3) $A \subset B$ *and* $B \subset C$ *implies that* $A \subset C$*.*

Associativity:

(2.1) $(A \cup B) \cup C = A \cup (B \cup C)$*;*

(2.2) $(A \cap B) \cap C = A \cap (B \cap C)$*.*

Commutativity:

(3.1) $A \cup B = B \cup A$*;*

(3.2) $A \cap B = B \cap A$*.*

Distributive law:

(4.1) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$*;*

(4.2) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$*.*

As an example, let us proof two items of the latter proposition.

Proof of (1.2) in Proposition 1.1. The strategy is to show that $A \subset B$ and $B \subset A$ implies that $A = B$ and $A = B$ implies that $A \subset B$ and $B \subset A$. Suppose that $A \subset B$ and $B \subset A$. By Definition 1.1, we need to show that this implies that A and B contain the same elements. Let $\omega \in A$, then $\omega \in B$, since $A \subset B$. On the other hand, take any $\omega \in B$. Since $B \subset A$, $\omega \in A$. Thus, any $\omega \in A$ is also an element of B and vice versa. This shows that $A \subset B$ and $B \subset A$ implies that $A = B$. For the other direction, let us assume that $A = B$. Then, given $\omega \in A$, it is true that $\omega \in B$ as well, thus by Definition 1.1, $A \subset B$. Also if $\omega \in B$, since $A = B$, it follows that $\omega \in A$. Thus again by Definition 1.1, $B \subset A$. Hence $A = B$ implies that $A \subset B$ and $B \subset A$. □

Proof of (4.1) in Proposition 1.1. We need to show that $A \cap (B \cup C) \subset (A \cap B) \cup (A \cap C)$ and $(A \cap B) \cup (A \cap C) \subset A \cap (B \cup C)$. Let $\omega \in A \cap (B \cup C)$. This implies that ω is an element of A and an element of $B \cup C$. By the definition of $B \cup C$, that means that $\omega \in B$ or $\omega \in C$. Still, $\omega \in A$, thus $\omega \in A$ and $\omega \in B$ or $\omega \in A$ and $\omega \in C$. This shows that $\omega \in (A \cap B) \cup (A \cap C)$ and hence $A \cap (B \cup C) \subset (A \cap B) \cup (A \cap C)$. For the other inclusion, let $\omega \in (A \cap B) \cup (A \cap C)$. Then, by the definition of $(A \cap B) \cup (A \cap C)$, $\omega \in A$ and $\omega \in B$ or $\omega \in A$ and $\omega \in C$. This means that ω is a member of $B \cup C$ and A. Hence, $\omega \in A \cap (B \cup C)$. □

Exercise 1.1. *Show (1.3) of Proposition 1.1.*

Exercise 1.2. *Let A and B be two sets. Show that $A \cap B \subset A$ and conclude that $A \subset B$ implies that $A \cap B = A$.*

Exercise 1.3. *Let A and B be two sets. Show that $B \subset A \cup B$ and conclude that $A \subset B$ implies that $A \cup B = B$.*

Exercise 1.4. *Let A and B be two sets. Show that $A \cup B = A \cup (B \setminus A)$.*

Example 1.6. *The set which has no elements is called the empty set and denoted with $\emptyset$.*

Proposition 1.2. *Given any subset A, $\emptyset \subset A$.*

Proof. Notice that we have used the definition that two sets are equal $A = B$ if and only if each element of A is also an element of B and vice versa. Assume for the moment that it is not true that for all subsets A, $\emptyset \subset A$. That means that there exists a subset A such that $\emptyset$ is not a subset of A. By Definition 1.1 that means that the empty set must contain an element which does not belong to A. This gives a contradiction and hence it is not true that there exists a subset A such that $\emptyset$ is not a subset of A, i.e., for any subsets A, $\emptyset \subset A$. □

Using Exercises 1.2 and 1.3, the latter result shows that $\emptyset \cap A = \emptyset$ and $\emptyset \cup A = A$.

Definition 1.2. *Let A and B be two sets. A and B are said to be disjoint if $A \cap B = \emptyset$. More generally, let $\{A_i : i \in I\}$ be any family of sets. $\{A_i : i \in I\}$ is said to be disjoint if $A_i \cap A_j = \emptyset$ for any $i \neq j$.*

Exercise 1.5. *Let A and B be two sets. Is it possible that $A \subset B$ (or vice versa) and A and B are disjoint.*

In the following, we list some properties of the difference between sets.

Proposition 1.3. *Let A, B and C be sets. Then,*

(i) $C \setminus (A \cap B) = (C \setminus A) \cup (C \setminus B)$*;*

(ii) $C \setminus (A \cup B) = (C \setminus A) \cap (C \setminus B)$;

(iii) $(B \setminus A) \cap C = (B \cap C) \setminus A$;

(iv) $(B \setminus A) \cup C = (B \cup C) \setminus (A \setminus C)$.

Notice that Exercise 1.4 is just a special case of item (iv) in Proposition 1.3 if we set $C = A$.

In what follows, it is often the case that a particular set Ω is given and one only considers subsets A of Ω. In that case, we use the notation $A^c = \Omega \setminus A$ for the complement of A in Ω.

Proposition 1.4. *Let A and B be subsets of Ω. Then,*

- $A \cup A^c = \Omega$;
- $A \cap A^c = \emptyset$;
- $A \setminus B = A \cap B^c$;
- $\emptyset^c = \Omega$;
- $\Omega^c = \emptyset$;
- $(A \subset B) \Rightarrow (B^c \subset A^c)$;
- $(A^c)^c = A$.

Further, it is true that

- $(A \cap B)^c = A^c \cup B^c$;
- $(A \cup B)^c = A^c \cap B^c$.

With reference to (1) and (2), we remark that the last two properties can be extended to unions and intersections of arbitrary families $\{A_i : i \in I\}$ of subsets of Ω:

$$\left(\bigcap_{i \in I} A_i\right)^c = \bigcup_{i \in I} A_i^c,$$

$$\left(\bigcup_{i \in I} A_i\right)^c = \bigcap_{i \in I} A_i^c.$$

We remark that since $\{A_i : i \in I\}$ consists only of subsets of Ω,

$$\bigcap_{i \in I} A_i = \{\omega \in \Omega : \omega \in A_i \ \forall\, i \in I\} \text{ and } \bigcup_{i \in I} A_i = \{\omega \in \Omega : \exists\, i \in I \text{ s.t. } \omega \in A_i\}.$$

1.2 The principle of induction

Let for any $n \in \mathbb{N}$, $S(n)$ be a statement. In order to proof that $S(n)$ is true for any $n \in \mathbb{N}$ we can adapt the following strategy:

Principle of induction To verify that $S(n)$ is true for any $n \in \mathbb{N}$ we check (I) and (II):

(I) $S(n)$ is true for $n = 1$ (base case);

(II) $S(n) \Rightarrow S(n+1)$ for any $n \in \mathbb{N}$ (induction step).

We note that if one seeks to proof $S(n)$ for $n \geq N$, $N \in \mathbb{N}$, then one needs to verify

(I) $S(N)$ is true (base case);

(II) $S(n) \Rightarrow S(n+1)$ for any $n \geq N$ (induction step).

Example 1.7. *Let us use a proof by induction to verify that for any $n \in \mathbb{N}$,*

$$1 + 2 + 3 + \cdots + n = \frac{n(n+1)}{2}.$$

First we verify the statement in the base case. Clearly, if $n = 1$, $1 = (1(1+1))/2$. For the induction step, we let $n \in \mathbb{N}$ be arbitrary and assume that $1 + 2 + \cdots + n = (n(n+1))/2$. Therefore,

$$1 + 2 + \cdots n + n + 1 = \frac{n(n+1)}{2} + n + 1 = \frac{n(n+1) + 2(n+1)}{2} = \frac{(n+1)(n+1+1)}{2},$$

which verifies the induction step and completes the argument.

Definition 1.3. *Given $n \in \mathbb{N} \cup \{0\}$, The number $n!$ is called n-factorial and given by*

$$n! = n(n-1) \cdot \ldots \cdot 1.$$

We use the convention that $0! = 1$.

Example 1.8. *We want to verify the following statement: Let $n \in \mathbb{N}$, then,*

$$\forall N \in \mathbb{N}\ \exists C > 0 \text{ s.t. } n! \geq CN^n.$$

In words, the statement reads as follows: Given any $n \in \mathbb{N}$ it is true that for any $N \in \mathbb{N}$ there exists a positive constant C s.t. $n!$ is greater or equal to CN^n. To proof it, let $n \in \mathbb{N}$ and $N \in \mathbb{N}$ be given. There are two cases, either $n \leq N$ or $n > N$. In the first case, if $n \leq N$, we can pick $C = 1/N^N$. Then,

$$n! \geq 1 \geq \frac{N^n}{N^N}.$$

Thus, in this case, no induction is needed. For the case $n > N$, we use a proof by induction on the new statement at base $n = N$:

$$\exists C > 0 \text{ s.t. } n! \geq CN^n \text{ for all } n \geq N.$$

By the first case, we have already verified the base step. Thus it remains to verify the induction step. We thus assume that for any given $n \geq N$, the latter statement is true. Then,

$$(n+1)! = (n+1)n! \geq (n+1)CN^n \geq (N+1)CN^n = CN^{n+1} + CN^n \geq CN^{n+1},$$

which completes the induction step. In summary, we started with an arbitrary $N \in \mathbb{N}$ and have shown that for both cases, $n \leq N$ and $n > N$, there exists C s.t. $n! \geq CN^n$. This shows the original statement for any $n \in \mathbb{N}$.

Exercise 1.6. *Verify that for any $n \in \mathbb{N}$,*

$$\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}.$$

1.3 Order structure of the real numbers

Definition 1.4. *Let $a < b$, $a, b \in \mathbb{R}$. Then,*

- $[a, b] = \{x \in \mathbb{R} \colon a \leq x \leq b\}$ *is called a closed interval;*
- $(a, b) = \{x \in \mathbb{R} \colon a < x < b\}$ *is called an open interval;*

- $[a, b) = \{x \in \mathbb{R} \colon a \leq x < b\}$ *is called a right-open interval;*
- $(a, b] = \{x \in \mathbb{R} \colon a < x \leq b\}$ *is called a left-open interval.*

A set $I \subset \mathbb{R}$ is said to be an interval if it is either closed, open, right-open or left-open.

Definition 1.5. *Let $a, b \in \mathbb{R}$. The unbounded real intervals are given by the sets:*

- $[a, \infty) = \{x \in \mathbb{R} \colon a \leq x < \infty\}$, $(a, \infty) = \{x \in \mathbb{R} \colon a < x < \infty\}$;
- $(-\infty, b] = \{x \in \mathbb{R} \colon -\infty < x \leq b\}$, $(-\infty, b) = \{x \in \mathbb{R} \colon -\infty < x < b\}$.

Definition 1.6. *Let $A \subset \mathbb{R}$. An element $s \in \mathbb{R}$ is called an upper (resp. lower) bound of A, if $x \leq s$ (resp. $x \geq s$) for all $x \in A$. If A has an upper (resp. lower) bound then we say that A is bounded from above (resp. below). If A is bounded from below and above, A is bounded.*

Definition 1.7. *Let $A \subset \mathbb{R}$ be a set. An element $s \in \mathbb{R}$ is called supremum of A (we write $s = \sup A$) if s is the smallest upper bound of A. That is, the following two items are satisfied:*

(i) *s is an upper bound of A;*

(ii) *Every number $s' < s$ is not an upper bound of A.*

Example 1.9. *Let $A = [0, 1)$. Then, 1 is an upper bound for $[0, 1)$, since $x < 1 \Rightarrow x \leq 1$ for any $x \in [0, 1)$. In order to show that $\sup[0, 1) = 1$, we need to verify that 1 is the smallest upper bound of $[0, 1)$, i.e., any $s' < 1$ can not be an upper bound of $[0, 1)$. To show this, it is sufficient to proof that there exists a real number q in $[0, 1) \cap (s', 1)$ (we provide an argument later). Then, $q \in [0, 1)$ with $q > s'$ and hence s' can not be an upper bound for $[0, 1)$. Thus, $\sup[0, 1) = 1$. Notice that $\sup[0, 1) \notin [0, 1)$, i.e., the supremum must not be an element of the set itself.*

Definition 1.8. *Let $A \subset \mathbb{R}$ be a set. An element $s \in \mathbb{R}$ is called infimum of A (we write $s = \inf A$) if s is the greatest lower bound of A. That is, the following two items are satisfied:*

(i) *s is a lower bound of A;*

(ii) *Every number $s' > s$ is not a lower bound of A.*

Example 1.10. *Let $A = [0, 1)$. Since $x \geq 0$ for any $x \in [0, 1)$, 0 is a lower bound of $[0, 1)$. As $0 \in [0, 1)$, $\inf[0, 1) = 0$.*

Definition 1.9. *Let $A \subset \mathbb{R}$. If $s = \sup A \in A$ (resp. $s = \inf A \in A$) we call s the maximum (resp. the minimum) of A.*

The following result is of general importance. It shows that for each nonempty subset of the real line which has an upper (resp. lower) bound, the supremum (resp. infimum) exists.

Proposition 1.5. *Let $A \subset \mathbb{R}$ s.t. $A \neq \emptyset$. Suppose that there exists an upper (resp. lower) bound for A. Then, $\sup A$ (resp. $\inf A$) exists.*

Example 1.11. *Let $A = [0, 1)$. We have seen that the minimum of $[0, 1)$ exists. However, the maximum of $[0, 1)$ does not exist, since $\sup[0, 1) = 1 \notin [0, 1)$. This makes sense, as $[0, 1)$ is right-open and there does not exist a maximal element of $[0, 1)$.*

In order to show that there exists a number in between any two distinct real numbers, the rational numbers $\mathbb{Q}$ are helpful.

Proposition 1.6 ($\mathbb{Q}$ is dense in $\mathbb{R}$)**.** *For any two real numbers $x_1, x_2 \in \mathbb{R}$ (say $x_1 < x_2$), there exists a rational number between x_1 and x_2, i.e., there exists $q \in \mathbb{Q}$ s.t. $x_1 < q < x_2$.*

To proof the above result, we rely on two fundamental results.

Proposition 1.7. *For any $x \in \mathbb{R}$ there exists $n \in \mathbb{N}$ s.t. $n > x$.*

Proposition 1.8. *Let $A \subset \mathbb{Z}$ s.t. $A \neq \emptyset$. If A has an upper (resp. lower) bound, then A has a maximum (resp. minimum).*

Proof of Proposition 1.6. By Proposition 1.7, since $x_2 \neq x_1$, let $n \in \mathbb{N}$ s.t.

$$n > \frac{1}{x_2 - x_1}.$$

Thus, $1/n < x_2 - x_1$. Let us define the following set

$$A = \{a \in \mathbb{Z} \colon a > nx_1\}.$$

By Proposition 1.7, $A \neq \emptyset$ Further, for any $a \in A$, $a > nx_1 \Rightarrow a \geq nx_1$. Hence, nx_1 is a lower bound for A. By Proposition 1.8, there exists a minimum m of A. Then, we must have that

$$m > nx_1 \text{ (since } m \in A) \text{ but } m - 1 \leq nx_1,$$

since otherwise we have $m - 1 \in A$ which contradicts that m is the smallest number which is strictly greater than nx_1. It follows that

$$x_1 < \frac{m}{n} = \frac{m-1}{n} + \frac{1}{n} \leq x_1 + \frac{1}{n} < x_1 + x_2 - x_1 = x_2.$$

Hence if we let $q = m/n$, the result follows. □

Example 1.12. *In Example 1.9, we have postponed an argument that there exists a real number in $[0, 1) \cap (s', 1)$. Clearly, only the case $s' > 0$ (i.e., $[0, 1) \cap (s', 1) = (s', 1)$) is of interest since otherwise $s' \leq 0$ and then s' is clearly no upper bound for $[0, 1)$. Using Proposition 1.6 we find $q \in \mathbb{Q} \subset \mathbb{R}$ s.t. $q \in (s', 1)$ and the result follows.*

Definition 1.10. *Let $A \subset \mathbb{R}$ s.t. $A \neq \emptyset$. We define,*

(i) $\sup A = \infty$ *if A has no upper bound;*

(ii) $\inf A = -\infty$ *if A has no lower bound.*

Proposition 1.9. *Let $A, B \subset \mathbb{R}$ s.t. $A \subset B$ $(A, B \neq \emptyset)$. Then,*

$$\inf A \geq \inf B \quad \textit{and} \quad \sup A \leq \sup B.$$

Proof. Let $m = \inf B$. Then, m is s.t. $m \leq x$ for any $x \in B$. Since $A \subset B$, every element of A is an element of B and thus $m \leq x$ for any $x \in A$. Then, by definition, $\inf A$ is the greatest lower bound of A and m was found to be a lower bound of A, thus $m \leq \inf A$. Let $m = \sup M$. Then, $x \leq M$ for any $x \in B$. In particular, since $A \subset B$, $x \leq M$ for any $x \in A$. By definition, $\sup A$ is the smallest upper bound of A, thus $\sup A \leq M$. □

Proposition 1.10. *Let $A \subset \mathbb{R}$ be a nonempty set. Then,*

- *if $\inf A > -\infty$, for any $\delta > 0$, there exists $x \in A$ s.t. $x < \inf A + \delta$;*
- *$\sup A < \infty$, for any $\delta > 0$, there exists $x \in A$ s.t. $x > \sup A - \delta$;*

Proof. Suppose that there exists $\delta > 0$ s.t. for any $x \in A$, $x \geq \inf A + \delta$. Then, $\inf A + \delta$ is a lower bound for A which is greater than $\inf A$. This is not possible. Similarly, suppose that there exists $\delta > 0$ s.t. for any $x \in A$ $x \leq \sup A - \delta$. Then, $\sup A - \delta$ is an upper bound for A which is smaller than $\sup A$. Again, this is not possible. □

Remark 1.1. *Clearly* $\mathbb{R}$ *is not bounded and hence* $\inf \mathbb{R} = -\infty$ *and* $\sup \mathbb{R} = \infty$*. Sometimes, it is convenient to adjoin* $\mathbb{R}$ *with the objects* $-\infty$ *and* ∞*, i.e., consider the set*

$$\mathbb{R} \cup \{-\infty, \infty\}.$$

This set is referred to as the extended real numbers (sometimes also written as $[-\infty, \infty]$*). We use the notation* $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$*. It is important to note that by definition,* $-\infty, \infty \notin \mathbb{R}$*, i.e., these objects are not numbers. However, for any real number* $x \in \mathbb{R}$*,* $x > -\infty$ *and* $x < \infty$*. Therefore, we assume that* $-\infty$ *and* ∞ *satisfy the relation* $-\infty < \infty$*. With that relation between* $-\infty$ *and* ∞*, we have that for any* $x \in \overline{\mathbb{R}}$*,* $-\infty \leq x \leq \infty$*. By now we have seen that* $-\infty$ *and* ∞ *appear in the context of unbounded sets. We will see latter, that* $-\infty$ *and* ∞ *also appear in the definition of diverging sequences. For future references, we also write* $\overline{\mathbb{R}}_+ = [0, \infty) \cup \{\infty\}$*.*

Example 1.13. *Let* $a, b \in \mathbb{R}$ *and assume that for any* $\varepsilon > 0$*,* $a \leq b + \varepsilon$*. Then,* $a \leq b$*. To see it, let* $B = \{x + \varepsilon \colon \varepsilon > 0, x \geq b\}$ *and* $A = \{x \colon x \geq a\}$*. We have that* $\inf B = b$ *and* $\inf A = a$*. Clearly,* $B \subset A$*. Hence, by the latter proposition,* $\inf A = a \leq b = \inf B$*. Similarly, if for any* $\varepsilon > 0$*,* $a \geq b - \varepsilon$*, we must have* $a \geq b$*. To see it, take* $B = \{x - \varepsilon \colon \varepsilon > 0, x \leq b\}$ *and* $A = \{x \colon x \leq a\}$*. Then* $\sup B = b$ *and* $\sup A = a$*. Hence, by the latter proposition, since* $B \subset A$ *we have that* $b \leq a$*. Notice, that the latter results do not change if we allow for* $a = \infty$ *or* $b = \infty$*, i.e.,* $a, b \in \overline{\mathbb{R}}$ *(cf. Remark 1.1). For example, in the case where* $a \leq b + \varepsilon$*, if* $a = \infty$*, then* $b = \infty$ *and hence* $a = b$*. If* $b = \infty$*, then either* $a = b$ *or* $a < b$*. Finally, we remark that if for any* $\varepsilon > 0$*,* $a < b + \varepsilon$*, then,* $a \leq b$ *as well (resp.* $a \geq b$ *if for any* $\varepsilon > 0$ $a > b - \varepsilon$*). This is because the statement* $a < b + \varepsilon$ *(resp.* $a > b - \varepsilon$*) implies that* $a \leq b + \varepsilon$ *(resp.* $a \geq b - \varepsilon$*).*

Exercise 1.7. *For each of the following sets, identify its infimum and supremum. Deduce whether the sets have a minimum or maximum.*

(a) $A = \{1/n \colon n \in \mathbb{N}\}$*;*

(b) $B = \mathbb{Q} \cap [0, 2)$*;*

(c) $C = \mathbb{Z} \cap (-\infty, 0]$*.*

Up to now $\inf A$ and $\sup A$ were only defined for $A \subset \mathbb{R}$. We extend the notions of infimum and supremum to the extended real numbers as follows:

Definition 1.11.

- *If* $A \subset \overline{\mathbb{R}}$ *s.t.* $A \subset \mathbb{R}$ *and* A *is bounded, then* $\sup A$ *and* $\inf A$ *are defined as in Definitions 1.7 and 1.8, respectively.*

- *If* $A \subset \overline{\mathbb{R}}$ *s.t. there exists no real number* s *which is s.t. for any* $x \in A$*,* $x \leq s$*, then* $\sup A = \infty$*. In particular, this is the case if* $\infty \in A$*.*

- *If there exists* $s \in \mathbb{R}$ *s.t.* $x \leq s$ *for any* $x \in A$*, then* $\sup A$ *is defined as follows: If* $A \subset \mathbb{R}$*, then* $\sup A$ *is defined as in Definition 1.7. Otherwise,* $A = \{-\infty\} \cup A^*$*,* $A^* \subset \mathbb{R}$*, and we define*

$$\sup A = \begin{cases} \sup A^*, & \text{if } A^* \neq \emptyset, \\ -\infty, & \text{otherwise.} \end{cases}$$

  In particular, with this definition, $\sup A$ *is the smallest upper bound of* A*.*

- *If* $A \subset \overline{\mathbb{R}}$ *s.t. there exists no real number* s *which is s.t. for any* $x \in A$*,* $x \geq s$*, then* $\inf A = -\infty$*. In particular, this is the case if* $-\infty \in A$*.*

- *If there exists $s \in \mathbb{R}$ s.t. $x \geq s$ for any $x \in A$, then* $\inf A$ *is defined as follows: If $A \subset \mathbb{R}$, then* $\inf A$ *is defined as in Definition 1.8. Otherwise, $A = A^* \cup \{\infty\}$, $A^* \subset \mathbb{R}$, and we define*

$$\inf A = \begin{cases} \inf A^*, & \text{if } A^* \neq \emptyset, \\ \infty, & \text{otherwise.} \end{cases}$$

In particular, with this definition, $\inf A$ *is the greatest lower bound of A.*

Proposition 1.11. $\sup \emptyset = -\infty$ *and* $\inf \emptyset = \infty$.

Proof. Suppose that there exists $s \in \overline{\mathbb{R}}$ which is not an upper bound for $\emptyset$. Then, there exists $a \in \emptyset$, s.t. $a > s$. Which is clearly not possible, since $\emptyset$ contains not a single element. Therefore, any $s \in \overline{\mathbb{R}}$ is in fact an upper bound for $\emptyset$. Therefore, $\sup \emptyset = \min(\overline{\mathbb{R}}) = -\infty$, the smallest upper bound of $\emptyset$. Similarly, $\inf \emptyset = \max(\overline{\mathbb{R}}) = \infty$, the largest lower bound of $\emptyset$. $\square$

1.4 Solution to exercises

Solution 1.1 (Solution to Exercise 1.1)**.** *We need to show that $\omega \in A \Rightarrow \omega \in C$. Thus, let $\omega \in A$. Since $A \subset B$, that implies that $\omega \in B$. Further, since $B \subset C$, any member of B is a member of C. In particular, $\omega \in C$. Since ω was an arbitrary element of A, this completes the argument.*

Solution 1.2 (Solution to Exercise 1.2)**.** *Let $\omega \in A \cap B$. By Definition 1.1, this means that $\omega \in A$ and $\omega \in B$. In particular, $\omega \in A$. Hence $A \cap B \subset A$. In order to verify the second claim, let us assume that $A \subset B$. We want to show that in this case $A \cap B = A$. Let $\omega \in A$, then since $A \subset B$, $\omega \in B$. In particular, $\omega \in A$ and $\omega \in B$. This shows that $A \subset A \cap B$. We already knwo that it is generally true that $A \cap B \subset A$. Thus, by (1.2) of Proposition 1.1, $A \subset B \Rightarrow A \cap B = A$.*

Solution 1.3 (Solution to Exercise 1.3)**.** *By Definition 1.1, $A \cup B$ must contain all the elements from B. To see this, suppose by contradiction that there exists $\omega \in B$ s.t. $\omega \notin A \cup B$. $\omega \notin A \cup B$ means ω is an element that is not in A and also not in B (otherwise it would be in one or the other). This is not true since $\omega \in B$. Hence $B \subset A \cup B$. Let us verify that if $A \subset B \Rightarrow A \cup B = B$. We have shown that in general $B \subset A \cup B$. Thus, it remains to show that if $A \subset B$, $A \cup B \subset B$. That is clear, since if $\omega \in A \cup B$, then $\omega \in A$ or $\omega \in B$. If $\omega \in B$, we are done. If $\omega \in A$, we know that $\omega \in B$ as well, since $A \subset B$ was assumed.*

Solution 1.4 (Solution to Exercise 1.4)**.** *We show that $A \cup B \subset A \cup (B \setminus A)$ and $A \cup (B \setminus A) \subset A \cup B$. Since $B \setminus A \subset B \subset A \cup B$ and $A \subset A \cup B$, it is clear that $A \cup (B \setminus A) \subset A \cup B$. Let $\omega \in A \cup B$, then $\omega \in A$ or $\omega \in B$. If $\omega \in A$, since $A \subset A \cup (B \setminus A)$, $\omega \in A \cup (B \setminus A)$. If $\omega \in B$, then either $\omega \notin A$, hence $\omega \in B \setminus A \subset A \cup (B \setminus A)$. Or $\omega \in A \cap B \subset A$ and hence $\omega \in A \cup (B \setminus A)$ as well.*

Solution 1.5 (Solution to Exercise 1.5)**.** *Yes, take $A = \emptyset$ and B any arbitrary set.*

Solution 1.6 (Solution to Exercise 1.6)**.** *We proof the claim be induction. The base step is clear. In order to verify the induction step, we choose $n \in \mathbb{N}$, and assume that*

$$\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}.$$

We have that

$$\begin{aligned}\sum_{k=1}^{n+1} k^2 &= \frac{n(n+1)(2n+1)}{6} + (n+1)^2 \\ &= \frac{n(n+1)(2n+1) + 6(n+1)^2}{6} \\ &= \frac{(n+1)(n(2n+1) + 6(n+1))}{6} \\ &= \frac{(n+1)(2n^2 + 7n + 6)}{6} \\ &= \frac{(n+1)(2n^2 + 2n + 5n + 6)}{6} \\ &= \frac{(n+1)((n+2)(2n+n) - n^2 - 4n + 5n + 6)}{6} \\ &= \frac{(n+1)((n+2)(2(n+1) + 1 + n - 3) - n^2 + n + 6)}{6} \\ &= \frac{(n+1)((n+2)(2(n+1)+1) + (n+2)(n-3) - n^2 + n + 6)}{6} \\ &= \frac{(n+1)((n+2)(2(n+1)+1) + n^2 - 3n + 2n - 6 - n^2 + n + 6)}{6} \\ &= \frac{(n+1)(n+2)(2(n+1)+1)}{6}.\end{aligned}$$

Solution 1.7 (Solution to Exercise 1.7)**.**

(a) We have that $x \leq 1$ for any $x \in A$. Thus, 1 is an upper bound for A. Further, $\sup A = 1$, i.e., 1 is the smallest upper bound of A. This is because $1 \in A$ and hence $s' < 1$ can not be a smaller upper bound of A. In particular, as $1 \in A$, the maximum of A is 1. It is true that $x \geq 0$ for any $x \in A$. Thus, 0 is a lower bound for A. Further, 0 is the largest lower bound of A. This is because for any $s' > 0$ there exists $n \in \mathbb{N}$ s.t. $1/n < s'$ (Proposition 1.7). Thus, it can not be the case that there exists a larger lower bound than 0. This shows that $\inf A = 0$. Notice that $0 \notin A$, hence A has no minimum.

(b) First, $0 = \inf B$ and in particular, $0 \in B$. Hence 0 is the minimum of B. Further, for any $x \in B$, $x \leq 2$. That is, 2 is an upper bound for B. By Proposition 1.6, 2 must be the smallest upper bound of B. Hence $\sup B = 2$. We note that since $2 \notin B$, B does not have a maximum.

(c) Clearly, there does not exists $s \in \mathbb{R}$ s.t. $x \geq s$ for any $x \in C$. This shows that $\inf B = -\infty$. In particular, C does not have a minimum. On the other hand, $x \leq 0$ for any $x \in C$. Thus, because $0 \in C$, $\sup C = 0$ and in particular, 0 is the maximum of C.

1.5 Additional exercises

Exercise 1.8. *Let I be an arbitrary set and $\{A_i \colon i \in I\}$ be a collection of subsets of a set Ω. Let $A \subset \Omega$. Show that*

$$A \subset A_i \; \forall\, i \in I \Leftrightarrow A \subset \bigcap_{i \in I} A_i.$$

Exercise 1.9. *Prove the following items of Proposition 1.4:*

- $A \setminus B = A \cap B^c$;
- $(A \subset B) \Rightarrow (B^c \subset A^c)$;
- $(A \cap B)^c = A^c \cup B^c$;
- $(A \cup B)^c = A^c \cap B^c$.

Exercise 1.10. *Let $n \in \mathbb{N}$ and A_i, $i = 1, \ldots, n$, be a collection of subsets of a set Ω. Show that*

$$\left(\bigcup_{i=1}^{n} A_i \right)^c = \bigcap_{i=1}^{n} A_i^c.$$

Exercise 1.11. *Let A be a set with n elements. Show that*

- *the number of permutations of the elements from A is $n!$;*
- *for any $0 \leq k \leq n$, the number of subsets of A having k elements is given by*

$$\frac{n!}{(n-k)!k!}.$$

Note 1: *If $A = \{\omega_1, \ldots, \omega_n\}$, then the vector $(\omega_{n_1}, \ldots, \omega_{n_n})$, $n_1, \ldots, n_n \in \{1, \ldots, n\}$, $n_i \neq n_j$, $i \neq j$, represents a permutation of the elements from A.*
Note 2: *The number $n!/((n-k)!k!)$ is denoted with $\binom{n}{k}$.*

Exercise 1.12. *Verify that*

$$\binom{n+1}{k+1} = \binom{n}{k} + \binom{n}{k+1}.$$

2 Introduction: Part II

2.1 Functions

Definition 2.1. *Let A and B be two sets. A function $f\colon A \to B$ is a rule which assigns to each element $a \in A$ exactly one element $f(a) \in B$.*

Definition 2.2. *Let $f\colon A \to B$ be a function.*

Image: *The image of f under $C \subset A$ is the set*

$$f(C) = \{f(a)\colon a \in C\}. \tag{3}$$

Preimage: *If $D \subset B$, the preimage of f under D is the set*

$$f^{-1}(D) = \{a \in A\colon f(a) \in D\}.$$

Example 2.1. *Let $f(x) = x^2$, $x \in \mathbb{R}$. We have that*

$$f(\mathbb{R}) = \{y\colon y \in [0, \infty)\}.$$

To see it, it is clear that $f(\mathbb{R}) \subset \{y\colon y \in [0, \infty)\}$. For the other inclusion, take $y \in \{y\colon y \in [0, \infty)\}$. If we set $x = \sqrt{y}$, then $x^2 = y$. Thus, $y \in f(\mathbb{R})$. Thus also $\{y\colon y \in [0, \infty)\} \subset f(\mathbb{R})$.

Example 2.2. *A function $f\colon A \to \mathbb{R}^k$ assigns to each element $a \in A$, a vector $f(a) \in \mathbb{R}^k$. We use the notation $f(a) = (f_1(a), \dots, f_k(a))$ for the value of f at a, where $f_i(a)$, $i = 1, \dots, k$, are referred to as the coordinate functions of f.*

To indicate that an element a is assigned to an element $f(a)$ we often use the notation $a \mapsto f(a)$ to define the assignment, i.e., the function.

Definition 2.3. *Let $f\colon A \to B$ be a function.*

Surjective: *f is called surjective, if $f(A) = B$.*

Injective: *f is called injective, if $a_1 \neq a_2 \Rightarrow f(a_1) \neq f(a_2)$.*

Bijective: *f is called bijective if it is surjective and injective.*

Example 2.3. *Let $f\colon \mathbb{R} \to [0, \infty)$, $f(x) = x^2$. We already know that f is surjective. However, it is not injective, since for $x_1 = -1$ and $x_2 = 1$, $f(x_1) = f(x_2) = 1$. This shows that $x \mapsto x^2$ is not bijective.*

Definition 2.4. *Let $f\colon A \to B$ be a function and $E \subset A$. The restriction of f to E is the function $f|_E\colon E \to B$ given by $f|_E(a) = f(a)$ for any $a \in E$.*

Example 2.4. *The function $f\colon \mathbb{R} \to [0, \infty)$, $f(x) = x^2$ is not bijective. However, its restriction to $[0, \infty)$ is.*

Definition 2.5. *Let $f\colon A \to B$ and $g\colon B \to C$ be two functions. The composition $g \circ f$ or $g(f)$ is defined pointwise, i.e.,*

$$(g \circ f)(a) = g(f)(a) = g(f(a)).$$

Thus, $g \circ f\colon A \to C$.

Definition 2.6. *Let $f\colon A \to B$ be a function. A function $g\colon B \to A$ is called an inverse of f if*

$$g(f(a)) = a \;\forall\, a \in A \quad \text{and} \quad f(g(b)) = b \;\forall\, b \in B.$$

If f has an inverse it is called invertible.

Proposition 2.1. *Let $f\colon A \to B$ be a function. If f is bijective, then it is invertible.*

Proof. Since f is surjective, $f(A) = B$. In particular, for any $b \in B$, there exists $a \in A$, s.t. $f(a) = b$. Thus, we can assign to each $b \in B$, a respective element $a \in A$ via the assignment $b \mapsto g(b) = a$, $b = f(a)$. Clearly, g assigns elements from B to elements from A. In order to show that the latter assignment is a function, the assignment rule must be unique in the sense that if $b = f(a) \in f(A)$ then g assigns b to exactly one element $g(b)$ of A. Suppose that this assignment rule is not unique, i.e., there exists $a^* \in A$, $a^* \neq a$ s.t. $g(b) = a^*$, $b = f(a)$ (two outputs for the same input). This can only happen if $f(a) = f(a^*)$, since if $f(a) \neq f(a^*)$, $g(f(a^*)) = a^* \neq a = g(f(a)) = g(b)$, i.e., $g(b) \neq a^*$. But the case $f(a) = f(a^*)$ for $a^* \neq a$ is not possible since f is surjective. Hence, it can not happen that there exists $a^* \in A$, $a^* \neq a$ s.t. $g(b) = a^*$. $\square$

Example 2.5. *Let $f(x) = \mathrm{e}^x$, $x \in \mathbb{R}$, i.e., f is the (natural) exponential function ($\mathrm{e} = \mathrm{e}^1$ is Euler's number). One can show that $f\colon \mathbb{R} \to (0,\infty)$ is bijective where the inverse is given by the natural logarithm $\log(y)\colon (0,\infty) \to \mathbb{R}$.*

Exercise 2.1. *Let $f(x) = 1 - e^{-x}$, $x \in [0,\infty)$. Is $f\colon [0,\infty) \to [0,1)$ invertible? If yes, what is its inverse?*

Example 2.6. *The trigonometric functions $\sin, \cos\colon \mathbb{R} \to \mathbb{R}$ are clearly not bijective. However, $\sin|_{[-\pi/2,\pi/2]}\colon [-\pi/2,\pi/2] \to [0,1]$ and $\cos|_{[0,\pi]}\colon [0,\pi] \to [-1,1]$ are bijective with inverse $\arcsin\colon [-1,1] \to [-\pi/2,\pi/2]$ and $\arccos\colon [-1,1] \to [0,\pi]$. In general, $\sin$ and $\cos$ satisfy the following addition formulas: Given any $\theta_1, \theta_2 \in \mathbb{R}$,*

$$\begin{aligned}
\sin(\theta_1 + \theta_2) &= \sin(\theta_1)\cos(\theta_2) + \sin(\theta_2)\cos(\theta_1)\\
\sin(\theta_1 - \theta_2) &= \sin(\theta_1)\cos(\theta_2) - \sin(\theta_2)\cos(\theta_1)\\
\cos(\theta_1 + \theta_2) &= \cos(\theta_1)\cos(\theta_2) - \sin(\theta_2)\sin(\theta_1)\\
\cos(\theta_1 - \theta_2) &= \cos(\theta_1)\cos(\theta_2) + \sin(\theta_2)\sin(\theta_1).
\end{aligned}$$

As a further example, the cotangent $\cot(\theta) = \cos(\theta)/\sin(\theta)$, $\theta \in (0,\pi)$ is s.t. $\cot\colon (0,\pi) \to \mathbb{R}$ is bijective with inverse $\operatorname{arccot}\colon \mathbb{R} \to (0,\pi)$.

Example 2.7. *Let*

$$U = \{(\rho,\theta) \in \mathbb{R}^2 \colon \rho > 0,\ 0 < \theta < 2\pi\} = (0,\infty) \times (0,2\pi).$$

Define the function

$$T(\rho,\theta) = (\rho\cos(\theta), \rho\sin(\theta)), \quad (\rho,\theta) \in U.$$

Set $V = \mathbb{R}^2 \setminus ([0,\infty) \times \{0\})$, i.e., V is $\mathbb{R}^2$ with the ray $[0,\infty) \times \{0\}$ removed. We notice that $T(U) \subset V$, since if $(x,y) \in T(U)$, then $x = \rho\cos(\theta)$ and $y = \rho\sin(\theta)$ for some $\rho > 0$ and $\theta \in (0,2\pi)$. Thus, if $(x,y) \in V^c$, then, by definition of V, $\sin(\theta) = 0$ and hence, $\theta = \pi$. This is not possible as it implies that $x = -\rho$ and then $(x,y) = (-\rho,0) \in V$. We verify that $T\colon U \to V$ is bijective, i.e., injective and surjective. To see that T is injective we show that for any two elements $x_1 = (\rho_1,\theta_1)$ and $x_2 = (\rho_2,\theta_2)$ in U, $T(x_1) = T(x_2)$ implies that $x_1 = x_2$. We note that $T(x_1) = T(x_2)$ implies that

$$\rho_1\cos(\theta_1) - \rho_2\cos(\theta_2) = 0 \tag{4}$$
$$\rho_1\sin(\theta_1) - \rho_2\sin(\theta_2) = 0. \tag{5}$$

Assume by contradiction that $x_1 \neq x_2$. If $\rho_1 \neq \rho_2$ but $\theta_1 = \theta_2$, we use (5) and conclude that $\sin(\theta_1) = \sin(\theta_2) = 0$. Since $\theta_1, \theta_2 \in (0,2\pi)$, this implies that $\theta_1 = \theta_2 = \pi$. Upon (4), we get that $-(\rho_1 - \rho_2) = 0$, which is a contradiction. For the remaining case assume that $\theta_1 \neq \theta_2$ (and either $\rho_1 = \rho_2$ or $\rho_1 \neq \rho_2$). Notice first that it is not possible that

$\sin(\theta_1) = \sin(\theta_2) = 0$, since this would imply that $\theta_1 = \theta_2 = \pi$. Further, it is not possible that $\cos(\theta_1) = \cos(\theta_2) = 0$, since then either $\theta_1 = \pi/2$ and $\theta_2 = (3\pi)/2$ or vice versa. This implies that either $\rho_1 + \rho_2 = 0$ or $-(\rho_1 + \rho_2) = 0$. Thus, as a first case, we assume that $\sin(\theta_1) \neq 0$ and $\cos(\theta_2) \neq 0$. Then, by (4) and (5),

$$\rho_2 = \rho_1 \frac{\cos(\theta_1)}{\cos(\theta_2)} \quad \textit{and} \quad \rho_1 = \rho_2 \frac{\sin(\theta_2)}{\sin(\theta_1)}.$$

The latter display is equivalent to

$$\sin(\theta_1)\cos(\theta_2) = \sin(\theta_2)\cos(\theta_1).$$

Upon the identity $\sin(\theta_1 - \theta_2) = \sin(\theta_1)\cos(\theta_2) - \sin(\theta_2)\cos(\theta_1)$, we conclude that $\sin(\theta_1 - \theta_2) = 0$. Since $\theta_1 \neq \theta_2$, this implies that $\theta_1 = \theta_2 + \pi$. Since $\sin(\theta_2 + \pi) = -\sin(\theta_2)$, we deduce from (5) that

$$-\rho_1 \sin(\theta_2) - \rho_2 \sin(\theta_2) = -\sin(\theta_2)(\rho_1 + \rho_2) = 0.$$

Thus, $\theta_2 = \pi$. This is not possible as it implies that $\theta_1 = 2\pi$ and we have assumed that $\theta_1 < 2\pi$. The remaining cases are similar and we thus conclude that T is indeed injective. In order to verify that T is also surjective we need to show that $T(U) = V$. We already know that $T(U) \subset V$, hence it remains to show the opposite inclusion. Let $(x, y) \in V$. First, it is not possible that $x = y = 0$, since $y = 0$ implies that $x < 0$. Assume first that $y \neq 0$. We define $\rho = \sqrt{x^2 + y^2}$ and $\theta = \operatorname{arccot}(x/y)$ (cf. Example 2.6). With this choice, $(\rho, \theta) \in U$ and $\rho\cos(\theta) = x$ and $\rho\sin(\theta) = y$. If $y = 0$, then again we define $\rho = \sqrt{x^2 + y^2} = |x| > 0$ and $\theta = \pi$. Again, $(\rho, \theta) \in U$ and $\rho\cos(\theta) = x$ and $\rho\sin(\theta) = y$.

Definition 2.7. *Let $f : I \to \mathbb{R}$, $I \subset \mathbb{R}$ be a function. f is called increasing (resp. decreasing) if $x_1 \leq x_2 \Rightarrow f(x_1) \leq f(x_2)$ (resp. $x_1 \leq x_2 \Rightarrow f(x_2) \leq f(x_1)$). f is strictly increasing (resp. strictly decreasing) if $x_1 < x_2 \Rightarrow f(x_1) < f(x_2)$ (resp. $x_1 < x_2 \Rightarrow f(x_2) < f(x_1)$). f is called monotonic (resp. strictly monotonic) if it is either increasing or decreasing (resp. strictly increasing or decreasing).*

Proposition 2.2. *Let $f : I \to \mathbb{R}$, $I \subset \mathbb{R}$ be a strictly monotonic function. Then, $f : I \to f(I)$ is a bijection. Further, if f is strictly increasing (resp. strictly decreasing) on I, then an inverse of f is strictly increasing (resp. strictly decreasing) on $f(I)$.*

Example 2.8. *Let $f : [0, \infty) \to [0, \infty)$ be the function $f(x) = x^2$. Then, using the same argument as in Example 2.1, we readily see that $f([0, \infty)) = [0, \infty)$. Further, if $x < y$, then $f(x) = x^2 < y^2 = f(y)$. Thus, f is strictly increasing. By Proposition 2.2, we know that f has inverse $f^{-1}(y)$, $y \in [0, \infty)$, which is also strictly increasing. Actually, in this case, we know that $f^{-1}(y) = \sqrt{y}$, $y \in [0, \infty)$, is the unique inverse of f, since $y = x^2$ has only one non-negative solution.*

Exercise 2.2. *Let $x, y \geq 0$. Show that $x < y \Rightarrow x^2 < y^2$.*

In order to verify whether a function f is monotone, the following proposition is helpful.

Proposition 2.3. *Let $a, b \in \mathbb{R}$ and $f : (a, b) \to \mathbb{R}$ be a function that is differentiable on (a, b). That is, the derivative f' of f exists on (a, b). Then:*

- *$f'(x) \geq 0 \; \forall x \in (a, b) \Rightarrow f$ is increasing on (a, b);*
- *$f'(x) > 0 \; \forall x \in (a, b) \Rightarrow f$ is strictly increasing on (a, b);*
- *$f'(x) \leq 0 \; \forall x \in (a, b) \Rightarrow f$ is decreasing on (a, b);*
- *$f'(x) < 0 \; \forall x \in (a, b) \Rightarrow f$ is strictly decreasing on (a, b).*

Remark 2.1. *Upon the latter proposition, the claim of Exercise 2.2 follows readily.*

Exercise 2.3. *Are the following functions monotone. If yes, are they increasing or decreasing (resp. strictly increasing or decreasing).*

(a) $f\colon (0,\infty) \to (0,\infty)$, $f(x) = 1/x$;

(b) $g\colon (0,\infty) \to \mathbb{R}$, $g(x) = \log(x)$;

(c) $h\colon \mathbb{R} \to [0,\infty)$, $h(x) = x^4$.

Definition 2.8. *Let $f, g\colon A \to \mathbb{R}$ be two real-valued functions.*

Sum: *$f + g$ is the function $a \mapsto (f+g)(a) = f(a) + g(a)$;*

Product: *fg is the function $a \mapsto (fg)(a) = f(a)g(a)$;*

Quotient: *If $g(a) \neq 0\ \forall\, a \in A$, then f/g is the function $a \mapsto (f/g)(a) = f(a)/g(a)$.*

Proposition 2.4. *Let $f\colon A \to B$ be a function. Let $B_* \subset B$. Then,*

(a) $f^{-1}(B_*^c) = f^{-1}(B_*)^c$.

Let I and J be some sets and $A_i \subset A$, $i \in I$, and $B_j \subset B$, $j \in J$, be a collection of sets from A and B, respectively. Then,

(b) $f(\cup_{i\in I} A_i) = \cup_{i\in I} f(A_i)$;

(c) $f^{-1}(\cup_{j\in J} B_j) = \cup_{j\in J} f^{-1}(B_j)$;

(d) $f^{-1}(\cap_{j\in J} B_j) = \cap_{j\in J} f^{-1}(B_j)$.

Definition 2.9. *Let $f\colon A \to \overline{\mathbb{R}}$ be a function and $E \subset A$. We use the notation*

$$\sup f(E) = \sup\{f(e)\colon e \in E\} = \sup_{e\in E} f(e),$$

for the supremum of $f(E)$. Similarly, we write

$$\inf f(E) = \inf\{f(e)\colon e \in E\} = \inf_{e\in E} f(e).$$

for the infimum of $f(E)$.

2.2 Cardinality of Sets

Definition 2.10. *Let A be a set. A is said to be finite if it contains a finite number of elements. Otherwise, if the number of elements in A is not finite, A is referred to as infinite.*

Example 2.9. *The sets $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$ and $\mathbb{R}$ are all infinite. There does not exist a number which counts the number of elements within the latter sets.*

Definition 2.11. *Let A and B be two arbitrary sets. A and B are said to have the same cardinality ($\#A = \#B$) if and only if there exists a bijection $f\colon A \to B$.*

In terms of a visual interpretation, if $\#A = \#B$, then we can draw a line from each element of A to an element of B (surjectivity) and it is not possible that a line emerges from two different elements of A to the same element of B.

Example 2.10. *If A and B are finite, then $\#A = \#B$ precisely means that A and B have the same number of elements ($\#A$ is given by the number of connecting lines between A and B). In particular, if we let $A = B$, $\#A$ gives the number of elements in A. This shows that $\#\emptyset = 0$, since $\emptyset$ is clearly finite.*

Definition 2.12. *Let A and B be two sets and $C \subset B$.*

- *If there exists a bijection $g\colon A \to C$ we write $\#A \leq \#B$;*
- *If there exists a bijection $g\colon A \to C$ but there exists no bijection $f\colon A \to B$, we write $\#A < \#B$.*

Example 2.11. *Let A and B be two sets. According to Definition 2.12, if $A \subset B$, $\#A \leq \#B$. This is because the identity map $g\colon A \to A$, $g(a) = a$, $a \in A$, is a bijection.*

Proposition 2.5. *We have that*

- *$\#\mathbb{N} = \#\mathbb{Z} = \#\mathbb{Q}$;*
- *$\#\mathbb{R} > \#\mathbb{N}$;*
- *any interval $I \subset \mathbb{R}$ is s.t. $\#I > \#\mathbb{N}$.*

Definition 2.13. *Let A be a set. If*

- *$\#A \leq \#\mathbb{N}$, then A is countable;*
- *$\#A > \#\mathbb{N}$, then A is uncountable.*

Proposition 2.6. *Let A be a set. If A is countable but not finite, then $\#A = \#\mathbb{N}$ and A is said to be countably infinite.*

Proposition 2.7. *Let $\{A_i\colon i \in \mathbb{N}\}$ be a collection of sets s.t. for any $i \in \mathbb{N}$, A_i is countable. Then the union $\cup_{i\in\mathbb{N}} A_i$ is countable as well.*

Proposition 2.8. *Let A_i, $i = 1, \ldots, n$, be countable sets. Then, their Cartesian product, $\prod_{i=1}^n A_i$ is countable.*

We remark that if $\{A_i\colon i \in I\}$ is some collection of sets, where I is some set, then if I is countable we might always set $I = \mathbb{N}$ or $I = \{1, \ldots, n\}$, $n \in \mathbb{N}$.

To conclude this section we state the following result:

Proposition 2.9. *Let $[a, b]$, $a < b \in \mathbb{R}$, be any closed interval. Let I be some countable set ($\#I \leq N$) and assume that there exists a family of open intervals (a_i, b_i), a_i, b_i, $i \in I$, s.t., $[a, b] \subset \cup_{i\in I}(a_i, b_i)$. Then, there exists $N \in \mathbb{N}$ and $i_1, \ldots, i_N \in I$, s.t. $[a, b] \subset \cup_{j=1}^N (a_{i_j}, b_{i_j})$.*

The latter result is known as the Heine-Borel theorem for intervals, it shows that any closed interval on the real line which is covered by a countable collection of open intervals can be covered by a finite sub-collection.

2.3 Euclidean distance

Definition 2.14. *Given two points $x, y \in \mathbb{R}^k$, the Euclidean distance between x and y is given by*

$$\|x - y\| = \sqrt{(x_1 - y_1)^2 + \cdots (x_k - y_k)^2}, \quad x = (x_1, \ldots, x_k),\ y = (y_1, \ldots, y_k).$$

Example 2.12. *If $k = 1$, then we write $\|x - y\| = |x - y|$, where the function $x \mapsto |x|$ maps $x \in \mathbb{R}$ to its absolute value*

$$|x| = \begin{cases} x, & \text{if } x \geq 0, \\ -x, & \text{if } x < 0. \end{cases}$$

Proposition 2.10. *For any $x, y, z \in \mathbb{R}^k$, The Euclidean distance satisfies:*

(i) $\|x - y\| \geq 0$ *and* $\|x - y\| = 0 \Leftrightarrow x = y$;

(ii) $\|x - y\| = \|y - x\|$ *(symmetry);*

(iii) $\|x + y\| \leq \|x\| + \|y\|$ *(triangular inequality);*

(iv) $\|x - y\| \geq \big|\|x\| - \|y\|\big|$ *(reverse triangular inequality).*

Notice that we use the notation 0 for the vector in $\mathbb{R}^k$, $k \in \mathbb{N}$, with all of its coordinates equal to zero.

Example 2.13. *In Proposition 1.6 of Example 1.4 we saw that there exists a rational number in between any two distinct real numbers. Let $\varepsilon > 0$ and $x \in \mathbb{R}$. Chose $n \in \mathbb{N}$ s.t. $1/n < \varepsilon$. Then, $y = x + 1/n$ is s.t. $|x - y| = 1/n < \varepsilon$. Still, $x \neq y$ and hence by Proposition 1.6 we find $q \in \mathbb{Q}$ s.t. $x < q < y$. In particular, $0 < q - x < y - x$ and hence $|q - x| < \varepsilon$. Since $\varepsilon > 0$ and $x \in \mathbb{R}$ were arbitrary, we can conclude that for any $\varepsilon > 0$ and $x \in \mathbb{R}$, there exists $q \in \mathbb{Q}$ s.t. $|x - q| < \varepsilon$.*

Definition 2.15. *An open (resp. closed) ball of radius $r > 0$ with center $y \in \mathbb{R}^k$ is denoted with*

$$B_r(y) = \{x \in \mathbb{R}^k \colon \|y - x\| < r\} \ (\textit{resp. } B_r[y] = \{x \in \mathbb{R}^k \colon \|y - x\| \leq r\})$$

Definition 2.16. *A set $U \subset \mathbb{R}^k$ is called open if for any point $x \in U$, there exists $\varepsilon > 0$, s.t. $B_\varepsilon(x) \subset U$. That is any point in U is the center of an open Ball contained in U.*

Example 2.14. *If U_1 and U_2 are two open subsets of $\mathbb{R}^k$, then $U_1 \cap U_2$ is open. By definition, if $x \in U_1 \cap U_2$, there exist $B_{\varepsilon_1}(x)$ and $B_{\varepsilon_2}(x)$ s.t. $B_{\varepsilon_1}(x) \subset U_1$ and $B_{\varepsilon_2}(x) \subset U_2$. Thus, set $\varepsilon = \min\{\varepsilon_1, \varepsilon_2\}$ and we obtain that $B_\varepsilon(x) \subset U_1 \cap U_2$.*

Example 2.15. *Let $a, b \in \mathbb{R}$. Then, the open interval (a, b) is an open set of $\mathbb{R}$. Let $x \in (a, b)$, i.e., $a < x < b$. Let $\varepsilon < \min\{|x - a|, |x - b|\}$. Then, take any $y \in (x - \varepsilon, x + \varepsilon) = B_\varepsilon(x)$. It follows that $x - y \leq |x - y| < |x - a| = x - a$ and hence, $y > a$. Also, $y - x \leq |x - y| < |x - b| = b - x$, i.e., $y < b$. Therefore, $y \in (a, b)$.*

Example 2.16. *An open ball $B_r(y) \subset \mathbb{R}^k$ is an open set. This is a consequence of the triangular inequality. Let $x \in B_r(y)$ be an arbitrary point and set $\delta = \|x - y\|$. Then, by definition of $B_r(y)$, $\delta < r$. Let $\varepsilon = r - \delta$. Then, $B_\varepsilon(x) \subset B_r(y)$, since for any $z \in B_\varepsilon(x)$, $\|z - y\| \leq \|z - x\| + \|x - y\| = \varepsilon + \delta = r$. Since $x \in B_r(y)$ was arbitrary, the result follows.*

Exercise 2.4. *Show that the open rectangle $\prod_{i=1}^k (a_i, b_i)$, a_i, b_i, $i = 1, \ldots, k$, is an open set of $\mathbb{R}^k$.*

We remark that continuous functions on Euclidean spaces are characterized in terms of open sets.

Definition 2.17. *A function $f \colon \mathbb{R}^m \to \mathbb{R}^k$ is continuous if for any open set $U \subset \mathbb{R}^k$, the set $f^{-1}(U)$ is open in $\mathbb{R}^m$.*

Example 2.17. *Let $f(x) = x$, $x \in \mathbb{R}$. Then, $f \colon \mathbb{R} \to \mathbb{R}$ is continuous. Take any $U \subset \mathbb{R}$ open. Then, $f^{-1}(U) = U$. Hence, $f^{-1}(U)$ is an open subset of $\mathbb{R}$.*

A motivation for Definition 2.17 is given in the appendix (Section A.4). In terms of functions defined on subsets of $\mathbb{R}^m$, continuity is characterized as follows (a proof is given in Section A.4):

Proposition 2.11. *Let $E \subset \mathbb{R}^m$, $E \neq \emptyset$. A function $f \colon E \to \mathbb{R}^k$ is continuous if and only if for any open set $U \subset \mathbb{R}^k$,*

$$f^{-1}(U) \in \{G \cap E \colon G \textit{ open in } \mathbb{R}^m\}.$$

We recall some classical examples of continuous functions.

Proposition 2.12. *The following functions are continuous:*

- $f\colon \mathbb{R}^k \to \mathbb{R}$, $f(x) = \sum_{i=1}^{k} x_k$, $x = (x_1, \ldots, x_k)$;
- $g\colon \mathbb{R}^k \to \mathbb{R}$, $g(x) = \prod_{i=1}^{k} x_k$, $x = (x_1, \ldots, x_k)$;
- $h\colon \mathbb{R} \setminus \{0\} \to \mathbb{R} \setminus \{0\}$, $h(x) = 1/x$.

For more details on the notion of continuity we refer to Section A.4.

Definition 2.18. *A set $V \subset \mathbb{R}^k$ is said to be closed if V^c is open.*

Example 2.18. *Given $r > 0$ and $y \in \mathbb{R}^k$, a closed ball $B_r[y] \subset \mathbb{R}^k$ is closed. To see it, we show that $B_r[y]^c = \mathbb{R}^k \setminus B_r[y]$ is open in $\mathbb{R}^k$. Let $x \in B_r[y]^c$, i.e., $\|x - y\| > r$. Set $\delta = \|x-y\| - r$. Then, $\delta > 0$. Consider the open ball $B_\delta(x)$. If we show that $B_\delta(x) \subset B_r[y]^c$, we are done. Hence, let $z \in B_\delta(x)$. Using the reverse triangular inequality (item (iv) of Proposition 2.10),*

$$\|z - y\| \geq \|x - y\| - \|z - x\| > \|x - y\| - \delta = r,$$

i.e., $z \in B_r[y]^c$ and we are done.

Example 2.19. *The set $V = \{(x, y) \in \mathbb{R}^2 \colon x \geq 0, y = 0\}$ is a closed set of $\mathbb{R}^2$. We show that V^c is open. Let $v = (v_1, v_2) \in V^c$. If $v_2 = 0$, then $v_1 < 0$ otherwise $v \in V$. Hence, if $v_2 = 0$, $v \in V^c$ implies that $B_r(v) \subset V^c$ with $r = |v_1|$. To see it, it is sufficient to verify that $z = (a, b) \in B_r(v)$ implies that $a < 0$ (this implies that $(a, b) \in V^c$). We notice that $z \in B_r(v)$ implies that*

$$(a - v_1)^2 + b^2 < v_1^2 \Leftrightarrow a^2 + b^2 < 2av_1.$$

Hence, by the previous inequality, since $v_1 < 0$ it can not be the case that $a \geq 0$. The remaining case is that $v_2 \neq 0$. Set $r = |v_2|$. We verify that $B_r(v) \subset V^c$. Let $z = (a, b) \in B_r(v)$. We verify that $b \neq 0$ (this shows that $(a, b) \in V^c$). We note that $z \in B_r(v)$ implies that $(a - v_1)^2 + (b - v_2)^2 < v_2^2$. But then, $b = 0$ is not possible, as it would imply that $(a - v_1)^2 < 0$. Hence, the set V^c is open and thus V is closed.

Definition 2.19. *A set $A \subset \mathbb{R}^k$ is said to be bounded if there exists $r > 0$ and $y \in A$ s.t. $A \subset B_r[y]$, i.e., A is contained in a closed ball.*

Definition 2.20. *Let $f\colon A \to \overline{\mathbb{R}}$ be a function and $E \subset A$, where A is some set. The function f is said to be bounded on E if there exists $0 \leq M < \infty$, s.t. $|f(a)| \leq M$ for any $a \in E$.*

2.4 Solution to exercises

Solution 2.1 (Solution to Exercise 2.1)**.** *Yes, f is invertible. The inverse of f is the function $g\colon [0, 1) \to [0, \infty)$, $g(y) = -\log(1 - y)$.*

Solution 2.2 (Solution to Exercise 2.2)**.** *We show first that if $a \in \mathbb{R}$ s.t. $a > 1$, then, for any $z \in (0, \infty)$, $az > z$. Write $a = 1 + \varepsilon$ with $\varepsilon = a - 1 > 0$. Then, $az = z + \varepsilon z$. Since $\varepsilon, z > 0$, $az > z$. Since $x < y$, $0 < (x - y)^2 = x^2 + y^2 - 2xy$. Then, since $x < y$, $2xy \geq 2x^2$. Notice that this is because $x \geq 0$. If $x = 0$, then $2xy = 0 = 2x^2$. Otherwise, if $x > 0$, then, since $x < y$, $y/x > 1$. Hence, $2xy = 2x^2x^{-1}y > 2x^2$. In particular, $0 < x^2 + y^2 - 2xy < x^2 + y^2 - 2x^2 = y^2 - x^2$. This solves the exercises.*

Solution 2.3 (Solution to Exercise 2.3)**.**

(a) *Given any* $0 < a < b < \infty$, $f'(x) = -1/x^2$, $x \in (a,b)$. *Thus,* $f'(x) < 0$ *for any* $x \in (a,b) \subset (0,\infty)$. *Since* a *and* b *were arbitrary, we easily verify that* f *must be strictly decreasing on* $(0,\infty)$. *If not, there exists* $x_1, x_2 \in (0,\infty)$ *with* $x_1 < x_2$ *and* $f(x_2) > f(x_1)$. *This is not possible, since we always find* $a < x_1 < x_2 < b$, *and upon Proposition 2.3, since* f *has strictly negative derivative on* (a,b), f *is strictly decreasing on* (a,b), *in particular* $f(x_2) < f(x_1)$.

(b) *Given* $a, b \in \mathbb{R}$, *we calculate* $g'(x) = 1/x$, $x \in (0,\infty)$. *Then, using the same reasoning as in (a), we conclude that* g *is strictly increasing.*

(c) h *is not monotone. To see it, we notice that* $h(-1) \geq h(x)$ *for any* $x \in [-1,1]$ *but* $h(x) > h(-1)$ *for any* $x \in (1,2)$.

Solution 2.4 (Solution to Exercise 2.4)**.** *We need to show that for any* $x \in \prod_{i=1}^k (a_i,b_i)$, *there exists* $\varepsilon > 0$, *s.t.* $B_\varepsilon(x) \subset \prod_{i=1}^k (a_i,b_i)$ *(cf. Definition 2.16). Let* $x = (x_1,\ldots,x_k) \in \prod_{i=1}^k (a_i,b_i)$ *and set* $m_i = \min\{|x_i - a_i|, |x_i - b_i|\}$, $i = 1,\ldots,k$. *Then, since for any* $i = 1,\ldots,k$, $m_i > 0$, *choose* $\varepsilon < \min\{m_i \colon i = 1,\ldots,k\}$. *Then, let* $y = (y_1,\ldots,y_k) \in B_\varepsilon(x)$. *For any* $i = 1,\ldots,k$, *we have that*

$$|y_i - x_i| \leq \|y - x\| < \varepsilon < m_i.$$

In particular, $x_i - y_i \leq |y_i - x_i| < |x_i - a_i| = x_i - a_i$ *and therefore,* $y_i > a_i$. *Also,* $y_i - x_i \leq |y_i - x_i| < |x_i - b_i| = b_i - x_i$ *and hence* $y_i < b_i$. *This shows that* $y_i \in (a_i,b_i)$ *for any* $i = 1,\ldots,k$.

2.5 Additional exercises

Exercise 2.5. *Show that for any* $x, y \in \mathbb{R}$ *and* $n \in \mathbb{N}$, $(x+y)^n = \sum_{k=0}^n \binom{n}{k} x^k y^{n-k}$.

Exercise 2.6. *Prove Proposition 2.4.*

Exercise 2.7. *Let* $f \colon \mathbb{N} \to \mathbb{Z}$ *be defined as follows:*

$$f(n) = \begin{cases} \frac{n}{2}, & \text{if } n \text{ is even,} \\ \frac{1-n}{2}, & \text{if } n \text{ is odd.} \end{cases}$$

Show that f *is bijective.*
Note: *This shows that* $\#\mathbb{N} = \#\mathbb{Z}$.

Exercise 2.8. *Let* A *and* B *be two sets. Show that if* A *and* B *are countable, then,* $A \cup B$ *is countable.*

Exercise 2.9. *Show that,*

(a) *if* $f \colon A \to B$ *and* $g \colon B \to C$ *are two functions s.t.* f *and* g *are bijective, then, the composition* $g \circ f \colon A \to C$ *is bijective as well;*

(b) $\#\mathbb{R} = \#(0,1)$.

3 Introduction: Part III

3.1 Real valued sequences

Definition 3.1. *A real-valued sequence is a function $f\colon \mathbb{N} \to \mathbb{R}$, i.e., $f(n) \in \mathbb{R}$ for any $n \in \mathbb{N}$. We use the notation $f = (a_n)_{n\in\mathbb{N}}$ for a real-valued sequence and $f(n) = a_n$ for the values of f at n.*

This section only treats real-valued sequences. Thus, for now, a sequence is a real-valued sequence.

Definition 3.2. *Let $(a_n)_{n\in\mathbb{N}}$ be a sequence. $(a_n)_{n\in\mathbb{N}}$ is said to be convergent if there exists a number $a \in \mathbb{R}$ s.t. for any $\varepsilon > 0$ there exists $N \in \mathbb{N}$ s.t. $|a_n - a| < \varepsilon$ for any $n \geq N$.*

The number a in Definition 3.2 is called the limit of $(a_n)_{n\in\mathbb{N}}$.

Example 3.1. *Let $a_n = 1/n$, $n \in \mathbb{N}$. We show that $(a_n)_{n\in\mathbb{N}}$ is convergent with limit 0. Let $\varepsilon > 0$ be an arbitrary strictly positive real number. According to Proposition 1.7, we pick $N \in \mathbb{N}$ s.t. $N > 1/\varepsilon$. Then, for any $n \geq N$,*

$$|a_n - 0| = \frac{1}{n} \leq \frac{1}{N} < \varepsilon.$$

Exercise 3.1. *If $a_n = c$ for any $n \in \mathbb{N}$ then $(a_n)_{n\in\mathbb{N}}$ is convergent with limit c*

Exercise 3.2. *Let $a_n = (-1)^n$, $n \in \mathbb{N}$. Is $(a_n)_{n\in\mathbb{N}}$ convergent? Try to only use Definition 3.2.*

Proposition 3.1. *If $(a_n)_{n\in\mathbb{N}}$ is convergent, then its limit a is unique and we write*

$$\lim_{n\to\infty} a_n = a \quad \text{or} \quad a_n \xrightarrow{n\to\infty} a.$$

In the following, we list some important results on real valued sequences.

Definition 3.3. *A sequence $(a_n)_{n\in\mathbb{N}}$ is said to be bounded if there exists $M > 0$ s.t. for any $n \in \mathbb{N}$, $|a_n| \leq M$. $(a_n)_{n\in\mathbb{N}}$ is said to be bounded from below (resp. above) if there exists $M \in \mathbb{R}$ s.t. $a_n \geq M$ (resp. $a_n \leq M$) for any $n \in \mathbb{N}$.*

Proposition 3.2. *If $(a_n)_{n\in\mathbb{N}}$ is convergent, then it is bounded.*

Definition 3.4. *Let $(a_n)_{n\in\mathbb{N}}$ be a sequence. $(a_n)_{n\in\mathbb{N}}$ is increasing (resp. decreasing) if $a_n \leq a_{n+1}\ \forall\, n \in \mathbb{N}$ (resp. $a_n \geq a_{n+1}\ \forall\, n \in \mathbb{N}$). $(a_n)_{n\in\mathbb{N}}$ is said to be monotonic if it is either increasing or decreasing.*

Definition 3.5. *If $(a_n)_{n\in\mathbb{N}}$ is increasing (resp. decreasing) with limit a, we write $a_n \uparrow a$ (resp. $a_n \downarrow a$).*

Proposition 3.3. *A bounded and monotonic sequence $(a_n)_{n\in\mathbb{N}}$ is convergent.*

Example 3.2. *Let $|r| < 1$, and consider $a_n = |r|^n$, $n \in \mathbb{N}$. For any $n \in \mathbb{N}$, $a_n < 1$ and*

$$|r|^{n+1} = |r|^n|r| \leq |r|^n.$$

Thus, $(a_n)_{n\in\mathbb{N}}$ is bounded and decreasing. By Proposition 3.3, there exists L s.t.,

$$\lim_{n\to\infty} a_n = L.$$

Proposition 3.4. *Let $(a_n)_{n\in\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}}$ be two convergent sequences s.t. $a_n \xrightarrow{n\to\infty} a$ and $b_n \xrightarrow{n\to\infty} b$. Then,* *(i)* $a_n + b_n \xrightarrow{n\to\infty} a + b$;

(ii) $a_n b_n \xrightarrow{n\to\infty} ab$;

(iii) $a_n/b_n \xrightarrow{n\to\infty} a/b$, *if* $b \neq 0$.

Proposition 3.5. *Let* $(a_n)_{n\in\mathbb{N}}$ *and* $(b_n)_{n\in\mathbb{N}}$ *be two convergent sequences s.t.* $a_n \xrightarrow{n\to\infty} a$ *and* $b_n \xrightarrow{n\to\infty} b$. *Assume that* $a_n \leq b_n$ *(resp.* $a_n \geq b_n$*) for any* $n \in \mathbb{N}$, *then* $a \leq b$ *(resp.* $a \geq b$*).*

Proposition 3.6. *Let* $(a_n)_{n\in\mathbb{N}}$ *and* $(b_n)_{n\in\mathbb{N}}$ *be two convergent sequences that converge to the same limit, i.e.,* $a_n \xrightarrow{n\to\infty} a$ *and* $b_n \xrightarrow{n\to\infty} a$. *Let* $(c_n)_{n\in\mathbb{N}}$ *be another sequence which is s.t. for any* $n \in \mathbb{N}$, $a_n \leq c_n \leq b_n$. *Then,* $c_n \xrightarrow{n\to\infty} a$.

Example 3.3. *Let* $|r| < 1$, *and consider* $a_n = r^n$, $n \in \mathbb{N}$. *We show that* $(a_n)_{n\in\mathbb{N}}$ *is convergent with limit* 0. *To do so, we consider* $(b_n)_{n\in\mathbb{N}}$, *with* $b_n = |r|^n$, $n \in \mathbb{N}$ *as in Example 3.2. We have that*

$$r^n = \begin{cases} r^n, & r \in [0,1), \\ (-1)^n |r|^n, & r \in (-1,0). \end{cases}$$

In particular, for any $n \in \mathbb{N}$,

$$-|r|^n \leq r^n \leq |r|^n.$$

Thus, by Propositions 3.4 and 3.6, it is sufficient to show that $(b_n)_{n\in\mathbb{N}}$ *converges to zero. We already know (see Example 3.2) that there exists* L *s.t.,*

$$\lim_{n\to\infty} b_n = L.$$

Now we have that

$$L = \lim_{n\to\infty} b_n = \lim_{n\to\infty} b_{n+1} = |r| \lim_{n\to\infty} b_n = |r|L.$$

Thus, since $r \notin \{-1, 1\}$, *it must be the case that* $L = 0$. *In conclusion,* $\lim_{n\to\infty} a_n = 0$.

Example 3.4. *In Example 2.13 we have seen that for any real number* x, *the Euclidean distance between* x *and elements from* $\mathbb{Q}$ *can be made arbitrary small. Let us show that for any* $x \in \mathbb{R}$, *there exists a sequence of rational numbers* $(q_n)_{n\in\mathbb{N}}$, $q_n \in \mathbb{Q}$, $n \in \mathbb{N}$, *s.t.* $q_n \uparrow x$. *By Proposition 1.6, for any* $\varepsilon > 0$, *there exists* $q \in \mathbb{Q}$, *s.t.* $x - \varepsilon < q \leq x$. *For* $n = 1$, *choose* $q_1 \in (x-1, x] \setminus (x - 1/2, x]$. *For* $n = 2$, *choose* $q_2 \in (x - 1/2, x] \setminus (x - 1/3, x]$ *and so on until we choose* q_n *in*

$$\left(x - \frac{1}{n}, x\right] \setminus \left(x - \frac{1}{n+1}, x\right].$$

Then, $q_n < q_{n+1}$ *and for any* $n \in \mathbb{N}$ *and*

$$x - \frac{1}{n} < q_n \leq x.$$

Using Proposition 3.6, this shows that $q_n \uparrow x$. *A similar argument shows there exists a sequence* $(q_n^*)_{n\in\mathbb{N}}$, $q_n^* \in \mathbb{Q}$, $n \in \mathbb{N}$, *s.t.* $q_n^* \downarrow x$.

Exercise 3.3. *Let*

$$a_n = \frac{n^2 + 3n^3 + n}{1 + n^4}, \quad n \in \mathbb{N}.$$

Is $(a_n)_{n\in\mathbb{N}}$ *convergent? If yes, what is its limit?*

Exercise 3.4. *Let $a_n = n!/n^n$, $n \in \mathbb{N}$. Is $(a_n)_{n\in\mathbb{N}}$ convergent? If yes, what is its limit?*

Definition 3.6. *Let $(a_n)_{n\in\mathbb{N}}$ be a sequence. We write:*

(i) $a_n \xrightarrow{n\to\infty} \infty$ (or $\lim_{n\to\infty} a_n = \infty$) if $\forall\, M \in \mathbb{R}\ \exists\, N \in \mathbb{N}$ s.t. $a_n \geq M\ \forall\, n \geq N$.

(ii) $a_n \xrightarrow{n\to\infty} -\infty$ (or $\lim_{n\to\infty} a_n = -\infty$) if $\forall\, M \in \mathbb{R}\ \exists\, N \in \mathbb{N}$ s.t. $a_n \leq M\ \forall\, n \geq N$.

If $\lim_{n\to\infty} a_n = \infty$ (resp. $\lim_{n\to\infty} a_n = -\infty$) we say that $(a_n)_{n\in\mathbb{N}}$ diverges to ∞ (resp. $-\infty$). We say that $(a_n)_{n\in\mathbb{N}}$ diverges if it either diverges to ∞ or $-\infty$.

Remark 3.1. *Let $(a_n)_{n\in\mathbb{N}}$ be a sequence. For now, if $\lim_{n\to\infty} a_n$ is well defined, i.e., $\lim_{n\to\infty} a_n \in \mathbb{R}$ (i.e., $(a_n)_{n\in\mathbb{N}}$ converges) or $(a_n)_{n\in\mathbb{N}}$ diverges, we understand $\lim_{n\to\infty} a_n$ as an element of the extended real line $\overline{\mathbb{R}}$ (cf. Remark 1.1). We write that $\lim_{n\to\infty} a_n$ exists, if $\lim_{n\to\infty} a_n \in \overline{\mathbb{R}}$. Further, if $\lim_{n\to\infty} a_n$ exists, then it is unique.*

Proposition 3.7. *Let $(a_n)_{n\in\mathbb{N}}$ be a monotone sequence. Then, $\lim_{n\to\infty} a_n$ exists. If $(a_n)_{n\in\mathbb{N}}$ is increasing and $(a_n)_{n\in\mathbb{N}}$ diverges, then it diverges to ∞. If $(a_n)_{n\in\mathbb{N}}$ is decreasing and $(a_n)_{n\in\mathbb{N}}$ diverges, then it diverges to $-\infty$.*

We notice that for increasing (resp. decreasing) sequences, Proposition 3.5 remains true even if the sequences do not converge.

Proposition 3.8. *Let $(a_n)_{n\in\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}}$ be two increasing (resp. decreasing) sequences. Then, if for any $n \in \mathbb{N}$, $a_n \leq b_n$, $\lim_{n\to\infty} a_n \leq \lim_{n\to\infty} b_n$.*

Exercise 3.5. *Prove Proposition 3.8.*

Definition 3.7. *Let $(a_i)_{i\in\mathbb{N}}$ be a sequence. The series*

$$\sum_{i\in\mathbb{N}} a_i = \sum_{i=1}^{\infty} a_i,$$

is understood as the sequence $(s_n)_{n\in\mathbb{N}}$, where $s_n = \sum_{i=1}^{n} a_i$, $n \in \mathbb{N}$. If $\lim_{n\to\infty} s_n$ exists we write $\lim_{n\to\infty} s_n = \sum_{i=1}^{\infty} a_i$ for the limit.

Proposition 3.9. *Let $\sum_{i\in\mathbb{N}} a_i$ be a series where $a_i \geq 0$ for any $i \in \mathbb{N}$. Then, either $\sum_{i\in\mathbb{N}} a_i < \infty$ or $\sum_{i\in\mathbb{N}} a_i = \infty$.*

Proof. This follows from Proposition 3.7. Notice that since $a_i \geq 0$ for any $i \in \mathbb{N}$, the sequence $(s_n)_{n\in\mathbb{N}}$, $s_n = \sum_{i=1}^{n} a_i$, is increasing. □

Example 3.5. *We have that*

$$\sum_{i\in\mathbb{N}} \frac{1}{i} = \infty,$$

i.e., the sequence $(s_n)_{n\in\mathbb{N}}$, $s_n = \sum_{i=1}^{n}(1/i)$, diverges to ∞. We will give an argument in the next section.

Example 3.6. *We have that*

$$\sum_{i\in\mathbb{N}} \frac{1}{i(i+1)} = 1.$$

We notice that

$$\frac{1}{i} - \frac{1}{i+1} = \frac{1}{i(i+1)}.$$

Then, for any $n \in \mathbb{N}$,

$$\begin{aligned} s_n &= \sum_{i=1}^{n} \frac{1}{i(i+1)} = \sum_{i=1}^{n} \left(\frac{1}{i} - \frac{1}{i+1}\right) \\ &= 1 - \frac{1}{1+1} + \frac{1}{2} - \frac{1}{2+1} + \frac{1}{3} - \frac{1}{3+1} + \cdots - \frac{1}{n-1+1} + \frac{1}{n} - \frac{1}{n+1} \\ &= 1 - \frac{1}{n+1}. \end{aligned}$$

Therefore, since $\lim_{n\to\infty} s_n = 1$, we have that $\sum_{i\in\mathbb{N}} 1/(i(i+1)) = 1$.

Example 3.7. *The functions $x \mapsto \mathrm{e}^x$, $x \mapsto \sin(x)$ and $x \mapsto \cos(x)$ are all defined in terms of a series:*

- $\mathrm{e}^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}$, $x \in \mathbb{R}$;
- $\sin(x) = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} x^{2k+1}$, $x \in \mathbb{R}$;
- $\cos(x) = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!} x^{2k}$, $x \in \mathbb{R}$.

To conclude this section, we list two useful results for series.

Proposition 3.10. *Let $\sum_{i\in\mathbb{N}} a_i$ be a series and $\sum_{i\in\mathbb{N}} b_i$ be a series s.t. $b_i \geq 0$ for any $i \in \mathbb{N}$ and $\sum_{i\in\mathbb{N}} b_i < \infty$. Suppose that $|a_i| \leq b_i$ for any $i \in \mathbb{N}$. Then $\sum_{i\in\mathbb{N}} a_i < \infty$.*

Proposition 3.11. *Let $I, J \subset \mathbb{N}$ and $f\colon I \times J \to \mathbb{R}$. For any $i, j \in \mathbb{N}$, set $a_{ij} = f(i,j)$. Thus, we obtain a doubly indexed sequence of real numbers $(a_{ij})_{(i,j)\in I\times J}$. Suppose that either $a_{ij} \geq 0$ for any $(i,j) \in I \times J$ or $\sum_{(i,j)\in I\times J} |a_{ij}| < \infty$. Then, $\sum_{(i,j)\in I\times J} a_{ij}$ is well defined and*

$$\sum_{(i,j)\in I\times J} a_{ij} = \sum_{i\in I}\left(\sum_{j\in J} a_{ij}\right) = \sum_{j\in J}\left(\sum_{i\in I} a_{ij}\right). \tag{6}$$

That is, we are allowed to change the order of summation. If $I = J$, we use the notation $\sum_{(i,j)\in I^2} a_{ij} = \sum_{i,j\in I} a_{ij}$ for the sum over all the pairs $(i,j) \in I^2$.

3.2 Subsequences: Limit inferior and limit superior

We remain in the setting of the previous Section, i.e., any sequence $(a_n)_{n\in\mathbb{N}}$ is a real-valued sequence according to Definition 3.1.

Definition 3.8. *Let $f = (a_n)_{n\in\mathbb{N}}$ be a sequence (cf. Definition 3.1). A subsequence of $(a_n)_{n\in\mathbb{N}}$ is a new sequence $g = (b_n)_{n\in\mathbb{N}}$, where $g = f \circ s$, with $s\colon \mathbb{N} \to \mathbb{N}$ s.t. $s(n) < s(n+1)$, i.e., for any $k \in \mathbb{N}$, $b_n = g(n) = f(s(n)) = a_{s(n)}$.*

Example 3.8. *Let $a_n = 1/n$, $n \in \mathbb{N}$. Then, $(a_{2n})_{n\in\mathbb{N}}$, is a subsequence of $(a_n)_{n\in\mathbb{N}}$.*

The following result is known as the Bolzano–Weierstrass theorem.

Proposition 3.12. *Let $(a_n)_{n\in\mathbb{N}}$ be a sequence. If $(a_n)_{n\in\mathbb{N}}$ is bounded, then there exists a subsequence of $(a_n)_{n\in\mathbb{N}}$ which is convergent.*

Example 3.9. *Let $a_n = (-1)^n$, $n \in \mathbb{N}$. We have seen that $(a_n)_{n\in\mathbb{N}}$ is not convergent. However, $(a_{2n})_{n\in\mathbb{N}}$ is a subsequence of $(a_n)_{n\in\mathbb{N}}$ with limit 1.*

An application of Proposition 3.12 is the following result (a proof is given in Section A.4).

Proposition 3.13. *Let $f\colon [a,b] \to \mathbb{R}$ be continuous, then there exists $x_M, x_m \in [a,b]$ s.t. $f(x_M) = \sup_{x\in[a,b]} f(x) = \max_{x\in[a,b]} f(x)$ and $f(x_m) = \inf_{x\in[a,b]} f(x) = \min_{x\in[a,b]} f(x)$, i.e., f attains its maximum and minimum in $[a,b]$. In particular, f is bounded.*

Definition 3.9. *Let $(a_n)_{n\in\mathbb{N}}$ be a sequence and $(a_{s(n)})_{n\in\mathbb{N}}$ be a subsequence of $(a_n)_{n\in\mathbb{N}}$ s.t. $\lim_{n\to\infty} a_{s(n)} = a$. Then, a is said to be an accumulation point of $(a_n)_{n\in\mathbb{N}}$.*

Example 3.10. *Let $a_n = (-1)^n$, $n \in \mathbb{N}$. Then, $(a_n)_{n\in\mathbb{N}}$ has two accumulation points, -1 and 1.*

Proposition 3.14. *Let a be an accumulation point of $(a_n)_{n\in\mathbb{N}}$. Then, for any $\varepsilon > 0$, there are infinitely a_n s.t. $a_n \in (a-\varepsilon, a+\varepsilon)$.*

Proof. Since $a = \lim_{n\to\infty} a_{s(n)}$, it follows that for any $\varepsilon > 0$, there exists $N \in \mathbb{N}$, s.t. for any $n \geq N$, $|a_{s(n)} - a| < \varepsilon \Rightarrow a_{s(n)} \in (a-\varepsilon, a+\varepsilon)$. □

Proposition 3.15. *Let $(a_n)_{n\in\mathbb{N}}$ be a sequence. If $(a_n)_{n\in\mathbb{N}}$ is convergent with limit a, then every subsequence of $(a_n)_{n\in\mathbb{N}}$ converges to a. That is, a convergent sequence has only one accumulation point.*

Proof. Clearly a is an accumulation point of $(a_n)_{n\in\mathbb{N}}$. Suppose by contradiction that $b \neq a$ is another accumulation point of $(a_n)_{n\in\mathbb{N}}$. Set $\delta = (a-b)/2$. Then $\delta > 0$ and hence, there exists $N \in \mathbb{N}$ s.t. $|a_n - a| < \delta$ for any $n \geq N$. Then,

$$|a_n - b| = |a_n - a + a - b| \geq \big||a_n - a| - |a-b|\big|,$$

by the reverse triangular inequality (cf. Proposition 2.10). Therefore, for any $n \geq N$,

$$|a_n - b| \geq \big||a_n - a| - 2\delta\big| = 2\delta - |a_n - a| > \delta.$$

This shows that there exists $\varepsilon > 0$ ($\varepsilon = \delta$) s.t. $|a_n - b| > \varepsilon$ for any $n \geq N$. Thus, for that particular ε, only at most finitely many a_n, are s.t. $a_n \in (b-\varepsilon, b+\varepsilon)$. This contradicts Proposition 3.14. Hence, b is not an accumulation point of $(a_n)_{n\in\mathbb{N}}$. □

Proposition 3.16. *Let $(a_n)_{n\in\mathbb{N}}$ be an increasing (resp. decreasing) sequence. Suppose that there exists a subsequence $(a_{s(n)})_{n\in\mathbb{N}}$ which is s.t. $\lim_{n\to\infty} a_{s(n)} = \infty$ (resp. $\lim_{n\to\infty} a_{s(n)} = -\infty$). Then, $\lim_{n\to\infty} a_n = \infty$ (resp. $\lim_{n\to\infty} a_n = -\infty$).*

Proof. Using Proposition 3.7, this follows directly from Proposition 3.15. If $\lim_{n\to\infty} a_n \neq \infty$ (resp. $\lim_{n\to\infty} a_n \neq -\infty$) it means that $(a_n)_{n\in\mathbb{N}}$ converges (cf. Proposition 3.7). In particular, there exists $a < \infty$ s.t. any accumulation point of $(a_n)_{n\in\mathbb{N}}$ is equal to a (cf. Proposition 3.15). This gives a contradiction with the assumption that $\lim_{n\to\infty} a_{s(n)} = \infty$ (resp. $\lim_{n\to\infty} a_{s(n)} = -\infty$). □

Example 3.11. *We show that $\sum_{i\in\mathbb{N}}(1/i) = \infty$ (cf. Example 3.5). Set, $s_n = \sum_{i=1}^{n}(1/i)$, $n \in \mathbb{N}$. Then, $(s_n)_{n\in\mathbb{N}}$ is increasing. Define the subsequence $(s_{(2^n-1)})_{n\in\mathbb{N}}$. Let $n \in \mathbb{N}$ and $k \in \{2,\dots,n\}$. We have that*

$$s_{(2^n-1)} = 1 + \sum_{k=2}^{n}\left(\sum_{i=2^{k-1}}^{2^k-1}\frac{1}{i}\right).$$

This is because, $\{2,\dots,2^n-1\} = \cup_{k=2}^{n}\{2^{k-1},\dots,2^k-1\}$. We can use induction to see it. If $n = 2$, then $\{2,3\} = \{2^{2-1}, 2^2-1\}$. Suppose that $\{2,\dots,2^n-1\} = \cup_{k=2}^{n}\{2^{k-1},\dots,2^k-1\}$. Then,

$$\begin{aligned}\{2,\dots,2^{n+1}-1\} &= \{2,\dots,2^n-1\} \cup \{2^n,\dots,2^{n+1}-1\} \\ &= \cup_{k=2}^{n}\{2^{k-1},\dots,2^k-1\} \cup \{2^n,\dots,2^{n+1}-1\} \\ &= \cup_{k=2}^{n+1}\{2^{k-1},\dots,2^k-1\}.\end{aligned}$$

For any $n \in \mathbb{N}$, the cardinality of $\{2^{k-1},\dots,2^k-1\}$ is $2^k - 1 - (2^{k-1}-1) = 2^k - 2^{k-1} = 2^{k-1}(2-1) = 2^{k-1}$. Hence, for any $n \in \mathbb{N}$, $\sum_{i=2^{k-1}}^{2^k-1}(1/i) \geq 2^{k-1}(1/2^k - 1) \geq 1/2$. Thus, for any $n \in \mathbb{N}$, $s_{(2^n-1)} \geq 1 + (n-1)/2$. Therefore, $(s_{(2^n-1)})_{n\in\mathbb{N}}$ can not be convergent and we conclude that $\lim_{n\to\infty} s_{(2^n-1)} = \infty$. Using Proposition 3.16, this shows that $\lim_{n\to\infty} s_n = \infty$ as well.

Proposition 3.17. *Suppose that* $(a_n)_{n\in\mathbb{N}}$ *is not bounded from below (resp. bounded from above). Then,* $\inf_{n\in\mathbb{N}} a_n = -\infty$ *(resp.* $\sup_{n\in\mathbb{N}} a_n = \infty$*).*

Exercise 3.6. *Prove Proposition 3.17.*

Exercise 3.7. *Let* $(a_n)_{n\in\mathbb{N}}$ *be a sequence bounded from below. Define the sequence*

$$m_n = \inf\{a_k : k \geq n\} = \inf_{k\geq n} a_k, \quad n \in \mathbb{N}.$$

Show that $(m_n)_{n\in\mathbb{N}}$ *is increasing.*

Exercise 3.8. *Let* $(a_n)_{n\in\mathbb{N}}$ *be a sequence bounded from above. Define the sequence*

$$M_n = \sup\{a_k : k \geq n\} = \sup_{k\geq n} a_k, \quad n \in \mathbb{N}.$$

Show that $(M_n)_{n\in\mathbb{N}}$ *is decreasing.*

Proposition 3.18. *Let* $(a_n)_{n\in\mathbb{N}}$ *and* $(m_n)_{n\in\mathbb{N}}$ *be as in Exercise 3.7. Then,*

$$\lim_{n\to\infty} m_n = \sup_{n\in\mathbb{N}} m_n = \sup_{n\in\mathbb{N}} \inf_{k\geq n} a_k.$$

Exercise 3.9. *Prove Proposition 3.18.*

Similarly, we have the following result.

Proposition 3.19. *Let* $(a_n)_{n\in\mathbb{N}}$ *and* $(M_n)_{n\in\mathbb{N}}$ *be as in Exercise 3.8. Then,*

$$\lim_{n\to\infty} M_n = \inf_{n\in\mathbb{N}} M_n = \inf_{n\in\mathbb{N}} \sup_{k\geq n} a_k.$$

Example 3.12. *Let* $a_n = (-1)^n$, $n \in \mathbb{N}$, *then* $m_n = -1$ *and* $M_n = 1$ *for any* $n \in \mathbb{N}$. *In particular,* $\lim_{n\to\infty} m_n = -1$ *and* $\lim_{n\to\infty} M_n = 1$.

We are in place to make the following definition:

Definition 3.10. *Let* $(a_n)_{n\in\mathbb{N}}$ *be a sequence. We define:*

$$\liminf_{n\to\infty} a_n = \begin{cases} \sup_{n\in\mathbb{N}} \inf_{k\geq n} a_k, & \textit{if } (a_n)_{n\in\mathbb{N}} \textit{ is bounded or bounded from below,} \\ -\infty, & \textit{otherwise,} \end{cases}$$

and

$$\limsup_{n\to\infty} a_n = \begin{cases} \inf_{n\in\mathbb{N}} \sup_{k\geq n} a_k, & \textit{if } (a_n)_{n\in\mathbb{N}} \textit{ is bounded or bounded from above,} \\ \infty, & \textit{otherwise.} \end{cases}$$

$\liminf_{n\to\infty} a_n$ *and* $\limsup_{n\to\infty} a_n$ *are referred to as limit inferior and limit superior of* $(a_n)_{n\in\mathbb{N}}$.

Exercise 3.10. *Suppose that* $\lim_{n\to\infty} a_n = -\infty$ *(resp.* $\lim_{n\to\infty} a_n = \infty$*), then,*

$$\liminf_{n\to\infty} a_n = -\infty = \limsup_{n\to\infty} a_n \ \textit{(resp.} \ \liminf_{n\to\infty} a_n = \infty = \limsup_{n\to\infty} a_n\textit{)}.$$

Proposition 3.20. *Let* $(a_n)_{n\in\mathbb{N}}$ *be a sequence. We have that*

$$\liminf_{n\to\infty} a_n \leq \limsup_{n\to\infty} a_n.$$

Proof. Given any $n \in \mathbb{N}$, $\inf_{k\geq n} a_k \leq \sup_{k\geq n} a_k$. Assume that $(a_n)_{n\in\mathbb{N}}$ is bounded. Then, $(\inf_{k\geq n} a_k)_{n\in\mathbb{N}}$ and $(\sup_{k\geq n} a_k)_{n\in\mathbb{N}}$ converge (cf. Proposition 3.3). We use Propositions 3.18 and 3.19 and obtain,

$$\liminf_{n\to\infty} a_n = \lim_{n\to\infty} \inf_{k\geq n} a_k \leq \lim_{n\to\infty} \sup_{k\geq n} a_k = \inf_{n\in\mathbb{N}} \sup_{k\geq n} a_k = \limsup_{n\to\infty} a_n.$$

Clearly, by Definition 3.10, $\liminf_{n\to\infty} a_n \leq \limsup_{n\to\infty} a_n$ in all the other cases. □

The following result gives another characterization of convergence:

Proposition 3.21. *Let $(a_n)_{n\in\mathbb{N}}$ be a bounded sequence. Then, $(a_n)_{n\in\mathbb{N}}$ is convergent with limit a if and only if*

$$\liminf_{n\to\infty} a_n = a = \limsup_{n\to\infty} a_n.$$

We use the following result to proof the latter proposition.

Proposition 3.22. *Assume that $(a_n)_{n\in\mathbb{N}}$ is a bounded sequence and define the set*

$$A = \{a \colon a \textit{ is an accumulation point of } (a_n)_{n\in\mathbb{N}}\}.$$

Then, $\min A = \liminf_{n\to\infty} a_n$ *and* $\max A = \limsup_{n\to\infty} a_n$.

Proof of Proposition 3.21. Assume that $(a_n)_{n\in\mathbb{N}}$ is convergent with limit a. It follows by Proposition 3.15 that the set of accumulation points of $(a_n)_{n\in\mathbb{N}}$ is given by $\{a\}$. Using Proposition 3.22, this shows that $\liminf_{n\to\infty} a_n = a = \limsup_{n\to\infty} a_n$. For the other direction, assume that $(a_n)_{n\in\mathbb{N}}$ is s.t. $\liminf_{n\to\infty} a_n = a = \limsup_{n\to\infty} a_n$. Then, we have that for any $n \in \mathbb{N}$,

$$m_n = \inf_{k\geq n} a_k \leq a_n \leq \sup_{k\geq n} a_k = M_n$$

Therefore, using Propositions 3.6, 3.18 and 3.19 we conclude that $(a_n)_{n\in\mathbb{N}}$ has limit a. □

A more general statement is the following.

Proposition 3.23. $\lim_{n\to\infty} a_n$ *exists if and only if*

$$\liminf_{n\to\infty} a_n = \lim_{n\to\infty} a_n = \limsup_{n\to\infty} a_n.$$

Proof. Using Proposition 3.21 and Exercise 3.10 it remains to show that if

$$\liminf_{n\to\infty} a_n = -\infty = \limsup_{n\to\infty} a_n \text{ (or } \liminf_{n\to\infty} a_n = \infty = \limsup_{n\to\infty} a_n),$$

then $\lim_{n\to\infty} a_n$ exists. Actually, if

$$\liminf_{n\to\infty} a_n = -\infty = \limsup_{n\to\infty} a_n,$$

then $\lim_{n\to\infty} a_n = -\infty$ as well. To see it we first notice that $(a_n)_{n\in\mathbb{N}}$ must be bounded from above, otherwise, $\limsup_{n\to\infty} a_n = \infty$, by definition. Hence, $(M_n)_{n\in\mathbb{N}}$ with $M_n = \sup_{k\geq n} a_k$ is decreasing (cf. Exercise 3.8). Also since $\liminf_{n\to\infty} a_n = -\infty$, $(a_n)_{n\in\mathbb{N}}$ is not bounded from below. Therefore, for any $M \in \mathbb{R}$, there exists $N \in \mathbb{N}$, s.t. for any $n \geq N$, $M_n \leq M$. Further, for any $n \in \mathbb{N}$, $a_n \leq \sup_{k\geq n} a_k$. In conclusion, we have shown that for any $M \in \mathbb{R}$, there exists $N \in \mathbb{N}$, s.t. $a_n \leq M$ for any $n \geq N$. This shows that $(a_n)_{n\in\mathbb{N}}$ diverges to $-\infty$ (cf. Definition 3.6). A similar argument shows that if

$$\liminf_{n\to\infty} a_n = \infty = \limsup_{n\to\infty} a_n,$$

then $\lim_{n\to\infty} a_n = \infty$. □

Example 3.13. *Let* $a_n = (-1)^n$, $n \in \mathbb{N}$, *then,* $\liminf_{n\to\infty} a_n = -1$ *and* $\limsup_{n\to\infty} a_n = 1$ *(cf. Example 3.12). Hence, using Proposition 3.21,* $(a_n)_{n\in\mathbb{N}}$ *can not be convergent (cf. Exercise 3.2).*

Exercise 3.11. *Let* $a_n = \cos(n\pi)$, $n \in \mathbb{N}$. *Find* $\liminf_{n\to\infty} a_n$ *and* $\limsup_{n\to\infty} a_n$.

We note that Proposition A.6 stated in Section A.3 of the appendix gives another useful characterization of convergence in terms of subsequences. It states that if $(a_n)_{n\in\mathbb{N}}$ is a real valued sequence s.t. for an arbitrary subsequence $(a_{s(n)})_{n\in\mathbb{N}}$ of $(a_n)_{n\in\mathbb{N}}$ one can extract a subsequence $(a_{t(s(n))})_{n\in\mathbb{N}}$ that converges to $a \in \mathbb{R}$, then the original sequence converges to a as well.

3.3 Vector-valued sequences

The previous section on real-valued sequences can easily be extended to the notion of vector-valued sequences.

Definition 3.11. *An* $\mathbb{R}^k$*-valued sequence is a function* $f\colon \mathbb{N} \to \mathbb{R}^k$, *where we write*

$$f(n) = (f_1(n), \dots, f_k(n)) = (a_1^n, \dots, a_k^n), \quad n \in \mathbb{N},$$

We use the notation $f = (a_n)_{n\in\mathbb{N}}$ *for a* $\mathbb{R}^k$*-valued sequence.*

We notice that the coordinate functions of an $\mathbb{R}^k$-valued sequence $(a_n)_{n\in\mathbb{N}}$ are real-valued sequences (cf. Definition 3.1). The space of all $\mathbb{R}^k$-valued sequences is denoted with $(\mathbb{R}^k)^{\mathbb{N}}$, $k \in \mathbb{N}$. If $k = 1$, then $(a_n)_{n\in\mathbb{N}}$ is a real-valued sequence. In particular, Definition 3.11 contains Definition 3.1. Upon the Euclidean metric $\|x - y\| = \sqrt{(x_1 - y_1)^2 + \cdots + (x_k - y_k)^2}$, $x = (x_1, \dots, x_k)$, $y = (y_1, \dots, y_k)$, we can introduce the notion of convergence for $\mathbb{R}^k$-valued sequence.

Definition 3.12. *Let* $(a_n)_{n\in\mathbb{N}} \in (\mathbb{R}^k)^{\mathbb{N}}$. $(a_n)_{n\in\mathbb{N}}$ *is said to be convergent if there exists a number* $a = (a_1, \dots, a_k) \in \mathbb{R}^k$ *s.t. for any* $\varepsilon > 0$ *there exists* $N \in \mathbb{N}$ *s.t.* $\|a_n - a\| < \varepsilon$ *for any* $n \geq N$.

If $(a_n)_{n\in\mathbb{N}} \in (\mathbb{R}^k)^{\mathbb{N}}$, we write $a_n \xrightarrow{n\to\infty} a$ (resp. $\lim_{n\to\infty} a_n = a$) to indicate that $(a_n)_{n\in\mathbb{N}}$ converges to a. The following result shows that in order to proof that an $\mathbb{R}^k$-valued sequence converges, it is enough to study the convergence of the individual coordinates.

Proposition 3.24. *Let* $(a_n)_{n\in\mathbb{N}} \in (\mathbb{R}^k)^{\mathbb{N}}$. *Then,*

$$(a_1^n, \dots, a_k^n) = a_n \xrightarrow{n\to\infty} a = (a_1, \dots, a_k) \;\Leftrightarrow\; a_i^n \xrightarrow{n\to\infty} a_i \;\forall\, i = 1, \dots, k.$$

The following result is called the sequence criterion for continuous functions (a proof is given in the appendix, Section A.4).

Proposition 3.25. *Let* $f\colon E \to \mathbb{R}^k$, $E \subset \mathbb{R}^m$ *and* $x \in E$. *Then, the following two statements are equivalent:*

(i) f *is continuous at* x;

(ii) $\forall\, (x_n)_{n\in\mathbb{N}} \subset E$ *with* $\lim_{n\to\infty} x_n = x$ *it follows that* $\lim_{n\to\infty} f(x_n) = f(x)$.

In the following, we will use the term sequence for $(a_n)_{n\in\mathbb{N}} \in (\mathbb{R}^k)^{\mathbb{N}}$, $k \in \mathbb{N}$, and it will be clear from the context whether $k = 1$ or $k > 1$.

3.4 Sequences of Functions

Definition 3.13. *In general, a sequence of functions, taking values in the extended real numbers, defined on some common set A, is a collection of functions $g_n\colon A \to \overline{\mathbb{R}}$, $n \in \mathbb{N}$. Then, given $E \subset \mathbb{N}$, the quantities $\inf_{n\in E} g_n$ and $\sup_{n\in E} g_n$ are defined pointwise on A, i.e.,*

$$(\inf_{n\in E} g_n)(x) = \inf_{n\in E} g_n(x) \text{ and } (\sup_{n\in E} g_n)(x) = \sup_{n\in E} g_n(x), \quad x \in A.$$

Hence, for any $x \in A$, $(\inf_{n\in E} g_n)(x), (\sup_{n\in E} g_n)(x) \in \overline{\mathbb{R}}$. Additionally, if for any $x \in A$, $\lim_{n\to\infty} g_n(x) \in \overline{\mathbb{R}}$, then also $\lim_{n\to\infty} g_n$ is defined pointwise, i.e.,

$$(\lim_{n\to\infty} g_n)(x) = \lim_{n\to\infty} g_n(x), \quad x \in A.$$

We say that the sequence of functions g_n, $n \in \mathbb{N}$, converges pointwise to a function $g\colon A \to \overline{\mathbb{R}}$, if for any $x \in A$, $\lim_{n\to\infty} g_n(x) = g(x)$.

Example 3.14. *Given $x \in [0, \pi]$, let $g_n(x) = \cos(nx)$, $n \in \mathbb{N}$. Then, g_n, $n \in \mathbb{N}$, is a sequence of functions. We have seen that $(\liminf_{n\to\infty} g_n)(\pi) = -1$ and $(\limsup_{n\to\infty} g_n)(\pi) = 1$. Therefore, using Proposition 3.21 it is not true that $x \mapsto \cos(nx)$, $n \in \mathbb{N}$, converges pointwise on $[0, \pi]$. Notice that for any $x \in [0, 2\pi]$, $(\liminf_{n\to\infty} g_n)(x) = \sup_{n\in\mathbb{N}}(\inf_{k\geq n} g_k)(x)$.*

3.5 Solution to exercises

Solution 3.1 (Solution to Exercise 3.1)**.** *Let $\varepsilon > 0$ and $a = c$. Then, for any $n \in \mathbb{N}$, $|a_n - c| = 0 < \varepsilon$. Thus, $(a_n)_{n\in\mathbb{N}}$ is convergent with limit c.*

Solution 3.2 (Solution to Exercise 3.2)**.** *The sequence $(a_n)_{n\in\mathbb{N}}$ is not convergent. To see it, assume by contradiction that $(a_n)_{n\in\mathbb{N}}$ is convergent with limit a. Suppose that $a \notin \{-1, 1\}$. Then, since $a \neq 1$ and $a \neq -1$, it follows that $|1-a| > 0$. Thus, if $\varepsilon < |1-a|$, for any $N \in \mathbb{N}$, $|a_n - a| = |1 - a| > \varepsilon$ for infinitely many n. Hence, it must be the case that $a \in \{-1, 1\}$. Again, this not possible, since if for example $a = 1$, we have that $|a_n - a| = 2$ for infinitely many n. A similar argument shows that $a = -1$ is also not possible. Hence, there is no chance that $(a_n)_{n\in\mathbb{N}}$ can be convergent.*

Solution 3.3 (Solution to Exercise 3.3)**.** *Yes, $(a_n)_{n\in\mathbb{N}}$ is convergent with limit 0. To see it, we write*

$$\frac{n^2 + 3n^3 + n}{1 + n^4} = \frac{\frac{n^2+3n^3+n}{n^4}}{\frac{1+n^4}{n^4}} = \frac{\frac{1}{n^2} + \frac{3}{n} + \frac{1}{n^3}}{\frac{1}{n^4} + 1}.$$

We then remark that for any $k \in \mathbb{N}$,

$$0 \leq \frac{1}{n^k} \leq \frac{1}{n}.$$

Thus, using Propositions 3.6 and 3.4,

$$\lim_{n\to\infty} a_n = \frac{0}{1} = 0.$$

Solution 3.4 (Solution to Exercise 3.4)**.** *Since*

$$0 \leq \frac{n!}{n^n} \leq \frac{1}{n},$$

it follows from Proposition 3.6 that $\lim_{n\to\infty} a_n = 0$.

Solution 3.5 (Solution to Exercise 3.5). *Let us first consider the case where $(a_n)_{n\in\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}}$ are both increasing. If both sequences converge, this is Proposition 3.5. The case where $(b_n)_{n\in\mathbb{N}}$ converges but $(a_n)_{n\in\mathbb{N}}$ does not is not possible by Proposition 3.7 and Proposition 3.2. If $(b_n)_{n\in\mathbb{N}}$ does not converge, it diverges to ∞ and we are left with two cases, either $(a_n)_{n\in\mathbb{N}}$ converges or it does not. In both cases it is clearly true that $\lim_{n\to\infty} a_n \leq \lim_{n\to\infty} b_n$, where we have equality when both series diverge (cf. Remark 3.1). The other case, i.e., when $(a_n)_{n\in\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}}$ are both decreasing is proved similarly. Here, it is not possible that $(a_n)_{n\in\mathbb{N}}$ converges but $(b_n)_{n\in\mathbb{N}}$ does not (cf. Propositions 3.7 and 3.2). If both sequences diverge, we have equality in the limit and if $(b_n)_{n\in\mathbb{N}}$ converges but $(a_n)_{n\in\mathbb{N}}$ diverges, we have strict inequality in the limit, i.e., $a < b$.*

Solution 3.6 (Solution to Exercise 3.6). *Suppose by contradiction that $\inf\{a_n : n \in \mathbb{N}\} > -\infty$, i.e., there exists a real number $M > 0$ s.t. $\inf\{a_n : n \in \mathbb{N}\} > -M$. Since $(a_n)_{n\in\mathbb{N}}$ is not bounded from below, then, for any integer $N \in \mathbb{N}$, there exists $n(N) \in \mathbb{N}$ s.t. $a_{n(N)} < -N$ (cf. Definition 3.3). Define the subsequence $a_{n(N)}$, $N \in \mathbb{N}$, where for any $N \in \mathbb{N}$, $a_{n(N)} < -N$. Clearly, $\{a_{n(N)} : N \in \mathbb{N}\} \subset \{a_n : n \in \mathbb{N}\}$, hence, by Proposition 1.9, for any integer $k \in \mathbb{N}$,*

$$-M < \inf\{a_n : n \in \mathbb{N}\} \leq \inf\{a_{n(N)} : n \in \mathbb{N}\} \leq a_{n(k)} < -k.$$

Hence, we have deduced that there exists a real number $M > 0$ s.t. for any integer $k \in \mathbb{N}$, $-M < -k$. This gives a contradiction with Proposition 1.7. Hence, $\inf\{a_n : n \in \mathbb{N}\} = -\infty$. A similar argument shows that if $(a_n)_{n\in\mathbb{N}}$ is not bounded from above, then $\sup\{a_n : n \in \mathbb{N}\} = \infty$.

Solution 3.7 (Solution to Exercise 3.7). *Let $n \in \mathbb{N}$. Write, $\{a_k : k \geq n\} = \{a_n\} \cup \{a_k : k \geq n+1\}$. Hence, $\{a_k : k \geq n+1\} \subset \{a_k : k \geq n\}$. It follows that $m_n \leq m_{n+1}$ (cf. Proposition 1.9).*

Solution 3.8 (Solution to Exercise 3.8). *Since $\{a_k : k \geq n+1\} \subset \{a_k : k \geq n\}$, it follows that $M_{n+1} \leq M_n$ (cf. Proposition 1.9).*

Solution 3.9 (Solution to Exercise 3.9). *If $S = \sup_{n\in\mathbb{N}} m_n < \infty$, define $b_n = S$ for any $n \in \mathbb{N}$. Then, $m_n \leq b_n$ for any $n \in \mathbb{N}$ and by Proposition 3.8, $\lim_{n\to\infty} m_n \leq S$. If $S = \infty$, clearly, $\lim_{n\to\infty} m_n \leq S$. For the other inequality, if $S < \infty$, by definition of $\sup_{n\in\mathbb{N}} m_n$ (cf. Proposition 1.10), there exists $n \in \mathbb{N}$, s.t. $m_n \geq S - \varepsilon$ for any $\varepsilon > 0$. Thus we can again use Proposition 3.8 to conclude that $\lim_{n\to\infty} m_n \geq S$ (take for example $b_n = S - (1/n)$). For the final case, suppose by contradiction that $S = \infty$ and $\lim_{n\to\infty} m_n < \infty$. By Proposition 3.2, this means that $(m_n)_{n\in\mathbb{N}}$ is bounded, i.e., there exists an upper bound for $\{m_n : n \in \mathbb{N}\}$. This is not the case, as $S = \sup\{m_n : n \in \mathbb{N}\} = \infty$ (cf. Definition 1.10). Hence, if $S = \infty$, then $\lim_{n\to\infty} m_n = \infty$.*

Solution 3.10 (Solution to Exercise 3.10). *If $\lim_{n\to\infty} a_n = -\infty$, then $(a_n)_{n\in\mathbb{N}}$ can not be bounded from below. Therefore, by Definition 3.10, $\liminf_{n\to\infty} a_n = -\infty$. Further, $(a_n)_{n\in\mathbb{N}}$ must be bounded from above, since there exists $N \in \mathbb{N}$, s.t. $a_n \leq 0$ for any $n \geq N$ (cf. Definition 3.6) and for any $n < N$, we have that $a_n \leq \max\{a_i : i = 1, \ldots, N-1\} = M$. In particular, $a_n \leq \max\{0, M\}$. Therefore, by Definition 3.10, $\limsup_{n\to\infty} a_n = \inf_{n\in\mathbb{N}} \sup_{k\geq n} a_k$. We know that $(M_n)_{n\in\mathbb{N}} = (\sup_{k\geq n} a_k)_{n\in\mathbb{N}}$ is decreasing and $\lim_{n\to\infty} M_n = \limsup_{n\to\infty} a_n$ (cf. Proposition 3.19). Suppose by contradiction that $(M_n)_{n\in\mathbb{N}}$ converges. Then, $(M_n)_{n\in\mathbb{N}}$ is bounded. That is, there exists $L \in \mathbb{R}$ s.t. $M_n > L$ for any $n \in \mathbb{N}$. Since $\lim_{n\to\infty} a_n = -\infty$, let $N \in \mathbb{N}$ s.t. $a_n \leq L$ for any $n \geq N$ (cf. Definition 3.6). Then, $\sup_{k\geq N} a_k = M_N \leq L$. This gives a contradiction, hence, $(M_n)_{n\in\mathbb{N}}$ must diverge and hence by Proposition 3.9, $\lim_{n\to\infty} M_n = \limsup_{n\to\infty} a_n = -\infty$. Similarly, one can show that if $\lim_{n\to\infty} a_n = \infty$, then $\liminf_{n\to\infty} a_n = \infty = \limsup_{n\to\infty} a_n$.*

Solution 3.11 (Solution to Exercise 3.11). *We have that $|a_n| \leq 1$ for any $n \in \mathbb{N}$. Thus, $(a_n)_{n\in\mathbb{N}}$ is bounded. Given any $n \in \mathbb{N}$, $\inf_{k\geq n} a_k = -1$ and hence $\liminf_{n\to\infty} a_n = -1$. Further, for any $n \in \mathbb{N}$, $\sup_{k\geq n} a_k = 1$. Therefore, $\limsup_{n\to\infty} a_n = 1$*

3.6 Additional exercises

Exercise 3.12. *Let* $(a_n)_{n\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}}$ *be a convergent sequence. Show that* $\lim_{n\to\infty} a_n$ *is unique, i.e., if* $\lim_{n\to\infty} a_n = a$ *and* $\lim_{n\to\infty} a_n = b$*, then* $a = b$*.*

Exercise 3.13. *Let* $(a_n)_{n\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}}$ *be a sequence. Show that:*

(i) If $a_n > 0$ *for any* $n \in \mathbb{N}$*, then if* $a_n \xrightarrow{n\to\infty} 0$ *it follows that* $1/a_n \xrightarrow{n\to\infty} \infty$*;*

(ii) If $a_n < 0$ *for any* $n \in \mathbb{N}$*, then if* $a_n \xrightarrow{n\to\infty} 0$ *it follows that* $1/a_n \xrightarrow{n\to\infty} -\infty$*.*

Exercise 3.14. *Let* $a \in \mathbb{R}$ *and* $|r| < 1$*. We consider the sequence*

$$s_n = \sum_{k=0}^{n} ar^k = a + ar + ar^2 + \cdots + ar^n, \quad n \in \mathbb{N}.$$

Show that $(s_n)_{n\in\mathbb{N}}$ *is convergent with limit* $a/(1-r)$*.*
Hint: *Compare* s_n *and* rs_n*.*

Exercise 3.15. *Let* $(a_n)_{n\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}}$ *be a bounded sequence. Show that*

$$\min\{a\colon a \text{ is an accumulation point of } (a_n)_{n\in\mathbb{N}}\} = \liminf_{n\to\infty} a_n.$$

Exercise 3.16. *Prove Proposition 3.24.*

4 Measurable sets: Part I

4.1 Measurable spaces

Definition 4.1 (σ-field)**.** *Let Ω be a nonempty set. A family of subsets $\mathcal{F}$ of Ω is called a σ-field on Ω if the following three items are satisfied:*

(i) $\Omega \in \mathcal{F}$;

(ii) $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$;

(iii) If $\{A_i \colon i \in \mathbb{N}\}$ is a collection of sets s.t. $A_i \in \mathcal{F}$ for any $i \in \mathbb{N}$, then $\cup_{i\in\mathbb{N}} A_i \in \mathcal{F}$.

Example 4.1. *Let $\Omega \neq \emptyset$ be an arbitrary set. Let*

$$\mathcal{F} = \{\emptyset, \Omega\}.$$

Then, $\mathcal{F}$ is a σ-field on Ω. Clearly, $\Omega \in \mathcal{F}$. Further, let $A \in \mathcal{F}$. Then, there are only two cases, either $A = \emptyset$ or $A = \Omega$. In each case, $A^c \in \mathcal{F}$. Consider a countable collection $\{A_i \colon i \in \mathbb{N}\} \subset \mathcal{F}$. This collection is composed only of the sets $A_i = \emptyset$ or $A_i = \Omega$, $i \in \mathbb{N}$. Thus (recall Exercise 1.3),

$$\bigcup_{i\in\mathbb{N}} A_i = \begin{cases} \Omega, & \text{if } \exists i \text{ s.t. } A_i = \Omega, \\ \emptyset, & \text{otherwise.} \end{cases}$$

This, items (i), (ii) and (iii) of Definition 4.1 are satisfied and hence $\mathcal{F}$ is a σ-field. We remark that $\mathcal{F}$ is referred to as the trivial σ-field. It is the smallest possible σ-field on Ω.

Example 4.2. *Let $\Omega \neq \emptyset$ be an arbitrary set. Let $\mathcal{F}$ be the family which consists of all possible subsets of Ω, i.e.,*

$$\mathcal{F} = \{A \colon A \subset \Omega\}.$$

Then, $\mathcal{F}$ is a σ-field on Ω. Since $\Omega \subset \Omega$, $\Omega \in \mathcal{F}$. Let $A \in \mathcal{F}$, then by definition, $A^c = \Omega \setminus A \subset \Omega$ and hence $A^c \in \mathcal{F}$. Let $\{A_i \colon i \in \mathbb{N}\} \subset \mathcal{F}$. Then, by definition,

$$\bigcup_{i\in\mathbb{N}} A_i = \{\omega \in \Omega \colon \exists i \text{ s.t. } \omega \in A_i\} \subset \Omega.$$

Hence, $\mathcal{F}$ is a σ-field. The given σ-field $\mathcal{F}$ is referred to as the power set of Ω and denoted with $\mathcal{P}(\Omega)$ (or 2^Ω). It is the largest possible σ-field on Ω.

Example 4.3. *Let Ω be an uncountable set. We consider the family*

$$\mathcal{F} = \{A \colon A \subset \Omega \text{ s.t. } A \text{ is countable or } A^c \text{ is countable}\}.$$

Then, $\mathcal{F}$ is a σ-field on Ω. We have that $\Omega^c = \emptyset$. We know that $\#\emptyset = 0$, in particular $\emptyset$ is countable. Thus, $\Omega \in \mathcal{F}$. Let $A \in \mathcal{F}$. Thus, A is countable or A^c is countable. Since $(A^c)^c = A$, $A^c \in \mathcal{F}$. Let $\{A_i \colon i \in \mathbb{N}\} \subset \mathcal{F}$. If there exists $j \in \mathbb{N}$ s.t. A_j is not countable, we have that

$$\left(\bigcup_{i\in\mathbb{N}} A_i\right)^c = \bigcap_{i\in\mathbb{N}} A_i^c = \bigcap_{\substack{i\in\mathbb{N}\\ i\neq j}} A_i^c \cap A_j^c \subset A_j^c.$$

Thus,

$$\#\left(\bigcup_{i\in\mathbb{N}} A_i\right)^c \leq \#A_j^c \leq \#\mathbb{N},$$

since A_j^c is countable. Hence $(\cup_{i\in\mathbb{N}}A_i)^c$ is countable as well and therefore an element of $\mathcal{F}$. If for any $i \in \mathbb{N}$, A_i is countable we rely on Proposition 2.7 and conclude that $\cup_{i\in\mathbb{N}}A_i$ must be countable as well and hence, $\cup_{i\in\mathbb{N}}A_i \in \mathcal{F}$. Notice that since Ω is uncountable it is not true that $\mathcal{F} = \mathcal{P}(\Omega)$ (cf. Example 4.2). As a simple example consider $\Omega = [0, 1)$, then $A = [0, 1/2)$ and $A^c = [1/2, 0)$ are not countable.

Exercise 4.1. *Let $\Omega \neq \emptyset$ be countable and define $\mathcal{F}$ as in Example 4.3. Is it true that $\mathcal{F} = \mathcal{P}(\Omega)$?*

Exercise 4.2. *Let Ω be a non empty set an $\mathcal{F}$ be a σ-field on Ω. Show that*

(a) if $\{A_i : i \in \mathbb{N}\} \subset \mathcal{F}$, then $\cap_{i\in\mathbb{N}}A_i \in \mathcal{F}$.

(b) if $A \in \mathcal{F}$ and $B \in \mathcal{F}$ then $A \setminus B \in \mathcal{F}$;

Exercise 4.3. *Let $A \subset \Omega$, $\Omega \neq \emptyset$. Show that*

$$\{\emptyset, A, A^c, \Omega\},$$

is a σ-field on Ω.

Example 4.4. *Let Ω be a non empty set an $\mathcal{F}$ be a σ-field on Ω. Let $\Omega_0 \subset \Omega$ s.t. $\Omega_0 \neq \emptyset$. Then, the collection $\mathcal{F} \cap \Omega_0$ defined by*

$$\mathcal{F} \cap \Omega_0 = \{A \cap \Omega_0 : A \in \mathcal{F}\},$$

is a σ-field on Ω_0. Clearly, $\Omega_0 = \Omega \cap \Omega_0 \in \mathcal{F} \cap \Omega_0$. If $B \in \mathcal{F} \cap \Omega_0$, then $B = A \cap \Omega_0$ for some $A \in \mathcal{F}$. Then,

$$\Omega_0 \setminus B = (\Omega_0 \setminus A) \cup (\Omega_0 \setminus \Omega_0) = \Omega_0 \setminus A = A^c \cap \Omega_0 \in \mathcal{F} \cap \Omega_0.$$

Suppose that $\{B_i : i \in \mathbb{N}\} \subset \mathcal{F} \cap \Omega_0$. Therefore, for any $i \in \mathbb{N}$, $B_i = A_i \cap \Omega_0$ for some $A_i \in \mathcal{F}$. Then, since $\cup_{i\in\mathbb{N}}B_i = (\cup_{i\in\mathbb{N}}A_i) \cap \Omega_0$, it follows that $\cup_{i\in\mathbb{N}}B_i \in \mathcal{F} \cap \Omega_0$.

Example 4.5. *Let Ω be an infinite set. Define the family*

$$\mathcal{G} = \{A : A \subset \Omega \text{ s.t. } A \text{ is finite or } A^c \text{ is finite}\}.$$

Then, $\mathcal{G}$ is not a σ-field on Ω. To see this, let $\{\omega_i : i \in \mathbb{N}\}$ be a countably infinite sequence of distinct points of Ω. This is possible, since Ω is not finite. Define the set $A = \{\omega_{2i} : i \in \mathbb{N}\}$. Let $A_i = \{\omega_{2i}\}$, $i \in \mathbb{N}$, be the singleton sets of A. Thus, $A = \cup_{i\in\mathbb{N}}A_i$. It is clear that $A_i \in \mathcal{G}$ for any $i \in \mathbb{N}$. It is also true that $A \notin \mathcal{G}$ since A and A^c are both not finite ($\{\omega_{2i+1} : i \in \mathbb{N}\} \subset A^c$). Therefore, $\mathcal{G}$ is not a σ-field on Ω.

Exercise 4.4. *Let $\Omega \neq \emptyset$ be finite and define $\mathcal{G}$ as in Example 4.5. Is it true that $\mathcal{G}$ is a σ-field on Ω?*

Example 4.6. *Let $\Omega = \mathbb{R}$ and consider the family*

$$\mathcal{R} = \{A : A = (a, b],\ a, b \in \mathbb{R}\} \cup \{\emptyset\}.$$

That is, the members of $\mathcal{R}$ are either the empty set or a left-open interval. The family $\mathcal{R}$ is not a σ-field on $\mathbb{R}$. To see it, if $A = (a, b] \in \mathcal{R}$, then $A^c = (-\infty, a] \cup (b, \infty) \notin \mathcal{R}$. For another way to see it, let $x \in \mathbb{R}$. Then (cf. Exercise 4.1) we must have

$$\bigcap_{n\in\mathbb{N}} (x - n^{-1}, x] \in \mathcal{R}.$$

But, we readily see that

$$\bigcap_{n\in\mathbb{N}} (x - n^{-1}, x] = \{x\} \notin \mathcal{R}. \tag{7}$$

In order to verify (7)*, we notice first that* $\{x\} \subset \cap_{n\in\mathbb{N}}(x - n^{-1}, x]$*. For the other inclusion, let* $y \in \cap_{n\in\mathbb{N}}(x - n^{-1}, x]$*. Then* $y \in (x - n^{-1}, x]$ *for any* $n \in \mathbb{N}$*. That is, for any* $n \in \mathbb{N}$*,*

$$x - \frac{1}{n} < y \leq x.$$

If we let $a_n = x - n^{-1}$*,* $n \in \mathbb{N}$*, and* $b_n = x$*,* $n \in \mathbb{N}$*, using Proposition 3.5, we have that*

$$\lim_{n\to\infty} a_n = x \leq y \leq x.$$

Hence, $y = x$ *and therefore* $\cap_{n\in\mathbb{N}}(x - n^{-1}, x] \subset \{x\}$*.*

The next result is of general importance as it shows that even though a family of subsets $\mathcal{G}$ might not be σ-field, one can always find a σ-filed which is the smallest possible σ-filed that contains $\mathcal{G}$.

Proposition 4.1. *Let* $\Omega \neq \emptyset$ *and* $\mathcal{G}$ *be a family of subsets of* Ω*. Then, there exists a* σ*-field* $\sigma(\mathcal{G})$ *which satisfies:*

(i) $\mathcal{G} \subset \sigma(\mathcal{G})$*;*

(ii) If $\mathcal{G} \subset \mathcal{U}$ *and* $\mathcal{U}$ *is a* σ*-field, then* $\sigma(\mathcal{G}) \subset \mathcal{U}$*.*

To prove Proposition 4.1 we rely on the following result:

Proposition 4.2. *Let* $\Omega \neq \emptyset$ *be a set. Let* $\mathcal{F}_i$*,* $i \in I$*, be a collection of* σ*-fields on* Ω *over an arbitrary set* I*. Then,*

$$\mathcal{F} = \bigcap_{i\in I} \mathcal{F}_i,$$

is a σ*-field on* Ω*.*

Exercise 4.5. *Prove Proposition 4.2.*

Proof of Proposition 4.1. We know that there exists at least one σ-field which contains $\mathcal{G}$, namely $\mathcal{P}(\Omega)$. Thus, we define

$$\sigma(\mathcal{G}) = \bigcap_{i\in I} \mathcal{F}_i,$$

where

$$\{\mathcal{F}_i : i \in I\} = \{\mathcal{U} : \mathcal{U} \text{ is a } \sigma\text{-field on } \Omega \text{ s.t. } \mathcal{G} \subset \mathcal{U}\}$$

Notice that the set I might not be countable. By Proposition 4.2, $\sigma(\mathcal{G})$ is a σ-field. It remains to show (i) and (ii) of Proposition 4.1. Since for each $i \in I$, $\mathcal{G} \subset \mathcal{F}_i$, it is not possible that $\mathcal{G}$ is not a subset of $\sigma(\mathcal{G})$ (see Exercise 1.8). With regard to (ii), let $g \in \sigma(\mathcal{G})$ and $\mathcal{U}$ be any σ-field which is s.t. $\mathcal{G} \subset \mathcal{U}$. Then, there exists $j \in I$ s.t. $\mathcal{U} = \mathcal{F}_j$. Since $g \in \sigma(\mathcal{G})$, $g \in \mathcal{F}_i$ for any $i \in I$. In particular, $g \in \mathcal{F}_j$. Thus, $\sigma(\mathcal{G}) \subset \mathcal{U}$. $\square$

The σ-field $\sigma(\mathcal{G})$ of Proposition 4.1 is referred to as the σ-field generated by $\mathcal{G}$.

Proposition 4.3. *Let* $\sigma(\mathcal{G})$ *be the* σ*-field generated by a family of subsets* $\mathcal{G}$ *of* Ω*. Let* $\mathcal{A}$ *be another family of subsets of* Ω*. Then,*

(a) if $\mathcal{A}$ *is a* σ*-field s.t.* $\mathcal{G} \subset \mathcal{A}$ *and* $\mathcal{A} \subset \sigma(\mathcal{G})$*, then* $\mathcal{A} = \sigma(\mathcal{G})$*.*

(b) $\mathcal{A} \subset \mathcal{G} \Rightarrow \sigma(\mathcal{A}) \subset \sigma(\mathcal{G})$*;*

(c) $\mathcal{A} \subset \mathcal{G} \subset \sigma(\mathcal{A}) \Rightarrow \sigma(\mathcal{A}) = \sigma(\mathcal{G})$*;*

Example 4.7. *Let $\Omega \neq \emptyset$ and let $\mathcal{G} = \{\emptyset\}$. Then, $\sigma(\mathcal{G}) = \{\emptyset, \Omega\}$, the trivial σ-field on Ω (cf. Example 4.1). By (a) of Proposition 4.3 it is enough to show that $\{\emptyset, \Omega\} \subset \sigma(\mathcal{G})$ since $\{\emptyset, \Omega\}$ is a σ-field that contains $\mathcal{G}$. It is clear that $\{\emptyset, \Omega\} \subset \sigma(\mathcal{G})$ since $\sigma(\mathcal{G})$ is a σ-field, hence it must contain Ω.*

Example 4.8. *Let $\Omega = \{1, 2, 3\}$ and define $\mathcal{G} = \{\{1\}\}$. Clearly, $\sigma(\mathcal{G})$ must contain $\{1\}$. Since $\sigma(\mathcal{G})$ is a σ-field, it must contain Ω and its complement $\emptyset$. Further, it must contain the complement of $\{1\}$, i.e., it contains $\{1\}^c = \{2, 3\}$. Thus, we claim that*

$$\sigma(\mathcal{G}) = \big\{\emptyset, \{1\}, \{2, 3\}, \Omega\big\}.$$

By (a) of Proposition 4.3, if $\big\{\emptyset, \{1\}, \{2, 3\}, \Omega\big\}$ is a σ-field on Ω, the claim is true. This is the case (cf. Exercise 4.3). Generally, upon Exercise 4.3, if $A \subset \Omega$, then,

$$\sigma(\{A\}) = \{\emptyset, A, A^c, \Omega\}.$$

We often omit the braces and use the notation $\sigma(A) = \sigma(\{A\})$.

Example 4.9. *Let $\Omega \neq \emptyset$ and*

$$\mathcal{F} = \{A \colon A \subset \Omega \text{ s.t. } A \text{ is countable or } A^c \text{ is countable}\},$$

i.e., $\mathcal{F}$ is the σ-field introduced in Example 4.3. We show that

$$\mathcal{F} = \sigma(\mathcal{G}),$$

where

$$\mathcal{G} = \big\{\{\omega\},\ \omega \in \Omega\big\}.$$

Clearly $\mathcal{G} \subset \mathcal{F}$ since each set $\{\omega\}$, $\omega \in \Omega$, has cardinality one. Thus, it remains to show that $\mathcal{F} \subset \sigma(\mathcal{G})$. Let $A \in \mathcal{F}$. Then either A or A^c is countable. Suppose that A is countable, then

$$A = \{\omega_i \colon i \in \mathbb{N}\},$$

for some collection of singletons $\omega_i \in \Omega$, $i \in \mathbb{N}$. Therefore, $A = \cup_{i\in\mathbb{N}}\{\omega_i\}$ and thus $A \in \sigma(\mathcal{G})$ since $\{\omega_i\} \in \sigma(\mathcal{G})$ for any $i \in \mathbb{N}$ and $\sigma(\mathcal{G})$ is a σ-field. If A^c is countable, by the latter argument, $A^c \in \sigma(\mathcal{G})$. But then, $(A^c)^c = A \in \sigma(\mathcal{G})$. The set $\mathcal{G}$ is referred to as the point-partition on Ω.

Exercise 4.6. *Let $\Omega \neq \emptyset$ and*

$$\mathcal{G} = \{A \colon A \subset \Omega \text{ s.t. } A \text{ is finite or } A^c \text{ is finite}\},$$

i.e., $\mathcal{G}$ is the family introduced in Example 4.5. Let $\mathcal{A} = \big\{\{\omega\} \colon \omega \in \Omega\big\}$. Show that

(a) $\mathcal{G} \subset \sigma(\mathcal{A})$;

(b) $\sigma(\mathcal{A}) = \mathcal{G}$ if Ω is finite.

Proposition 4.4. *Let Ω be a non empty set an $\mathcal{F}$ be a σ-field on Ω. Let $\Omega_0 \subset \Omega$ s.t. $\Omega_0 \neq \emptyset$. Define $\mathcal{F} \cap \Omega_0$ as in Example 4.4 and assume $\mathcal{F} = \sigma(\mathcal{G})$ for some family $\mathcal{G}$ of subsets of Ω. Then,*

$$\mathcal{F} \cap \Omega_0 = \sigma(\{G \cap \Omega_0 \colon G \in \mathcal{G}\}).$$

Proof. Given any $G \in \mathcal{G}$, since $\mathcal{F} = \sigma(\mathcal{G})$, it follows that $G \cap \Omega_0 \in \mathcal{F} \cap \Omega_0$. That is $\mathcal{F} \cap \Omega_0$ is a σ-field on Ω_0 that contains $\{G \cap \Omega_0 \colon G \in \mathcal{G}\}$. Since $\sigma(\{G \cap \Omega_0 \colon G \in \mathcal{G}\})$ is the smallest such σ-field, it follows that $\sigma(\{G \cap \Omega_0 \colon G \in \mathcal{G}\}) \subset \mathcal{F} \cap \Omega_0$. To make the notation easier, we write $\sigma(\{G \cap \Omega_0 \colon G \in \mathcal{G}\}) = \mathcal{F}_0$. Hence, it remains to show that $\mathcal{F} \cap \Omega_0 \subset \mathcal{F}_0$. Suppose that $A \in \mathcal{F}$ implies that $A \in \{A \subset \Omega \colon A \cap \Omega_0 \in \mathcal{F}_0\}$. Then, given any $B \in \mathcal{F} \cap \Omega_0$, i.e., $B = A \cap \Omega_0$ for $A \in \mathcal{F}$, it follows that $B \in \mathcal{F}_0$. Therefore, it is sufficient to show that $\mathcal{F} \subset \{A \subset \Omega \colon A \cap \Omega_0 \in \mathcal{F}_0\}$. We set $\mathcal{A} = \{A \subset \Omega \colon A \cap \Omega_0 \in \mathcal{F}_0\}$. We notice first that $\mathcal{G} \subset \mathcal{A}$, since any set $G \in \mathcal{G}$ is s.t. $G \cap \Omega_0 \in \mathcal{F}_0$ by definition. If we show that $\mathcal{A}$ is a σ-field on Ω, we are done, since then $\mathcal{F} = \sigma(\mathcal{G}) \subset \mathcal{A}$. We notice that by definition $\mathcal{F}_0$ is a σ-field on Ω_0, i.e., it contains Ω_0. Thus, $\Omega \in \mathcal{A}$ since $\Omega \cap \Omega_0 = \Omega_0 \in \mathcal{F}_0$. If $A \in \mathcal{A}$, then $A \cap \Omega_0 \in \mathcal{F}_0$ and hence $\Omega_0 \setminus A \cap \Omega_0 = \Omega_0 \setminus A \in \mathcal{F}_0$. Hence, $\Omega \setminus A$ is s.t. $(\Omega \setminus A) \cap \Omega_0 = \Omega_0 \setminus A \in \mathcal{F}_0$. Therefore, $\Omega \setminus A \in \mathcal{A}$. Suppose that $\{A_i \colon i \in \mathbb{N}\} \subset \mathcal{A}$. Then, $A_i \cap \Omega_0 \in \mathcal{F}_0$ for any $i \in \mathbb{N}$. Therefore, $\cup_{i\in\mathbb{N}}(A_i \cap \Omega_0) = (\cup_{i\in\mathbb{N}} A_i) \cap \Omega_0 \in \mathcal{F}_0$. That is, $\cup_{i\in\mathbb{N}} A_i \in \mathcal{A}$. □

Example 4.10. *Let $\Omega = \mathbb{R}$ and $\mathcal{R}$ be the family of left-open intervals with the empty set adjoined (cf. Example 4.6). The σ-field $\mathfrak{B}(\mathbb{R}) = \sigma(\mathcal{R})$ is referred to as the Borel σ-field on $\mathbb{R}$.*

Exercise 4.7. *Let $\mathcal{R}' = \{(-\infty, x] \colon x \in \mathbb{R}\}$. Show that $\mathfrak{B}(\mathbb{R}) = \sigma(\mathcal{R}')$.*

Example 4.11. *Let $\Omega = \mathbb{R}^k$, $k \in \mathbb{N}$. Given $a_i, b_i \in \mathbb{R}$, $i = 1, \dots, k$, a set*

$$\prod_{i=1}^{k} (a_i, b_i],$$

is referred to as a rectangle on $\mathbb{R}^k$. Define

$$\mathcal{R}_k = \big\{A \colon A = \prod_{i=1}^{k} (a_i, b_i],\ a_i, b_i \in \mathbb{R},\ i = 1, \dots, k\big\} \cup \{\emptyset\},$$

i.e., the family of rectangles in $\mathbb{R}^k$. Then, the σ-field $\mathfrak{B}(\mathbb{R}^k) = \sigma(\mathcal{R}_k)$ is referred to as the Borel σ-field on $\mathbb{R}^k$.

Exercise 4.8. *Let*

$$\mathcal{R}'_k = \{(-\infty, x_1] \times \cdots (-\infty, x_k] \colon x = (x_1, \dots, x_k) \in \mathbb{R}^k\} \cup \{\emptyset\}.$$

Show that $\mathfrak{B}(\mathbb{R}^k) = \sigma(\mathcal{R}'_k)$.

Having in mind Example 4.4, we define the Borel σ-field restricted to a subset of $\mathbb{R}^k$ as follows:

Definition 4.2. *Let $E \subset \mathbb{R}^k$, $E \neq \emptyset$. The Borel σ-field on E is defined by*

$$\mathfrak{B}(E) = \mathfrak{B}(\mathbb{R}^k) \cap E = \{A \cap E \colon A \in \mathfrak{B}(\mathbb{R}^k)\}.$$

Then, using Proposition 4.4, we obtain that

Proposition 4.5. *Given any $E \subset \mathbb{R}^k$ s.t. $E \neq \emptyset$,*

$$\mathfrak{B}(E) = \sigma(\{G \cap E \colon G \in \mathcal{R}_k\}).$$

Definition 4.3. *Let $\Omega \neq \emptyset$ and $\mathcal{F}$ be a σ-field on Ω. The pair $(\Omega, \mathcal{F})$ is referred to as a measurable space. If $A \in \mathcal{F}$, then A is said to be measurable. If $\mathcal{A} \subset \mathcal{F}$ and $\mathcal{A}$ is a σ-field on Ω, $\mathcal{A}$ is referred to as a sub-σ-field on Ω.*

4.2 Solution to exercises

Solution 4.1 (Solution to Exercise 4.1)**.** *Yes, if Ω is countable, then $\mathcal{F} = \mathcal{P}(\Omega)$. This follows from the fact that in this case, any set from $\mathcal{P}(\Omega)$ is countable and hence a member of $\mathcal{F}$.*

Solution 4.2 (Solution to Exercise 4.2)**.** *We have seen that (cf. Exercise 1.10) for any $n \in \mathbb{N}$,*

$$\bigcap_{i=1}^{n} A_i = \left(\bigcup_{i=1}^{n} A_i^c \right)^c.$$

Hence, (a) follows from (ii) and (iii) of Definition 4.1. With regard to (b), using Proposition 1.4, $A \setminus B = A \cap B^c$ and thus (b) follows from (a) and (ii) of Definition 4.1.

Solution 4.3 (Solution to Exercise 4.3)**.** *Let $\mathcal{F} = \{\emptyset, A, A^c, \Omega\}$. We need to verify items (i), (ii) and (iii) of Definition 4.1. Items (i) and (ii) are clearly satisfied. Since $\mathcal{F}$ only contains 4 sets, we can argue case by case. Let $k \in \{2, 3\}$ and*

$$\bigcap_{i_k=1}^{k} A_{i_k},$$

be any intersection of subsets $A_{i_k} \in \mathcal{F}$. Clearly, for any k, if $A_{i_k} = \emptyset$ for some i_k, $\cap_{i_k=1}^{k} A_{i_k} = \emptyset \in \mathcal{F}$. Thus, we have that for any $k = 2, 3$,

$$\bigcap_{i_k=1}^{k} A_{i_k} = \begin{cases} \emptyset, & \text{if } \exists\, i_k \text{ s.t. } A_{i_k} = \emptyset, \\ \bigcap_{i_k=1}^{k} A_{i_k}^1, & \text{otherwise,} \end{cases}$$

where $A_{i_k}^1$ are updated in the sense that $A_{i_k}^1 \in \{A, A^c, \Omega\}$. If $k = 2$, i.e., we intersect two subsets of $\{A, A^c, \Omega\}$,

$$\bigcap_{i_k=1}^{k} A_{i_k}^1 = \begin{cases} A, & \text{if } A_{i_k}^1 \in \{A, \Omega\},\ i_k = 1, 2 \\ A^c, & \text{if } A_{i_k}^1 \in \{A^c, \Omega\},\ i_k = 1, 2 \\ \emptyset, & \text{if } A_{i_k}^1 \in \{A^c, A\},\ i_k = 1, 2. \end{cases}$$

Otherwise, if $k = 3$, $\cap_{i_k=1}^{k} A_{i_k} = \emptyset$, since $A \cap A^c = \emptyset$. This shows that any possible intersection of sets from $\mathcal{F}$ is again a member of $\mathcal{F}$. Therefore, since for any $k \in \{2, 3\}$,

$$\bigcup_{i_k=1}^{k} A_{i_k} = \left(\bigcap_{i_k=1}^{k} A_{i_k}^c \right)^c,$$

item (iii) is satisfied.

Solution 4.4 (Solution to Exercise 4.4)**.** *Yes, if Ω is finite, then $\mathcal{G} = \mathcal{P}(\Omega)$. This follows from the fact that in this case, any set from $\mathcal{P}(\Omega)$ is finite and hence a member of $\mathcal{G}$.*

Solution 4.5 (Solution to Exercise 4.5)**.** *We need to verify items (i), (ii) and (iii) of Definition 4.1. Since $\Omega \in \mathcal{F}_i$ for any $i \in I$, $\Omega \in \mathcal{F} = \cap_{i \in I} \mathcal{F}_i$ by definition. Also, if $A \in \mathcal{F}$, then $A \in \mathcal{F}_i$ for any $i \in I$. In particular, $A^c \in \mathcal{F}_i$ for any $i \in I$. Thus, $A^c \in \mathcal{F}$. If $\{A_n \colon n \in \mathbb{N}\} \subset \mathcal{F}$ we have that $\{A_n \colon n \in \mathbb{N}\} \subset \mathcal{F}_i$ for any $i \in I$ (cf. Exercise 1.8). Thus, $\cup_{i \in I} A_i \in \mathcal{F}_i$ for any $i \in I$. Hence, $\cup_{i \in I} A_i \in \mathcal{F}$.*

Solution 4.6 (Solution to Exercise 4.6)**.** *Item (a) follows from the fact that $\sigma(\mathcal{A})$ is a σ-field: If $A \in \mathcal{G}$, then since either A or A^c is finite, either of the two is a countable union of elements from $\mathcal{A}$, i.e., $A \in \sigma(\mathcal{A})$. With regard to (b), it remains to show that $\sigma(\mathcal{A}) \subset \mathcal{G}$ if Ω is finite. This is certainly true since in this case $\mathcal{G} = \mathcal{P}(\Omega)$.*

Solution 4.7 (Solution to Exercise 4.7). *We show that $\sigma(\mathcal{R})$ contains $\mathcal{R}'$ ($\Rightarrow \sigma(\mathcal{R}') \subset \sigma(\mathcal{R})$) and $\sigma(\mathcal{R}')$ contains $\mathcal{R}$ ($\Rightarrow \sigma(\mathcal{R}) \subset \sigma(\mathcal{R}')$). Let $x \in \mathbb{R}$, then, $(-\infty, x] = \cup_{n\in I}(-n, x] \in \sigma(\mathcal{R})$, where $I = \{n \in \mathbb{N}\colon -n < x\}$. Let $(a, b] \in \mathcal{R}$, then $(a, b]^c = (-\infty, a] \cup (b, \infty) = (-\infty, a] \cup (-\infty, b]^c \in \sigma(\mathcal{R}')$.*

Solution 4.8 (Solution to Exercise 4.8). *The inequality $\sigma(\mathcal{R}'_k) \subset \sigma(\mathcal{R}_k)$ follows from the fact that for any $x = (x_1, \ldots, x_k) \in \mathbb{R}^k$,*

$$(-\infty, x_1] \times \ldots \times (-\infty, x_n] = \bigcup_{n\in I} \Big((-n, x_1] \times \ldots \times (-n, x_n] \Big),$$

where $I = \{n \in \mathbb{N}\colon -n < \min\{x_i\colon i = 1, \ldots, k\}\}$. For the other inequality, let $A = \prod_{i=1}^k (a_i, b_i] \in \mathcal{R}_k$. Consider the sets $(-\infty, b_i]$, $i = 1, \ldots, k$. Define the sets

$$S_i = (-\infty, b_1] \times \cdots \times \underbrace{(-\infty, a_i]}_{Position\ i} \times \cdots \times (-\infty, b_k], \quad i = 1, \ldots, k.$$

Then,

$$A = \Big(\prod_{i=1}^k (-\infty, b_i] \Big) \setminus \Big(\bigcup_{i=1}^k S_i \Big).$$

To see it, we notice that by Proposition 1.3,

$$\Big(\prod_{i=1}^k (-\infty, b_i] \Big) \setminus \Big(\bigcup_{i=1}^k S_i \Big) = \bigcap_{i=1}^k \Big(\Big(\prod_{i=1}^k (-\infty, b_i] \Big) \setminus S_i \Big)$$

Then, we have that for any $i = 1, \ldots, k$,

$$\Big(\prod_{i=1}^k (-\infty, b_i] \Big) \setminus S_i = \Big(\prod_{i=1}^k (-\infty, b_i] \Big) \cap S_i^c = (-\infty, b_1] \times \cdots \times \underbrace{(a_i, b_i]}_{Position\ i} \times \cdots \times (-\infty, b_k].$$

Thus,

$$\bigcap_{i=1}^k \Big((-\infty, b_1] \times \cdots \times \underbrace{(a_i, b_i]}_{Position\ i} \times \cdots \times (-\infty, b_k] \Big) = A,$$

as desired. Now,

$$\Big(\prod_{i=1}^k (-\infty, b_i] \Big) \setminus \Big(\bigcup_{i=1}^k S_i \Big) \in \sigma(\mathcal{R}'_k).$$

Hence, $\sigma(\mathcal{R}'_k)$ is another σ-field which contains sets of the form A. We know that $\sigma(\mathcal{R}_k)$ is the smallest such σ-field. Hence, $\sigma(\mathcal{R}_k) \subset \sigma(\mathcal{R}'_k)$.

4.3 Additional exercises

Exercise 4.9. *Let $f\colon X \to Y$ be a function and $\mathfrak{B}$ be a σ-field on Y. Show that*

$$\sigma(f) = \{f^{-1}(B)\colon B \in \mathfrak{B}\},$$

is a σ-field on X. $\sigma(f)$ is referred to as the σ-field generated by f.

Exercise 4.10. *Find an example of a set $\Omega \neq \emptyset$ and two σ-fields $\mathcal{F}_1$ and $\mathcal{F}_2$ on Ω s.t. the union $\mathcal{F}_1 \cup \mathcal{F}_2$ is not a σ-field on Ω. This shows that in contrast to the intersection of σ-fields (cf. Proposition 4.2), the union of σ-fields is not necessarily a σ-field.*
Hint: *Example 4.8.*

Exercise 4.11. *Let* $\Omega = [0,1]$ *and* $\mathcal{G} = \{[0,1/2],[1/2,1]\}$*. Find* $\sigma(\mathcal{G})$*.*

Exercise 4.12. *Let* $\Omega = \mathbb{R}$*, equipped with the Borel* σ*-field* $\mathfrak{B}(\mathbb{R})$ *(cf. Example 4.10). Show that*

$$\sigma(\mathcal{R}_*) = \mathfrak{B}(\mathbb{R}),$$

where

$$\mathcal{R}_* = \{A\colon A = [a,b),\ a,b \in \mathbb{R}\} \cup \{\emptyset\},$$

i.e., the right-open intervals with the empty set adjoined.

Exercise 4.13. *We remain in the setting of Exercise 4.12. Show that*

$$\mathfrak{B}(\mathbb{R}) = \sigma(\mathcal{R}_{**}),$$

where

$$\mathcal{R}_{**} = \{A\colon A = (q_1,q_2],\ q_1,q_2 \in \mathbb{Q}\} \cup \{\emptyset\},$$

i.e., the left-open intervals with rational end-points and the empty set adjoined.
Hint: *Recall Example 3.4.*

5 Measurable sets: Part II

5.1 Measure spaces

Definition 5.1. *Let $(\Omega, \mathcal{F})$ be a measurable space. A function $\mu \colon \mathcal{F} \to \overline{\mathbb{R}}_+$ is said to be a measure on $\mathcal{F}$ if the following two items are satisfied:*

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

(ii) Given any disjoint family $\{A_i \colon i \in \mathbb{N}\}$ of measurable sets (i.e., $\{A_i \colon i \in \mathbb{N}\} \subset \mathcal{F}$),

$$\mu\Big(\bigcup_{i\in\mathbb{N}} A_i\Big) = \sum_{i\in\mathbb{N}} \mu(A_i).$$

Example 5.1. *Let $(\Omega, \mathcal{F})$ be a measurable space and $x \in \Omega$ be a given point of Ω. Define the function*

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

Then, $A \mapsto \delta_x(A)$ is a measure on $\mathcal{F}$. From the definition of δ_x, we immediately see that $\delta_x(\emptyset) = 0$. Let $\{A_i \colon i \in \mathbb{N}\} \subset \mathcal{F}$ be disjoint and set $A = \cup_{i\in\mathbb{N}} A_i$. If $x \notin A$ (i.e., $x \in \cap_{i\in\mathbb{N}} A_i^c$), there does not exists i s.t. $x \in A_i$, hence $\delta_x(A) = 0 = \sum_{i\in\mathbb{N}} \delta_x(A_i)$. Otherwise, if $x \in A$, since $\{A_i \colon i \in \mathbb{N}\}$ is disjoint, there exists a unique $j \in \mathbb{N}$, s.t. $x \in A_j$. Hence also in this case

$$\delta_x(A) = 1 = \delta_x(A_j) = \sum_{i\in\mathbb{N}} \delta_x(A_i).$$

Example 5.2. *We consider the measurable space $(\Omega, \mathcal{P}(\Omega))$, where Ω is a finite set and $\mathcal{P}(\Omega)$ is the power set of Ω (cf. Example 4.2). Define $\mu(A) = \#A$, $A \in \mathcal{P}(\Omega)$. Then, μ is a measure on $\mathcal{F}$. Clearly, for any $A \in \mathcal{P}(\Omega)$, $\mu(A) \geq 0$. We need to verify (i) and (ii) of Definition 5.1. Since $\#\emptyset = 0$, (i) is satisfied. Let $\{A_i \colon i \in \mathbb{N}\} \subset \mathcal{P}(\Omega)$ be disjoint. Naturally, since $\{A_i \colon i \in \mathbb{N}\}$ is a disjoint family of sets, the cardinality of its union is the sum of the individual set cardinalities. Let us show it. Clearly, since Ω is finite, the sets A_i, $i \in \mathbb{N}$, and $A = \cup_{i\in\mathbb{N}} A_i \subset \Omega$ are all finite sets. Let I_1 be s.t. $A_i = \emptyset$ for any $i \in I_1$ and $I_2 = \mathbb{N} \setminus I_1$, i.e., $A_i \neq \emptyset$ for any $i \in I_2$. Then, since A is finite, I_2 must be finite as well. That is, $A_i \neq \emptyset$ only for finitely many i. Then, for any $i \in I_2$, let $A_i = \{x_{i_1}, \ldots, x_{i_{n_i}}\}$, $n_i \in \mathbb{N}$. Since $\{A_i \colon i \in I_2\}$ is disjoint,*

$$\mu\Big(\bigcup_{i\in I_2} A_i\Big) = \#\bigcup_{i\in I_2} A_i = \sum_{i\in I_2} \#\{x_{i_k} \colon k = 1, \ldots, n_i\} = \sum_{i\in I_2} \mu(A_i).$$

In conclusion, we obtain that

$$\begin{aligned}
\mu(A) &= \mu\Big(\bigcup_{i\in I_1} A_i \cup \bigcup_{i\in I_2} A_i\Big) \\
&= \mu\Big(\bigcup_{i\in I_2} A_i\Big) \\
&= \sum_{i\in I_2} \mu(A_i) \\
&= \sum_{i\in I_1} \mu(A_i) + \sum_{i\in I_2} \mu(A_i) = \sum_{i\in\mathbb{N}} \mu(A_i).
\end{aligned}$$

We can generalize the previous example.

Example 5.3. *Consider the measurable space $(\Omega, \mathcal{P}(\Omega))$, where $\mathcal{P}(\Omega)$ is the power set of Ω and Ω is not necessarily finite. Define*

$$\mu(A) = \begin{cases} \#A, & \text{if } A \in \mathcal{P}(\Omega) \text{ s.t. } A \text{ is finite,} \\ \infty & \text{otherwise.} \end{cases}$$

Then, μ is a measure on $\mathcal{P}(\Omega)$. We have already seen in the previous example that (i) of Definition 5.1 is satisfied. Thus, let $\{A_i : i \in \mathbb{N}\} \subset \mathcal{P}(\Omega)$ be disjoint. Suppose that there exists $j \in \mathbb{N}$ s.t. A_j is not finite. Then,

$$\mu\Big(\bigcup_{i\in\mathbb{N}} A_i\Big) = \infty = \sum_{i\in\mathbb{N}} \mu(A_i).$$

Thus, it remains to show (ii) for the case where A_i is finite for any $i \in \mathbb{N}$. Given $i \in \mathbb{N}$, let $A_i = \{x_{i_1}, \dots, x_{i_{n_i}}\}$, $n_i \in \mathbb{N}$. Set $U_n = \cup_{i=1}^n A_i$ and $A = \cup_{i\in\mathbb{N}} A_i$. As in the previous example, since the family $\{A_i : i \in \mathbb{N}\}$ is disjoint, we have that for any $n \in \mathbb{N}$,

$$\mu(U_n) = \#\{x_{i_k} : k = 1, \dots, n_i,\ 1 \le i \le n\} = \sum_{i=1}^n \#\{x_{i_k} : k = 1, \dots, n_i\} = \sum_{i=1}^n \mu(A_i).$$

We show that

$$\lim_{n\to\infty} \mu(U_n) = \mu(A), \tag{8}$$

then, (ii) of Definition 5.1 is verified. There are two cases, either A is finite or not. If A is finite, we repeat the arguments from Example 5.1 and obtain

$$\mu(A) = \sum_{i\in\mathbb{N}} \mu(A_i).$$

Thus, if we show (8) for the case wehen A is not finite, we are done. If A is not finite, then, for any $N \in \mathbb{N}$, there exists an integer $n \ge N$, s.t. A_n is not the empty set. In particular, for any $N \in \{n \in \mathbb{N} : n \ge 2\}$ there exists an integer $n \ge N$ s.t. $\#U_n - \#U_{n-1} = \#A_n = m_n$ for some positive integer m_n. Hence, we let $n_1 \in \mathbb{N}$ be s.t.

$$\#U_{n_1} = \#U_{n_1-1} + m_{n_1}, \quad n_1 \ge 1 + 1.$$

Then, we let n_2 be s.t.

$$\#U_{n_2} = \#U_{n_2-1} + m_{n_2}, \quad n_2 \ge n_1 + 1.$$

We continue like that and let n_k, $k \in \mathbb{N}$, be s.t.

$$\#U_{n_k} = \#U_{n_k-1} + m_{n_k}, \quad n_k \ge n_{k-1} + 1.$$

Now we note that for any $n \in \mathbb{N}$, $\#U_n \le \#U_{n+1}$, i.e., $(\#U_n)_{n\in\mathbb{N}}$ is increasing. Therefore,

$$\begin{aligned} \#U_{n_k} = \#U_{n_k-1} + m_{n_k} \ge \#U_{n_{k-1}+1-1} + m_{n_k} &= \#U_{n_{k-1}} + m_{n_k} \\ &= \#U_{n_{k-1}-1} + m_{n_{k-1}} + m_{n_k}. \end{aligned}$$

This shows that for any $k \in \mathbb{N}$,

$$\#U_{n_k} \ge \#U_{n_1-1} + \sum_{i=0}^{k-1} m_{n_{k-i}}.$$

Hence, given any $M \in \mathbb{R}$, if we let $k \in \mathbb{N}$ be large enough s.t. $\#U_{n_1-1} + \sum_{i=0}^{k-1} m_{n_{k-i}} \geq M$, we have that $\#U_n \geq M$ for any $n \geq n_k$. Finding such an integer k is certainly possible, since $\#U_{n_1-1}$ is a finite number and

$$\sum_{i=0}^{k-1} m_{n_{k-i}} \geq \underbrace{1 + \cdots + 1}_{k\text{-times}} = k.$$

In Particular, for any $M \in \mathbb{R}$ there exists $N = n_k \in \mathbb{N}$ s.t. $\#U_n \geq M$ for any $n \geq N$. Hence, by Definition 3.6, $\lim_{n\to\infty} \#U_n = \lim_{n\to\infty} \mu(U_n) = \infty$. Clearly, by definition of μ, since A is not finite, $\mu(A) = \infty$. This shows (8) and completes the argument. The measure μ is called the counting measure.

Let us also consider a more general version of Example 5.1.

Example 5.4. *Let $(\Omega, \mathcal{F})$ be a measurable space, I be a countable set and $E = \{x_i : i \in I\} \subset \Omega$ be a collection of points in Ω. Assume that α_i, $i \in I$, are s.t. $\alpha_i \in [0, \infty)$ for any $i \in I$. Define*

$$\mu(A) = \sum_{i \in I} \alpha_i \delta_{x_i}(A) = \sum_{x \in E} \alpha_x \delta_x(A), \quad A \in \mathcal{F},$$

where for any $i \in I$,

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

Then, μ is a measure on $\mathcal{F}$. It is clear that (i) of Definition 5.1 is satisfied. Let $\{A_k : k \in \mathbb{N}\} \subset \mathcal{F}$ be disjoint and set $A = \cup_{k\in\mathbb{N}} A_k$. We need to verify that

$$\mu(A) = \sum_{k \in \mathbb{N}} \mu(A_k).$$

Certainly, this is true if for any $i \in I$, $x_i \notin A$. Assume that there exists $i \in I$ s.t. $x_i \in A$. Let us label these i's, i.e., i_1 is s.t. $x_{i_1} \in A$, i_2 is s.t. $x_{i_2} \in A$ and so on to obtain a family $\{x_{i_j} : j \in J\}$, where the set J is s.t. $J \subset I$ and for any $j \in J$, $x_{i_j} \in A$. We assume that any $x \in \{x_i : i \in I\} \setminus \{x_{i_j} : j \in J\}$ is s.t. $x \notin A$. Otherwise, we can always enlarge the set J. Then, since $\{A_k : k \in \mathbb{N}\}$ is disjoint, for each $j \in J$, there exists a unique $k_j \in \mathbb{N}$ s.t. $x_{i_j} \in A_{k_j}$. Therefore,

$$\begin{aligned}
\mu(A) &= \sum_{i \in I} \alpha_i \delta_{x_i}(A) \\
&= \sum_{x \in \{x_{i_j} : j \in J\}} \alpha_x \delta_x(A) + \sum_{x \notin \{x_{i_j} : j \in J\}} \alpha_x \delta_x(A) \\
&= \sum_{j \in J} \alpha_{i_j} \delta_{x_{i_j}}(A) \\
&= \sum_{j \in J} \alpha_{i_j} \delta_{x_{i_j}}(A_{k_j}) = \sum_{j \in J} \alpha_{i_j}.
\end{aligned}$$

Then, we notice that if $k \in \mathbb{N}$,

$$\begin{aligned}
\sum_{i \in I} \alpha_i \delta_{x_i}(A_k) &= \sum_{x \in \{x_{i_j} : j \in J\}} \alpha_x \delta_x(A_k) + \sum_{x \notin \{x_{i_j} : j \in J\}} \alpha_x \delta_x(A_k) \\
&= \sum_{x \in \{x_{i_j} : j \in J\}} \alpha_x \delta_x(A_k) \\
&= \sum_{j \in J} \alpha_{i_j} \delta_{x_{i_j}}(A_k).
\end{aligned}$$

Hence, since for any $j \in J$, $\sum_{k\in\mathbb{N}} \delta_{x_{i_j}}(A_k) = \delta_{x_{i_j}}(A_{k_j}) = 1$,

$$\begin{aligned}\sum_{k\in\mathbb{N}} \mu(A_k) &= \sum_{k\in\mathbb{N}} \Big(\sum_{i\in I} \alpha_i \delta_{x_i}(A_k) \Big) \\ &= \sum_{k\in\mathbb{N}} \Big(\sum_{j\in J} \alpha_{i_j} \delta_{x_{i_j}}(A_k) \Big) \\ &= \sum_{j\in J} \alpha_{i_j} \Big(\sum_{k\in\mathbb{N}} \delta_{x_{i_j}}(A_k) \Big) \\ &= \sum_{j\in J} \alpha_{i_j} \delta_{x_{i_j}}(A_{k_j}) = \sum_{j\in J} \alpha_{i_j}.\end{aligned}$$

It is important to remark that since $\alpha_{i_j}\delta_{x_{i_j}}(A_k) = a_{k,j} \geq 0$, *the double sum* $\sum_{(k,j)\in\mathbb{N}\times J} a_{k,j}$ *is well defined and we are allowed to interchange the order of summation (cf. Proposition 3.11). Clearly, if* I *is finite,* $\mu(A) < \infty$ *for any* $A \in \mathcal{F}$. *If* I *is countably infinite, we have that* $\#I = \#\mathbb{N}$ *and hence, for any* $A \in \mathcal{F}$, $\mu(A) = \sum_{n\in\mathbb{N}} \alpha_n \delta_{x_n}(A)$. *Therefore, since* $\alpha_i \delta_{x_i}(A) \leq \alpha_i$ *and* $\alpha_i \geq 0$, *if follows from Proposition 3.10 that* $\mu(A) < \infty$ *for any* $A \in \mathcal{F}$ *if* $\sum_{i\in I} \alpha_i < \infty$.

Exercise 5.1. *Consider the measurable space* $(\mathbb{R}, \mathfrak{B}(\mathbb{R}))$, *where* $\mathfrak{B}(\mathbb{R})$ *is the Borel* σ*-field on* $\mathbb{R}$. *Define* $\mu(B) = \sum_{n\in B\cap\mathbb{N}} 2^{-n}$, $B \in \mathfrak{B}(\mathbb{R})$. *Is* μ *a measure on* $\mathfrak{B}(\mathbb{R})$*?*

Exercise 5.2. *Let* $E = \{0, 1\}$, $p \in (0, 1)$ *and set* $p_0 = 1 - p$ *and* $p_1 = p$. *Define the function*

$$P(B) = \sum_{x\in E\cap B} p_x, \quad B \in \mathfrak{B}(\mathbb{R}).$$

That is,

$$P(B) = \begin{cases} 0, & \text{if } 0 \notin B \text{ and } 1 \notin B, \\ 1-p, & \text{if } 0 \in B \text{ and } 1 \notin B, \\ p, & \text{if } 0 \notin B \text{ and } 1 \in B, \\ 1, & \text{if } 0 \in B \text{ and } 1 \in B. \end{cases}$$

Is P *is a measure on* $\mathfrak{B}(\mathbb{R})$*?*

Example 5.5. *Consider the measurable space* $(\mathbb{R}, \mathfrak{B}(\mathbb{R}))$, *where* $\mathfrak{B}(\mathbb{R}))$ *is the Borel* σ*-field on* $\mathbb{R}$ *(cf. Example 4.6 and 4.10). Then, there exists a unique measure* λ *on* $(\mathbb{R}, \mathfrak{B}(\mathbb{R}))$ *s.t. for any left-open interval* $(a, b]$, $a, b \in \mathbb{R}$, $\lambda((a, b])$ *returns the length of* $(a, b]$, *i.e.,* $\lambda((a, b]) = b - a$. *The measure* λ *is referred to as the Lebesgue measure on* $\mathfrak{B}(\mathbb{R})$. *An explicit construction of* λ *is given in the next section.*

In the following we list some general properties of measures.

Proposition 5.1. *Let* $(\Omega, \mathcal{F})$ *be a measurable space and* μ *be a measure on* $\mathcal{F}$. *Then,*

(i) given $n \in \mathbb{N}$ *and* $\{A_i : 1 \leq i \leq n\} \subset \mathcal{F}$ *disjoint, it follows that*

$$\mu\Big(\bigcup_{i=1}^{n} A_i \Big) = \sum_{i=1}^{n} \mu(A_i);$$

(ii) if $A, B \in \mathcal{F}$ *s.t.* $A \subset B$ *it follows that* $\mu(A) \leq \mu(B)$*;*

(iii) if $A, B \in \mathcal{F}$ *s.t.* $A \subset B$ *and* $\mu(A) < \infty$ *it follows that* $\mu(B \setminus A) = \mu(B) - \mu(A)$*;*

(iv) given $A, B \in \mathcal{F}$, $\mu(A) + \mu(B) = \mu(A \cup B) + \mu(A \cap B)$*;*

(v) if $\{A_i \colon i \in \mathbb{N}\} \subset \mathcal{F}$ *is s.t.* $A_i \subset A_{i+1}$*,*

$$\mu\Big(\bigcup_{i=1}^{n} A_i\Big) = \mu(A_n) \uparrow \mu\Big(\bigcup_{i\in\mathbb{N}} A_i\Big);$$

(vi) if $\{A_i \colon i \in \mathbb{N}\} \subset \mathcal{F}$ *is s.t.* $\mu(A_1) < \infty$ *and* $A_{i+1} \subset A_i$*,*

$$\mu\Big(\bigcap_{i=1}^{n} A_i\Big) = \mu(A_n) \downarrow \mu\Big(\bigcap_{i\in\mathbb{N}} A_i\Big).$$

(vii) if $\{A_i \colon n \in \mathbb{N}\} \subset \mathcal{F}$*,*

$$\mu\Big(\bigcup_{i\in\mathbb{N}} A_i\Big) \le \sum_{i\in\mathbb{N}} \mu(A_i);$$

Proof of item (v) of Proposition 5.1. Let $B_1 = A_1$, $B_2 = A_2 \setminus A_1$ and so on until we set for any $i \in \mathbb{N}$, $B_i = A_i \setminus A_{i-1}$. It is true that for any $i \neq j$, $B_i \cap B_j = \emptyset$. Either, $i < j$, and then

$$B_i = A_i \setminus A_{i-1} \subset A_i \subset A_{j-1} \subset (A_j^c \cup A_{j-1}) = (A_j \cap A_{j-1}^c)^c = (A_j \setminus A_{j-1})^c = B_j^c,$$

or $j > i$ and then $B_j \subset B_i^c$. Thus, The family $\{B_i \colon i \in \mathbb{N}\}$ is disjoint. By item (ii) of Definition 5.1,

$$\mu\Big(\bigcup_{i\in\mathbb{N}} B_i\Big) = \sum_{i\in\mathbb{N}} \mu(B_i),$$

and hence,

$$\sum_{i=1}^{n} \mu(B_i) \uparrow \mu\Big(\bigcup_{i\in\mathbb{N}} B_i\Big).$$

Further, for any $n \in \mathbb{N}$, $\cup_{i=1}^n B_i = \cup_{i=1}^n A_i$. To see it, we use an argument by induction. By definition, $B_1 = A_1$. Then,

$$\begin{aligned}\bigcup_{i=1}^{n+1} B_i &= \bigcup_{i=1}^{n} A_i \cup B_{n+1} = \bigcup_{i=1}^{n} A_i \cup A_{n+1} \setminus A_n \\ &= \bigcup_{i=1}^{n-1} A_i \cup A_n \cup A_{n+1} \setminus A_n = \bigcup_{i=1}^{n-1} A_i \cup A_n \cup A_{n+1}.\end{aligned}$$

Therefore, since for any $i \in \mathbb{N}$, $A_i \subset A_{i+1}$, it follows that $\cup_{i=1}^n B_i = A_n$. In conclusion, since

$$\bigcup_{i\in\mathbb{N}} A_i = \bigcup_{n\in\mathbb{N}} A_n = \bigcup_{n\in\mathbb{N}} \Big(\bigcup_{i=1}^{n} B_i\Big) = \bigcup_{i\in\mathbb{N}} B_i,$$

we have under application of item (i) of Proposition 5.1,

$$\mu(A_n) = \mu\Big(\bigcup_{i=1}^{n} B_i\Big) = \sum_{i=1}^{n} \mu(B_i) \uparrow \mu\Big(\bigcup_{i\in\mathbb{N}} B_i\Big) = \mu\Big(\bigcup_{i\in\mathbb{N}} A_i\Big).$$

$\square$

Proof of item (vii) of Proposition 5.1. Let $B_1 = A_1$, $B_2 = A_2 \setminus A_1$, $B_3 = A_3 \setminus A_1 \cup A_2$, and so on s.t. for any $i \in \mathbb{N}$,

$$B_i = A_i \setminus \bigcup_{k=1}^{i-1} A_k.$$

We notice that for $i \neq j$, $B_i \cap B_j = \emptyset$ since either $i < j$ and hence

$$B_i = A_i \setminus \cup_{k=1}^{i-1} A_k \subset A_i \subset \cup_{k=1}^{j-1} A_k \cup A_j^c = B_j^c,$$

or $j < i$ and then $B_j \subset B_i^c$. Hence, the family $\{B_i \colon i \in \mathbb{N}\}$ is disjoint. Further, we note that for any $n \in \mathbb{N}$,

$$\bigcup_{i=1}^{n} B_i = \bigcup_{i=1}^{n} A_i.$$

To see it, we can use an argument by induction. We have that $A_1 = B_1$. Then, assume that $\cup_{i=1}^{n} B_i = \cup_{i=1}^{n} A_i$. It follows that

$$\bigcup_{i=1}^{n+1} B_i = \left(\bigcup_{i=1}^{n} B_i\right) \cup B_{n+1} = \left(\bigcup_{i=1}^{n} A_i\right) \cup B_{n+1} = \bigcup_{i=1}^{n} A_i \cup \left(A_{n+1} \setminus \bigcup_{i=1}^{n} A_i\right) = \bigcup_{i=1}^{n+1} A_i.$$

Therefore, by item (i) of Proposition 5.1, we have that for any $n \in \mathbb{N}$,

$$\mu\left(\bigcup_{i=1}^{n} A_i\right) = \mu\left(\bigcup_{i=1}^{n} B_i\right) = \sum_{i=1}^{n} \mu(B_i) \leq \sum_{i=1}^{n} \mu(A_i), \tag{9}$$

where we used that for any $1 \leq i \leq n$, $B_i \subset A_i$ and hence $\mu(B_i) \leq \mu(A_i)$ by (ii) of Proposition 5.1. To conclude, we define $C_n = \cup_{i=1}^{n} A_i$ and obtain $C_n \subset C_{n+1}$ and hence

$$\mu\left(\bigcup_{i=1}^{n} A_i\right) \uparrow \mu\left(\bigcup_{i \in \mathbb{N}} A_i\right),$$

by (v) of Proposition 5.1. This concludes the argument (cf. Proposition 3.8). □

Exercise 5.3. *Prove items (i), (ii), (iii) and (iv) of Proposition 5.1.*

Remark 5.1. *We remark that item (iv) of Proposition 5.1 can be stated more generally for a finite collection of sets with finite measure (the details are given in Section B.1).*

Definition 5.2. *Let $(\Omega, \mathcal{F})$ be a measurable space and μ be a measure on $\mathcal{F}$. The triple $(\Omega, \mathcal{F}, \mu)$ is referred to as a measure space.*

5.2 Semirings

Definition 5.3. *Let Ω be a nonempty set. A family of subsets $\mathcal{A}$ of Ω is said to be a semiring on Ω if*

(i) $\emptyset \in \mathcal{A}$*;*

(ii) $A, B \in \mathcal{A} \Rightarrow A \cap B \in \mathcal{A}$*;*

(iii) if $A, B \in \mathcal{A}$ *and* $A \subset B$*, then there exists a disjoint collection of sets* $\{C_1, \ldots, C_n\} \subset \mathcal{A}$ *s.t.* $B \setminus A = \cup_{k=1}^{n} C_k$*.*

Example 5.6. *We have seen in Example 4.6 that the family of left-open intervals*

$$\mathcal{R} = \{A\colon\ A = (a,b],\ a,b \in \mathbb{R}\} \cup \{\emptyset\},$$

is not a σ-field on $\mathbb{R}$. Still, it is a semiring on $\mathbb{R}$. By definition, $\emptyset \in \mathcal{R}$. Let $A, B \in \mathcal{R}$, i.e., $A = (a_1, a_2]$ and $B = (b_1, b_2]$. If $A \cap B = \emptyset$, then $A \cap B \in \mathcal{R}$. Otherwise,

$$A \cap B = (\max\{a_1, b_1\}, \min\{a_2, b_2\}].$$

Thus, item (ii) of Definition 5.3 is satisfied. With regard to (iii), let $A, B \in \mathcal{R}$ s.t. $A \subset B$. If $A = B = \emptyset$, then $A \setminus B = \emptyset$ and with $C = \emptyset$, the result follows. Similarly, if $A = B$, the result follows with $C = \emptyset$. If $B \neq \emptyset$ but $A = \emptyset$, then $B \setminus A = B$ and with $C = B$, the result follows. Thus, suppose that $A, B \neq \emptyset$, $A \subset B$ and $A \neq B$, i.e., $b_1 < a_1$ and $a_2 < b_2$. We have that $B \setminus A = (b_1, a_1] \cup (a_2, b_2]$. Hence, with $C_1 = (b_1, a_1]$ and $C_2 = (a_2, b_2]$, the result follows.

Example 5.7. *Let Ω be an infinite set and consider the family*

$$\mathcal{G} = \{A\colon\ A \subset \Omega \text{ s.t. } A \text{ is finite or } A^c \text{ is finite}\}.$$

Then, $\mathcal{G}$ is not a σ-field on Ω (cf. Example 4.5). However, $\mathcal{G}$ is a semiring on Ω.

5.3 Solution to exercises

Solution 5.1 (Solution to Exercise 5.1)**.** *Notice that we can write for any $B \in \mathfrak{B}(\mathbb{R})$,*

$$\mu(B) = \sum_{n \in B \cap \mathbb{N}} 2^{-n}\delta_n(B) = \sum_{n \in B \cap \mathbb{N}} 2^{-n}\delta_n(B) + \sum_{n \in B^c \cap \mathbb{N}} 2^{-n}\delta_n(B) = \sum_{n \in \mathbb{N}} 2^{-n}\delta_n(B).$$

Thus, μ is a measure on $\mathfrak{B}(\mathbb{R})$ (cf. Example 5.4). We remark that upon Exercise 3.14, we have that $\sum_{n\in\mathbb{N}} 2^{-n} = 2$. That is, $\mu(A) < \infty$ for any $A \subset \mathfrak{B}(\mathbb{R})$.

Solution 5.2 (Solution to Exercise 5.2)**.** *We set $\alpha_0 = 1 - p$ and $\alpha_1 = p$ and obtain that for any $B \in \mathfrak{B}(\mathbb{R})$,*

$$\sum_{x \in E} \alpha_x \delta_x(B) = \sum_{x \in E \cap B} p_x.$$

Thus, P is a measure on $\mathfrak{B}(\mathbb{R})$ (cf. Example 5.4). We remark that $\sum_{x\in E} p_x = 1$.

Solution 5.3 (Solution to Exercise 5.3)**.**

(i) We define $A_i = \emptyset$ for any $i > n$ and the statement is verified as a consequence of (i) and (ii) in Definition 5.1.

(ii) In general, $A \cup B = A \cup (B \setminus A)$. Sine $A \subset B$, it follows that $B = A \cup (B \setminus A)$. Hence by (i), $\mu(B) = \mu(A) + \mu(B \setminus A)$. Since $\mu(B \setminus A) \geq 0$, the result follows. Notice that $\mu(A) = \infty$ is possible.

(iii) As with the previous solution, $\mu(B) = \mu(A) + \mu(B \setminus A)$. Now, since $\mu(A) < \infty$, we obtain $\mu(B) - \mu(A) = \mu(B \setminus A)$.

(iv) We write $A \cup B = (A \setminus B) \cup (B \setminus A) \cup (A \cap B)$. Then, since $\{A \setminus B, B \setminus A, A \cap B\}$ is disjoint, it follows that

$$\mu(A \cup B) = \mu(A \setminus B) + \mu(B \setminus A) + \mu(A \cap B)$$

Similarly, since $A \cap B$ is a subset of both, A and B, we write

$$A = (A \setminus (A \cap B)) \cup (A \cap B) = (A \setminus B) \cup (A \cap B).$$

and

$$B = (B \setminus (A \cap B)) \cup (A \cap B) = (B \setminus A) \cup (A \cap B).$$

Hence,

$$\mu(A) + \mu(B) = \mu(A \setminus B) + \mu(B \setminus A) + 2\mu(A \cap B) = \mu(A \cup B) + \mu(A \cap B).$$

5.4 Additional exercises

Exercise 5.4. *Suppose that $E \subset \mathbb{R}$ is a countable set. Let $f\colon E \to [0, \infty)$ be a function and define*

$$\mu(B) = \sum_{x \in E \cap B} f(x), \quad B \in \mathfrak{B}(\mathbb{R}).$$

Verify that μ is a measure on $\mathfrak{B}(\mathbb{R})$.

Exercise 5.5. *Let $E = \{0, 1, \dots, n\}$, $n \in \mathbb{N}$, and $p \in (0, 1)$. For any $k \in E$, let*

$$p_k = \binom{n}{k} p^k (1-p)^{n-k}.$$

Define the function

$$P(B) = \sum_{k \in E \cap B} p_k, \quad B \in \mathfrak{B}(\mathbb{R}).$$

Is P a measure on $\mathfrak{B}(\mathbb{R})$?

Exercise 5.6. *Let $(\Omega, \mathcal{F})$ be a measurable space and $\mu\colon \mathcal{F} \to \overline{\mathbb{R}}_+$ be a function s.t. $\mu(\emptyset) = 0$. Suppose that*

- *For any finite disjoint collection $\{A_i\colon i = 1, \dots, n\} \subset \mathcal{F}$, $\mu(\cup_{i=1}^n A_i) = \sum_{i=1}^n \mu(A_i)$ (μ is finitely additive on $\mathcal{F}$);*
- *for any collection $\{A_i\colon i \in \mathbb{N}\} \subset \mathcal{F}$, $\mu(\cup_{i=1}^\infty A_i) \le \sum_{i=1}^\infty \mu(A_i)$ (μ is countable subadditive on $\mathcal{F}$).*

Show that μ is a measure on $\mathcal{F}$.

Exercise 5.7. *Show that the family*

$$\mathcal{G} = \{A\colon A \subset \Omega \text{ s.t. } A \text{ is finite or } A^c \text{ is finite}\},$$

of Example 5.7 is a semiring on Ω.

Exercise 5.8. *Prove item (vi) of Proposition 5.1.*

6 Measurable sets: Part III

A selection of omitted proofs of this chapter are found in Section B.2 of the appendix.

6.1 The Lebesgue measure

The main result of this chapter is the following proposition:

Proposition 6.1. *There exists a measure λ on $\mathfrak{B}(\mathbb{R})$ (referred to as the Lebesgue measure on $\mathfrak{B}(\mathbb{R})$) which is s.t.*

$$\text{for any left-open interval } (a,b],\ \lambda \text{ returns its length, i.e., } \lambda((a,b]) = b-a. \tag{10}$$

Further, λ is the unique measure on $\mathfrak{B}(\mathbb{R})$ which satisfies (10).

The intention of this chapter is to explore the tools that provide arguments for the latter proposition.

6.2 Measure extensions

Proposition 6.2. *Let $(a,b]$, $a < b \in \mathbb{R}$, be any left-open interval. Let I be countable and $(a_i,b_i]$, $i \in I$, be s.t., $(a,b] \subset \cup_{i\in I}(a_i,b_i]$, then*

$$b-a \leq \sum_{i\in I}(b_i - a_i). \tag{11}$$

If the collection $\{(a_i,b_i]\colon i \in I\}$ is disjoint we also have the following result (Exercise 6.10).

Proposition 6.3. *Let $(a,b]$, $a < b \in \mathbb{R}$, be any left-open interval. Let I be countable and $\{(a_i,b_i]\colon i \in I\}$ be a disjoint collection of left-open intervals s.t. $\cup_{i\in I}(a_i,b_i] \subset (a,b]$. Then*

$$\sum_{i\in I}(b_i - a_i) \leq b-a.$$

Definition 6.1. *Let $\Omega \neq \emptyset$ be a set and $\mathcal{A}$ be a collection of subsets from Ω. Let $A \in \mathcal{P}(\Omega)$ be any subset of Ω. A collection $\{U_i\colon i \in I\}$ is said to be a covering of A by sets from $\mathcal{A}$ if $\{U_i\colon i\in I\} \subset \mathcal{A}$ and $A \subset \cup_{i\in I}U_i$. A covering $\{U_i\colon i \in I\}$ of A by sets from $\mathcal{A}$ is referred to as countable (resp. finite) if I is countable (resp. finite). We write $C_{\mathcal{A}}(A)$ for the set which contains all the countable coverings of A by sets from $\mathcal{A}$, i.e.,*

$$C_{\mathcal{A}}(A) = \{\xi\colon \xi \text{ is a countable covering of } A \text{ by sets from } \mathcal{A}\}.$$

Example 6.1. *Consider the setting of Example 4.6 and let $\Omega = \mathbb{R}$ and $\mathcal{R}$ be the family of left-open intervals with the empty set adjoined:*

$$\mathcal{R} = \{A\colon A = (a,b],\ a,b \in \mathbb{R}\} \cup \{\emptyset\}.$$

Let $B_r(x)$ be any open ball with center $x \in \mathbb{R}$ and radius $r > 0$. That is, $B_r(x) = (x-r, x+r)$ is an open interval with endpoints $a = x-r$ and $b = x+r$. Consider the set $\xi_1 = \{(a,x],(x,b]\}$. Then, $\xi_1 \in C_{\mathcal{R}}((a,b))$. As another example, let for $n \in \mathbb{N}$,

$$U_i^n = \left(a + \frac{2ri}{2^n}, a + \frac{2r(i+1)}{2^n}\right], \quad i = 0,\dots 2^n-1.$$

Then, $\xi_2 = \{U_i^n\colon i = 0,\dots 2^n-1\} \in C_{\mathcal{R}}((a,b))$ for any $n \in \mathbb{N}$. As a final example, define

$$U_k = \left(\frac{a}{2^k}, \frac{b}{2^k}\right], \quad k \in \mathbb{N}\cup\{0\}.$$

Then, $\xi_3 = \{U_k : k \in \mathbb{N} \cup \{0\}\} \in C_{\mathcal{R}}((a,b))$. Each of the coverings ξ_1, ξ_2 and ξ_3 of (a,b) by sets from $\mathcal{R}$ offers an approach to quantify the length of (a,b) by summing up the respective lengths of the sets from $\mathcal{R}$. Given $A \in \mathcal{P}(\mathbb{R})$, we define the function $v_\ell(\xi) = \sum_{U\in\xi} \ell(U)$, $\xi \in C_{\mathcal{R}}(A)$ where $\ell \colon \mathcal{R} \to [0,\infty)$ is s.t.

$$\ell(U) = \begin{cases} b-a, & \text{if } U = (a,b], \\ 0, & \text{if } U = \emptyset. \end{cases}$$

As an example, we have that $v_\ell(\xi_1) = x - a + b - x = b - a$. Notice also, that

$$\begin{aligned} v_\ell(\xi_2) &= \sum_{i=0}^{2^n-1} \frac{2r(i+1)-i}{2^n} \\ &= \frac{2r}{2^n} + \frac{4r}{2^n} - \frac{2r}{2^n} + \frac{6r}{2^n} - \frac{4r}{2^n} + \cdots + \frac{2r(2^n-1)}{2^n} + 2r - \frac{2r(2^n-1)}{2^n} \\ &= 2r = b - a. \end{aligned}$$

Exercise 6.1. *Verify that $v_\ell(\xi_3) = 2(b-a)$.*

In the following we show that

$$\inf\{v_\ell(\xi) \colon \xi \in C_{\mathcal{R}}((a,b])\} = \inf_{\xi\in C_{\mathcal{R}}((a,b])} v_\ell(\xi) = b - a, \tag{12}$$

i.e., $b - a$ is a lower bound for the values of $v_\ell(\xi)$, $\xi \in C_{\mathcal{R}}((a,b])$.

Exercise 6.2. *Verify that $\inf_{\xi\in C_{\mathcal{R}}((a,b])} v_\ell(\xi) \le b - a$.*

Upon the later exercise, it remains to show that $b - a \le \inf_{\xi\in C_{\mathcal{R}}((a,b])} v_\ell(\xi)$. Let ξ be any countable covering of $(a,b]$ by sets from $\mathcal{R}$. That is, $\xi = \{U_i : i \in I\}$, with $U_i = (a_i, b_i]$ or $U_i = \emptyset$, $i \in I$, where I is countable. Since $\ell(\emptyset) = 0$, we assume without loss of generality that $U_i = (a_i, b_i]$ for any $i \in I$. Therefore, we have that $(a,b] \subset \cup_{i\in I}(a_i,b_i]$ and $v_\ell(\xi) = \sum_{i\in I}(b_i - a_i)$. Since I is countable, either I is finite or $\#I = \#\mathbb{N}$. Using Proposition 6.2 we obtain,

$$b - a \le \sum_{i\in I}(b_i - a_i) = v_\ell(\xi).$$

It follows that $b - a \le \inf_{\xi\in C_{\mathcal{R}}((a,b])} v_\ell(\xi)$. We also saw that there exists $\xi \in C_{\mathcal{R}}((a,b])$ s.t. $b - a = v_\ell(\xi)$. Hence, the latter infimum is actually a minimum. In conclusion we have proven the following result.

Proposition 6.4. *Given any left-open interval $(a,b]$, $\min_{\xi\in C_{\mathcal{R}}((a,b])} v_\ell(\xi) = b - a$.*

We build on the latter result and define the function

$$\ell^*(A) = \inf_{\xi\in C_{\mathcal{R}}(A)} v_\ell(\xi), \quad A \in \mathcal{P}(\mathbb{R}).$$

Exercise 6.3. *Verify that $\ell^*(\{a\}) = 0$ for any point $a \in \mathbb{R}$.*

The given examples show that the function $\mathcal{P}(\mathbb{R}) \ni A \mapsto \ell^*(A)$ is in alignment with our intuitive understanding of the length of an interval. However, we will see in the following that ℓ^* is not a measure on $\mathcal{P}(\mathbb{R})$. Nevertheless, we will show that it is always possible to restrict ℓ^* to $\sigma(\mathcal{R}) \subset \mathcal{P}(\mathbb{R})$ to obtain a measure λ on $\sigma(\mathcal{R}) = \mathfrak{B}(\mathbb{R})$ which agrees with ℓ on $\mathcal{R}$. This measure λ will be referred to as the Lebesgue measure on $\mathfrak{B}(\mathbb{R})$ (cf. Example 5.5). The function ℓ^* will be termed an outer measure, in accordance with the property that restricted to a smaller σ-field it becomes a measure.

Leaving the framework of the latter example, we introduce the following general definition.

Definition 6.2. *Let Ω be a nonempty set. An outer measure μ^* is a function $\mu^*\colon \mathcal{P}(\Omega) \to \overline{\mathbb{R}}_+$ that satisfies the following properties:*

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

(ii) $A, B \in \mathcal{P}(\Omega)$, *s.t.* $A \subset B \Rightarrow \mu^*(A) \leq \mu^*(B)$;

(iii) μ^* *is countable subadditive on* $\mathcal{P}(\Omega)$, *i.e., for any collection* $\{A_n : n \in \mathbb{N}\} \subset \mathcal{P}(\Omega)$, $\mu^*(\cup_{n=1}^{\infty} A_n) \leq \sum_{n=1}^{\infty} \mu^*(A_n)$.

Then, the following result offers a general classification of the function ℓ^* covered in Example 6.1:

Proposition 6.5. *Let $\mathcal{A} \subset \mathcal{P}(\Omega)$ where Ω is some nonempty set. Suppose that $\emptyset \in \mathcal{A}$ and $\rho\colon \mathcal{A} \to \overline{\mathbb{R}}_+$ is a function s.t. $\rho(\emptyset) = 0$. Let $v_\rho(\xi) = \sum_{U \in \xi} \rho(U)$, $\xi \in C_{\mathcal{A}}(A)$. Then, the function,*

$$\mathcal{P}(\Omega) \ni A \mapsto \rho^*(A) = \inf_{\xi \in C_{\mathcal{A}}(A)} v_\rho(\xi),$$

is an outer measure.

Exercise 6.4. *Let ℓ be as in Example 6.1. Verify that $\ell^*(\mathbb{R}) = \infty$.*

Definition 6.3. *Let $\mu^*\colon \mathcal{P}(\Omega) \to \overline{\mathbb{R}}_+$ be a function. Define the set*

$$\mathcal{M}(\mu^*) = \{A \in \mathcal{P}(\Omega) \colon \mu^*(A \cap E) + \mu^*(A^c \cap E) = \mu^*(E) \ \forall E \in \mathcal{P}(\Omega)\}.$$

We say that A is μ^*-measurable if $A \in \mathcal{M}(\mu^*)$.

We obtain the following result.

Proposition 6.6. *Let μ^* be an outer measure on $\mathcal{P}(\Omega)$. Then,*

(I) $\mathcal{M}(\mu^*)$ *is a σ-field;*

(II) The restriction $\mu^*|_{\mathcal{M}(\mu^*)}$ *is a measure.*

The main result of this section is the following:

Proposition 6.7. *Let $\mathcal{A} \subset \mathcal{P}(\Omega)$ be a semiring and $\rho\colon \mathcal{A} \to \overline{\mathbb{R}}_+$ be a function which is s.t.,*

- $\rho(\emptyset) = 0$;
- *ρ is finitely additive on $\mathcal{A}$, i.e., for any finite disjoint collection $\{A_i : i = 1, \dots, n\} \subset \mathcal{A}$, $n \in \mathbb{N}$, with $\cup_{i=1}^n A_i \in \mathcal{A}$, $\rho(\cup_{i=1}^n A_i) = \sum_{i=1}^n \rho(A_i)$;*
- *ρ is countable subadditive on $\mathcal{A}$, i.e., for any collection $\{A_i : i \in \mathbb{N}\} \subset \mathcal{A}$ with $\cup_{i=1}^{\infty} A_i \in \mathcal{A}$, $\rho(\cup_{i=1}^{\infty} A_i) \leq \sum_{i=1}^{\infty} \rho(A_i)$.*

Then, ρ extends to a measure on $\sigma(\mathcal{A})$. That is, there exists a measure $\rho_\uparrow\colon \sigma(\mathcal{A}) \to \overline{\mathbb{R}}_+$ which is s.t. $\rho_\uparrow(A) = \rho(A)$ for any $A \in \mathcal{A}$.

Exercise 6.5. *Let ρ be as in Proposition 6.7. Show that ρ is monotone on $\mathcal{A}$, i.e., $A, B \in \mathcal{A}$ s.t. $A \subset B \Rightarrow \rho(A) \leq \rho(B)$.*

We remark that in the statement of Proposition 6.7, since $\mathcal{A}$ is not necessarily a σ-field, to make sense of the condition that ρ is finitely additive (resp. countable subadditive) on $\mathcal{A}$, it is required to demand that $\cup_{i=1}^n A_i \in \mathcal{A}$ (resp. $\cup_{i=1}^{\infty} A_i \in \mathcal{A}$).

Example 6.2. *We remain in the setting of Example 6.1, where* $\ell\colon \mathcal{R} \to [0,\infty)$ *is s.t.*

$$\ell(U) = \begin{cases} b-a, & \textit{if } U = (a,b], \\ 0, & \textit{if } U = \emptyset, \end{cases}$$

and for $A \in \mathcal{P}(\mathbb{R})$*,*

$$\ell^*(A) = \inf_{\xi \in C_{\mathcal{R}}(A)} v_\ell(\xi), \quad v_\ell(\xi) = \sum_{U\in\xi} \ell(U), \quad \xi \in C_{\mathcal{R}}(A),$$

with

$$C_{\mathcal{R}}(A) = \{\xi\colon \xi \textit{ is a countable covering of } A \textit{ by sets from } \mathcal{R}\}.$$

Proposition 6.5 shows that ℓ^* *is an outer measure on* $\mathcal{P}(\mathbb{R})$*. By Proposition 6.6,* $\mathcal{M}(\ell^*)$ *is in fact a* σ*-field on* $\mathbb{R}$ *and the restriction of the outer measure* ℓ^* *to* $\mathcal{M}(\ell^*)$ *written as* $\ell^*|_{\mathcal{M}(\ell^*)}$ *is a measure on* $\mathcal{M}(\ell^*)$*. In addition, a set* $A \subset \mathbb{R}$ *is called Lebesgue measurable, if* $A \in \mathcal{M}(\ell^*)$*, i.e., if A is* ℓ^**-measurable. We also know that the family of left-open intervals* $\mathcal{R}$ *is a semiring (cf. Example 5.6). Then, let* $\{A_i, i\in\mathbb{N}\} \subset \mathcal{R}$*,* $A_i = (a_i, b_i]$*, be disjoint and s.t.,* $\cup_{i\in\mathbb{N}}(a_i,b_i] = A \in \mathcal{R}$*. Using Propositions 6.2 and 6.3, we conclude that* $\ell(A) = \sum_{i\in\mathbb{N}} \ell(A_i)$*. Hence,* ℓ *is additive on* $\mathcal{R}$ *and hence clearly finitely additive on* $\mathcal{R}$*. If* $\{A_i, i\in\mathbb{N}\} \subset \mathcal{R}$ *is not disjoint, then still, Propositions 6.2 shows that* $\ell(A) \le \sum_{i\in\mathbb{N}} \ell(A_i)$*. Hence,* ℓ *satisfies all the conditions of Proposition 6.7. Therefore, there exists a measure* $\ell_\uparrow$ *on* $\sigma(\mathcal{R})$ *which is s.t. for any* $A \in \mathcal{R}$*,* $\ell_\uparrow(A) = \ell(A)$*. The measure* $\ell_\uparrow$ *is called the Lebesgue measure on* $\sigma(\mathcal{R})$ *and we use the notation* $\ell_\uparrow = \lambda$ *(cf. Example 5.5). In fact, we have that* $\lambda = \ell^*|_{\sigma(\mathcal{R})}$*, i.e., the Lebesgue measure on* $\mathfrak{B}(\mathbb{R})$ *equals the restriction of the outer measure* ℓ^* *to* $\mathfrak{B}(\mathbb{R})$ *(cf. the proof of Proposition 6.7). We also have that* $\mathfrak{B}(\mathbb{R}) \subset \mathcal{M}(\ell^*)$ *(cf. the proof of Proposition 6.7). Thus, every Borel set is also Lebesgue measurable. One can show that* $\mathcal{M}(\ell^*) \setminus \mathfrak{B}(\mathbb{R}) \neq \emptyset$*, i.e., there are sets that are Lebesgue measurable but not Borel measurable. Even further, the following result shows that* $\mathcal{P}(\mathbb{R}) \setminus \mathcal{M}(\ell^*) \neq \emptyset$*.*

Proposition 6.8. *There are sets* $V \subset \mathbb{R}$ *which are not Lebesgue measurable.*

We omit a proof and refer to [1] or [2].

Definition 6.4. *Let* $(\Omega, \mathcal{F})$ *be a measurable space and* μ *be a measure on* $\mathcal{F}$*. Let* $\mathcal{A} \subset \mathcal{F}$*. The measure* μ *is called* σ*-finite on* $\mathcal{A}$ *if* $\Omega = \cup_{i\in I} A_i$ *for some countable collection of sets* $\{A_i\colon i \in I\} \subset \mathcal{A}$ *which are s.t.* $\mu(A_i) < \infty$ *for any* $i \in I$*.*

A proof of the following result is given in Section B.3 of the appendix.

Proposition 6.9. *Let* $\Omega \neq \emptyset$ *be a set and assume that* $\mathcal{A} \subset \mathcal{P}(\Omega)$ *is s.t.* $A, B \in \mathcal{A} \Rightarrow A \cap B \in \mathcal{A}$*. Let* μ_1 *and* μ_2 *be two measures on* $\sigma(\mathcal{A})$ *where at least one of the measures* μ_1 *and* μ_2 *is a* σ*-finite measure on* $\mathcal{A}$*. If* $\mu_1(A) = \mu_2(A)$ *for any* $A \in \mathcal{A}$*, then* $\mu_1(A) = \mu_2(A)$ *for any* $A \in \sigma(\mathcal{A})$*, i.e., the two measures agree on* $\sigma(\mathcal{A})$*.*

Proposition 6.10. *The Lebesgue measure* λ *on* $\mathfrak{B}(\mathbb{R})$ *is the unique measure on* $\mathfrak{B}(\mathbb{R})$ *which is s.t. for any left-open interval* $(a,b]$*,* $\lambda((a,b]) = b-a$*.*

Remark 6.1. *We remark that upon Example 6.2 and the latter proposition, we have proven Proposition 6.1.*

6.3 The Lebesgue measure on real coordinate spaces

Let $\Omega = \mathbb{R}^k$, $k \in \mathbb{N}$, and consider the function

$$\ell_k(A) = \begin{cases} \prod_{i=1}^k (b_i - a_i) & \text{if } A = \prod_{i=1}^k (a_i, b_i] \\ 0, & \text{otherwise,} \end{cases}$$

defined on the set

$$\mathcal{R}_k = \left\{A\colon A = \prod_{i=1}^{k}(a_i, b_i],\ a_i, b_i \in \mathbb{R},\ i = 1, \dots, k\right\} \cup \{\emptyset\},$$

i.e., the family of rectangles in $\mathbb{R}^k$ (cf. Example 4.11). That is, ℓ_k $k \in \mathbb{N}$, returns length ($k = 1$), area ($k = 2$), volume ($k = 3$) and hypervolume ($k \geq 4$). Then, Proposition 6.1 is generalized as follows:

Proposition 6.11. *There exists a measure λ_k on $\mathfrak{B}(\mathbb{R}^k) = \sigma(\mathcal{R}_k)$ (referred to as the (k-dimensional) Lebesgue measure on $\mathfrak{B}(\mathbb{R}^k)$) which is s.t.*

$$\text{for any rectangle } A = \prod_{i=1}^{k}(a_i, b_i],\ \lambda_k \text{ satisfies } \lambda_k(A) = \ell_k(A). \tag{13}$$

Further, λ_k is the unique measure on $\mathfrak{B}(\mathbb{R}^k)$ which satisfies (13).

6.4 Solution to exercises

Solution 6.1 (Solution to Exercise 6.1)**.** *We have that*

$$v_\ell(\xi_3) = \sum_{k=0}^{\infty} \frac{(b-a)}{2^k} = (b-a)\sum_{k=0}^{\infty}\left(\frac{1}{2}\right)^k = 2(b-a).$$

Solution 6.2 (Solution to Exercise 6.2)**.** *This is clearly the case, as an example, take $\xi_* = \{(a, b]\}$, then $\xi_* \in C_{\mathcal{R}}((a, b])$ and $v_\ell(\xi_*) = b - a$. Hence,*

$$\inf_{\xi \in C_{\mathcal{R}}((a,b])} v_\ell(\xi) \leq b - a.$$

Solution 6.3 (Solution to Exercise 6.3)**.** *Let $a \in \mathbb{R}$. We know that $\ell^*(A) \geq 0$ for any $A \in \mathcal{P}(\mathbb{R})$. In particular, $\ell^*(\{a\}) \geq 0$. Define for any $\varepsilon > 0$,*

$$\xi_a^\varepsilon = \left\{\left(a - \frac{\varepsilon}{2^n}, a\right] : n \in \mathbb{N}\right\} \in C_{\mathcal{R}}(\{a\}).$$

We have that $v_\ell(\xi_a^\varepsilon) = \varepsilon$. Hence, since ε was arbitrary, $v_\ell(\xi_a^\varepsilon) = 0$. This shows that $\ell^*(\{a\}) \leq 0$ and hence $\ell^*(\{a\}) = 0$.

Solution 6.4 (Solution to Exercise 6.4)**.** *By Proposition 6.5, ℓ^* is an outer measure on $\mathcal{P}(\mathbb{R})$. In particular, it is monotone. That is, $A \subset B \Rightarrow \ell^*(A) \leq \ell^*(B)$. Also, because of Proposition 6.4, for any $n \in \mathbb{N}$, $\ell^*((-n/2, n/2]) = n$. This shows that for any $n \in \mathbb{N}$, $n = \ell^*((-n/2, n/2]) \leq \ell^*(\mathbb{R})$ and hence $\ell^*(\mathbb{R})$ can not be finite.*

Solution 6.5 (Solution to Exercise 6.5)**.** *Let $A, B \in \mathcal{A}$ s.t. $A \subset B$. Then, since $\mathcal{A}$ is a semiring, there exists a disjoint collection $\{C_i : i = 1, \dots, n\} \subset \mathcal{A}$, $n \in \mathbb{N}$, s.t. $B \setminus A = \cup_{i=1}^{n} C_i$. In particular, $B = A \cup (B \setminus A)$ and hence, since ρ is assumed to be finitely additive on $\mathcal{A}$,*

$$\rho(B) = \rho\left(A \cup \left(\bigcup_{i=1}^{n} C_k\right)\right) = \rho(A) + \rho\left(\bigcup_{i=1}^{n} C_k\right) \geq \rho(A).$$

6.5 Additional exercises

Exercise 6.6. *Argue that each of the following sets is a Borel set, i.e., a member of* $\mathfrak{B}(\mathbb{R})$ *and deduce its Lebesgue measure:*

- $\mathbb{R}$*;*
- $\{a\}$, $a \in \mathbb{R}$ *(singleton sets);*
- (a, b), $[a, b)$ *and* $[a, b]$, *where* $a, b \in \mathbb{R}$.

Exercise 6.7. *Prove Proposition 6.10.*

Exercise 6.8. *Show that* ℓ^* *is not finitely additive on* $\mathcal{P}(\mathbb{R})$, *i.e., there exists a disjoint collection* $\{A_i : i = 1, \ldots, n\} \subset \mathcal{P}(\mathbb{R})$ *s.t.* $\ell^*(\cup_{i=1}^n A_i) \neq \sum_{i=1}^n \ell^*(A_i)$. *Conclude that* ℓ^* *is not a measure on* $\mathcal{P}(\mathbb{R})$.
Hint: *Proposition 6.8.*

Exercise 6.9. *Prove Proposition 6.3.*
Hint: *Induction.*

Exercise 6.10. *Let* $\mathcal{U} = \{U : U \subset \mathbb{R}^k \text{ open}\}$, *i.e.,* $\mathcal{U}$ *contains all the open sets of* $\mathbb{R}^k$ *(cf. Definition 2.16). Show that* $\mathfrak{B}(\mathbb{R}^k) = \sigma(\mathcal{U})$.

$\mathfrak{B}(\mathbb{R}^k)$ **contains** $\mathcal{U}$**:**

Step1: *Define*

$$\mathcal{R}_k(\mathbb{Q}) = \big\{A : A = \prod_{i=1}^k (a_i, b_i],\ a_i, b_i \in \mathbb{Q},\ i = 1, \ldots, k\big\}.$$

Verify that $\mathcal{R}_k(\mathbb{Q})$ *is countable.*

Step2: *Let* $U \in \mathcal{U}$ *be nonempty. Argue that for any* $x \in U$, *there exists* $R_x \in \mathcal{R}_k(\mathbb{Q})$ *s.t.* $x \in R_x$ *and* $R_x \subset U$.
Hint: *Recall that since* U *is open, it follows that for any* $x \in U$, *there exists an open ball* $B_{\varepsilon_x}(x)$ *s.t.* $B_{\varepsilon_x}(x) \subset U$. *Hence, it is sufficient to find*

$$R_x = (a_1, b_1] \times \cdots \times (a_k, b_k] \in \mathcal{R}_k(\mathbb{Q})$$

s.t. $x \in R_x$ *and for* $i = 1, \ldots, k$, $b_i - a_i < r(\varepsilon_x, k)$, *where* $r(\varepsilon_x, k)$ *is chosen s.t.* $\|x - y\| < \varepsilon_x$ *for any* $y \in R_x$.

Step3: *Write* U *as a countable union of rectangles from* $\mathcal{R}_k(\mathbb{Q})$.

$\sigma(\mathcal{U})$ **contains** $\mathcal{R}_k$**:** *Show that any element of* $\mathcal{R}_k$ *can be written as a countable intersection of open rectangles (cf. Exercise 2.4).*

7 Measurable functions

The omitted proofs of this chapter are found in Section B.4 of the appendix.

7.1 The concept of measurable functions

Definition 7.1. *Let* $(\Omega, \mathcal{F})$ *and* $(\Omega^*, \mathcal{F}^*)$ *be two measurable spaces (cf. Definition 4.3). A function* $f\colon \Omega \to \Omega^*$ *is said to be measurable* $\mathcal{F}/\mathcal{F}^*$ *if for any* $A^* \in \mathcal{F}^*$, $f^{-1}(A^*) \in \mathcal{F}$.

Proposition 7.1. *Let* $(\Omega, \mathcal{F})$ *and* $(\Omega^*, \mathcal{F}^*)$ *be two measurable spaces and* $f\colon \Omega \to \Omega^*$ *be a function. Suppose that* $\mathcal{F}^* = \sigma(\mathcal{G})$ *and for any* $G \in \mathcal{G}$, $f^{-1}(G) \in \mathcal{F}$. *Then,* f *is* $\mathcal{F}/\mathcal{F}^*$ *measurable.*

Proof. It is enough to show that $\Sigma^* = \{A^*\colon f^{-1}(A^*) \in \mathcal{F}\}$ is a σ-field on Ω^*. We have that $\Omega^* \in \Sigma^*$, since $f^{-1}(\Omega^*) = \Omega$ and $\mathcal{F}$ is a σ-field, i.e., it contains Ω. Let $A^* \in \Sigma^*$. Then, $f^{-1}((A^*)^c) = f^{-1}(A^*)^c$ (cf. Proposition 2.4) and hence $(A^*)^c \in \Sigma^*$. Let $\{A_i^*\colon i \in \mathbb{N}\} \subset \Sigma^*$. Then, since $f^{-1}(\cup_{i=1}^{\infty} A_i^*) = \cup_{i=1}^{\infty} f^{-1}(A_i^*)$ (cf. Proposition 2.4), $\cup_{i=1}^{\infty} A_i^* \in \Sigma^*$. □

Example 7.1. *Consider the measurable space* $(\mathbb{R}^m, \mathfrak{B}(\mathbb{R}^m))$ *and* $(\mathbb{R}^k, \mathfrak{B}(\mathbb{R}^k))$. *Let* $f\colon \mathbb{R}^m \to \mathbb{R}^k$ *be continuous (cf. Definition 2.17). Then,* f *is measurable* $\mathfrak{B}(\mathbb{R}^m)/\mathfrak{B}(\mathbb{R}^k)$. *To see it, let* $\mathcal{U} = \{U\colon U \text{ open in } \mathbb{R}^k\}$. *Then, since* f *is continuous, for any* $U \in \mathcal{U}$, $f^{-1}(U)$ *is an open set of* $\mathbb{R}^m$. *Since* $\mathfrak{B}(\mathbb{R}^m)$ *contains the open sets of* $\mathbb{R}^m$, *it follows that for any* $U \in \mathcal{U}$, $f^{-1}(U) \in \mathfrak{B}(\mathbb{R}^m)$. *We know that* $\mathfrak{B}(\mathbb{R}^k) = \sigma(\mathcal{U})$ *(cf. Exercise 6.10). Hence, by Proposition 7.1,* f *is measurable.*

Definition 7.2. *A function* $f\colon \mathbb{R}^m \to \mathbb{R}^k$ *is called Borel function if it is measurable* $\mathfrak{B}(\mathbb{R}^m)/\mathfrak{B}(\mathbb{R}^k)$.

Upon Example 7.2 we have proven the following result.

Proposition 7.2. *Any continuous function* $f\colon \mathbb{R}^m \to \mathbb{R}^k$ *is a Borel function.*

Remark 7.1. *Suppose that* $f\colon E \to \mathbb{R}^k$, *where* $E \subset \mathbb{R}^m$, $E \neq \emptyset$. *Using Proposition 4.4 and Exercise 6.10, we know that*

$$\mathfrak{B}(E) = \sigma(\{G \cap E\colon G \text{ open in } \mathbb{R}^m\}).$$

Further, by Proposition 2.11, if $f\colon E \to \mathbb{R}^k$ *is continuous, then, for any* $U \subset \mathbb{R}^k$ *open,* $f^{-1}(U) \in \{G \cap E\colon G \text{ open in } \mathbb{R}^m\}$. *Hence, by Proposition 7.1,* f *is* $\mathfrak{B}(E)/\mathfrak{B}(\mathbb{R}^k)$ *measurable.*

Using the structure of the Borel σ-field, the next result is helpful to verify that a function $f\colon \Omega \to \mathbb{R}$ is $\mathcal{F}/\mathfrak{B}(\mathbb{R})$ measurable.

Proposition 7.3. *Let* $(\Omega, \mathcal{F})$ *be a measurable space and* $f\colon \Omega \to \mathbb{R}$ *be a real-valued function. Suppose that* $\{\omega \in \Omega\colon f(\omega) \leq x\} \in \mathcal{F}$ *for any* $x \in \mathbb{R}$, *then* f *is* $\mathcal{F}/\mathfrak{B}(\mathbb{R})$ *measurable.*

Proof. Notice that for any $x \in \mathbb{R}$, $\{\omega \in \Omega\colon f(\omega) \leq x\} = f^{-1}((-\infty, x])$. Since $\mathfrak{B}(\mathbb{R}) = \sigma(\{(-\infty, x]\colon x \in \mathbb{R}\})$ (cf. Exercise 4.7), f is $\mathcal{F}/\mathfrak{B}(\mathbb{R})$ measurable (cf. Proposition 7.1). □

Exercise 7.1. *Let* $(\Omega, \mathcal{F})$ *be a measurable space and* $f(\omega) = \alpha$ *for any* $\omega \in \Omega$, *where* $\alpha \in \mathbb{R}$. *Show that* f *is* $\mathcal{F}/\mathfrak{B}(\mathbb{R})$ *measurable.*

Example 7.2. *In Example 5.1, we have considered the measure* $A \mapsto \delta_\omega(A)$ *for fixed* $\omega \in \Omega$. *If now* $A \subset \Omega$ *is fixed and* ω *is variable, then, the function*

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

is referred to as the indicator function of the set A. Then, if $A \in \mathcal{F}$, $\omega \mapsto \mathbb{1}_A(\omega)$ is $\mathcal{F}/\mathfrak{B}(\mathbb{R})$ measurable. To see it, we notice that for any $x \in \mathbb{R}$,

$$\{\omega \in \Omega \colon \mathbb{1}_A(\omega) > x\} = \begin{cases} \emptyset, & \text{if } x \geq 1, \\ A, & \text{if } 0 \leq x < 1, \\ \Omega & \text{if } x < 0. \end{cases}$$

Since $A \in \mathcal{F}$, $\{\omega \in \Omega \colon \mathbb{1}_A(\omega) > x\} \in \mathcal{F}$ for any $x \in \mathbb{R}$. Thus, $\{\omega \in \Omega \colon \mathbb{1}_A(\omega) \leq x\} \in \mathcal{F}$ for any $x \in \mathbb{R}$. By Proposition 7.3, $\omega \mapsto \mathbb{1}_A(\omega)$ is $\mathcal{F}/\mathfrak{B}(\mathbb{R})$ measurable.

Example 7.3. *Let $\Omega = \{h, t\}$ and $\mathcal{F} = \mathcal{P}(\{h, t\}) = \{\emptyset, \{h\}, \{t\}, \{h, t\}\}$. Then, $\{h\} \in \mathcal{P}(\{h, t\})$. Thus,*

$$f(\omega) = \begin{cases} 1, & \text{if } \omega = h, \\ 0, & \text{if } \omega = t, \end{cases}$$

is $\mathcal{P}(\{h, t\})/\mathfrak{B}(\mathbb{R})$ measurable.

Example 7.4. *Let $(\Omega, \mathcal{F})$ be a measurable space and $f_i \colon \Omega \to \mathbb{R}$, $i = 1, \ldots, k$, be $\mathcal{F}/\mathfrak{B}(\mathbb{R})$ measurable. Then, $f = \min\{f_i \colon i = 1, \ldots, k\}$ is $\mathcal{F}/\mathfrak{B}(\mathbb{R})$ measurable. To see it, we notice that for any $x \in \mathbb{R}$ and $i = 1, \ldots, k$, since f_i is $\mathcal{F}/\mathfrak{B}(\mathbb{R})$ measurable, $f_i^{-1}((x, \infty)) \in \mathcal{F}$. Therefore, for any $x \in \mathbb{R}$,*

$$f^{-1}((x, \infty)) = \{\omega \in \Omega \colon f(\omega) > x\} = \cap_{i=1}^k \{\omega \in \Omega \colon f_i(\omega) > x\} = \cap_{i=1}^k f_i^{-1}((x, \infty)) \in \mathcal{F}.$$

Therefore, for any $x \in \mathbb{R}$, $(f^{-1}((x, \infty)))^c = f^{-1}((-\infty, x]) \in \mathcal{F}$. Using Proposition 7.3, f is $\mathcal{F}/\mathfrak{B}(\mathbb{R})$ measurable.

The following result shows that the measurability of a vector valued function is characterized in terms of the measurability of the respective coordinate functions.

Proposition 7.4. *Let $(\Omega, \mathcal{F})$ be a measurable space and $f \colon \Omega \to \mathbb{R}^k$, i.e.,*

$$f(\omega) = (f_1(\omega), \ldots, f_k(\omega)).$$

Then, f is $\mathcal{F}/\mathfrak{B}(\mathbb{R}^k)$ measurable if and only if for any $i = 1, \ldots, k$, $f_i \colon \Omega \to \mathbb{R}$ is $\mathcal{F}/\mathfrak{B}(\mathbb{R})$ measurable.

The latter result offers a helpful tool to show that if $f_1, \ldots, f_k \colon \Omega \to \mathbb{R}$ are $\mathcal{F}/\mathfrak{B}(\mathbb{R})$ measurable functions, then, the functions $\sum_{i=1}^k f_i$ and $\prod_{i=1}^k f_i$ are $\mathcal{F}/\mathfrak{B}(\mathbb{R})$ measurable as well. In order to arrive there, the conclusion of the following exercise is valuable.

Exercise 7.2. *Let $(\Omega, \mathcal{F})$ and $(\Omega^*, \mathcal{F}^*)$ be two measurable spaces and $f \colon \Omega \to \Omega^*$ be a function. Let $(\Omega^{**}, \mathcal{F}^{**})$ be a third measurable space and $f^* \colon \Omega^* \to \Omega^{**}$ be another function. Show that if f is $\mathcal{F}/\mathcal{F}^*$ measurable and f^* is $\mathcal{F}^*/\mathcal{F}^{**}$ measurable, then the composition $f^*(f) \colon \Omega \to \Omega^{**}$ is $\mathcal{F}/\mathcal{F}^{**}$ measurable.*

Proposition 7.5. *Let $(\Omega, \mathcal{F})$ be a measurable space and $f_i \colon \Omega \to \mathbb{R}$, $i = 1, \ldots, k$, be $\mathcal{F}/\mathfrak{B}(\mathbb{R})$ measurable. Suppose that $g \colon \mathbb{R}^k \to \mathbb{R}$ is $\mathfrak{B}(\mathbb{R}^k)/\mathfrak{B}(\mathbb{R})$ measurable. Then,*

$$\omega \mapsto g((f_1(\omega), \ldots, f_k(\omega))),$$

is $\mathcal{F}/\mathfrak{B}(\mathbb{R})$ measurable.

Notice that we usually avoid the double bracket and simply write

$$\omega \mapsto g((f_1(\omega), \ldots, f_k(\omega))) = g(f_1(\omega), \ldots, f_k(\omega)).$$

Proof of Proposition 7.5. Since $f_i \colon \Omega \to \mathbb{R}$, $i = 1, \ldots, k$, are $\mathcal{F}/\mathfrak{B}(\mathbb{R})$ measurable, we use Proposition 7.4 and deduce that the map

$$\omega \mapsto (f_1(\omega), \ldots, f_k(\omega)),$$

is $\mathcal{F}/\mathfrak{B}(\mathbb{R}^k)$ measurable. Then, since g is $\mathfrak{B}(\mathbb{R}^k)/\mathfrak{B}(\mathbb{R})$ measurable, we rely on Exercise 7.2 and verify that

$$\omega \mapsto g(f_1(\omega), \ldots, f_k(\omega)),$$

is $\mathcal{F}/\mathfrak{B}(\mathbb{R})$ measurable. $\square$

A direct consequence of the latter result is the following (cf. Example 7.1).

Proposition 7.6. *Let $(\Omega, \mathcal{F})$ be a measurable space and $f_i \colon \Omega \to \mathbb{R}$, $i = 1, \ldots, k$, be $\mathcal{F}/\mathfrak{B}(\mathbb{R})$ measurable. Then, if $g \colon \mathbb{R}^k \to \mathbb{R}$ is continuous,*

$$\omega \mapsto g(f_1(\omega), \ldots, f_k(\omega)),$$

is $\mathcal{F}/\mathfrak{B}(\mathbb{R})$ measurable.

Example 7.5. *Let $(\Omega, \mathcal{F})$ be a measurable space and $f_i \colon \Omega \to \mathbb{R}$, $i = 1, \ldots, k$, be $\mathcal{F}/\mathfrak{B}(\mathbb{R})$ measurable. Then, $\sum_{i=1}^k f_i$ is $\mathcal{F}/\mathfrak{B}(\mathbb{R})$ measurable (cf. Proposition 2.12).*

Example 7.6. *Let $(\Omega, \mathcal{F})$ be a measurable space and $f_i \colon \Omega \to \mathbb{R}$, $i = 1, \ldots, k$, be $\mathcal{F}/\mathfrak{B}(\mathbb{R})$ measurable. Then, $\prod_{i=1}^k f_i$ is $\mathcal{F}/\mathfrak{B}(\mathbb{R})$ measurable (cf. Proposition 2.12).*

Exercise 7.3. *Let $(\Omega, \mathcal{F})$ be a measurable space and $f_i \colon \Omega \to \mathbb{R}$, $i = 1, \ldots, k$, be $\mathcal{F}/\mathfrak{B}(\mathbb{R})$ measurable. Let $c_1, \ldots, c_k \in \mathbb{R}$. Then, the function*

$$\omega \mapsto f(\omega) = \sum_{i=1}^k c_i f_i(\omega),$$

is $\mathcal{F}/\mathfrak{B}(\mathbb{R})$ measurable.

Remark 7.2. *Let $(\Omega, \mathcal{F})$ be a measurable space and $f \colon \Omega \to \mathbb{R}$ be $\mathcal{F}/\mathfrak{B}(\mathbb{R})$ measurable s.t. $f(\Omega) \subset E$, where $E \in \mathfrak{B}(\mathbb{R})$. Suppose that $g \colon E \to \mathbb{R}$ is continuous. Then, the composition $g(f) \colon \Omega \to \mathbb{R}$ is $\mathcal{F}/\mathfrak{B}(\mathbb{R})$ measurable. To see it, we notice that for any $B \in \mathfrak{B}(\mathbb{R})$,*

$$g(f)^{-1}(B) = \{\omega \colon g(f(\omega)) \in B\} = \{\omega \colon f(\omega) \in g^{-1}(B)\} = f^{-1}(g^{-1}(B)).$$

Therefore, since by assumption f is $\mathcal{F}/\mathfrak{B}(\mathbb{R})$ measurable, it remains to check that $g^{-1}(B) \in \mathfrak{B}(\mathbb{R})$. Upon Remark 7.1 we know that since g is continuous, it is $\mathfrak{B}(E)/\mathfrak{B}(\mathbb{R})$ measurable. In particular, $g^{-1}(B) \in \mathfrak{B}(E)$. Further, since $E \in \mathfrak{B}(\mathbb{R})$, it follows that $\mathfrak{B}(E) \subset \mathfrak{B}(\mathbb{R})$ since for any $A \in \mathfrak{B}(\mathbb{R})$, $A \cap E \in \mathfrak{B}(\mathbb{R})$ (recall Definition 4.2). In conclusion, $g^{-1}(B) \in \mathfrak{B}(\mathbb{R})$.

Definition 7.3. *A function $f \colon \Omega \to \mathbb{R}$ is called simple if there exists $n \in \mathbb{N}$, $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ and sets $A_1, \ldots, A_n \subset \Omega$ s.t.*

$$f(\omega) = \sum_{i=1}^n \alpha_i \mathbb{1}_{A_i}(\omega), \quad \omega \in \Omega.$$

That is, a simple function is a finite linear combination of indicator functions.

Example 7.7. *Let $(\Omega, \mathcal{F})$ be a measurable space and f be a simple function on Ω, i.e., $f(\omega) = \sum_{i=1}^n \alpha_i \mathbb{1}_{A_i}(\omega)$. Then, if $A_i \in \mathcal{F}$ for any $i = 1, \ldots, n$, f is $\mathcal{F}/\mathfrak{B}(\mathbb{R})$ measurable.*

Example 7.8. *Let $n \in \mathbb{N}$ and $\Omega = \{\omega\colon \omega = (\omega_1, \dots, \omega_n)\colon \omega_i \in \{0,1\},\ i = 1, \dots, n\}$. We write $\Omega = A_0 \cup A_1 \cup \cdots A_n$, where*

$$A_k = \{\omega \in \Omega\colon \sum_{i=1}^{n} \omega_i = k\}, \quad k = 0, \dots, n.$$

Define

$$f(\omega) = \sum_{k=0}^{n} k \mathbb{1}_{A_k}(\omega).$$

Then, since $A_k \in \mathcal{P}(\Omega)$ for any $k = 0, \dots, n$, f is $\mathcal{P}(\Omega)/\mathfrak{B}(\mathbb{R})$ measurable.

Definition 7.4. *Let $(\Omega, \mathcal{F})$ be a measurable space and $f\colon \Omega \to \mathbb{R}$ be a simple function on Ω, i.e., $f = \sum_{i=1}^{n} \alpha_i \mathbb{1}_{A_i}$. f is called standard if $\cup_{i=1}^{n} A_i = \Omega$ and $\{A_1, \dots, A_n\} \subset \mathcal{F}$ is disjoint. If f is standard, we say that f is a simple function in standard form.*

The following result shows that simple functions can be written in standard form.

Proposition 7.7. *Let $(\Omega, \mathcal{F})$ be a measurable space and $f(\omega) = \sum_{i=1}^{n} \alpha_i \mathbb{1}_{A_i}(\omega)$, $\omega \in \Omega$, be a simple function on Ω s.t. $A_i \in \mathcal{F}$ for any $i = 1, \dots, n$. Then, there exists a standard simple function $g\colon \Omega \to \mathbb{R}$ s.t. $f(\omega) = g(\omega)$ for any $\omega \in \Omega$.*

7.2 Functions taking values in the extended real numbers

In the context of measurable functions, it is helpful to work with the extended real numbers $\overline{\mathbb{R}}$. Thus, we extend the definition of a measurable function to allow for function $f\colon \Omega \to \overline{\mathbb{R}}$. This will be useful in the context of limits of measurable functions.

Definition 7.5. *Let $(\Omega, \mathcal{F})$ be a measurable space and $f\colon \Omega \to \overline{\mathbb{R}}$. We say that f is $\mathcal{F}$ measurable if for any $A \in \mathfrak{B}(\mathbb{R})$, $\{\omega \in \Omega\colon f(\omega) \in A\} \in \mathcal{F}$ and $\{\omega \in \Omega\colon f(\omega) = -\infty\} \in \mathcal{F}$ and $\{\omega \in \Omega\colon f(\omega) = \infty\} \in \mathcal{F}$.*

Remark 7.3. *To combine terminology, if $f\colon \Omega \to \mathbb{R}$, then f is said to be $\mathcal{F}$ measurable if it is $\mathcal{F}/\mathfrak{B}(\mathbb{R})$ measurable, i.e., in this case there is no need to bother about the sets $\{\omega \in \Omega\colon f(\omega) = -\infty\}$ and $\{\omega \in \Omega\colon f(\omega) = -\infty\}$. Notice that if $f(\omega) \in \mathbb{R}$ for any $\omega \in \Omega$, we anyway have that $\{\omega \in \Omega\colon f(\omega) = -\infty\} = \{\omega \in \Omega\colon f(\omega) = -\infty\} = \emptyset \in \mathcal{F}$. Hence, any results on $\mathcal{F}$ measurable functions $f\colon \Omega \to \overline{\mathbb{R}}$ apply directly to $\mathcal{F}/\mathfrak{B}(\mathbb{R})$ measurable functions $f\colon \Omega \to \mathbb{R}$.*

Remark 7.4. *If $(\Omega, \mathcal{F})$ be a measurable space and $f\colon \Omega \to \mathbb{R}^k$, $k \geq 1$, then the statement f is $\mathcal{F}$ measurable always means that f is $\mathcal{F}/\mathfrak{B}(\mathbb{R}^k)$ measurable. As an example, if the measurable space $(\Omega, \mathcal{F})$ is given by $(\mathbb{R}^m, \mathfrak{B}(\mathbb{R}^m))$, i.e., $\mathcal{F} = \mathfrak{B}(\mathbb{R}^m)$ and $f\colon \mathbb{R}^m \to \mathbb{R}^k$, then f is $\mathfrak{B}(\mathbb{R}^m)$ measurable if it is $\mathfrak{B}(\mathbb{R}^m)/\mathfrak{B}(\mathbb{R}^k)$ measurable.*

Proposition 7.8. *Let $(\Omega, \mathcal{F})$ be a measurable space and $f, g\colon \Omega \to \overline{\mathbb{R}}$ be two $\mathcal{F}$ measurable functions. Then, $\{\omega \in \Omega\colon f(\omega) = g(\omega)\} \in \mathcal{F}$.*

Exercise 7.4. *Let $(\Omega, \mathcal{F})$ be a measurable space and $f\colon \Omega \to \overline{\mathbb{R}}$ be $\mathcal{F}$ measurable. Show that the function $\omega \mapsto cf(\omega)$, $c \in \mathbb{R}$, is $\mathcal{F}$ measurable.*

7.3 Sequences of measurable functions

Proposition 7.9. *Let $(\Omega, \mathcal{F})$ be a measurable space and $f_n\colon \Omega \to \overline{\mathbb{R}}$, $n \in \mathbb{N}$, be a sequence of functions s.t. f_n is $\mathcal{F}$ measurable for any $n \in \mathbb{N}$. Then,*

(i) given $E \subset \mathbb{N}$, the functions $\sup_{n \in E} f_n$ and $\inf_{n \in E} f_n$ are $\mathcal{F}$ measurable;

(ii) *The functions* $\liminf_{n\to\infty} f_n$ *and* $\limsup_{n\to\infty} f_n$ *are* $\mathcal{F}$ *measurable;*

(iii) *If for any* $\omega \in \Omega$, $\lim_{n\to\infty} f_n(\omega)$ *exists, then* $\omega \mapsto (\lim_{n\to\infty} f_n)(\omega)$ *is* $\mathcal{F}$ *measurable;*

(iv) *We have that* $\{\omega \in \Omega \colon (f_n(\omega))_{n\in\mathbb{N}} \text{ converges}\} \in \mathcal{F}$*;*

(v) *Let* $f \colon \Omega \to \overline{\mathbb{R}}$ *be* $\mathcal{F}$ *measurable, then,* $\{\omega \in \Omega \colon f_n(\omega) \xrightarrow{n\to\infty} f(\omega)\} \in \mathcal{F}$.

The following result shows that any non-negative $\mathcal{F}$ measurable function can be understood as a monotone limit of non-negative standard simple functions.

Proposition 7.10. *Let* $(\Omega, \mathcal{F})$ *be a measurable space* $f \colon \Omega \to \overline{\mathbb{R}}$ *be a* $\mathcal{F}$ *measurable function s.t.* $f(\omega) \geq 0$ *for any* $\omega \in \Omega$*. Then, there exists a sequence of standard simple functions* $f_n \colon \Omega \to [0, \infty)$*,* $n \in \mathbb{N}$*, s.t. for any* $\omega \in \Omega$*,* $f_n(\omega) \uparrow f(\omega)$*.*

Definition 7.6. *Let* $f \colon \Omega \to \overline{\mathbb{R}}$ *be a function. We define the positive part of* f *as the function*

$$f^+ = \max\{f, 0\},$$

and the negative part of f *as*

$$f^- = \max\{-f, 0\},$$

Exercise 7.5. *Show that*

(a) *for any* $\omega \in \Omega$*,* $f^+(\omega) \geq 0$ *and* $f^-(\omega) \geq 0$*;*

(b) $f(\omega) = f^+(\omega) - f^-(\omega)$*;*

(c) $|f(\omega)| = f^+(\omega) + f^-(\omega)$*.*

Verify further that if $(\Omega, \mathcal{F})$ *is a measurable space and* $f \colon \Omega \to \overline{\mathbb{R}}$ *is a* $\mathcal{F}$ *measurable function, then its positive and negative parts are* $\mathcal{F}$ *measurable.*

Upon the latter exercise, we have the following result.

Proposition 7.11. *Let* $(\Omega, \mathcal{F})$ *be a measurable space and* $f \colon \Omega \to \overline{\mathbb{R}}$ *be a* $\mathcal{F}$ *measurable function. Then, there exists a sequence of standard simple functions* $(f_n)_{n\in\mathbb{N}}$ *s.t. for any* $\omega \in \Omega$*,* $\lim_{n\to\infty} f_n(\omega) = f(\omega)$*.*

Proof. We write $f = f^+ - f^-$. Since both, f^+ and f^- are non-negative and $\mathcal{F}$ measurable, we use Proposition 7.10 to find simple functions $(f_n^+)_{n\in\mathbb{N}}$ and $(f_n^-)_{n\in\mathbb{N}}$ s.t. $f_n^+(\omega) \uparrow f^+(\omega)$ and $f_n^-(\omega) \uparrow f^-(\omega)$. Therefore, for any $\omega \in \Omega$, $f_n^+(\omega) - f_n^-(\omega) \xrightarrow{n\to\infty} f^+(\omega) - f^-(\omega) = f(\omega)$. $\square$

We note that the statement of Exercise 7.3 can be generalized.

Proposition 7.12. *Let* $(\Omega, \mathcal{F})$ *be a measurable space and* $f_i \colon \Omega \to \overline{\mathbb{R}}$*,* $i = 1, \ldots, k$*, be* $\mathcal{F}$ *measurable. Then, the functions*

- $\omega \mapsto \prod_{i=1}^k f_i(\omega)$*;*
- $\omega \mapsto \sum_{i=1}^k c_i f_i(\omega)$*,* $c_1, \ldots, c_k \in \mathbb{R}$*;*

are $\mathcal{F}$ *measurable.*

Proof. We may see this result as a consequence of Propositions 7.11 and 7.9. Since for any $i = 1, \ldots, k$, f_i is $\mathcal{F}$ measurable, we write $f_i(\omega) = \lim_{n\to\infty} f_i^n(\omega)$, $\omega \in \Omega$, where $(f_i^n)_{n\in\mathbb{N}}$ is a sequence of $\mathcal{F}$ measurable functions (cf. Propositions 7.11). Therefore, by item (iii) of Propositions 7.9,

$$\omega \mapsto \lim_{n\to\infty} \left(\prod_{i=1}^k f_i^n \right)(\omega) = \prod_{i=1}^k f_i(\omega),$$

and

$$\omega \mapsto \lim_{n\to\infty} \left(\sum_{i=}^{k} c_i f_i^n \right)(\omega) = \sum_{i=1}^{k} c_i f_i(\omega),$$

are $\mathcal{F}$ measurable. □

7.4 Minimal measurability

Let $f(\omega) = (f_1(\omega), \ldots, f_k(\omega))$, $\omega \in \Omega$, be a function defined on some set Ω, taking values in $\mathbb{R}^k$. We recall (cf. Exercise 4.9) that the σ-field generated by f is given by

$$\sigma(f) = \{f^{-1}(B)\colon B \in \mathfrak{B}(\mathbb{R}^k)\}.$$

Let

$$I = \{\mathcal{F}\colon \mathcal{F} \text{ is a } \sigma\text{-field on } \Omega \text{ s.t. } \sigma(f) \subset \mathcal{F}\}.$$

If $\mathcal{F} \in I$, then f is $\mathcal{F}$ measurable, since for any $B \in \mathfrak{B}(\mathbb{R}^k)$, $f^{-1}(B) \in \sigma(f) \subset \mathcal{F}$. Let

$$\Sigma = \bigcap_{\mathcal{F}\in I} \mathcal{F}.$$

We know that Σ is a σ-field (cf. Proposition 4.2) and by definition, we know that

$$\Sigma = \sigma(\sigma(f)) = \sigma(f),$$

since $\sigma(f)$ is a σ-field. Further, since $\sigma(f)$ is a σ-field that contains $\sigma(f)$, $\sigma(f) \in I$. Consider the set

$$J = \{\mathcal{F}\colon \mathcal{F} \text{ is a } \sigma\text{-field on } \Omega \text{ and } f \text{ is } \mathcal{F} \text{ measurable}\}.$$

We have that $I \subset J$. Take $\mathcal{F} \in J$, hence f is $\mathcal{F}$ measurable. Then, $\sigma(f) \subset \mathcal{F}$. Hence, $I = J$. In conclusion, $\sigma(f)$ is the smallest σ-field on Ω s.t. f is $\sigma(f)$ measurable since any other $\mathcal{A} \in J = I$, is s.t. $\sigma(f) = \Sigma = \cap_{\mathcal{F}\in I}\mathcal{F} \subset \mathcal{A}$. We notice that any function $h\colon \Omega \to \mathbb{R}$ that is $\sigma(f)$ measurable is s.t. for any $A \in \mathfrak{B}(\mathbb{R})$ there exists $B \in \mathfrak{B}(\mathbb{R}^k)$ s.t.

$$\{\omega \in \Omega\colon g(\omega) \in A\} = \{\omega \in \Omega\colon f(\omega) \in B\}.$$

The next result shows that the condition that $h\colon \Omega \to \mathbb{R}$ is $\sigma(f)$ measurable is equivalent to the condition that $h = g(f)$, for some function $g\colon \mathbb{R}^k \to \mathbb{R}$.

Proposition 7.13. *Let $f(\omega) = (f_1(\omega), \ldots, f_k(\omega))$, $\omega \in \Omega$, be a function defined on some set Ω, taking values in $\mathbb{R}^k$. Let $\sigma(f)$ be the σ-field generated by f and $h\colon \Omega \to \mathbb{R}$ be a function. Then, h is $\sigma(f)$ measurable if and only if there exists a function $g\colon \mathbb{R}^k \to \mathbb{R}$ which is $\mathfrak{B}(\mathbb{R}^k)$ measurable and s.t. $h(\omega) = g(f(\omega))$.*

Example 7.9. *Let $\Omega = \{h, t\}$, $\mathcal{F} = \mathcal{P}(\{h, t\})$ and f_1 be the $\mathcal{P}(\{h, t\})/\mathfrak{B}(\mathbb{R})$ measurable map*

$$f_1(\omega) = \begin{cases} 1, & \text{if } \omega = h, \\ 0, & \text{if } \omega = t, \end{cases}$$

given in Example 7.3. Further, let

$$f_2(\omega) = \begin{cases} 1, & \text{if } \omega = t, \\ 0, & \text{if } \omega = h. \end{cases}$$

That is, $f_1 = \mathbb{1}_H$ and $f_2 = \mathbb{1}_T$, where $H = \{\omega \in \Omega\colon \omega = h\}$ and $T = \{\omega \in \Omega\colon \omega = t\}$. Define the map $h = \mathbb{1}_H + \mathbb{1}_T$. Then, h is $\sigma(f)$ measurable, where $f = (f_1, f_2)$. This follows from Proposition 7.13, with $g(x) = x_1 + x_2$, $x = (x_1, x_2) \in \mathbb{R}^2$, which is $\mathfrak{B}(\mathbb{R}^k)$ measurable since it is continuous (cf. Propositions 2.12 and 7.2).

7.5 Solution to exercises

Solution 7.1 (Solution to Exercise 7.1). *This follows from the fact that for any $x \in \mathbb{R}$,*

$$\{\omega \in \Omega \colon f(\omega) > x\} \in \{\emptyset, \Omega\} \subset \mathcal{F}.$$

Solution 7.2 (Solution to Exercise 7.2). *We need to show that for any $A^{**} \in \mathcal{F}^{**}$,*

$$f^*(f)^{-1}(A^{**}) \in \mathcal{F}.$$

Therefore, let $A^{**} \in \mathcal{F}^{**}$. We have that

$$\begin{aligned} f^*(f)^{-1}(A^{**}) &= \{\omega \in \Omega \colon f^*(f)(\omega) \in A^{**}\} \\ &= \{\omega \in \Omega \colon f^*(f(\omega)) \in A^{**}\} \\ &= \{\omega \in \Omega \colon f(\omega) \in (f^*)^{-1}(A^{**})\} \\ &= f^{-1}\big((f^*)^{-1}(A^{**})\big) \end{aligned}$$

Then, since by assumption, f^* is $\mathcal{F}^*/\mathcal{F}^{**}$ measurable, it follows that $(f^*)^{-1}(A^{**}) \in \mathcal{F}^*$. Further, since f is $\mathcal{F}/\mathcal{F}^*$ measurable, we conclude that $f^{-1}((f^*)^{-1}(A^{**})) \in \mathcal{F}$.

Solution 7.3 (Solution to Exercise 7.3). *Since for any $i = 1, \ldots, k$, $\omega \mapsto c_i$ and $\omega \mapsto f_i(\omega)$ are $\mathcal{F}/\mathfrak{B}(\mathbb{R})$ measurable, $\omega \mapsto c_i f_i(\omega)$ is $\mathcal{F}/\mathfrak{B}(\mathbb{R})$ measurable (cf. Example 7.6). Therefore, f is $\mathcal{F}/\mathfrak{B}(\mathbb{R})$ measurable (cf. Example 7.5).*

Solution 7.4 (Solution to Exercise 7.4). *If $c = 0$, then $\{\omega \in \Omega \colon cf(\omega) = 0\} = \Omega$. Thus, for any set $A \subset \overline{\mathbb{R}}$, the set $\{\omega \in \Omega \colon cf(\omega) \in A\}$ is either the empty set or Ω. Suppose that $c \neq 0$. Then,*

$$\{\omega \in \Omega \colon cf(\omega) = -\infty\} = \{\omega \in \Omega \colon f(\omega) = -\infty\} \in \mathcal{F},$$

and

$$\{\omega \in \Omega \colon cf(\omega) = \infty\} = \{\omega \in \Omega \colon f(\omega) = \infty\} \in \mathcal{F},$$

since f is $\mathcal{F}$ measurable. Consider the function

$$f^*(\omega) = cf(\omega)\mathbb{1}_F(\omega), \quad \omega \in \Omega,$$

where $F = \{\omega \in \Omega \colon cf(\omega) \in \mathbb{R}\}$. We notice that $F \in \mathcal{F}$, since

$$F^c = \{\omega \in \Omega \colon cf(\omega) = -\infty\} \cup \{\omega \in \Omega \colon cf(\omega) = \infty\}.$$

Let $x \in \mathbb{R}$ and define $z(\omega) = 0$ for any $\omega \in \Omega$, i.e., the function that is constant and equal to zero on Ω. We have that

$$\begin{aligned} \{\omega \in \Omega \colon f^*(\omega) \leq x\} &= \big(\{\omega \in \Omega \colon f^*(\omega) \leq x\} \cap F\big) \cup (\{\omega \in \Omega \colon f^*(\omega) \leq x\} \cap F^c) \\ &= \big(\{\omega \in \Omega \colon cf(\omega) \leq x\} \cap F\big) \cup (\{\omega \in \Omega \colon z(\omega) \leq x\} \cap F^c) \\ &= \big(\{\omega \in \Omega \colon f(\omega) \leq x/c\} \cap F\big) \cup (\{\omega \in \Omega \colon z(\omega) \leq x\} \cap F^c) \in \mathcal{F}, \end{aligned}$$

since f is $\mathcal{F}$ measurable, $F \in \mathcal{F}$ and z is $\mathcal{F}$ measurable. This shows that f^* is $\mathcal{F}/\mathfrak{B}(\mathbb{R})$ measurable (cf. Proposition 7.3). Therefore, given $A \in \mathfrak{B}(\mathbb{R})$,

$$\begin{aligned} \{\omega \in \Omega \colon cf(\omega) \in A\} &= \big(\{\omega \in \Omega \colon f^*(\omega) \in A\} \cap F\big) \cup (\{\omega \in \Omega \colon cf(\omega) \in A\} \cap F^c) \\ &= \{\omega \in \Omega \colon f^*(\omega) \in A\} \cap F \in \mathcal{F}. \end{aligned}$$

This shows that $\omega \mapsto cf(\omega)$ is $\mathcal{F}$ measurable.

Solution 7.5 (Solution to Exercise 7.5). *By definition, $f^+(\omega) = \max\{f(\omega), 0\} \geq 0$ and $f^-(\omega) = \max\{-f(\omega), 0\} \geq 0$. Regarding (b), let $\omega \in \Omega$, then, if $f(\omega) < 0$, $-f(\omega) > 0$, and hence $f^+(\omega) - f^-(\omega) = f(\omega)$. Similarly, if $f(\omega) > 0$, then $f^+(\omega) - f^-(\omega) = f(\omega)$. If $f(\omega) = 0$, then $f^+(\omega) = f^-(\omega) = 0$. To see (c), notice that for any $\omega \in \Omega$,*

$$f^+(\omega) + f^-(\omega) = \max\{f(\omega), 0\} + \max\{-f(\omega), 0\} = \begin{cases} f(\omega), & \text{if } f(\omega) \geq 0, \\ -f(\omega), & \text{if } f(\omega) < 0. \end{cases}$$

Finally, by item (i) of Proposition 7.9, if f is $\mathcal{F}$ measurable, f^+ and f^- are $\mathcal{F}$ measurable as well (cf. Exercise 7.4).

7.6 Additional exercises

Exercise 7.6. *Let $(\Omega, \mathcal{F})$ be a measurable space and $f \colon \Omega \to \mathbb{R}$ be $\mathcal{F}$ measurable. Verify that the following functions are $\mathcal{F}$ measurable:*

- $\omega \mapsto \mathrm{e}^{f(\omega)}$;
- $\omega \mapsto 1/f(\omega)$ *(provided that for any $\omega \in \Omega$, $f(\omega) \neq 0$);*
- $\omega \mapsto \log(f(\omega))$ *(provided that for any $\omega \in \Omega$, $f(\omega) > 0$).*

Exercise 7.7. *Let $(\Omega, \mathcal{F})$ be a measurable space and $f_1 \colon \Omega \to \mathbb{R}$ and $f_2 \colon \Omega \to \mathbb{R}$ be two $\mathcal{F}/\mathfrak{B}(\mathbb{R})$ measurable functions s.t. $f_1(\Omega) \subset E_1$ and $f_2(\Omega) \subset E_2$, where $E_1, E_2 \in \mathfrak{B}(\mathbb{R})$. Let $g_1 \colon E_1 \to \mathbb{R}$ and $g_2 \colon E_2 \to \mathbb{R}$ be continuous. Verify that the function*

$$f(\omega) = (g_1(f_1)(\omega), g_2(f_2)(\omega)), \quad \omega \in \Omega,$$

is $\mathcal{F}/\mathfrak{B}(\mathbb{R}^2)$ measurable.

Exercise 7.8. *Let $(\Omega, \mathcal{F})$ and $(\Omega^*, \mathcal{F}^*)$ be two measurable spaces and $g \colon \Omega \to \Omega^*$ be $\mathcal{F}/\mathcal{F}^*$ measurable. Let μ be a measure on $\mathcal{F}$. Define the function*

$$\mu g^{-1}(A^*) = \mu(g^{-1}(A^*)) = \mu(\{\omega \in \Omega \colon g(\omega) \in A^*\}), \quad A^* \in \mathcal{F}^*.$$

Show that μg^{-1} is a measure on $\mathcal{F}^*$.

Exercise 7.9. *Let f and h be as in Proposition 7.13. Suppose that there exists $g \colon \mathbb{R}^k \to \mathbb{R}$ which is $\mathfrak{B}(\mathbb{R}^k)$ measurable and s.t. $h(\omega) = g(f(\omega))$. Then, h is $\sigma(f)$ measurable.*
Hint: *f is $\sigma(f)$ measurable.*

Exercise 7.10. *Let f and h be as in Proposition 7.13. Suppose that h is a simple function in standard form (cf. Definition 7.4 with $\mathcal{F} = \sigma(f)$). Show that there exists $g \colon \mathbb{R}^k \to \mathbb{R}$ which is $\mathfrak{B}(\mathbb{R}^k)$ measurable and s.t. $h(\omega) = g(f(\omega))$.*
Hint: *Consider the sets $B_i = \{\omega \in \Omega \colon h(\omega) = \alpha_i\}$, $i = 1 \ldots, n$, where $\alpha_i \in \mathbb{R}$, $i = 1 \ldots, n$, are the n values that h can take.*

8 Integration: Part I

We rely on the following conventions regarding infinity:

$$\begin{aligned}
&x+\infty=\infty+x=\infty, \quad x-\infty=-\infty+x=-\infty, \quad x\in\mathbb{R},\\
&x\cdot\infty=\infty\cdot x=\infty, \quad x\cdot(-\infty)=(-\infty)\cdot x=-\infty, \quad x>0,\\
&0\cdot\infty=\infty\cdot 0=0,\\
&\infty\cdot\infty=\infty.
\end{aligned}$$

The omitted proofs of this chapter are found in Section B.5 of the appendix.

8.1 The integral for nonnegative functions

If $f\colon\Omega\to\overline{\mathbb{R}}$ is s.t. $f(\omega)\geq 0$ for any $\omega\in\Omega$, f is said to be nonnegative.

Definition 8.1. *Let Ω be a set. A partition of Ω is a disjoint collection $\{A\colon A\in P\}$, $P\subset\mathcal{P}(\Omega)$, s.t. $\cup_{A\in P}A=\Omega$. That is, a partition of Ω is a disjoint collection of subsets of Ω whose union is Ω. If ξ is a partition of Ω, a set $A\in\xi$ is referred to as an atom of ξ. A partition ξ of Ω is said to be finite, if it contains a finite number of atoms.*

Example 8.1. *Let $\Omega=\{0,1,\ldots,N\}$, $N\in\mathbb{N}$. Then, $\xi=\{\{\omega\}\colon\omega\in\Omega\}$ is a partition of Ω.*

Definition 8.2. *Let $(\Omega,\mathcal{F})$ be a measurable space. We use the notation $Z_0^{\mathcal{F}}(\Omega)=Z_0^{\mathcal{F}}$ for the set which contains all the finite partitions of Ω with atoms form $\mathcal{F}$. That is,*

$$Z_0^{\mathcal{F}}=\{\xi\colon \xi \textit{ is a finite partition of } \Omega \textit{ s.t. for any } A\in\xi,\ A\in\mathcal{F}\}.$$

Definition 8.3. *Let $(\Omega,\mathcal{F},\mu)$ be a measure space and $f\colon\Omega\to\overline{\mathbb{R}}$ be nonnegative and $\mathcal{F}$ measurable. Then, we define*

$$S_\mu^f(\xi)=\sum_{A\in\xi}\Big(\inf_{\omega\in A}f(\omega)\Big)\mu(A), \quad \xi\in Z_0^{\mathcal{F}},$$

and

$$\int_\Omega f(\omega)\mu(d\omega)=\sup_{\xi\in Z_0^{\mathcal{F}}}S_\mu^f(\xi).$$

Upon the latter definition, we deduce the integral for a (nonnegative) standard simple function (cf. Definition 7.4).

Proposition 8.1. *Let $(\Omega,\mathcal{F},\mu)$ be a measure space and $f(\omega)=\sum_{i=1}^N\alpha_i\mathbb{1}_{A_i}$, where $N\in\mathbb{N}$, $\alpha_i\in[0,\infty)$, $i=1,\ldots,N$, and $\{A_i\colon i=1,\ldots,N\}\in Z_0^{\mathcal{F}}$. That is, f is a simple function in standard form, with nonnegative coefficients α_i. We have that*

$$\int_\Omega f(\omega)\mu(d\omega)=\sum_{i=1}^N\alpha_i\mu(A_i).$$

Example 8.2. *Consider the measure space $(\mathbb{R},\mathfrak{B}(\mathbb{R}),\lambda)$, i.e., the real numbers equipped with the Borel σ-field and the Lebesgue measure (cf. Example 6.2). Let $-\infty<a=a_0<a_1<\cdots<a_N=b<\infty$ and consider the partition*

$$\xi=\big\{(-\infty,a_0]\big\}\cup\big\{(a_{i-1},a_i]\colon i=1,\ldots,N\big\}\cup\big\{(a_N,\infty)\big\}.$$

Define

$$f(x)=\begin{cases}\sum_{i=1}^N\alpha_i\mathbb{1}_{(a_{i-1},a_i]}(x), & \textit{if } x\in(a,b],\\ 0, & \textit{otherwise.}\end{cases}$$

Using the convention that $0 \cdot \infty = 0$ (see also Exercise 6.5), we obtain that

$$\int_{\mathbb{R}} f(x)\lambda(dx) = \sum_{i=1}^{N} \alpha_i(a_i - a_{i-1}),$$

i.e., $\int_{\mathbb{R}} f(x)\lambda(dx)$ gives the "area under the curve" of f.

Exercise 8.1. *Let $(\Omega, \mathcal{F}, \mu)$ be a measure space and $f, g\colon \Omega \to \overline{\mathbb{R}}$ be two nonnegative and $\mathcal{F}$ measurable functions s.t. for any $\omega \in \Omega$, $f(\omega) \le g(\omega)$. Show that $\int_\Omega f(\omega)\mu(d\omega) \le \int_\Omega g(\omega)\mu(d\omega)$.*

The following two results are important tools in integration theory. The first is known as the monotone convergence theorem and the second shows that the integral of nonnegative functions is linear.

Proposition 8.2. *Let $(\Omega, \mathcal{F}, \mu)$ be a measure space and $f_n\colon \Omega \to \overline{\mathbb{R}}$, $n \in \mathbb{N}$, be a sequence of nonnegative $\mathcal{F}$ measurable functions s.t. for any $\omega \in \Omega$ $f_n(\omega) \uparrow f(\omega)$ for some $f\colon \Omega \to \overline{\mathbb{R}}$. Then,*

$$\int_\Omega f_n(\omega)\mu(d\omega) \uparrow \int_\Omega f(\omega)\mu(d\omega).$$

Proposition 8.3. *Let $(\Omega, \mathcal{F}, \mu)$ be a measure space, $f, g\colon \Omega \to \overline{\mathbb{R}}$ be two nonnegative and $\mathcal{F}$ measurable functions. Given $\alpha, \beta \in [0, \infty)$ we have that*

$$\int_\Omega (\alpha f + \beta g)(\omega)\mu(d\omega) = \alpha \int_\Omega f(\omega)\mu(d\omega) + \beta \int_\Omega g(\omega)\mu(d\omega).$$

As a consequence of the latter two propositions we have the following result:

Proposition 8.4. *Let $(\Omega, \mathcal{F}, \mu)$ be a measure space and $f_i\colon \Omega \to \overline{\mathbb{R}}$, $i \in \mathbb{N}$, be a sequence of nonnegative $\mathcal{F}$ measurable functions, then*

$$\int_\Omega \Big(\sum_{i\in\mathbb{N}} f_i\Big)(\omega)\mu(d\omega) = \sum_{i\in\mathbb{N}} \Big(\int_\Omega f_i(\omega)\mu(d\omega)\Big).$$

Proof. Since for any $i \in \mathbb{N}$, $f_i(\omega) \ge 0$, we have that $\sum_{i=1}^n f_i(\omega) \uparrow \sum_{i\in\mathbb{N}} f_i(\omega)$. Notice that $(\sum_{i=1}^n f_i(\omega))_{n\in\mathbb{N}}$ is a sequence of nonnegative $\mathcal{F}$ measurable functions. By Proposition 8.2,

$$\int_\Omega \Big(\sum_{i=1}^{n} f_i\Big)(\omega)\mu(d\omega) \uparrow \int_\Omega \Big(\sum_{i\in\mathbb{N}} f_i\Big)(\omega)\mu(d\omega).$$

Further, by Proposition 8.3,

$$\int_\Omega \Big(\sum_{i=1}^{n} f_i\Big)(\omega)\mu(d\omega) = \sum_{i=1}^{n} \Big(\int_\Omega f_i(\omega)\mu(d\omega)\Big) \uparrow \sum_{i\in\mathbb{N}} \Big(\int_\Omega f_i(\omega)\mu(d\omega)\Big).$$

Since a limit of a real valued sequence is unique, the result follows. □

Remark 8.1. *We remark that the monotone convergence theorem (cf. Proposition 8.2), allows for another interpretation of the integral for nonnegative functions. If $(\Omega, \mathcal{F}, \mu)$ is a measure space and $f\colon \Omega \to \overline{\mathbb{R}}$ is a nonnegative and $\mathcal{F}$ measurable function, then, we have seen in Proposition 7.10 that it is always possible to approximate f by a sequence of (nonnegative) standard simple functions, i.e., for any $\omega \in \Omega$, $f_n(\omega) \uparrow f(\omega)$ where for any $n \in \mathbb{N}$, f_n is a simple function in standard form. Hence, upon the monotone convergence theorem, we obtain,*

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

where the latter convergence is monotone. Hence, the integral for a nonnegative and $\mathcal{F}$ measurable function can be understood as the monotone limit of the integral of simple functions.

Definition 8.4. *Let $(\Omega, \mathcal{F}, \mu)$ be a measure space. Suppose that for any $\omega \in \Omega$, $S(\omega)$ is a statement on Ω. We say S is true μ almost everywhere (a.e.) if $\mu(\{\omega \colon S(\omega) \text{ is false}\}) = 0$.*

Example 8.3. *Let $(\Omega, \mathcal{F}, \mu)$ be a measure space and $f\colon \Omega \to \overline{\mathbb{R}}$ be $\mathcal{F}$ measurable and nonnegative. If $f = 0$ μ a.e., then $\int_\Omega f(\omega)\mu(d\omega) = 0$. To see it take any $\xi \in Z_0^{\mathcal{F}}$. Let $A \in \xi$. Suppose that $A \cap \{\omega \colon f(\omega) = 0\} \neq \emptyset$, then $\inf_{\omega \in A} f(\omega) = 0$. Otherwise, if $A \cap \{\omega \colon f(\omega) = 0\} = \emptyset$,*

$$\mu(A) = \mu(A \cap \{\omega \colon f(\omega) \neq 0\}) \leq \mu(\{\omega \colon f(\omega) \neq 0\}) = 0.$$

Hence, $S_\mu^f(\xi) = 0$.

We can derive further properties of the integral for nonnegative functions.

Proposition 8.5. *Let $(\Omega, \mathcal{F}, \mu)$ be a measure space. Assume that $f, g\colon \Omega \to \overline{\mathbb{R}}$ be two nonnegative and $\mathcal{F}$ measurable functions.*

(i) If $\mu(\{\omega \colon f(\omega) > 0\}) > 0$, then $\int_\Omega f(\omega)\mu(d\omega) > 0$;

(ii) If $\int_\Omega f(\omega)\mu(d\omega) < \infty$, then $f < \infty$ μ a.e.;

(iii) If $f \leq g$ μ a.e., then $\int_\Omega f(\omega)\mu(d\omega) \leq \int_\Omega g(\omega)\mu(d\omega)$;

(iv) If $f = g$ μ a.e., then $\int_\Omega f(\omega)\mu(d\omega) = \int_\Omega g(\omega)\mu(d\omega)$.

Exercise 8.2. *Let $(\Omega, \mathcal{F}, \mu)$ be a measure space and $f_n\colon \Omega \to \overline{\mathbb{R}}$, $n \in \mathbb{N}$, be a sequence of nonnegative $\mathcal{F}$ measurable functions s.t. $f_n \uparrow f$ μ a.e. for some $f\colon \Omega \to \overline{\mathbb{R}}$. Show that $\int_\Omega f_n(\omega)\mu(d\omega) \uparrow \int_\Omega f(\omega)\mu(d\omega)$.*

8.2 Integrable functions

We recall the definition of the positive (f^+) and negative (f^-) parts of a function (cf. Definition 7.6, recall also Exercise 7.5).

Definition 8.5. *Let $(\Omega, \mathcal{F}, \mu)$ be a measure space and $f\colon \Omega \to \overline{\mathbb{R}}$ be a $\mathcal{F}$ measurable function. The integral of f is defined by*

$$\int_\Omega f(\omega)\mu(d\omega) = \int_\Omega f^+(\omega)\mu(d\omega) - \int_\Omega f^-(\omega)\mu(d\omega),$$

unless $\int_\Omega f^+(\omega)\mu(d\omega) = \int_\Omega f^-(\omega)\mu(d\omega) = \infty$, in which case $\int_\Omega f(\omega)\mu(d\omega)$ is not defined. If both, $\int_\Omega f^+(\omega)\mu(d\omega) < \infty$ and $\int_\Omega f^-(\omega)\mu(d\omega) < \infty$, f is said to be integrable.

Remark 8.2. *If $(\Omega, \mathcal{F}, \mu)$ is a measure space and f is as in the latter definition, then the assumption that f is integrable is defined upon the measure μ, i.e., if one wants to further refer to the measure of integration one specifies that f is integrable with respect to μ.*

Exercise 8.3. *Let $(\Omega, \mathcal{F}, \mu)$ be a measure space and $f, g\colon \Omega \to \overline{\mathbb{R}}$ be $\mathcal{F}$ measurable. Suppose that f is integrable and $f = g$ μ a.e. Show that g is integrable and $\int_\Omega f(\omega)\mu(d\omega) = \int_\Omega g(\omega)\mu(d\omega)$.*

Another characterization of integrability is the following:

Proposition 8.6. *Let $(\Omega, \mathcal{F}, \mu)$ be a measure space and $f\colon \Omega \to \overline{\mathbb{R}}$ be $\mathcal{F}$ measurable. Then, f is integrable if and only if $\int_\Omega |f(\omega)|\mu(d\omega) < \infty$.*

Proof. By definition, f is integrable if $\int_\Omega f^+(\omega)\mu(d\omega) < \infty$ and $\int_\Omega f^-(\omega)\mu(d\omega) < \infty$. This is equivalent to

$$\int_\Omega f^+(\omega)\mu(d\omega) + \int_\Omega f^-(\omega)\mu(d\omega) = \int_\Omega f^+(\omega) + f^-(\omega)\mu(d\omega) = \int_\Omega |f(\omega)(\omega)|\mu(d\omega) < \infty.$$

Recall also Exercise 7.5. Notice also that f integrable and $\int_\Omega |f(\omega)(\omega)|\mu(d\omega) = \infty$ gives a contradiction. Hence, if f is integrable, it must follow that $\int_\Omega |f(\omega)(\omega)|\mu(d\omega) < \infty$. $\square$

Exercise 8.4. *Let* $(\Omega, \mathcal{F}, \mu)$ *be a measure space and* $f, g \colon \Omega \to \overline{\mathbb{R}}$ *be* $\mathcal{F}$ *measurable. Show that*

(a) if $|f| \le |g|$ *a.e., and* g *is integrable, then* f *is integrable as well;*

(b) if $\mu(\Omega) < \infty$ *and* f *is bounded on* Ω*, then* f *is integrable.*

The following is a general version of (iii) in Proposition 8.5.

Proposition 8.7. *Let* $(\Omega, \mathcal{F}, \mu)$ *be a measure space and* $f, g \colon \Omega \to \overline{\mathbb{R}}$ *be* $\mathcal{F}$ *measurable. If* f *and* g *are integrable and* $f \le g$ *a.e., then,* $\int_\Omega f(\omega)\mu(d\omega) \le \int_\Omega g(\omega)\mu(d\omega)$.

The extension of Proposition 8.3 reads as follows:

Proposition 8.8. *Let* $(\Omega, \mathcal{F}, \mu)$ *be a measure space and* $f, g \colon \Omega \to \overline{\mathbb{R}}$ *be* $\mathcal{F}$ *measurable. If* f *and* g *are integrable, then for any* $\alpha, \beta \in \mathbb{R}$, $\alpha f + \beta g$ *is integrable and*

$$\int_\Omega (\alpha f + \beta g)(\omega)\mu(d\omega) = \alpha \int_\Omega f(\omega)\mu(d\omega) + \beta \int_\Omega g(\omega)\mu(d\omega).$$

Exercise 8.5. *Let* $(\Omega, \mathcal{F}, \mu)$ *be a measure space and* $A \in \mathcal{F}$ *be s.t.* $\mu(A) < \infty$*. Let* $f = \sum_{i=1}^N \alpha_i \mathbb{1}_{A_i}$, $N \in \mathbb{N}$, $\alpha_i \in \mathbb{R}$, $i = 1, \dots, N$*, be a simple function where* $\{A_i \colon i = 1, \dots, N\} \subset \mathcal{F}$ *is disjoint and* $\cup_{i=1}^N A_i = A$*. Show that*

$$\int_A f(\omega)\mu(d\omega) = \sum_{i=1}^N \alpha_i \mu(A_i).$$

Exercise 8.6. *Let* $(\Omega, \mathcal{F}, \mu)$ *be a measure space and* $f, g \colon \Omega \to \overline{\mathbb{R}}$ *be* $\mathcal{F}$ *measurable and integrable. Show that*

$$\left| \int_\Omega f(\omega)\mu(d\omega) - \int_\Omega g(\omega)\mu(d\omega) \right| \le \int_\Omega |f(\omega) - g(\omega)|\mu(d\omega).$$

8.3 Fatou's lemma and Lebesgue's dominated convergence theorem

The following is known as Fatou's lemma.

Proposition 8.9. *Let* $(\Omega, \mathcal{F}, \mu)$ *be a measure space and* $f_n \colon \Omega \to \overline{\mathbb{R}}$, $n \in \mathbb{N}$*, be a sequence of nonnegative and* $\mathcal{F}$ *measurable functions. Then,*

$$\int_\Omega \liminf_{n\to\infty} f_n(\omega)\mu(d\omega) \le \liminf_{n\to\infty} \int_\Omega f_n(\omega)\mu(d\omega).$$

Fatou's lemma is used to prove Lebesgue's dominated convergence (cf. Section B.5 of the appendix):

Proposition 8.10. *Let* $(\Omega, \mathcal{F}, \mu)$ *be a measure space and* $f_n \colon \Omega \to \overline{\mathbb{R}}$, $n \in \mathbb{N}$*, be a sequence of* $\mathcal{F}$ *measurable functions. Suppose that there exist* $f, g \colon \Omega \to \overline{\mathbb{R}}$ *s.t.* $f_n \to f$ *a.e. and for any* $n \in \mathbb{N}$, $|f_n| \le g$ *a.e. where* g *is integrable. Then* f *is integrable and*

$$\int_\Omega f_n(\omega)\mu(d\omega) \xrightarrow{n\to\infty} \int_\Omega f(\omega)\mu(d\omega).$$

In the following we discuss some applications of the latter proposition. As a first consequence, we can further extend Proposition 8.4.

Proposition 8.11. *Let* $(\Omega, \mathcal{F}, \mu)$ *be a measure space and* $f_i \colon \Omega \to \overline{\mathbb{R}}$, $i \in \mathbb{N}$*, be a sequence of* $\mathcal{F}$ *measurable functions. If* $\lim_{n\to\infty} \sum_{i=1}^n f_i$ *exists a.e. and there exists integrable* g *s.t.* $|\sum_{i=1}^n f_i| \le g$ *a.e., then* $\sum_{i\in\mathbb{N}} f_i$ *and* f_i, $i \in \mathbb{N}$*, are integrable and*

$$\sum_{i\in\mathbb{N}} \left(\int_\Omega f_i(\omega)\mu(d\omega) \right) = \int_\Omega \left(\sum_{i\in\mathbb{N}} f_i \right)(\omega)\mu(d\omega).$$

Proof. Since $\lim_{n\to\infty}\sum_{i=1}^n f_i$ exists a.e., we have that $\sum_{i=1}^n f_i(\omega) \xrightarrow{n\to\infty} \sum_{i\in\mathbb{N}} f_i(\omega)$ a.e. By Proposition 8.10, we can interchange limit and integration, i.e.,

$$\int_\Omega \Big(\sum_{i=1}^n f_i\Big)(\omega)\mu(d\omega) \xrightarrow{n\to\infty} \int_\Omega \Big(\sum_{i\in\mathbb{N}} f_i\Big)(\omega)\mu(d\omega).$$

In particular, $\lim_{n\to\infty}\int_\Omega(\sum_{i=1}^n f_i)(\omega)\mu(d\omega)$ exists. Hence, since

$$\lim_{n\to\infty}\int_\Omega \Big(\sum_{i=1}^n f_i\Big)(\omega)\mu(d\omega) = \lim_{n\to\infty}\sum_{i=1}^n \Big(\int_\Omega f_i(\omega)\mu(d\omega)\Big) = \sum_{i\in\mathbb{N}} \Big(\int_\Omega f_i(\omega)\mu(d\omega)\Big),$$

the result follows. □

Another consequence is the following result:

Proposition 8.12. *Let $(\Omega,\mathcal{F},\mu)$ be a measure space and $f\colon U\times\Omega\to\mathbb{R}$ be a function where $U\subset\mathbb{R}^k$. Assume that*

(i) for any $u\in U$ $\omega\mapsto f(u,\omega)$ is $\mathcal{F}$ measurable;

(ii) for any $u_0\in U$, $u\mapsto f(u,\omega)$ is continuous in u_0 μ a.e.;

(iii) there exists a nonnegative and μ integrable $g\colon\Omega\to\mathbb{R}$ s.t. for any $u\in U$, $|f(u,\omega)|\le g(\omega)$, μ a.e.

Then, $F(u)=\int_\Omega f(u,\omega)\mu(d\omega)$, $u\in U$, is s.t. $F\colon U\to\mathbb{R}$ and F is continuous on U.

Proof. Using Exercise 8.4, (iii) implies that for any $u\in U$, $\omega\mapsto f(u,\omega)$ is μ integrable. Hence, $F\colon U\to\mathbb{R}$. Let u_n, $n\in\mathbb{N}$, be s.t. $u_n\xrightarrow{n\to\infty}u_0$, $u_0\in U$. Then, (ii) implies that $f(u_n,\omega)\xrightarrow{n\to\infty}f(u_0,\omega)$, μ a.e. on Ω (cf. Proposition 3.25). Therefore, we apply Proposition 8.10, and conclude that $F(u_n)\xrightarrow{n\to\infty}F(u_0)$, as well. Since $u_0\in U$ was arbitrary, F is continuous on U. □

Example 8.4. *Consider the measure space $(\mathbb{R},\mathfrak{B}(\mathbb{R}),\lambda)$, where λ is the Lebesgue measure on $\mathfrak{B}(\mathbb{R})$. Let $\varphi\colon\mathbb{R}\to\mathbb{R}$ be $\mathfrak{B}(\mathbb{R})$ measurable and integrable, i.e., $\int_\mathbb{R}|\varphi(x)|\lambda(dx)<\infty$. The Fourier transform $\hat\varphi$ of φ is defined as*

$$\hat\varphi(u)=\int_\mathbb{R} \mathrm{e}^{iux}\varphi(x)\lambda(dx),$$

where $i^2=-1$, the imaginary number. Let $f(u,x)=\mathrm{e}^{iux}\varphi(x)$, $u,x\in\mathbb{R}$. Then, given any $u\in\mathbb{R}$, $x\mapsto\mathrm{e}^{iux}$ is $\mathfrak{B}(\mathbb{R})$ measurable since it is continuous on $\mathbb{R}$. Therefore, $x\mapsto\mathrm{e}^{iux}\varphi(x)$ is $\mathfrak{B}(\mathbb{R})$ measurable (cf. Example 7.6). Then, by the continuity of $u\mapsto\mathrm{e}^{iux}$, for any $x\in\mathbb{R}$, the function $u\mapsto f(u,x)$ is continuous on $\mathbb{R}$. Further, it follows with $|\mathrm{e}^{iux}|=1$, that for any $u,x\in\mathbb{R}$, $|f(u,x)|\le|\varphi(x)|$, where by assumption $\int_\mathbb{R}|\varphi(x)|\lambda(dx)<\infty$. Therefore, we apply Proposition 8.12, and conclude that the Fourier transform of φ is s.t. $\hat\varphi\colon\mathbb{R}\to\mathbb{R}$ and is continuous on $\mathbb{R}$.

Example 8.5. *Consider $(\mathbb{R},\mathfrak{B}(\mathbb{R}),\lambda)$ as in the previous example. Let $\varphi\colon\mathbb{R}\to\mathbb{R}$ be $\mathfrak{B}(\mathbb{R})$ measurable and integrable. Suppose that $h\colon\mathbb{R}\to\mathbb{R}$ is bounded and continuous. Let*

$$h*\varphi(u)=\int_\mathbb{R} h(u-x)\varphi(x)\lambda(dx),\quad u\in\mathbb{R}.$$

Similar to the previous example, we apply Proposition 8.12 and readily see that $h*\varphi$ is continuous and bounded on $\mathbb{R}$.

In order to verify that integration can be interchanged with differentiation, Lebesgue's dominated convergence theorem is a key tool.

Proposition 8.13. *Let $(\Omega, \mathcal{F}, \mu)$ be a measure space and $f\colon (a,b)\times\Omega \to \mathbb{R}$ be a function where $(a,b)\subset\mathbb{R}$ is an open interval. Let $u_0 \in (a,b)$. Suppose that*

(i) for any $u\in(a,b)$ $\omega \mapsto f(u,\omega)$ is $\mathcal{F}$ measurable and integrable with respect to μ;

(ii) There exists a set $F\in\mathcal{F}$ s.t. $\mu(F^c)=0$ and for any $\omega\in F$, the map $u\mapsto f(u,\omega)$ is differentiable in u_0 with derivative $(\partial f/\partial u)(u_0,\omega)$ (that is, $u\mapsto f(u,\omega)$ is differentiable in u_0 μ a.e.);

(iii) there exists a $\mathcal{F}$ measurable and nonnegative function $g\colon \Omega\to\mathbb{R}$ s.t. $\int_\Omega g(\omega)\mu(d\omega)<\infty$ and for any $u\in(a,b)$ $(u\neq u_0)$ we have that μ a.e.,

$$|f(u,\omega)-f(u_0,\omega)|\le g(\omega)|u-u_0|.$$

Then, the function $F(u)=\int_\Omega f(u,\omega)\mu(d\omega)$ is differentiable in u_0 with derivative

$$F'(u_0)=\int_\Omega (\partial f/\partial u)(u_0,\omega)\mu(d\omega),$$

i.e., we are allowed to interchange integration with differentiation.

Proof. Let $(u_n)_{n\in\mathbb{N}}$ be a sequence in $(a,b)\setminus\{u_0\}$ s.t. $u_n \xrightarrow{n\to\infty} u_0$ and define

$$\varphi_n(u_0,\omega)=\frac{f(u_n,\omega)-f(u_0,\omega)}{u_n-u_0},\quad \omega\in\Omega.$$

Let f_{n,u_0} and f_{u_0} be the map $\omega\mapsto\varphi_n(u_0,\omega)$ and $\omega\mapsto(\partial f/\partial u)(u_0,\omega)\mathbb{1}_F(\omega)$, respectively. By item (ii), we have that μ a.e., $f_{n,u_0}\xrightarrow{n\to\infty} f_{u_0}$. Further, by item (iii), for any $n\in\mathbb{N}$,

$$|f_{n,u_0}|\le g(\omega),\quad \mu \text{ a.e.},$$

and $\int_\Omega |g(\omega)|\mu(d\omega)<\infty$. Hence, we are in position to apply Proposition 8.10, and deduce that

$$\lim_{n\to\infty}\frac{F(u_n)-F(u_0)}{u_n-u_0}=\lim_{n\to\infty}\int_\Omega f_{n,u_0}(\omega)\mu(d\omega)=\int_\Omega(\partial f/\partial u)(u_0,\omega)\mu(d\omega).$$

□

We remark that stronger conditions as given in Proposition 8.13 can be of interest for practical applications.

Proposition 8.14. *Let $(\Omega,\mathcal{F},\mu)$ and $f\colon(a,b)\times\Omega\to\mathbb{R}$ be as in Proposition 8.13. Suppose that*

(i) for any $u\in(a,b)$ $\omega\mapsto f(u,\omega)$ is $\mathcal{F}$ measurable and integrable with respect to μ;

(ii) for any $\omega\in\Omega$, the map $f_\omega(u)=f(u,\omega)$, $u\in(a,b)$ is differentiable (cf. Definition A.9) with derivative f'_ω that satisfies

$$|f'_\omega(u_0)|\le g(\omega),\quad u_0\in(a,b),$$

where $g\colon\Omega\to\mathbb{R}$ is nonnegative and $\mathcal{F}$ measurable and s.t. $\int_\Omega g(\omega)\mu(d\omega)<\infty$.

Then, the function $F(u)=\int_\Omega f(u,\omega)\mu(d\omega)$, $u\in(a,b)$, is differentiable with derivative

$$F'(u_0)=\int_\Omega(\partial f/\partial u)(u_0,\omega)\mu(d\omega),\quad u_0\in(a,b).$$

Proof. Clearly, item (ii) of Proposition 8.13 is s satisfied. We show that also (iii) holds. It is a consequence of the mean value theorem (cf. Proposition A.22). Let $u_0 \in (a, b)$. Then, for any $u \in (a, b)$ $(u \neq u_0)$, either $u > u_0$ or $u < u_0$. If $u > u_0$, we apply the mean value theorem and obtain $m_1 \in (u_0, u)$ s.t. for any $\omega \in \Omega$,

$$f_\omega(u) - f_\omega(u_0) = f'_\omega(m_1)(u - u_0).$$

If $u < u_0$, we obtain $m_2 \in (u, u_0)$ s.t.

$$f_\omega(u_0) - f_\omega(u) = f'_\omega(m_2)(u_0 - u).$$

In particular, for any $\omega \in \Omega$,

$$|f_\omega(u) - f_\omega(u_0)| \leq \max\{f'_\omega(m_1), f'_\omega(m_2)\}|u - u_0| \leq g(\omega)|u - u_0|,$$

and hence (iii) of Proposition 8.13 is satisfied. This completes the proof. □

8.4 Integration over measurable sets

Definition 8.6. *Let $(\Omega, \mathcal{F}, \mu)$ be a measure space and $f : \Omega \to \overline{\mathbb{R}}$ be a $\mathcal{F}$ measurable function. The integral of f over a set $A \in \mathcal{F}$ is defined as*

$$\int_A f(\omega)\mu(d\omega) = \int_\Omega (\mathbb{1}_A f)(\omega)\mu(d\omega).$$

Exercise 8.7. *Let $(\Omega, \mathcal{F}, \mu)$ be a measure space and $f : \Omega \to \overline{\mathbb{R}}$ be $\mathcal{F}$ measurable and either nonnegative or integrable. Show that if $\mu(A) = 0$, $A \in \mathcal{F}$, then $\int_A f(\omega)\mu(d\omega) = 0$.*

Exercise 8.8. *Let $(\Omega, \mathcal{F}, \mu)$ be a measure space and $F \in \mathcal{F}$. Show that*

$$\mathcal{F} \ni A \mapsto \mu_F(F \cap A),$$

is a measure on $\mathcal{F}$ and for any nonnegative and $\mathcal{F}$ measurable function $f : \Omega \to \overline{\mathbb{R}}$,

$$\int_\Omega f(\omega)\mu_F(d\omega) = \int_F f(\omega)\mu(d\omega).$$

Exercise 8.9. *Let $(\Omega, \mathcal{F}, \mu)$ be a measure space and $f : \Omega \to \overline{\mathbb{R}}$ be a $\mathcal{F}$ measurable function. Suppose that either f is nonnegative or integrable and let $\{A_i : i \in I\} \subset \mathcal{F}$ be disjoint, where $I \subset \mathbb{N}$. Show that*

$$\int_{\cup_{i\in I} A_i} f(\omega)\mu(d\omega) = \sum_{i \in I}\Big(\int_{A_i} f(\omega)\mu(d\omega)\Big).$$

Example 8.6. *Consider the measure space $(\mathbb{N}, \mathcal{P}(\mathbb{N}), \mu)$, where μ is the counting measure and $\mathcal{P}(\mathbb{N})$ is the power set of $\mathbb{N}$ (cf. Example 5.3). We have that for any nonnegative $\mathcal{P}(\mathbb{N})$ measurable function $f : \mathbb{N} \to \mathbb{R}$,*

$$\int_{\mathbb{N}} f(k)\mu(dk) = \sum_{k\in\mathbb{N}} f(k).$$

To see it, we define the sequence of functions

$$f_n(k) = (f\mathbb{1}_{\{0,\ldots,n\}})(k) = \begin{cases} f(k), & \text{if } 0 \leq k \leq n, \\ 0, & \text{otherwise.} \end{cases}$$

Then, $(f_n)_{n\in\mathbb{N}}$ is a sequence of nonnegative $\mathcal{P}(\mathbb{N})$ measurable functions. Further, given any $k \in \mathbb{N}$, $f_n(k) \uparrow f(k)$. By Proposition 8.2,

$$\int_{\mathbb{N}} f_n(k)\mu(dk) \uparrow \int_{\mathbb{N}} f(k)\mu(dk).$$

We write

$$\mathbb{N} = \{1\} \cup \ldots \cup \{n\} \cup (\mathbb{N} \setminus \{1, \ldots, n\}).$$

Therefore, using Exercise 8.9, we obtain,

$$\begin{aligned}
\int_{\mathbb{N}} f_n(k)\mu(dk) &= \int_{\{1\}} f_n(k)\mu(dk) + \cdots + \int_{\{n\}} f_n(k)\mu(dk) + \int_{\mathbb{N}\setminus\{1,\ldots,n\}} f_n(k)\mu(dk) \\
&= \int_{\mathbb{N}} f_n(k)\mathbb{1}_{\{1\}}(k)\mu(dk) + \cdots + \int_{\mathbb{N}} f_n(k)\mathbb{1}_{\{n\}}(k)\mu(dk) \\
&= f_n(1)\int_{\mathbb{N}} \mathbb{1}_{\{1\}}\mu(dk) + \cdots + f_n(n)\int_{\mathbb{N}} \mathbb{1}_{\{1\}}\mu(dk) \\
&= f_n(1)\mu(\{1\}) + \cdots f_n(n)\mu(\{n\}) \\
&= \sum_{k=1}^{n} f_n(k) = \sum_{k=1}^{n} f(k).
\end{aligned}$$

Hence, $\sum_{k\in\mathbb{N}} f(k) = \int_{\mathbb{N}} f(k)\mu(dk)$.

Example 8.7. *Let $f\colon \mathbb{N} \times \mathbb{N} \to \mathbb{R}$, $f(i,j) = a_{ij}$ with $a_{ij} \geq 0$ for any $(i,j) \in \mathbb{N} \times \mathbb{N}$. We consider the measure space $(\mathbb{N}, \mathcal{P}(\mathbb{N}), \mu)$, where μ is the counting measure and $\mathcal{P}(\mathbb{N})$ is the power set of $\mathbb{N}$ (cf. Example 5.3). Fix any $j \in \mathbb{N}$, and define the function $f_j(i) = f(i,j)$, $i \in \mathbb{N}$, i.e., $f_j\colon \mathbb{N} \to \mathbb{R}$ is a sequence of real numbers. Clearly, given any $j \in \mathbb{N}$, $f_j^{-1}(B) \in \mathcal{P}(\mathbb{N})$, $B \in \mathfrak{B}(\mathbb{R})$. That is, f_j, $j \in \mathbb{N}$, is a sequence of nonnegative $\mathcal{P}(\mathbb{N})$ measurable functions. Let $S(i) = \sum_{j\in\mathbb{N}} f_j(i)$, $i \in \mathbb{N}$. Notice that as the limit of a sequence of measurable functions $(\sum_{j=1}^{n} f_j(i))$, S is $\mathcal{P}(\mathbb{N})$ measurable (cf. Proposition 7.9). Using Example 8.6, we know that*

$$\int_{\mathbb{N}} S(i)\mu(di) = \sum_{i\in\mathbb{N}} S(i) = \sum_{i\in\mathbb{N}}\left(\sum_{j\in\mathbb{N}} f_j(i)\right) = \sum_{i\in\mathbb{N}}\left(\sum_{j\in\mathbb{N}} a_{ij}\right).$$

Also, by Proposition 8.4, using Example 8.6 again,

$$\int_{\mathbb{N}} S(i)\mu(di) = \sum_{j\in\mathbb{N}}\left(\int_{\mathbb{N}} f_j(i)\mu(di)\right) = \sum_{j\in\mathbb{N}}\left(\sum_{i\in\mathbb{N}} f_j(i)\right) = \sum_{j\in\mathbb{N}}\left(\sum_{i\in\mathbb{N}} a_{ij}\right).$$

Which shows that

$$\sum_{i\in\mathbb{N}}\left(\sum_{j\in\mathbb{N}} a_{ij}\right) = \sum_{j\in\mathbb{N}}\left(\sum_{i\in\mathbb{N}} a_{ij}\right).$$

Example 8.8. *Let $(\Omega, \mathcal{F}, \mu)$ be a measure space and $f\colon \Omega \to \overline{\mathbb{R}}$ be $\mathcal{F}$ measurable and integrable. Suppose that I is a countable set and $\mu = \sum_{i\in I} \mu_i$, where μ_i is a collection of measures on $\mathcal{F}$. Then,*

$$\int_{\Omega} f(\omega)\mu(d\omega) = \sum_{i\in I}\left(\int_{\Omega} f(\omega)\mu_i(d\omega)\right).$$

To see it, suppose first that f is nonnegative. If $f(\omega) = \mathbb{1}_A(\omega)$, $A \in \mathcal{F}$, then

$$\int_{\Omega} f(\omega)\mu(d\omega) = \mu(A) = \sum_{i\in I} \mu_i(A) = \sum_{i\in I}\left(\int_{\Omega} f(\omega)\mu_i(d\omega)\right).$$

Now suppose that $f = \sum_{k=1}^{N} \alpha_k \mathbb{1}_{A_k}$ *is a nonnegative simple function (*f *is assumed to be in standard from, cf. Proposition 7.7). By Proposition 8.3, we obtain:*

$$\int_\Omega f(\omega)\mu(d\omega) = \int_\Omega \Big(\sum_{k=1}^{N} \alpha_k \mathbb{1}_{A_k}(\omega)\Big)\mu(d\omega) = \sum_{k=1}^{N} \alpha_k \Big(\int_\Omega \mathbb{1}_{A_k}\mu(d\omega)\Big)$$
$$= \sum_{k=1}^{N} \alpha_k \Big(\sum_{i\in I} \mu_i(A_k)\Big) = \sum_{i\in I}\Big(\sum_{k=1}^{N} \alpha_k \mu_i(A_k)\Big) = \sum_{i\in I}\Big(\int_\Omega f(\omega)\mu_i(d\omega)\Big).$$

Notice that we are allowed to interchange the order of summation by Proposition 3.11. If f *is* $\mathcal{F}$ *measurable and nonnegative function, we rely on Proposition 7.10 and find a sequence of nonnegative standard simple functions* $(f_n)_{n\in\mathbb{N}}$ *s.t. for any* $\omega \in \Omega$, $f_n(\omega) \uparrow f(\omega)$. *Using the previous case, we deduce that*

$$\int_\Omega f(\omega)\mu(d\omega) = \lim_{n\to\infty}\Big(\int_\Omega f_n(\omega)\mu(d\omega)\Big) = \lim_{n\to\infty}\Big(\sum_{i\in I}\Big(\int_\Omega f_n(\omega)\mu_i(d\omega)\Big)\Big).$$

Given any $n \in \mathbb{N}$, *write* $f_n = \sum_{k=1}^{N_n} \alpha_{k_n} \mathbb{1}_{A_{k_n}}$. *Then, on the right-hand side of the latter display it reads*

$$\lim_{n\to\infty}\Big(\sum_{i\in I}\Big(\sum_{k=1}^{N_n} \alpha_{k_n}\mu_i(A_{k_n})\Big)\Big)$$

Set $g_n^i = \sum_{k=1}^{N_n} \alpha_{k_n}\mu_i(A_{k_n})$, $i \in I$, $n \in \mathbb{N}$. *By Proposition 8.2, we know that for any* $i \in I$, $g_n^i \uparrow \int_\Omega f(\omega)\mu_i(d\omega)$. *Suppose that* I *is finite and for any* $i \in I$, $\int_\Omega f(\omega)\mu_i(d\omega) < \infty$. *Then, it follows from Proposition 3.4, that*

$$\lim_{n\to\infty}\Big(\sum_{i\in I} g_n^i\Big) = \sum_{i\in I}\Big(\int_\Omega f(\omega)\mu_i(d\omega)\Big), \tag{14}$$

and the result follows. If there exists $i \in I$, *s.t.* $g_n^i \uparrow \int_\Omega f(\omega)\mu_i(d\omega) = \infty$, *then, both sides of (14) are equal to* ∞ *and the result remains true. Thus, if* I *is finite, the statement is verified. If* I *is countably infinite,* $\#I = \#\mathbb{N}$, *and upon Example 8.7, we write* $\sum_{m=1}^{\infty} g_n^m = \int_{\mathbb{N}} g_n^m c(dm)$, *where* c *is the counting measure on* $\mathcal{P}(\mathbb{N})$. *Therefore, as a consequence of Proposition 8.2, we obtain*

$$\lim_{n\to\infty}\Big(\sum_{m=1}^{\infty} g_n^m\Big) = \lim_{n\to\infty}\Big(\int_{\mathbb{N}} g_n^m c(dm)\Big) = \int_{\mathbb{N}}\Big(\int_\Omega f(\omega)\mu_m(d\omega)\Big)c(dm)$$
$$= \sum_{m=1}^{\infty}\Big(\int_\Omega f(\omega)\mu_m(d\omega)\Big) = \sum_{i\in I}\Big(\int_\Omega f(\omega)\mu_i(d\omega)\Big).$$

For the remaining case, we write $f = f^+ - f^-$ *and apply the result to* f^+ *and* f^-.

8.5 Solution to exercises

Solution 8.1 (Solution to Exercise 8.1)**.** *This follows from the fact that* $S_\mu^f(\xi) \le S_\mu^g(\xi)$ *for any* $\xi \in Z_0^{\mathcal{F}}$ *(cf. Definition 1.11).*

Solution 8.2 (Solution to Exercise 8.2)**.** *Let* $A = \{\omega\colon f_n(\omega) \uparrow f(\omega)\}$. *By assumption,* $\mu(A^c) = 0$. *Define* $f_n^*(\omega) = f_n(\omega)\mathbb{1}_A(\omega)$, $f^*(\omega) = f(\omega)\mathbb{1}_A(\omega)$, $\omega \in \Omega$. *Then, for any* $\omega \in \Omega$, $f_n^*(\omega) \uparrow f^*(\omega)$. *Hence, using Proposition 8.2, it follows that* $\int_\Omega f_n^*(\omega)\mu(d\omega) \uparrow \int_\Omega f^*(\omega)\mu(d\omega)$. *Further, for any* $n \in \mathbb{N}$, $f_n = f_n^*$ μ *a.e. since* $\{\omega\colon f_n \ne f_n^*\} \subset A^c$. *Similarly,* $f = f^*$ μ *a.e. Therefore, using item (iv) of Proposition 8.5,*

$$\int_\Omega f_n(\omega)\mu(d\omega) = \int_\Omega f_n^*(\omega)\mu(d\omega) \uparrow \int_\Omega f^*(\omega)\mu(d\omega) = \int_\Omega f(\omega)\mu(d\omega).$$

Solution 8.3 (Solution to Exercise 8.3). *Since f is integrable, $\int_\Omega f^+(\omega)\mu(d\omega) < \infty$ and $\int_\Omega f^-(\omega)\mu(d\omega) < \infty$. Notice that by definition of f^+ and f^-, $\{\omega\colon f^+(\omega) = g^+(\omega)\} \cap \{\omega\colon f^-(\omega) = g^-(\omega)\} = \{\omega\colon f(\omega) = g(\omega)\}$. So that $\{\omega\colon f(\omega) \neq g(\omega)\} = \{\omega\colon f^+(\omega) \neq g^+(\omega)\} \cup \{\omega\colon f^-(\omega) \neq g^-(\omega)\}$. Hence, $f^+ = g^+$ and $f^- = g^-$ μ a.e. Thus, the result follows from item (iv) of Proposition 8.5.*

Solution 8.4 (Solution to Exercise 8.4).

(a) First of all, $\omega \mapsto |f(\omega)|$ and $\omega \mapsto |g(\omega)|$ are nonnegative and $\mathcal{F}$ measurable. Using item (iii) of Proposition 8.5, it follows that

$$\int_\Omega |f(\omega)(\omega)|\mu(d\omega) \leq \int_\Omega |g(\omega)(\omega)|\mu(d\omega),$$

where the latter integral is finite by Proposition 8.6. Therefore, using Proposition 8.6 again, f is integrable.

(b) We recall that f bounded on Ω means that there exists $0 \leq M < \infty$ s.t. $|f(\omega)| \leq M$ for any $\omega \in \Omega$ (cf. Definition 2.20). Then, using the result of Exercise 8.1, we obtain

$$\int_\Omega |f(\omega)(\omega)|\mu(d\omega) \leq M \int_\Omega \mu(d\omega) = M\mu(\Omega) < \infty.$$

Solution 8.5 (Solution to Exercise 8.5). *Given any $\omega \in \Omega$,*

$$|f(\omega)| \leq |\max\{\alpha_i\colon i = 1, \dots, N\}|.$$

Hence, the function $\mathbb{1}_A f$ is integrable since $\mu(A) < \infty$. Then, we use Proposition 8.8, and obtain

$$\begin{aligned}
\int_A f(\omega)\mu(d\omega) &= \sum_{i=1}^n \alpha_i \Big(\int_\Omega \mathbb{1}_A \mathbb{1}_{A_i}(\omega)\mu(d\omega)\Big) \\
&= \sum_{i=1}^n \alpha_i \Big(\int_\Omega \mathbb{1}_{A_i}(\omega)\mu(d\omega)\Big) \\
&= \sum_{i=1}^n \alpha_i \mu(A_i),
\end{aligned}$$

where we used that $A_i \subset A$ and $\mathbb{1}_{A_i}$ is nonnegative for any $i = 1, \dots, N$.

Solution 8.6 (Solution to Exercise 8.6). *Let $h\colon \Omega \to \overline{\mathbb{R}}$ be any $\mathcal{F}$ measurable and integrable function. Given any $\omega \in \Omega$, $-|h(\omega)| \leq h(\omega) \leq |h(\omega)|$. Hence, by Proposition 8.7, we obtain $\int_\Omega h(\omega)\mu(d\omega) \leq \int_\Omega |h(\omega)|\mu(d\omega)$ and $\int_\Omega -h(\omega)\mu(d\omega) \leq \int_\Omega |h(\omega)|\mu(d\omega)$. There are two cases, either $\int_\Omega h(\omega)\mu(d\omega) \geq 0$, then,*

$$\left| \int_\Omega h(\omega)\mu(d\omega)\right| = \int_\Omega h(\omega)\mu(d\omega) \leq \int_\Omega |h(\omega)|\mu(d\omega),$$

or $\int_\Omega h(\omega)\mu(d\omega) < 0$, then (cf. Proposition 8.3),

$$\left| \int_\Omega h(\omega)\mu(d\omega)\right| = -\int_\Omega h(\omega)\mu(d\omega) = \int_\Omega -h(\omega)\mu(d\omega) \leq \int_\Omega |h(\omega)|\mu(d\omega).$$

To conclude, we set $h = f - g$.

Solution 8.7 (Solution to Exercise 8.7). *Since $\{\omega\colon f\mathbb{1}_A \neq 0\} \subset A$, it follows that $f\mathbb{1}_A$ is zero a.e. Thus, if f is nonnegative we rely on item (iv) of Proposition 8.5 and conclude. If f is integrable, we use Exercise 8.3 and arrive at the same conclusion.*

Solution 8.8 (Solution to Exercise 8.8). *It is clear that μ_F is a measure on $\mathcal{F}$. Let $f = \mathbb{1}_A$, $A \in \mathcal{F}$. We have that*

$$\begin{aligned}\int_\Omega f(x)\mu_F(d\omega) &= \int_A \mu_F(d\omega)\\ &= \mu_F(A) = \mu(F\cap A) = \int_{F\cap A}\mu(d\omega) = \int_\Omega \mathbb{1}_{F\cap A}(x)\mu(d\omega)\\ &= \int_\Omega \mathbb{1}_F(\omega)\mathbb{1}_A(\omega)\mu(d\omega) = \int_F f(x)\mu(d\omega).\end{aligned}$$

If $f = \sum_{i=1}^N \alpha_i \mathbb{1}_{A_i}$ is a nonnegative simple function (f is assumed to be in standard from, cf. Proposition 7.7), then by Proposition 8.3,

$$\int_\Omega f(x)\mu_F(d\omega) = \sum_{i=1}^N \alpha_i\Big(\int_\Omega \mathbb{1}_{A_i}(\omega)\mu_F(d\omega)\Big) = \sum_{i=1}^N \alpha_i\Big(\int_F \mathbb{1}_{A_i}(\omega)\mu(d\omega)\Big) = \int_F f(x)\mu(d\omega).$$

Let f be any $\mathcal{F}$ measurable and nonnegative function. We rely on Proposition 7.10 and find a sequence of nonnegative standard simple functions $(f_n)_{n\in\mathbb{N}}$ s.t. for any $\omega\in\Omega$, $f_n(\omega)\uparrow f(\omega)$. By Proposition 8.2, it follows from the previous case that

$$\int_\Omega f(x)\mu_F(d\omega) = \lim_{n\to\infty}\Big(\int_\Omega f_n(x)\mu_F(d\omega)\Big) = \lim_{n\to\infty}\int_F f_n(x)\mu(d\omega) = \int_F f(x)\mu(d\omega).$$

Solution 8.9 (Solution to Exercise 8.9). *We notice first that if I is finite, then the result follows from Proposition 8.8. Hence we assume that $I=\mathbb{N}$. Suppose that f is nonnegative. Using Definition 8.6, we have that*

$$\int_{\cup_{i\in\mathbb{N}}A_i} f(\omega)\mu(d\omega) = \int_\Omega \mathbb{1}_{\cup_{i\in\mathbb{N}}A_i}(\omega)f(\omega)\mu(d\omega) = \int_\Omega\Big(\sum_{i\in\mathbb{N}}\mathbb{1}_{A_i}(\omega)f(\omega)\Big)\mu(d\omega),$$

since $\{A_i : i\in\mathbb{N}\}$ is disjoint. Then, we use Proposition 8.4 and obtain

$$\int_{\cup_{i\in\mathbb{N}}A_i} f(\omega)\mu(d\omega) = \sum_{i\in\mathbb{N}}\Big(\int_\Omega \mathbb{1}_{A_i}(\omega)f(\omega)\mu(d\omega)\Big) = \sum_{i\in\mathbb{N}}\Big(\int_{A_i} f(\omega)\mu(d\omega)\Big).$$

Suppose now that f is integrable. Define $f_i = f\mathbb{1}_{A_i}$, $i\in\mathbb{N}$. We have that for any $\omega\in\Omega$, $\lim_{n\to\infty}\sum_{i=1}^n f_i(\omega) = f\mathbb{1}_{\cup_{i\in\mathbb{N}}A_i}(\omega)$. In particular,

$$\Big|\sum_{i=1}^n f_i(\omega)\Big| \le |f|\mathbb{1}_{\cup_{i\in\mathbb{N}}A_i}(\omega) \le \big|f(\omega)\big|.$$

Therefore, by Proposition 8.11,

$$\int_{\cup_{i\in\mathbb{N}}A_i} f(\omega)\mu(d\omega) = \sum_{i\in\mathbb{N}}\Big(\int_{A_i} f(\omega)\mu(d\omega)\Big).$$

8.6 Additional exercises

Exercise 8.10. *Let $(\Omega,\mathcal{F})$ be a measurable space and $x\in\Omega$ be given. Define the measure*

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

on $\mathcal{F}$ (cf. Example 5.1). Let $f\colon\Omega\to\overline{\mathbb{R}}$ be nonnegative and $\mathcal{F}$ measurable. Show that

$$\int_\Omega f(\omega)\delta_x(d\omega) = f(x).$$

Hint: *Calculate* $\int_\Omega f(\omega)\delta_x(d\omega)$ *if* f *is a nonnegative simple function in standard form and rely on Propositions 7.10 and 8.2.*
Extension: *Using positive and negative parts of* f*, it can be shown that it is enough to demand that* $f\colon \Omega \to \overline{\mathbb{R}}$ *is* $\mathcal{F}$ *measurable to obtain* $\int_\Omega f(\omega)\delta_x(d\omega) = f(x)$*.*

Exercise 8.11. *Calculate the integral in each of the following cases:*

(a) $\int_\mathbb{R} \big(\sum_{n\in\mathbb{N}} \mathbb{1}_{(0,1/2^n]}(x)\big)\lambda(dx)$*, where* λ *is the Lebesgue measure on* $\mathfrak{B}(\mathbb{R})$*;*

(b) $\int_\mathbb{N} x^2\lambda(dx)$*, where* λ *is the Lebesgue measure on* $\mathfrak{B}(\mathbb{R})$*;*

(c) $\int_\mathbb{N} 1/n\,\mu(dn)$*, where* μ *is the counting measure on the power set of* $\mathbb{N}$*;*

(d) $\int_\mathbb{R} y^2\mu(dy)$*, where* μ *is the measure on* $\mathfrak{B}(\mathbb{R})$ *given by (cf. Example 5.4),*

$$\mu(B) = \sum_{x=0}^{N} p_x\delta_x(B), \quad \delta_x(B) = \begin{cases} 1, & \text{if } x \in B, \\ 0, & \text{if } x \notin B, \end{cases}$$

with $N \in \mathbb{N}$*, and* $0 \le p_x \le 1$ *for any* $x \in \{1, \dots, N\}$*.*

Exercise 8.12. *Let* $(\Omega, \mathcal{F}, \mu)$ *be a measure space and* $f_n\colon \Omega \to \overline{\mathbb{R}}$*,* $n \in \mathbb{N}$*, be a sequence of* $\mathcal{F}$ *measurable and integrable functions s.t.* $\sup_{n\in\mathbb{N}} \int_\Omega f_n(\omega)\mu(d\omega) < \infty$*. Show that if for any* $\omega \in \Omega$*,* $f_n(\omega) \uparrow f(\omega)$*, then* f *is integrable and*

$$\int_\Omega f_n(\omega)\mu(d\omega) \uparrow \int_\Omega f(\omega)\mu(d\omega).$$

Hint: *Consider the sequence of nonnegative and* $\mathcal{F}$ *measurable functions* $f_n - f_1$*.*

Exercise 8.13. *Let* $(\Omega, \mathcal{F}, \mu)$ *be a measure space. Assume that* $\mathcal{M}$ *is another* σ*-field on* Ω *and* $\mu|_\mathcal{M}$ *is the restriction of* μ *to* $\mathcal{M}$*. Suppose that* $f\colon \Omega \to \overline{\mathbb{R}}$ *is* $\mathcal{M}$ *measurable. Show that if either* f *is nonnegative or integrable with respect to* μ*, then,*

$$\int_\Omega f(\omega)\mu(d\omega) = \int_\Omega f(\omega)\mu|_\mathcal{M}(d\omega).$$

Exercise 8.14. *Consider* $(\mathbb{R}, \mathfrak{B}(\mathbb{R}), \lambda)$*, where* λ *is the Lebesgue measure on* $\mathfrak{B}(\mathbb{R})$*. Let* $f(x) = 0$ *for any* $x \in \mathbb{R}$*. Find a sequence of* $\mathfrak{B}(\mathbb{R})$ *measurable functions s.t.* $f_n(x) \xrightarrow{n\to\infty} f(x)$ *for any* $x \in \mathbb{R}$ *and* $\int_\mathbb{R} f_n(x)\lambda(dx) \xrightarrow{n\to\infty} \infty$*.*

A Results from analysis

A.1 On infima and suprema: arithmetic set operations

We derive further results on the infimum and supremum of subsets of the real numbers. With regard to arithmetics on sets, we make the following definition:

Definition A.1. *Let $\emptyset \neq A \subset \mathbb{R}$. Then,*

(1.) $-A = \{-x \colon x \in A\}$;

(2.) $cA = \{ca \colon a \in A\}$, $c \in \mathbb{R}$;

(3.) $A + B = \{a + b \colon a \in A, b \in B\}$.

Proposition A.1. *Let $\emptyset \neq A \subset \mathbb{R}$. be bounded from above.*

(i) If A is bounded from above, then $-A$ is bounded from below and $\inf -A = -\sup A$;

(ii) if A is bounded from below, then $-A$ is bounded from above and $\sup -A = -\inf A$.

Proof. We only verify item (i) (the argument for item (ii) is similar). We notice first that since A is bounded from above, we rely on Proposition 1.5 and know that $\sup A \in \mathbb{R}$. Hence, since $\sup A$ is an upper bound for A, it follows that for any $x \in A$, $x \leq \sup A$. In particular, for any $x \in A$. $-x \geq -\sup A$. Thus, $-\sup A$ is a lower bound of $-A$. By Proposition 1.5, the infimum of $-A$ exists as an element of the real numbers. We thus write $L = \inf -A$. Then, we show that $L = -\sup A$. Clearly, since L is the infimum of $-A$ and $-\sup A$ is a lower bound of $-A$ it follows that $L \geq -\sup A$ (L is the greatest lower bound). On the other hand, given any $x \in A$, $-x \geq L$, it follows that $x \leq -L$ and hence $-L$ is an upper bound for A. This leads to $\sup A \leq -L$ and hence $-\sup A \geq L$. □

Proposition A.2. *Let $\emptyset \neq A \subset \mathbb{R}$ and $c > 0$.*

(i) If $\sup(cA) < \infty$, then $\sup A < \infty$ and $\sup(cA) = c \sup A$;

(ii) if $\inf(cA) > -\infty$, then $\inf A > -\infty$ and $\inf(cA) = c \inf(A)$.

Proof. Given any $a \in A$, $ca \leq \sup(cA)$. Hence, since $c > 0$, $a \leq \sup(cA)/c$. Thus, since by assumption $\sup(cA) < \infty$ and $a \in A$ was arbitrary, it follows that $\sup A \leq \sup(cA)/c < \infty$. Notice that the latter inequality shows that $c \sup A \leq \sup(cA)$. Also, since $c > 0$, $ca \leq c \sup A$ for any $a \in A$, hence, $\sup(cA) \leq c \sup A$ and we conclude that $\sup(cA) = c \sup A$. This shows (i). The argument for (ii) is similar. □

Proposition A.3. *Let $A, B \subset \mathbb{R}$, A and B not empty.*

(i) If $\sup(A + B) < \infty$, then $\sup A < \infty$, $\sup B < \infty$ and $\sup(A + B) = \sup A + \sup B$;

(ii) If $\inf(A + B) > -\infty$, then $\inf A > -\infty$, $\inf B > -\infty$ and $\inf(A + B) = \inf A + \inf B$.

Proof. We show item (i). Notice that since $\sup(A + B) < \infty$, it follows that for any $a \in A$, $a + b \leq \sup(A + B)$ and hence $a \leq \sup(A + B) - b$. Therefore, since $a \in A$ was arbitrary, $\sup A \leq \sup(A + B) - b < \infty$ and we conclude that $\sup A < \infty$. A similar argument shows that $\sup B < \infty$. We readily see that $\sup(A + B) \leq \sup A + \sup B$. Now either $\sup(A + B) < \sup A + \sup B$ or $\sup(A + B) = \sup A + \sup B$. We show that if we assume that $\sup(A + B) < \sup A + \sup B$, then we arrive at a contradiction. Hence, suppose that $\sup(A+B) < \sup A+\sup B$. Let $\delta = \sup A+\sup B-\sup(A+B) > 0$. By Proposition 1.10, there exists $a \in A$ s.t. $a > \sup A - (\delta/2)$. Similarly, there exists $b \in B$ s.t. $b > \sup B - (\delta/2)$. Hence,

$$a + b > \sup A + \sup B - \delta = \sup(A + B),$$

which is a contradiction. Therefore, $\sup(A + B) = \sup A + \sup B$. The argument for (ii) is similar. □

A.2 On limit inferior and limit superior

We derive further properties of the limit inferior and limit superior.

Proposition A.4. *Let $(b_n)_{n\in\mathbb{N}}$ be a real valued sequence s.t. $\lim_{n\to\infty} b_n = b$. Then,*

(i) $\liminf_{n\to\infty}(b_n + a_n) = b + \liminf_{n\to\infty} a_n$ *and* $\liminf_{n\to\infty}(-a_n) = -\limsup_{n\to\infty} a_n$*;*

(ii) $\limsup_{n\to\infty}(b_n + a_n) = b + \limsup_{n\to\infty} a_n$ *and* $\limsup_{n\to\infty}(-a_n) = -\liminf_{n\to\infty} a_n$*.*

Proof. We only show (i), since the arguments for (ii) are similar. Assume first that $(a_n)_{n\in\mathbb{N}}$ is bounded from below. We have that for any $n \in \mathbb{N}$ and any given $j \geq n$, $\inf_{k\geq n} b_k + \inf_{k\geq n} a_k \leq b_j + a_j$. Therefore, $\inf_{k\geq n} b_k + \inf_{k\geq n} a_k \leq \inf_{k\geq n}(b_k + a_k)$. Hence,

$$b + \liminf_{n\to\infty} a_n = \lim_{n\to\infty} (\inf_{k\geq n} b_k + \inf_{k\geq n} a_k) \leq \lim_{n\to\infty} \inf_{k\geq n} (b_k + a_k) = \liminf_{n\to\infty}(b_n + a_n).$$

If $(a_n)_{n\in\mathbb{N}}$ is not bounded from below, then by definition, $-\infty = b + \liminf_{n\to\infty} a_n = \liminf_{n\to\infty}(b_n + a_n)$ (where we used the convention that $-\infty + x = -\infty$ for $x \in \mathbb{R}$). For the other inequality we notice that by the previous inequality,

$$\liminf_{n\to\infty} a_n = \liminf_{n\to\infty}(b_n + a_n - b_n) \geq \liminf_{n\to\infty}(b_n + a_n) + \liminf_{n\to\infty}(-b_n) = \liminf_{n\to\infty}(b_n + a_n) - b.$$

Thus also $\liminf_{n\to\infty}(b_n + a_n) \leq b + \liminf_{n\to\infty} a_n$. Let us show that $\liminf_{n\to\infty}(-a_n) = -\limsup_{n\to\infty} a_n$. We assume first that $(a_n)_{n\in\mathbb{N}}$ is bounded from above. By Proposition A.1, we know that $\{-a_n : n \in \mathbb{N}\}$ is bounded from below and $\inf_{k\geq n}(-a_k) = -\sup_{k\geq n} a_k$. Hence,

$$\liminf_{n\to\infty}(-a_n) = \lim_{n\to\infty} \inf_{k\geq n}(-a_k) = -\lim_{n\to\infty} \sup_{k\geq n} a_k = -\limsup_{n\to\infty} a_n.$$

If $(a_n)_{n\in\mathbb{N}}$ is not bounded from above, then $\{-a_n : n \in \mathbb{N}\}$ is not bounded from below and by definition of the limit inferior and limit superior $\liminf_{n\to\infty}(-a_n) = -\limsup_{n\to\infty} a_n = -\infty$. □

Proposition A.5. *Let $f_n : A \to \overline{\mathbb{R}}$, $n \in \mathbb{N}$, be a sequence of functions. Define the sequences of functions $g_n = \inf_{k\geq n} f_k$ and $h_n = \sup_{k\geq n} f_k$, $n \in \mathbb{N}$. Then,*

(i) for any $x \in A$, $(g_n(x))_{n\in\mathbb{N}}$ is increasing and $g_n(x) \uparrow \liminf_{n\to\infty} f_n(x)$;

(ii) for any $x \in A$, $(h_n(x))_{n\in\mathbb{N}}$ is decreasing and $h_n(x) \downarrow \limsup_{n\to\infty} f_n(x)$.

Proof.

(i) Let $x \in A$. Suppose that $(g_n(x))$ is bounded from below. Then, using Exercise 3.7 and Proposition 3.18, $(g_n(x))$ is increasing with limit $\liminf_{n\to\infty} f_n(x)$. If $(g_n(x))$ is not bounded from below, still, for any $n \in \mathbb{N}$, $g_{n+1}(x) = \inf_{k\geq n+1} f_k(x) \geq \inf_{k\geq n} f_k(x) = g_n(x)$, i.e., $(g_n(x))$ is increasing. Suppose that there exists $N \in \mathbb{N}$ s.t. $(g_N(x)) > -\infty$. Then, we rely on the sequence $g^*_n(x) = g_{n-1+N}(x)$, $n \in \mathbb{N}$, and repeat the arguments given in the proof of Proposition 3.18 to show that $\lim_{n\to\infty} g^*_n(x) = \sup_{n\in\mathbb{N}} g^*_n(x)$. Therefore, $\lim_{n\to\infty} g_n(x) = \sup_{n\in\mathbb{N}} g^*_n(x)$. Since $\sup_{n\in\mathbb{N}} g_n(x) \leq \sup_{n\in\mathbb{N}} g^*_n(x)$, this shows that $\lim_{n\to\infty} g_n(x) \geq \sup_{n\in\mathbb{N}} g_n(x)$. The other inequality, $\lim_{n\to\infty} g_n(x) \leq \sup_{n\in\mathbb{N}} g_n(x)$, is obtained from the arguments already given in the proof of Proposition 3.18. Finally, if for any $n \in \mathbb{N}$, $g_n(x) = -\infty$, $\lim_{n\to\infty} g_n(x) = -\infty = \liminf_{n\to\infty} f_n(x)$ (cf. Definition 3.10).

(ii) We may repeat a similar argument as in (i).

□

A.3 On convergent sequences

The following is known as the subsequence criterion for convergent sequences.

Proposition A.6. *Let $(a_n)_{n\in\mathbb{N}}$ be a real-valued sequence. Consider the follwong assumption:*

(A) for any subsequence $(a_{s(n)})_{n\in\mathbb{N}}$ of $(a_n)_{n\in\mathbb{N}}$ there exists a subsequence $(a_{t(s(n))})_{n\in\mathbb{N}}$ of $(a_{s(n)})_{n\in\mathbb{N}}$ s.t. $a_{t(s(n))} \xrightarrow{n\to\infty} a$.

Then, if (A) holds, $a_n \xrightarrow{n\to\infty} a$.

Before we prove the latter proposition, we deduce the following:

Proposition A.7. *Let $(a_n)_{n\in\mathbb{N}}$ be a real-valued sequence s.t. for any $n \in \mathbb{N}$, $|a_n| > n$. Then, $(a_n)_{n\in\mathbb{N}}$ has no convergent subsequence.*

Proof. This follows from the fact that any subsequence $(a_{s(n)})_{n\in\mathbb{N}}$ of $(a_n)_{n\in\mathbb{N}}$ is not bounded. Notice that if $(a_{s(n)})_{n\in\mathbb{N}}$ is a subsequence of $(a_n)_{n\in\mathbb{N}}$, then for any $n \in \mathbb{N}$, $|a_{s(n)}| > s(n)$. Since $s(n) < s(n+1)$ for any $n \in \mathbb{N}$ the set $\{s(n) \colon n \in \mathbb{N}\}$ is countably infinite and hence not bounded. Thus, $\sup\{s(n) \colon n \in \mathbb{N}\} = \infty$. Therefore, for any $M \in \mathbb{R}$, there exists $n \in \mathbb{N}$ s.t. $s(n) > M$. Since $|a_{s(n)}| > s(n)$, it follows in particular that $|a_{s(n)}| > M$. Therefore, $(a_{s(n)})_{n\in\mathbb{N}}$ is not bounded (cf. Definition 3.3). □

Proof of Proposition A.6. First, we prove that (A) implies that $(a_n)_{n\in\mathbb{N}}$ is bounded. To see it, suppose by contradiction that $(a_n)_{n\in\mathbb{N}}$ is not bounded. Then, find $n_1 \in \mathbb{N}$ s.t. $1 < |a_{n_1}| \le 1 + N_1$ for some $N_1 \in \mathbb{N} \setminus \{1\}$. Since $(a_n)_{n\in\mathbb{N}}$ is not bounded, there exists $n_2 \in \mathbb{N}$ s.t. $2 < 1 + N_1 < |a_{n_2}| < 1 + N_1 + N_2$ for some $N_2 \in \mathbb{N} \setminus \{1\}$. We continue like this and obtain a subsequence $(b_k)_{k\in\mathbb{N}} = (a_{s(k)})_{k\in\mathbb{N}}$, $s(k) = n_k$, $k \in \mathbb{N}$ where for any $k \in \mathbb{N}$, $|b_k| > k$. By Proposition A.7, $(b_k)_{k\in\mathbb{N}}$ can not have any convergent subsequence. This contradicts (A). Hence, (A) implies that $(a_n)_{n\in\mathbb{N}}$ is bounded. We verify that (A) implies that $\lim_{n\to\infty} a_n = a$. Again, we consider an argument by contradiction and suppose that (A) is true but $\lim_{n\to\infty} a_n \neq a$. If $\lim_{n\to\infty} a_n \neq a$, (A) implies that $(a_n)_{n\in\mathbb{N}}$ can not be convergent, since if $(a_n)_{n\in\mathbb{N}}$ was convergent with limit L, then any subsequence of $(a_n)_{n\in\mathbb{N}}$ must converge to the same limit L and then (A) implies that $L = \lim_{n\to\infty} a_n = a$. In summary, under the assumption that (A) holds and $\lim_{n\to\infty} a_n \neq a$, $(a_n)_{n\in\mathbb{N}}$ is bounded but does not converge. Using Propositions 3.20 and 3.21, this implies that $m = \liminf_{n\to\infty} a_n < \limsup_{n\to\infty} a_n = M$, where $m = \lim_{n\to\infty} m_n$ and $M = \lim_{n\to\infty} M_n$ with $m_n = \inf\{a_k \colon k \ge n\}$ and $M_n = \sup\{a_k \colon k \ge n\}$, respectively (cf. Propositions 3.18 and 3.19). Thus, if we write $a_{s(n)} = m_n$, $n \in \mathbb{N}$, and $a_{\tilde{s}(n)} = M_n$, $n \in \mathbb{N}$, we identify two subsequences $(a_{s(n)})_{n\in\mathbb{N}}$ and $(a_{\tilde{s}(n)})_{n\in\mathbb{N}}$ of $(a_n)_{n\in\mathbb{N}}$ with limit m and M, respectively. But then, for any subsequence $(a_{t(s(n))})_{n\in\mathbb{N}}$ and $(a_{\tilde{t}(\tilde{s}(n))})_{n\in\mathbb{N}}$ of $(a_{s(n)})_{n\in\mathbb{N}}$ and $(a_{\tilde{s}(n)})_{n\in\mathbb{N}}$, respectively, it follows that $\lim_{n\to\infty} a_{t(s(n))} = m$ and $\lim_{n\to\infty} a_{\tilde{t}(\tilde{s}(n))} = M$. Since $m \neq M$, (A) is violated and hence we arrive at a contradiction. We conclude that $\lim_{n\to\infty} a_n = a$. □

A.4 Continuity

For this section, we use the notation $\|\cdot\|_m$ and $\|\cdot\|_k$ for the Euclidean distance on $\mathbb{R}^m$ and $\mathbb{R}^k$, respectively (cf. Definition 2.14).

Definition A.2. *Let $f \colon E \to \mathbb{R}^k$, $E \subset \mathbb{R}^m$. Then, f is continuous at a point $x \in E$, if for any $\varepsilon > 0$, there exists $\delta > 0$ s.t. for any $y \in E$, if $\|x - y\|_m < \delta$, then $\|f(x) - f(y)\|_k < \varepsilon$.*

Definition A.3. *A function $f \colon E \to \mathbb{R}^k$, $E \subset \mathbb{R}^m$ is referred to as continuous (or continuous on E) if it is continuous at any point $x \in E$.*

Example A.1. *Let $m = k = 1$ and consider $f(x) = x$, $x \in \mathbb{R}$. Then, clearly $f \colon \mathbb{R} \to \mathbb{R}$ is continuous. For any $\varepsilon > 0$, let $\delta = \varepsilon$, then $|x - y| < \delta$ implies that $|f(x) - f(y)| < \varepsilon$.*

Example A.2. *Let $c \in \mathbb{R}$ and $f(x) = c$ for any $x \in \mathbb{R}$. Then, clearly $f\colon \mathbb{R} \to \mathbb{R}$ is continuous since for any $\varepsilon > 0$, $|f(x) - f(y)| = 0 < \varepsilon$.*

The following is known as the intermediate value theorem, it supports our natural understanding of a continuous function defined on a closed interval.

Proposition A.8. *Let $f\colon [a,b] \to \mathbb{R}$ be continuous s.t. $f(a) \neq f(b)$. Suppose that γ is a value s.t.*

$$\gamma \in \big[\min\{f(a), f(b)\}, \max\{f(a), f(b)\}\big],$$

i.e., γ is between $f(a)$ and $f(b)$, then there exists $s \in [a,b]$ s.t. $f(s) = \gamma$. That is, any value γ between $f(a)$ and $f(b)$ is attained by f.

Proof. Let $\gamma \in \big[\min\{f(a), f(b)\}, \max\{f(a), f(b)\}\big]$. Assume that that $f(a) < f(b)$. Suppose that $\gamma \in (f(a), f(b))$, since if $\gamma = f(a)$ or $\gamma = f(b)$, the result follows. Define

$$A = \{x \in [a,b] \colon f(x) < \gamma\}.$$

A is not empty, since $f(a) < \gamma$. Clearly, since A is bounded, $s = \sup A < \infty$. Further, $s \in [a,b]$. There are three cases, $f(s) < \gamma$, $f(s) > \gamma$ or $f(s) = \gamma$. We show that only $f(s) = \gamma$ is possible. Suppose first that that $f(s) < \gamma$, i.e., $s \in A$. In particular, $s < b$ since $s = b$ is not possible ($s = b$ implies with $s \in A$ that $\gamma > f(b)$). Then, since f is continuous on $[a,b]$ and $s \in [a,b]$, f is continuous at s, i.e., there exists $\delta > 0$ s.t. for any $x \in [a,b]$, if $|x - s| < \delta$, then $|f(x) - f(s)| < \gamma - f(s)$ (recall that $\gamma - f(s) > 0$). In particular, $f(x) - f(s) \leq |f(x) - f(s)| < \gamma - f(s)$ and hence, if $x \in [a,b]$ s.t. $|x - s| < \delta$, we obtain $f(x) < \gamma$. Since $s < b$, we let $\varepsilon > 0$ s.t. $s + \varepsilon \leq b$. Set $x = s + \min\{(\delta/2), \varepsilon\}$. Then, $x \in [a,b]$ and $|x - s| < \delta$ and hence $f(x) < \gamma$. This is a contradiction with $s = \sup A$. Hence, the case $f(s) < \gamma$ is not possible. Assume now $f(s) > \gamma$. Using the same reasoning as before, there exists $\delta > 0$ s.t. for any $x \in [a,b]$, if $|x - s| < \delta$ then $f(x) > \gamma$. Then, since $s < \infty$, by Proposition 1.10, there exists $x \in A$ s.t. $x > s - \delta$. Since $s \geq x$, that particular x is s.t. $f(x) > \gamma$ and $f(x) < \gamma$, which is again a contradiction. Hence, we obtain $f(s) = \gamma$ and the proposition is proven. □

We prove the sequence criterion for continuous functions stated in Section 3.3.

Proof of Proposition 3.25. We first show that item (i) implies item (ii). Hence, let f be continuous at x. Take any $(x_n)_{n\in\mathbb{N}} \subset E$ which is s.t. $\lim_{n\to\infty} x_n = x$. Since f is continuous at x, for any $\varepsilon > 0$, there exists $\delta > 0$ s.t. $\|f(x) - f(y)\|_k < \varepsilon$ if $\|x - y\|_m < \delta$. Since $\lim_{n\to\infty} x_n = x$, there exists $N \in \mathbb{N}$, s.t. $\|x - x_n\|_m < \delta$ for any $n \geq N$. Therefore, $\|f(x) - f(x_n)\|_k < \varepsilon$ for any $n \geq N$. Since $\varepsilon > 0$ was arbitrary, the result follows. For the other direction, assume that item (ii) is true. Assume by contradiction that f is not continuous at x. Hence, there exists $\varepsilon > 0$ s.t. for any $\delta > 0$, there exists $y \in E$ with $\|x - y\|_m < \delta$ but $\|f(x) - f(y)\|_k \geq \varepsilon$. Let $\delta = 1/n$, $n \in \mathbb{N}$. For any $n \in \mathbb{N}$, let y_n be s.t. $\|x - y_n\|_m < 1/n$ and $\|f(x) - f(y_n)\|_k \geq \varepsilon$. Then $\lim_{n\to\infty} y_n = y$ but $\lim_{n\to\infty} f(y_n) \neq f(x)$. This contradicts item (ii), and hence the proof is complete. □

Proposition A.9. *Let $f = (f_1, \ldots, f_k)\colon E \to \mathbb{R}^k$, $E \subset \mathbb{R}^m$. Then, f is continuous at $x \in E$ if and only if f_i is continuous at x for any $i = 1, \ldots, k$.*

Proof. This follows from Propositions 3.25 and 3.24. □

Proof of Proposition 2.12. Let $f\colon \mathbb{R}^k \to \mathbb{R}$, $f(x) = \sum_{i=1}^k x_k$, $x = (x_1, \ldots, x_k)$. Assume that $(a_n)_{n\in\mathbb{N}} \in (\mathbb{R}^k)^{\mathbb{N}}$, $a_n = (a_1^n, \ldots, a_k^n)$, s.t. $a_n \xrightarrow{n\to\infty} x$ where $x \in \mathbb{R}^k$. Then, by Proposition 3.24, $\lim_{n\to\infty} a_i^n = x_i$ for any $i = 1, \ldots, k$. Hence, using Proposition 3.4,

$$f(a_n) = \sum_{i=1}^k a_i^n \xrightarrow{n\to\infty} \sum_{i=1}^k x_k = f(x).$$

Therefore, by Proposition 3.25, f is continuous at x. Since x was arbitrary, $f\colon \mathbb{R}^k \to \mathbb{R}$ is continuous. A similar argument shows that g und h are continuous. $\square$

Proposition A.10. *Let $f\colon A \to B$, $A \subset \mathbb{R}^m$ and $B \subset \mathbb{R}^k$. Further, let $g\colon B \to C$, $C \subset \mathbb{R}^l$. If f is continuous at $x \in A$ and g is continuous at $f(x)$, then $g(f)\colon A \to C$ is continuous at x.*

Proof. Let $(x_n)_{n\in\mathbb{N}} \in A^{\mathbb{N}}$ s.t. $\lim_{n\to\infty} x_n = x$. Since f is continuous at x, it follows that $\lim_{n\to\infty} f(x_n) = f(x)$. Then, since g is continuous at $f(x)$, $\lim_{n\to\infty} g(f)(x_n) = g(f)(x)$. $\square$

Example A.3. *Let $f_i\colon E \to \mathbb{R}$, $E \subset \mathbb{R}^m$, $i = 1, \dots, k$, be continuous. Then, $f = \sum_{i=1}^k f_i\colon E \to \mathbb{R}$ and $g = \prod_{i=1}^k f_i\colon E \to \mathbb{R}$ are continuous as well. To see it, let $x \in E$. We first use Proposition A.9 and conclude that $x \mapsto \psi(x) = (f_1(x), \dots, f_k(x))$ is continuous at x. Further, by Proposition 2.12, given any $y = (y_1, \dots, y_k) \in \mathbb{R}^k$, $y \mapsto s(y) = \sum_{i=1}^k y_i$ and $y \mapsto p(y) = \prod_{i=1}^k y_i$ are continuous at y. Therefore, using Proposition A.10, it follows that $f(x) = s(\psi)(x)$ and $g(x) = p(\psi)(x)$ are continuous at x. Since $x \in E$ was arbitrary, the result follows.*

Proposition A.11. *Let $f\colon \mathbb{R}^m \to \mathbb{R}^k$. The following are equivalent:*

(i) f is continuous;

(ii) for any open set $U \subset \mathbb{R}^k$, $f^{-1}(U)$ is open in $\mathbb{R}^m$.

Proof. To avoid confusion, we write $B_r^m(a)$ and $B_r^k(b)$ for an open ball with center $a \in \mathbb{R}^m$ and $b \in \mathbb{R}^k$, respectively. We first show that (i) implies (ii). Therefore, let $U \subset \mathbb{R}^k$ be an open set. Let $x \in f^{-1}(U)$, i.e., $f(x) \in U$. Since U is open, there exists $\varepsilon > 0$, s.t. $B_\varepsilon^k(f(x)) \subset U$. Since f is continuous at x, there exists $\delta > 0$ s.t. if $y \in B_\delta^m(x)$, then $f(y) \in B_\varepsilon^k(f(x))$. Hence, $f(B_\delta^m(x)) \subset B_\varepsilon^k(f(x))$. Hence, $B_\delta^m(x) \subset f^{-1}(B_\varepsilon^k(f(x))) \subset f^{-1}(U)$. Therefore, $f^{-1}(U)$ is open in $\mathbb{R}^m$. We show that (ii) implies (i). Let $x \in \mathbb{R}^m$, $\varepsilon > 0$ and consider the open set $B_\varepsilon^k(f(x))$. Then, by assumption, $f^{-1}(B_\varepsilon^k(f(x)))$ is open in $\mathbb{R}^m$. Clearly, $x \in f^{-1}(B_\varepsilon^k(f(x)))$. Hence, there exists $\delta > 0$, s.t. $B_\delta^m(x) \subset f^{-1}(B_\varepsilon^k(f(x)))$. That is, $f(B_\delta^m(x)) \subset B_\varepsilon^k(f(x))$. This implies that f is continuous at x. Since $x \in \mathbb{R}^m$ was arbitrary, the result follows. $\square$

In accordance with Definition 2.16, we can define open sets in E, where $E \subset \mathbb{R}^m$.

Definition A.4. *Let $E \subset \mathbb{R}^m$, $E \neq \emptyset$. The Euclidean distance $\|\cdot\|_m$ restricted to E is denoted with $\|\cdot\|_E$. Given $x \in E$, we write $B_r^E(x) = \{y \in E\colon \|y - x\|_E < r\}$ for an open ball in E of radius $r > 0$ with center x. A set $V \subset E$ is open if for any $x \in V$, there exists $\varepsilon > 0$ s.t. $B_\varepsilon^E(x) \subset V$.*

The latter definition leads to the following result.

Proposition A.12. *Let $E \subset \mathbb{R}^m$, $E \neq \emptyset$. A function $f\colon E \to \mathbb{R}^k$ is continuous if and only if for any $U \subset \mathbb{R}^k$ open, $f^{-1}(U)$ is an open set in E.*

Proof. We repeat the arguments given in Proposition A.11 (replace open balls in $\mathbb{R}^m$ with open balls in E) and the result follows. $\square$

We prove Propositions 2.11 and 3.13.

Proof of Proposition 2.11. Using Proposition A.12, we show that a set $V \subset E$ is open if and only if there exists $G \subset \mathbb{R}^m$ which is open in $\mathbb{R}^m$ and s.t. $V = G \cap E$. Let $V \subset E$. Suppose that there exists $G \subset \mathbb{R}^m$, open in $\mathbb{R}^m$, s.t. $V = G \cap E$. Let $x \in V$, then $x \in G \cap E$, and since G is open and $x \in G$, there exists $B_\varepsilon^m(x) \subset G$. Thus, $B_\varepsilon^m(x) \cap E = B_\varepsilon^E(x) \subset G \cap E = V$. This shows that V is open in E. For the other direction, let $V \subset E$ be open in E. Then, given any $x \in V$, there exists $B^E_{\varepsilon_x}(x)$ s.t. $B^E_{\varepsilon_x}(x) \subset V$. We know that for any $x \in V$, $B^m_{\varepsilon_x}(x)$ is open in $\mathbb{R}^m$ (cf. Example 2.16). Define

$$G = \bigcup_{x \in V} B^m_{\varepsilon_x}(x).$$

It is clear that G is open in $\mathbb{R}^m$, since given any $y \in G$, there exists $x \in V$, s.t. $y \in B^m_{\varepsilon_x}(x) \subset G$. Certainly $V \subset G$ and hence $V = V \cap E \subset G \cap E$. On the other hand, $G \cap E = \cup_{x \in V} B^E_{\varepsilon_x}(x) \subset V$, since for any $x \in V$, $B^E_{\varepsilon_x}(x) \subset V$. Therefore, $V = G \cap E$. □

Proof of Proposition 3.13. We only show the existence of x_M since the arguments for the minimum are similar. Let $S = \sup_{x \in \mathbb{R}} f(x)$. There are two cases, either $S < \infty$ or $S = \infty$. If $S = \infty$, let $(x_n)_{n \in \mathbb{N}}$ be a sequence s.t. for any $n \in \mathbb{N}$ $x_n \in [a, b]$ and $f(x_n) \geq n$ (for example, given any $n \in \mathbb{N}$, set $x_n = \min\{x \in [a, b] \colon f(x) \geq n\}$). Since $x_n \in [a, b]$ for any $n \in \mathbb{N}$, it follows that $(x_n)_{n \in \mathbb{N}}$ is bounded and hence by Proposition 3.12, there exists a subsequence $(x_{s(n)})_{n \in \mathbb{N}}$ s.t. $\lim_{n \to \infty} x_{s(n)} = \xi \in [a, b]$. Notice that $\xi \in [a, b]$, since $a \leq x_n \leq b$ for any $n \in \mathbb{N}$. In particular, since f is continuous on $[a, b]$, it follows that $\lim_{n \to \infty} f(x_{s(n)}) = f(\xi)$. On the other hand, since for any $n \in \mathbb{N}$, $f(x_n) \geq n$, it follows that $\lim_{n \to \infty} f(x_{s(n)}) = \infty$. Which gives a contradiction. Hence it must be the case that $S < \infty$. In this case, we set $a_n = S - 1/n$, $n \in \mathbb{N}$, and by Proposition 1.10, there exists $(x_n)_{n \in \mathbb{N}}$ s.t. for any $n \in \mathbb{N}$ $x_n \in [a, b]$ and $f(x_n) > S_n$. Again, we apply Proposition 3.12, and find a subsequence $(x_{s(n)})_{n \in \mathbb{N}}$ s.t. $\lim_{n \to \infty} x_{s(n)} = x_M \in [a, b]$. Since f is continuous on $[a, b]$, $\lim_{n \to \infty} f(x_{s(n)}) = f(x_M)$. Then, we have that for any $n \in \mathbb{N}$,

$$f(x_{s(n)}) \geq S - \frac{1}{s(n)},$$

and hence, $\lim_{n \to \infty} f(x_{s(n)}) \geq S$. Clearly, $\lim_{n \to \infty} f(x_{s(n)}) \leq S$ as well and hence $S = f(x_M) = \max_{x \in [a,b]} f(x)$. To see that f is bounded (cf. Definition 2.20), we apply Proposition A.10 and note that $x \mapsto |f(x)|$ is continuous on $[a, b]$. Therefore, for any $x \in [a, b]$, $|f(x)| \leq |f(x_M)|$. □

Example A.4. *We list some examples of continuous functions:*

- $x \mapsto \mathrm{e}^x$ *as a function from* $\mathbb{R}$ *to* $\mathbb{R}$;
- $x \mapsto \log(x)$ *as a function from* $(0, \infty)$ *to* $\mathbb{R}$;
- $x \mapsto \sin(x)$ *as a function from* $\mathbb{R}$ *to* $\mathbb{R}$;
- $x \mapsto \cos(x)$ *as a function from* $\mathbb{R}$ *to* $\mathbb{R}$.

A.5 Limit points

Definition A.5. *Let $E \subset \mathbb{R}^m$ be a nonempty set. A point $a \in \mathbb{R}^m$ is said to be a limit or accumulation point of E if there exists a vector-valued sequence $(x_n)_{n \in \mathbb{N}}$ which satisfies $x_n \in E \setminus \{a\}$ for any $n \in \mathbb{N}$ and $x_n \xrightarrow{n \to \infty} a$.*

For this section, unless stated otherwise, any sequence is a real-valued sequence according to Definition 3.1.

Definition A.6. *Let $E \subset \mathbb{R}$ be a nonempty set and $f \colon E \to \mathbb{R}$ be a function. If $a \in \mathbb{R}$ is a limit point of E, we write $\lim_{x \to a} f(x) = L$ if there exists $L \in \mathbb{R}$ s.t. for any sequence $(x_n)_{n \in \mathbb{N}}$ which satisfies $x_n \in E \setminus \{a\}$ for any $n \in \mathbb{N}$ and $x_n \xrightarrow{n \to \infty} a$, it follows that $f(x_n) \xrightarrow{n \to \infty} L$.*

Proposition A.13. *Let $E \subset \mathbb{R}$ be a nonempty set and $f \colon E \to \mathbb{R}$ be a function. Then, $\lim_{x \to a} f(x) = L$ if and only if for any $\varepsilon > 0$, there exists $\delta > 0$ s.t. for any $x \in E \setminus \{a\}$, if $|x - a| < \delta$, $|f(x) - L| < \varepsilon$.*

Proof. This is essentially the proof of the sequence criterion of continuous function where the limit $f(x)$ is replaced with L (cf. the proof of Proposition 3.25). $\square$

Definition A.7. *Let $E \subset \mathbb{R}$ be a nonempty set and $f : E \to \mathbb{R}$ be a function. If $a \in \mathbb{R}$ is a limit point of $E \cap (-\infty, a)$, we write $\lim_{x\uparrow a} f(x) = L_l$ if there exists $L_l \in \mathbb{R}$ s.t. for any sequence $(x_n)_{n\in\mathbb{N}}$ which is s.t.*

$$\{x_n : n \in \mathbb{N}\} \subset E \cap (-\infty, a) \text{ and } x_n \xrightarrow{n\to\infty} a,$$

it follows that $f(x_n) \xrightarrow{n\to\infty} L_l$. Similarly, if $a \in \mathbb{R}$ is a limit point of $E \cap (a, \infty)$, we use the notation $\lim_{x\downarrow a} f(x) = L_r$ if there exists $L_r \in \mathbb{R}$ s.t. for any sequence $(x_n)_{n\in\mathbb{N}}$ which is s.t.

$$\{x_n : n \in \mathbb{N}\} \subset E \cap (a, \infty) \text{ and } x_n \xrightarrow{n\to\infty} a,$$

it follows that $f(x_n) \xrightarrow{n\to\infty} L_r$. If they exists, L_l and L_r are referred to as the left-hand and right-hand limit of f as x approaches a.

If we adapt the proof of Proposition 3.25 accordingly, we readily verify the following result:

Proposition A.14. *Let $E \subset \mathbb{R}$ be a nonempty set and $f : E \to \mathbb{R}$ be a function. Then, $\lim_{x\uparrow a} f(x) = L_l$ if and only if for any $\varepsilon > 0$, there exists $\delta_l > 0$ s.t. for any $x \in E \setminus \{a\}$, $x \in (a - \delta_l, a)$ implies that $|f(x) - L_l| < \varepsilon$. Similarly, $\lim_{x\downarrow a} f(x) = L_r$ if and only if for any $\varepsilon > 0$, there exists $\delta_r > 0$ s.t. for any $x \in E \setminus \{a\}$, $x \in (a, a + \delta_l)$ implies that $|f(x) - L_r| < \varepsilon$.*

Proposition A.15. *Let $E \subset \mathbb{R}$ be a nonempty set and $f : E \to \mathbb{R}$ be a function. Then, $\lim_{x\to a} f(x) = L$ if and only if*

$$\lim_{x\uparrow a} f(x) = L = \lim_{x\downarrow a} f(x).$$

Proof. Let δ_l and δ_r be as in Proposition A.14 and apply Proposition A.13 with $\delta = \min\{\delta_l, \delta_r\}$. $\square$

Upon the definition of continuity (cf. Definition A.2), Proposition A.15 shows the following:

Proposition A.16. *Let $E \subset \mathbb{R}$ be a nonempty set and $f : E \to \mathbb{R}$ be a function. Then, f is continuous at a point $a \in E$, if and only if $\lim_{x\uparrow a} f(x) = f(a) = \lim_{x\downarrow a} f(x)$.*

Proposition A.17. *Let $E \subset \mathbb{R}$ be a nonempty set and $f : E \to \mathbb{R}$ be a function. Then, $\lim_{x\uparrow a} f(x) = L_l$ if and only if*

$$\forall\, (x_n)_{n\in\mathbb{N}} \text{ s.t. } x_n \in E \setminus \{a\}\ \forall n \in \mathbb{N} \text{ and } x_n \uparrow a \text{ it follows that } f(x_n) \xrightarrow{n\to\infty} L_l. \tag{15}$$

Similarly, $\lim_{x\downarrow a} f(x) = L_r$ if and only if

$$\forall\, (x_n)_{n\in\mathbb{N}} \text{ s.t. } x_n \in E \setminus \{a\}\ \forall n \in \mathbb{N} \text{ and } x_n \downarrow a \text{ it follows that } f(x_n) \xrightarrow{n\to\infty} L_r. \tag{16}$$

Proof. We only show that (15) implies $\lim_{x\uparrow a} f(x) = L_l$ and vice versa. The argument for the right-hand limit is the same. By Definition A.7, if $\lim_{x\uparrow a} f(x) = L_l$, then, (15) holds. Thus, suppose that (15) holds and let $(x_n)_{n\in\mathbb{N}}$ be a sequence s.t.

$$\{x_n : n \in \mathbb{N}\} \subset E \cap (-\infty, a) \text{ and } x_n \xrightarrow{n\to\infty} a.$$

Let $(x_{s(n)})_{n\in\mathbb{N}}$ be any subsequence of $(x_n)_{n\in\mathbb{N}}$. Since $x_n \xrightarrow{n\to\infty} a$, it follows that $x_{s(n)} \xrightarrow{n\to\infty} a$. In particular, $(x_{s(n)})_{n\in\mathbb{N}}$ is bounded. Define $x_{t(s(n))} = \inf\{x_{s(k)} : k \geq n\}$, $n \in \mathbb{N}$. Then, $(x_{t(s(n))})_{n\in\mathbb{N}}$ is increasing and s.t. $x_{t(s(n))} \uparrow a$ (cf. Exercise 3.7). Thus, by (15), $\lim_{n\to\infty} f(x_{t(s(n))}) = L_l$. Therefore, we have shown that for any subsequence $(f(x_{s(n)}))_{n\in\mathbb{N}}$ of $(f(x_n))_{n\in\mathbb{N}}$ there exists a subsequence $(f(x_{t(s(n))}))_{n\in\mathbb{N}}$ s.t. $\lim_{n\to\infty} f(x_{t(s(n))}) = L_l$. By Proposition A.6, $\lim_{n\to\infty} f(x_n) = L_l$ and hence (15) implies that $\lim_{x\uparrow a} f(x) = L_l$. $\square$

Definition A.8. *Let $E \subset \mathbb{R}$ be a nonempty set and $f: E \to \mathbb{R}$ be a function. If E is not bounded from above, we write $\lim_{x\to\infty} f(x) = L$ if there exists $L \in \mathbb{R}$ s.t. for any sequence $(x_n)_{n\in\mathbb{N}}$ which is s.t. $\{x_n : n \in \mathbb{N}\} \subset E$ and $x_n \xrightarrow{n\to\infty} \infty$, it follows that $f(x_n) \xrightarrow{n\to\infty} L$. Similarly, if E is not bounded from below, we write $\lim_{x\to-\infty} f(x) = L$ if there exists $L \in \mathbb{R}$ s.t. for any sequence $(x_n)_{n\in\mathbb{N}}$ which is s.t. $\{x_n : n \in \mathbb{N}\} \subset E$ and $x_n \xrightarrow{n\to\infty} -\infty$, it follows that $f(x_n) \xrightarrow{n\to\infty} L$.*

Proposition A.18. *Let $E \subset \mathbb{R}$ be a nonempty set and $f: E \to \mathbb{R}$ be a function. If E is not bounded from above, then $\lim_{x\to\infty} f(x) = L$ if and only if for any $\varepsilon > 0$ there exists a real number $M > 0$ s.t. for any $x \in E$ with $x > M$, $|f(x) - L| < \varepsilon$. Similarly, if E is not bounded from below, $\lim_{x\to-\infty} f(x) = L$ if and only if for any $\varepsilon > 0$ there exists a real number $M > 0$ s.t. for any $x \in E$ with $x < -M$, $|f(x) - L| < \varepsilon$.*

Proof. Suppose that $\lim_{x\to\infty} f(x) = L$ and there exists $\varepsilon > 0$ s.t. for any $M > 0$ there exists $x \in E$ s.t. $x > M$ with $|f(x) - L| \geq \varepsilon$. Let $n_1 = 1$ and obtain $n_1 < x_1 \leq n_2$ which is s.t. $|f(x_1) - L| \geq \varepsilon$. Then, find $n_2 < x_2 \leq n_3$ which also satisfies $|f(x_2) - L| \geq \varepsilon$. If we continue like this, we obtain a sequence $(x_n)_{n\in\mathbb{N}}$ which is s.t. for any $n \in \mathbb{N}$, $|f(x_n) - L| \geq \varepsilon$. Since $x_n \xrightarrow{n\to\infty} \infty$ this contradicts $\lim_{x\to\infty} f(x) = L$. Therefore, $\lim_{x\to\infty} f(x) = L$ implies that for any $\varepsilon > 0$ there exists a real number $M > 0$ s.t. for any $x \in E$ s.t. $x > M$, $|f(x) - L| < \varepsilon$. For the other direction, suppose that for any $\varepsilon > 0$ there exists a real number $M > 0$ s.t. for any $x \in E$ with $x > M$, $|f(x) - L| < \varepsilon$. Let $\varepsilon > 0$ and consider $\{x_n : n \in \mathbb{N}\} \subset E$ s.t. $x_n \xrightarrow{n\to\infty} \infty$. Since $x_n \xrightarrow{n\to\infty} \infty$, there exists $N \in \mathbb{N}$, s.t. for any $n \geq N$, $x_n > M$. Thus, by assumption, $|f(x_n) - L| < \varepsilon$, i.e., $f(x_n) \xrightarrow{n\to\infty} L$. The argument for $\lim_{x\to-\infty} f(x) = L$ is similar and we consider the proposition as proven. □

Similar to Proposition A.17, we have the following result:

Proposition A.19. *Suppose that $f: E \to \mathbb{R}$ is a function where $E \subset \mathbb{R}$ is a nonempty set. Suppose that E is not bounded from above. Then, $\lim_{x\to\infty} f(x) = L$ if and only if*

$$\forall\, (x_n)_{n\in\mathbb{N}} \text{ s.t. } x_n \in E \ \forall n \in \mathbb{N} \text{ and } x_n \uparrow \infty \text{ it follows that } f(x_n) \xrightarrow{n\to\infty} L. \tag{17}$$

Similarly, if E is not bounded from below, $\lim_{x\to-\infty} f(x) = L$ if and only if

$$\forall\, (x_n)_{n\in\mathbb{N}} \text{ s.t. } x_n \in E \ \forall n \in \mathbb{N} \text{ and } x_n \downarrow -\infty \text{ it follows that } f(x_n) \xrightarrow{n\to\infty} L. \tag{18}$$

Proof. We only show that (18) implies that $\lim_{x\to-\infty} f(x) = L$ and vice versa (the remaining argument is similar). Suppose that E is not bounded from below. Clearly, if $\lim_{x\to-\infty} f(x) = L$, then by Definition A.8, (18) is satisfied. Hence, suppose that (18) is satisfied. Assume by contradiction that there exists $\varepsilon > 0$ s.t. for any $M > 0$ there exists $x \in E$ s.t. $x < -M$ with $|f(x) - L| \geq \varepsilon$. Let $n_1 = -1$ and obtain $n_2 \leq x_1 < n_1$ which is s.t. $|f(x_1) - L| \geq \varepsilon$. Then, find $n_3 \leq x_2 < n_2$ which also satisfies $|f(x_2) - L| \geq \varepsilon$. If we continue like this, we obtain a sequence $(x_n)_{n\in\mathbb{N}}$ which is s.t. $x_n \downarrow -\infty$ but $\lim_{n\to\infty} f(x_n) \neq L$. This contradicts (18). Hence, for any $\varepsilon > 0$ there exists a real number $M > 0$ s.t. for any $x \in E$ with $x < -M$, $|f(x) - L| < \varepsilon$. By Proposition A.18 this implies that $\lim_{x\to-\infty} f(x) = L$. □

Remark A.1. *Notice that if $E \subset \mathbb{R}^m$, $a \in \mathbb{R}^m$ is a limit point of E and $f: E \to \mathbb{R}^k$ is a function, $\lim_{x\to a} f(x) = L$, $L \in \mathbb{R}^k$, is defined as in Definition A.6 with $(x_n)_{n\in\mathbb{N}}$ vector-valued.*

A.6 Differentiability in one variable and the mean value theorem

Definition A.9. *Let $f: [a, b] \to \mathbb{R}$ be a function. The derivative of f at $x_0 \in (a, b)$ is defined as the limit*

$$f'(x_0) = \lim_{x\to x_0} \frac{f(x) - f(x_0)}{x - x_0} \quad \left(= \lim_{h\to 0} \frac{f(x_0 + h) - f(x_0)}{h} \right).$$

f is referred to as differentiable if $f'(x_0) \in \mathbb{R}$ for any $x_0 \in (a, b)$, i.e., $f': (a, b) \to \mathbb{R}$.

Remark A.2. *If $f\colon [a,b] \to \mathbb{R}$ is differentiable, then f is continuous on (a,b). In particular, if f is continuous in a and b, f is continuous on the entire $[a,b]$.*

Proposition A.20. *Let $f\colon [a,b] \to \mathbb{R}$ be a differentiable function. Suppose that $x_0 \in (a,b)$ is a maximum point (resp. minimum point) of f, i.e., $f(x) \leq f(x_0)$ for any $x \in [a,b]$ (resp. $f(x) \geq f(x_0)$ for any $x \in [a,b]$). Then, $f'(x_0) = 0$.*

Proof. Suppose that $x_0 \in (a,b)$ is a maximum point of f. By Proposition A.15, we know that

$$\lim_{x \uparrow x_0} \frac{\overbrace{f(x) - f(x_0)}^{\leq 0}}{\underbrace{x - x_0\}}_{\leq 0}} = f'(x_0) = \lim_{x \downarrow x_0} \frac{\overbrace{f(x) - f(x_0)}^{\leq 0}}{\underbrace{x - x_0\}}_{\geq 0}}.$$

Hence, $f'(x_0) = 0$. The same argument works if $x_0 \in (a,b)$ is a minimum point of f. ☐

The next result is known as Rolle's theorem.

Proposition A.21. *Let $f\colon [a,b] \to \mathbb{R}$ be continuous and differentiable. Suppose that $f(a) = f(b)$. Then, there exists $x_0 \in (a,b)$ s.t. $f'(x_0) = 0$.*

Proof. Since $f(a) = f(b)$, there are three cases:

(i) for any $x \in (a,b)$ $f(x) = f(b)$;

(ii) there exists $x \in (a,b)$ s.t. $f(x) > f(b)$;

(iii) there exists $x \in (a,b)$ s.t. $f(x) < f(b)$.

In case (i), f is constant on (a,b), i.e., $f'(x) = 0$ for any $x \in (a,b)$. If case (ii) is true, then, since f is continuous on $[a,b]$, there exists $x_0 \in [a,b]$ s.t. $f(x_0) \geq f(y)$ for any $y \in [a,b]$ (cf. Proposition 3.13). We notice that $x_0 = a$ or $x_0 = b$ is not possible since this would imply that $f(x_0) = f(a) = f(b) < x$. Thus, the maximum point is s.t. $x_0 \in (a,b)$. By Proposition A.20, $f'(x_0) = 0$. Finally suppose that case (iii) holds. By Proposition 3.13 again, let x_0 be a minimum point of $[a,b]$. Again, $x_0 \in (a,b)$ and hence, $f'(x_0) = 0$. ☐

The mean value theorem (or Lagrange theorem) reads as follows:

Proposition A.22. *Let $f\colon [a,b] \to \mathbb{R}$ be continuous and differentiable. Then, there exists $m \in (a,b)$ s.t.*

$$f(b) - f(a) = f'(m)(b-a).$$

Proof. Define

$$g(x) = f(x) - \frac{x-a}{b-a}(f(b) - f(a)).$$

We notice that g is s.t. $g(a) = g(b)$, g is continuous on $[a,b]$ and the derivative of g exists for any point $x_0 \in (a,b)$. Thus, by Proposition A.21, there exists $m \in (a,b)$ s.t. $g'(m) = 0$. That is,

$$0 = g'(m) = f'(m) - \frac{f(b) - f(a)}{b-a}.$$

☐

Example A.5. *Given any $x \in \mathbb{R}$, $1 + x \leq \mathrm{e}^x$. Clearly, the inequality becomes an equality if $x = 0$. Let $x > 0$ and consider the interval $[0, x]$. By the mean value theorem, there exists $m \in (0, x)$ s.t.*

$$\mathrm{e}^x - \mathrm{e}^0 = \mathrm{e}^x - 1 = \mathrm{e}^m (x - 0) = \mathrm{e}^m x.$$

Since $m > 0$, $\mathrm{e}^m > 1$. Therefore, the previous display shows that $\mathrm{e}^x - 1 > x \Leftrightarrow 1 + x < \mathrm{e}^x$. If $x < 0$, then, again by the mean value theorem, there exists $m \in (x, 0)$ s.t. $1 - \mathrm{e}^x = -\mathrm{e}^m x$. Then, since $m \in (x, 0)$ and $-x > 0$, we obtain with $0 < \mathrm{e}^m < 1$ that $1 - \mathrm{e}^x < -x \Leftrightarrow 1 + x < \mathrm{e}^x$.

A.7 Differentiability in several variables

Definition A.10. *A function $L\colon \mathbb{R}^m \to \mathbb{R}^k$ is linear if for any $v_1, v_2 \in \mathbb{R}^m$ and for any $\lambda_1, \lambda_2 \in \mathbb{R}$,*

$$L(\lambda_1 v_1 + \lambda_2 v_2) = \lambda_1 L(v_1) + \lambda_2 L(v_2).$$

Suppose that $f\colon [a, b] \to \mathbb{R}$ is differentiable in $x_0 \in (a, b)$, i.e., the limit $f'(x_0)$ exists according to Definition A.9. Define the map: $L_{x_0}(h) = f'(x_0)h$, $h \in \mathbb{R}$. Then, $L_{x_0}\colon \mathbb{R} \to \mathbb{R}$ is linear and

$$\lim_{h \to 0} \frac{f(x_0 + h) - f(x_0) - L_{x_0}(h)}{h} = 0. \tag{19}$$

On the other hand if there exists a linear map $L_{x_0}\colon \mathbb{R} \to \mathbb{R}$ s.t. f satisfies (19) for some $x_0 \in (a, b)$, then,

$$\lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h} = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0) - L_{x_0}(h)}{h} + L_{x_0}(1) = L_{x_0}(1),$$

that is f is differentiable in x_0 according to Definition A.9. This shows that the following definition is equivalent to Definition A.9.

Definition A.11. *A function $f\colon [a, b] \to \mathbb{R}$ is differentiable in $x_0 \in (a, b)$ if there exists a linear map $L_{x_0}\colon \mathbb{R} \to \mathbb{R}$ s.t. (19) is satisfied.*

In general, differentiability is defined as follows (again $\|\cdot\|_m$ and $\|\cdot\|_k$ denote the Euclidean distance on $\mathbb{R}^m$ and $\mathbb{R}^k$, respectively):

Definition A.12. *Let $U \subset \mathbb{R}^m$ be an open set. A function $f\colon U \to \mathbb{R}^k$ is differentiable in $x_0 \in U$ if there exists a linear map $L_{x_0}\colon \mathbb{R}^m \to \mathbb{R}^k$ s.t.*

$$\lim_{h \to 0} \frac{f(x_0 + h) - f(x_0) - L_{x_0}(h)}{\|h\|_m} = 0. \tag{20}$$

The map f is referred to as differentiable on U if f is differentiable for any $x_0 \in U$. Further, the linear map L_{x_0} is said to be differential of f in x_0.

Remark A.3. *Let $U \subset \mathbb{R}^m$ be an open set and $f\colon U \to \mathbb{R}^k$ be differentiable in $x_0 \in U$. Then the linear map L_{x_0} in (20) is unique. That is, if L_{x_0} and $\tilde{L}_{x_0}$ are two linear maps that satisfy (20), then $L_{x_0} = \tilde{L}_{x_0}$. To see it, if L_{x_0} and $\tilde{L}_{x_0}$ satisfy (20), then for any $v \in \mathbb{R}^m$, $v \neq 0$, by linearity and (20),*

$$L_{x_0}(v/\|v\|_m) - \tilde{L}_{x_0}(v/\|v\|_m) = 0.$$

Thus, $L_{x_0} = \tilde{L}_{x_0}$.

Example A.6. *Let* $f(x) = Ax + b$, $A \in \mathbb{R}^{k\times m}$, $b \in \mathbb{R}^k$. *Let* $L(x) = Ax$, $x \in \mathbb{R}^m$. *Given any* $x_0 \in \mathbb{R}^m$, *we have that*

$$\lim_{h\to 0} \frac{f(x_0+h) - f(x_0) - L(h)}{\|h\|_m} = 0.$$

Therefore, $f\colon \mathbb{R}^m \to \mathbb{R}^k$ *is differentiable with differential* L *(cf. Remark A.3).*

The following proposition is of central importance:

Proposition A.23. *Let* $U \subset \mathbb{R}^m$ *be an open set and* $f = (f_1, \ldots, f_k)\colon U \to \mathbb{R}^k$ *be a map. Assume that* $x_0 \in U$. *Then,* f *is differentiable in* x_0 *if and only if any function* $f_i\colon U \to \mathbb{R}$, $i = 1, \ldots, k$, *is differentiable in* x_0. *Further, if* f *is differentiable in* x_0, *the differential of* f *in* x_0 *is given by the map*

$$L_{x_0} = (L^1_{x_0}, \ldots, L^k_{x_0}),$$

where $L^i_{x_0}\colon \mathbb{R}^m \to \mathbb{R}$ *is the differential of* f_i *in* x_0, $i = 1, \ldots, k$.

Proof. Suppose first that f is differentiable in x_0, i.e., there exists a linear map $L_{x_0}\colon \mathbb{R}^m \to \mathbb{R}^k$ s.t. (20) is satisfied. Define the linear map

$$L^i_{x_0}(v) = \sum_{j=1}^m a_{ij} v_j, \quad i = 1, \ldots, k,$$

where $A = (a_{ij})_{1\le i\le k, 1\le j\le m}$ is the matrix representation of L_{x_0}. Notice that this implies that

$$Av = (L^1_{x_0}(v), \ldots, L^k_{x_0}(v)), \quad v \in \mathbb{R}^m.$$

Further, for any $i = 1, \ldots, k$,

$$|f_i(x_0+v) - f_i(x_0) - L^i_{x_0}(v)| \le \|f(x_0+v) - f(x_0) - L_{x_0}(v)\|_k.$$

Thus, by (20),

$$\lim_{v\to 0} \frac{|f_i(x_0+v) - f_i(x_0) - L^i_{x_0}(v)|}{\|v\|_m} = 0,$$

and in particular (recall that $t \mapsto |t| = \sqrt{t^2}$ is continuous),

$$\lim_{v\to 0} \frac{f_i(x_0+v) - f_i(x_0) - L^i_{x_0}(v)}{\|v\|_m} = 0.$$

For the other direction, suppose that for any $i = 1, \ldots, k$, f_i is differentiable in x_0 with differential $L^i_{x_0}$ in x_0. Then, we define the map

$$L_{x_0}(v) = (L^1_{x_0}(v), \ldots, L^k_{x_0}(v)), \quad v \in \mathbb{R}^m,$$

and notice that $L_{x_0}\colon \mathbb{R}^m \to \mathbb{R}^k$ is linear since $L^i_{x_0}$, $i = 1, \ldots, k$, are linear. Then, since for any $i = 1, \ldots, k$,

$$\lim_{v\to 0} \frac{f_i(x_0+v) - f_i(x_0) - L^i_{x_0}(v)}{\|v\|_m} = 0,$$

it follows with

$$\frac{\|f(x_0+v) - f(x_0) - L_{x_0}(v)\|_k}{\|v\|_m} = \sqrt{\sum_{i=1}^k \left(\frac{f_i(x_0+v) - f_i(x_0) - L^i_{x_0}(v)}{\|v\|_m}\right)^2},$$

that

$$\lim_{v\to 0} \frac{\|f(x_0+v)-f(x_0)-L_{x_0}(v)\|_k}{\|v\|_m} = 0.$$

In particular,

$$\lim_{v\to 0} \frac{f(x_0+v)-f(x_0)-L_{x_0}(v)}{\|v\|_m} = 0,$$

and hence f is differentiable in x_0 with differential L_{x_0}. $\square$

In the following, we aim to find an explicit description of the matrix representation of the differential of a differentiable map. This matrix will be referred to as the Jacobian matrix.

Definition A.13. *Let $U \subset \mathbb{R}^m$ be open, $x_0 \in U$ and $f\colon U \to \mathbb{R}$ be a function. The directional derivative of f in direction $v \neq 0$ is defined as the limit (if it exists)*

$$\partial_v f(x_0) = \frac{\partial f}{\partial v}(x_0) = \lim_{t\to 0} \frac{f(x_0+tv)-f(x_0)}{t}.$$

Proposition A.24. *Let $U \subset \mathbb{R}^m$ be an open set and $f\colon U \to \mathbb{R}$ be differentiable in $x_0 \in U$. Then, for any $v \neq 0$, the directional derivative $\partial_v f(x_0)$ exists and is given by*

$$\frac{\partial f}{\partial v}(x_0) = L_{x_0}(v).$$

Proof. By assumption, f is differentiable in x_0, hence there exists a linear map $L_{x_0}\colon \mathbb{R}^m \to \mathbb{R}$ s.t. (20) is satisfied. Then, given any $v \in \mathbb{R}^m$, $v \neq 0$, we obtain

$$\begin{aligned}\frac{f(x_0+tv)-f(x_0)}{\|tv\|_m} &= \frac{f(x_0+tv)-f(x_0)-L_{x_0}(tv)}{\|tv\|_m} + \frac{L_{x_0}(tv)}{\|tv\|_m} \\ &= \frac{f(x_0+tv)-f(x_0)-L_{x_0}(tv)}{\|tv\|_m} + \frac{L_{x_0}(tv)}{\|tv\|_m}.\end{aligned}$$

By (20), it follows that

$$\lim_{t\downarrow 0} \frac{f(x_0+tv)-f(x_0)}{\|tv\|_m} = \frac{L_{x_0}(v)}{\|v\|_m}.$$

Thus,

$$\lim_{t\downarrow 0} \frac{f(x_0+tv)-f(x_0)}{t} = \|v\|_m \lim_{t\downarrow 0} \frac{f(x_0+tv)-f(x_0)}{\|tv\|_m} = L_{x_0}(v).$$

Similarly,

$$\lim_{t\uparrow 0} \frac{f(x_0+tv)-f(x_0)}{t} = L_{x_0}(v),$$

and the result follows. $\square$

Definition A.14. *Let $U \subset \mathbb{R}^m$ be open, $x_0 \in U$ and $f\colon U \to \mathbb{R}$ be a function. Let $e_1, \ldots, e_m$, be the standard basis of $\mathbb{R}^m$, i.e.,*

$$e_j = (0, \ldots, 0, \underbrace{1}_{\text{position } j}, 0, \ldots, 0), \quad j = 1, \ldots, m.$$

The directional derivatives of f in direction e_j, $j = 1, \ldots, m$, are referred to as the partial derivatives of f in x_0. We use the notation

$$\frac{\partial f}{\partial e_j}(x_0) = \frac{\partial f}{\partial x_j}(x_0) = \partial_{x_j} f(x_0), \quad j = 1, \ldots, m.$$

Remark A.4. *Let $U \subset \mathbb{R}^m$ be an open set and $x_0 \in U$. Suppose that $f\colon U \to \mathbb{R}$ is differentiable in x_0. Then, by Proposition A.24, the partial derivatives $\partial_{x_j} f(x_0)$ exists for any $j = 1, \dots, m$. In particular, for any $j = 1, \dots, m$,*

$$\frac{\partial f}{\partial x_j}(x_0) = L_{x_0}(e_j).$$

Write $v = \sum_{j=1}^m v_j e_j$, where $v_j \in \mathbb{R}$ and $e_1, \dots, e_m$, is the standard basis of $\mathbb{R}^m$. By linearity, we obtain that

$$L_{x_0}(v) = \sum_{j=1}^m v_j L_{x_0}(e_j) = \sum_{j=1}^m v_j \frac{\partial f}{\partial x_j}(x_0). \tag{21}$$

The identity (21) is the representation of the differential of f in x_0 in terms of the partial derivatives.

Definition A.15. *Let $f\colon U \to \mathbb{R}^k$, $U \subset \mathbb{R}^m$ open. Suppose that the partial derivatives $\partial_{x_j} f_i(x_0)$ exist for any $j = 1, \dots, m$ and $i = 1, \dots, k$. Then, the matrix*

$$J_f(x_0) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x_0) & \frac{\partial f_1}{\partial x_2}(x_0) & \dots & \frac{\partial f_1}{\partial x_m}(x_0) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_k}{\partial x_1}(x_0) & \frac{\partial f_k}{\partial x_2}(x_0) & \dots & \frac{\partial f_k}{\partial x_m}(x_0) \end{pmatrix}$$

is referred to as the Jacobian matrix of f in x_0.

We have all the tools to easily verify the following:

Proposition A.25. *Let $f\colon U \to \mathbb{R}^k$, $U \subset \mathbb{R}^m$ open. Suppose that f is differentiable in $x_0 \in U$ with differential L_{x_0}. Then, the matrix representation of the linear map L_{x_0} is given by the Jacobian matrix $J_f(x_0)$.*

Proof. Suppose that $v \neq 0$ (for $v = 0$, the result is trivial). By Proposition A.23 and (21), we have that

$$L_{x_0}(v) = (L^1_{x_0}(v), \dots, L^k_{x_0}(v)) = \left(\sum_{j=1}^m v_j \frac{\partial f_1}{\partial x_j}(x_0), \dots, \sum_{j=1}^m v_j \frac{\partial f_k}{\partial x_j}(x_0) \right) = J_f(x_0)v.$$

□

Remark A.5. *One can show that $f\colon U \to \mathbb{R}^k$, $U \subset \mathbb{R}^m$ open, is differentiable in $x_0 \in U$ if for any $v \in U$, the partial derivatives $\partial_{x_j} f_i(v)$, $j = 1, \dots, m$, $i = 1, \dots, k$, exist and are continuous in x_0. This provides a useful criterion to verify the differentiability of f in x_0.*

B Measure and integration

B.1 Inclusion-exclusion principle

The following is known as the inclusion-exclusion formula:

Proposition B.1. *Let $(\Omega, \mathcal{F})$ be a measurable space and μ be a measure on $\mathcal{F}$. Assume that $A_1, \dots, A_n \in \mathcal{F}$ are s.t. $\mu(A_i) < \infty$ for any $i = 1, \dots, n$. We have that*

$$\mu\left(\bigcup_{i=1}^n A_i \right) = \sum_{k=1}^n \left((-1)^{k-1} \sum_{I \in A_k^n} \mu(A_I) \right), \tag{22}$$

where

$$A_k^n = \{I \subset \{1, \dots, n\} \colon \#I = k\}, \quad k = 1, \dots, n, \quad n \in \mathbb{N},$$

and for any $I \in A_k^n$, $\mu(A_I) = \mu(\cap_{i \in I} A_i)$.

Proof of Proposition B.1. If $n = 1$, (22) is trivial. If $n = 2$, (22) states that

$$\mu(A_1 \cup A_2) = \mu(A_1) + \mu(A_2) - \mu(A_1 \cap A_2),$$

which is true since by item (iv) of Proposition 5.1, $\mu(A_1 \cup A_2) + \mu(A_1 \cap A_2) = \mu(A_1) + \mu(A_2)$ and since $\mu(A_1)$ and $\mu(A_2)$ are assumed to be finite, we can subtract $\mu(A_1 \cap A_2)$ on both sides of the latter equation. By induction, assume that (22) holds for $n \in \mathbb{N}$. We have that

$$\begin{aligned}
\mu\left(\bigcup_{i=1}^{n+1} A_i\right) &= \mu\left(\bigcup_{i=1}^{n} A_i\right) + \mu(A_{n+1}) - \mu\left(\bigcup_{i=1}^{n} (A_i \cap A_{n+1})\right) \\
&= \sum_{k=1}^{n} \left((-1)^{k-1} \sum_{I \in A_k^n} \mu(A_I)\right) + \mu(A_{n+1}) \\
&\quad - \left(\sum_{k=1}^{n} \left((-1)^{k-1} \sum_{I \in A_k^n} \mu(A_I \cap A_{n+1})\right)\right). \qquad (23)
\end{aligned}$$

Then, for any $k = 1, \ldots, n$, $A_k^n \subset A_k^{n+1}$, and therefore,

$$\begin{aligned}
&\sum_{k=1}^{n} \left((-1)^{k-1} \sum_{I \in A_k^n} \mu(A_I))\right) - \sum_{k=1}^{n+1} \left((-1)^{k-1} \sum_{I \in A_k^{n+1}} \mu(A_I))\right) \\
&= -\left(\sum_{k=1}^{n} \left((-1)^{k-1} \sum_{I \in (A_k^{n+1} \setminus A_k^n)} \mu(A_I))\right)\right).
\end{aligned}$$

We notice that the latter sum is equal to

$$\begin{aligned}
&-\mu(A_{n+1}) - \left(\sum_{k=1}^{n} \left((-1)^{k+1-1} \sum_{I \in (A_k^{n+1} \setminus A_k^n)} \mu(A_I))\right)\right) \\
&= -\mu(A_{n+1}) + \left(\sum_{k=1}^{n} \left((-1)^{k-1} \sum_{I \in (A_k^{n+1} \setminus A_k^n)} \mu(A_I))\right)\right).
\end{aligned}$$

Further, for any $I \in A_k^{n+1} \setminus A_k^n$, $A_I = A_J \cap A_{n+1}$ for $J \in A_k^n$. Hence, the latter sum reads as

$$-\mu(A_{n+1}) + \left(\sum_{k=1}^{n} \left((-1)^{k-1} \sum_{I \in A_k^n} \mu(A_I \cap A_{n+1}))\right)\right).$$

If we add and subtract $\sum_{k=1}^{n+1}((-1)^{k-1} \sum_{I \in A_k^{n+1}} \mu(A_I)))$ from (23), we obtain that

$$\mu\left(\bigcup_{i=1}^{n+1} A_i\right) = \sum_{k=1}^{n+1} \left((-1)^{k-1} \sum_{I \in A_k^{n+1}} \mu(A_I))\right).$$

□

B.2 On measure extensions

Proof of Proposition 6.2. The finite case, i.e., there exists $N \in \mathbb{N}$ s.t.

$$\cup_{i \in I}(a_i, b_i] = \cup_{i=1}^{N}(a_i, b_i],$$

follows by induction. The base step of the induction is clear, if $(a,b] \subset (c,d]$, then $c \leq a < b \leq d$ and hence $b-a \leq d-c$. For the induction step assume that (11) holds for $N-1$ intervals. Let $(a,b] \subset \cup_{i=1}^{N}(a_i,b_i]$. We want to show that $b-a \leq \sum_{i=1}^{N}(b_i-a_i)$. Notice first that we can always assume that $b_1 \leq b_2 \leq \cdots \leq b_N$. If not, we can just consider a relabeling and the union would remain unchanged. Assume first that $b \notin (a_N,b_N]$. Then, $b \leq a_N$ since $b > b_N$ is not possible. To see this, assume by contradiction that $b > b_N$. Then, since $b_1 \leq b_2 \leq \cdots \leq b_N$, $b \notin (a_i,b_i]$ for any $i = 1,\dots,N$. Since $b \in (a,b] \subset \cup_{i=1}^{N}(a_i,b_i]$, this is not possible. Hence, $b \notin (a_N,b_N] \Rightarrow b \leq a_N$. Hence, $(a,b] \subset \cup_{i=1}^{N-1}(a_i,b_i]$ since if $y \in (a,b]$, $y \leq b \leq a_N$ and hence $y \notin (a_N,b_N]$. By the induction hypothesis, the result follows. Thus, in the remaining we assume that $b \in (a_N,b_N]$. If $a_N \leq a$, then $a_N \leq a < b \leq b_N$ and the result follows. Hence, assume that $a < a_N$. Then, $(a,a_N] \subset \cup_{i=1}^{N-1}(a_i,b_i]$. This is because $y \in (a,a_N]$ implies that $y \notin (a_N,b_N]$. Further $y \in (a,a_N]$ implies that $a < y \leq a_N < b$ ($b \in (a_N,b_N]$) and hence $(a,a_N] \subset (a,b]$. Since $(a,b]$ is a subset of $\cup_{i=1}^{N}(a_i,b_i]$ it follows that $y \in (a_i,b_i]$ for some $i \neq N$, i.e., $(a,a_N] \subset \cup_{i=1}^{N-1}(a_i,b_i]$. By the induction hypothesis, $\sum_{i=1}^{N-1}(b_i-a_i) \geq a_N-a$. Therefore, $\sum_{i=1}^{N}(b_i-a_i) \geq a_N-a+b_N-a_N \geq a_N-a+b-a_N = b-a$. We use the Heine-Borel theorem for intervals (cf. Proposition 2.9) to prove the infinite case, i.e., $\cup_{i\in I}(a_i,b_i] = \cup_{i=1}^{\infty}(a_i,b_i]$. Suppose that $(a,b] \subset \cup_{i=1}^{\infty}(a_i,b_i]$. Let $\varepsilon > 0$ be s.t. $b-a > \varepsilon$. This is possible since $b \neq a$. Clearly, the family of intervals $(a_i, b_i+\varepsilon 2^{-i})$, $i \in \mathbb{N}$, are s.t.

$$[a+\varepsilon, b] \subset \bigcup_{i\in\mathbb{N}}(a_i, b_i+\varepsilon 2^{-i}).$$

By Proposition 2.9 it follows that there exists $i_1,\dots,i_N$, s.t.

$$[a+\varepsilon, b] \subset \bigcup_{k=1}^{N}(a_{i_k}, b_{i_k}+\varepsilon 2^{-i_k}).$$

Hence, by the finite case,

$$\begin{aligned} b-a+\varepsilon &\leq \sum_{k=1}^{N}(b_{i_k}-a_{i_k}+\varepsilon 2^{-i_k}) = \sum_{k=1}^{N}(b_{i_k}-a_{i_k}) + \varepsilon\sum_{k=1}^{N}2^{-i_k} \\ &\leq \sum_{i=1}^{\infty}(b_i-a_i) + \varepsilon\sum_{i=1}^{\infty}2^{-i} = \sum_{i=1}^{\infty}(b_i-a_i) + \frac{\varepsilon}{2}\sum_{i=0}^{\infty}2^{-i}. \end{aligned}$$

By Exercise 3.14, we obtain that $b-a+\varepsilon \leq \sum_{i=1}^{\infty}(b_i-a_i)+\varepsilon$. This completes the argument. $\square$

Proof of Proposition 6.5. Given any $\xi \in C_{\mathcal{A}}(A)$, $A \in \mathcal{P}(\Omega)$, $v_\rho(\xi) \geq 0$. It follows that $\rho^*(A) \geq 0$ for any $A \in \mathcal{P}(\Omega)$. By assumption, $\emptyset \in \mathcal{A}$. Hence we can take $\xi = \{\emptyset\}$ and have that ξ is a covering of $\emptyset$ by sets from $\mathcal{A}$. Further, since $\rho(\emptyset) = 0$, it follows that $v_\rho(\xi) = 0$. This shows that $\rho^*(\emptyset) \leq 0$. Since $\rho^*(A) \geq 0$ for any $A \in \mathcal{P}(\Omega)$, it follows that $\rho^*(\emptyset) = 0$. Let $A, B \in \mathcal{P}(\Omega)$ s.t. $A \subset B$. Since $A \subset B$, we have that

$$\{v_\rho(\xi) \colon \xi \in C_{\mathcal{A}}(B)\} \subset \{v_\rho(\xi) \colon \xi \in C_{\mathcal{A}}(A)\}.$$

This shows that $\inf_{\xi\in C_{\mathcal{A}}(A)} v_\rho(\xi) \leq \inf_{\xi\in C_{\mathcal{A}}(B)} v_\rho(\xi)$ (cf. Proposition 1.9). To complete the argument, it remains to show that ρ^* is countable subadditive on $\mathcal{P}(\Omega)$. Let $\{A_n \colon n \in \mathbb{N}\} \subset \mathcal{P}(\Omega)$. We need to show that

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

Clearly, if there exists $n \in \mathbb{N}$ s.t. $\rho^*(A_n) = \infty$, $\rho^*(\cup_{n=1}^{\infty}A_n) \leq \sum_{n=1}^{\infty}\rho^*(A_n)$. Thus suppose that for any $n \in \mathbb{N}$, $\rho^*(A_n) < \infty$. Given $n \in \mathbb{N}$, let $\xi_\varepsilon^n = \{U_{n_k} \colon k \in \mathbb{N}\} \in C_{\mathcal{A}}(A_n)$ be a covering of A_n by sets from $\mathcal{A}$ s.t.

$$v_\rho(\xi_\varepsilon^n) < \rho^*(A_n) + \frac{\varepsilon}{2^n}.$$

This is possible since $\inf_{\xi \in C_{\mathcal{A}}(A_n)} v_\rho(\xi)$ exists as an element of the real numbers (cf. Proposition 1.10). Since, $\{U_{n_k} : k \in \mathbb{N}, n \in \mathbb{N}\}$ is a covering of $\cup_{n\in\mathbb{N}} A_n$ by sets from $\mathcal{A}$, it follows that

$$\rho^*\Big(\bigcup_{n\in\mathbb{N}} A_n\Big) \le \sum_{n=1}^{\infty} v_\rho(\xi_\varepsilon^n) < \sum_{n=1}^{\infty} \rho^*(A_n) + \varepsilon.$$

Since ε was arbitrary, the result follows (cf. Example 1.13). $\square$

Proposition B.2. *Let μ^* be an outer measure on $\mathcal{P}(\Omega)$. Suppose that $\{A_i : i \in I\} \subset \mathcal{M}(\mu^*)$ is disjoint, where I is either finite or countably infinite. Then, for any $E \in \mathcal{P}(\Omega)$,*

$$\mu^*\Big(\Big(\bigcup_{i\in I} A_i\Big) \cap E\Big) = \sum_{i\in I} \mu^*(A_i \cap E).$$

Proof. Suppose first that I is finite, i.e., $\{A_i : i = 1,\dots,n\} \subset \mathcal{M}(\mu^*)$, $n \in \mathbb{N}$. We prove by induction that for any $n \in \mathbb{N}$, $\mu^*((\cup_{i=1}^n A_i) \cap E) = \sum_{i=1}^n \mu^*(A_i \cap E)$. If $n = 1$, then the result follows immediately. If $n = 2$, and $A_1 \cup A_2 = \Omega$, we have that $A_2 = A_1^c$ and since $A_1 \in \mathcal{M}(\mu^*)$, it follows that

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

Suppose that $A_1 \cup A_2$ is a proper subset of Ω, i.e., $\Omega \setminus (A_1 \cup A_2) \neq \emptyset$. Since A_1 and A_2 are disjoint and $A_1^c \subset A_2$, we have that $((E\cap(A_1\cup A_2))\cap A_1) = E\cup A_1$ and $((E\cap(A_1\cup A_2))\cap A_1^c) = E \cup A_2$. Hence, since $A_1 \in \mathcal{M}(\mu^*)$, it follows that

$$\underbrace{\mu^*\big((E \cap (A_1 \cup A_2)) \cap A_1\big)}_{=\mu^*(E\cup A_1)} + \underbrace{\mu^*\big((E \cap (A_1 \cup A_2)) \cap A_1^c\big)}_{=\mu^*(E\cup A_2)} = \mu^*(E \cap (A_1 \cup A_2)). \tag{24}$$

Assume that $\mu^*((\cup_{i=1}^{n-1} A_i) \cap E) = \sum_{i=1}^{n-1} \mu^*(A_i \cap E)$. Using (24), we have that

$$\begin{aligned}\mu^*\Big(\Big(\bigcup_{i=1}^{n} A_i\Big) \cap E\Big) &= \mu^*\Big(\Big(\bigcup_{i=1}^{n-1} A_i \cup A_n\Big) \cap E\Big)\\ &= \mu^*\Big(\Big(\bigcup_{i=1}^{n-1} A_i\Big) \cap E\Big) + \mu^*(A_n \cap E).\end{aligned}$$

Then, by the induction hypothesis, the result follows. Assume now that I is countably infinite, i.e., $\cup_{i\in I} A_i = \cup_{i\in\mathbb{N}} A_i$. Since μ^* is an outer measure, it satisfies (ii) of Definition 6.2. It follows that

$$\mu^*\Big(\Big(\bigcup_{i\in\mathbb{N}} A_i\Big) \cap E\Big) \ge \mu^*\Big(\Big(\bigcup_{i=1}^{n} A_i\Big) \cap E\Big) = \sum_{i=1}^{n} \mu^*(A_i \cap E).$$

If we let $n \to \infty$, we obtain that

$$\mu^*\Big(\Big(\bigcup_{i\in\mathbb{N}} A_i\Big) \cap E\Big) \ge \sum_{i\in\mathbb{N}} \mu^*(A_i \cap E).$$

This completes the argument since the other inequality follows from the fact that μ^* is countable subadditive on $\mathcal{P}(\Omega)$. $\square$

Proof of Proposition 6.6. We first notice that under the assumption that (I) is true, i.e., $\mathcal{M}(\mu^*)$ is a σ-field, item (II) is to show that for any disjoint collection $\{A_i : i \in \mathbb{N}\} \subset \mathcal{M}(\mu^*)$, $\mu^*(\cup_{i\in\mathbb{N}} A_i) = \sum_{i\in\mathbb{N}} \mu^*(A_i)$. This follows immediately from Proposition B.2 if we take $E = \Omega$. Hence, it remains to show that $\mathcal{M}(\mu^*)$ is a σ-field. We need to verify that $\mathcal{M}(\mu^*)$ satisfies the items of Definition 4.1. Let $E \in \mathcal{P}(\Omega)$. We have that

$$\mu^*(\Omega \cap E) + \mu^*(\Omega^c \cap E) = \mu^*(E).$$

Thus, $\Omega \in \mathcal{M}(\mu^*)$. Let $A \in \mathcal{M}(\mu^*)$, then

$$\mu^*(A^c \cap E) + \mu^*((A^c)^c \cap E) = \mu^*(E),$$

since $A \in \mathcal{M}(\mu^*)$. Thus, items (i) and (ii) of Definition 4.1 are clearly satisfied. We notice that for any given $A \in \mathcal{P}(\Omega)$,

$$E = \Omega \cap E = (A \cup A^c) \cap E = (A \cap E) \cup (A^c \cap E).$$

Hence, since μ^* is an outer measure and therefore countable subadditive on $\mathcal{P}(\Omega)$, it follows that for any $A \in \mathcal{P}(\Omega)$, $\mu^*(E) \le \mu^*(A \cap E) + \mu^*(A^c \cap E)$. Thus, A is μ^* measurable if A is s.t. for any $E \in \mathcal{P}(\Omega)$, $\mu^*(E) \ge \mu^*(A \cap E) + \mu^*(A^c \cap E)$. Hence, if μ^* is an outer measure,

$$\mathcal{M}(\mu^*) = \{A \in \mathcal{P}(\Omega) \colon \mu^*(A \cap E) + \mu^*(A^c \cap E) \le \mu^*(E)\ \forall\, E \in \mathcal{P}(\Omega)\}.$$

We show that if $A, B \in \mathcal{M}(\mu^*)$, then $A \cup B \in \mathcal{M}(\mu^*)$. Since $A \in \mathcal{M}(\mu^*)$,

$$\begin{aligned}\mu^*\big((A \cup B) \cap E\big) &= \mu^*\big(A \cap \big((A \cup B) \cap E\big)\big) + \mu^*\big(A^c \cap \big((A \cup B) \cap E\big)\big)\\ &\le \mu^*\big(A \cap E\big) + \mu^*\big(A^c \cap B \cap E\big).\end{aligned}$$

It follows that

$$\begin{aligned}&\mu^*\big((A \cup B) \cap E\big) + \mu^*\big((A \cup B)^c \cap E\big)\\ &\le \mu^*\big(A \cap E\big) + \mu^*\big(A^c \cap B \cap E\big) + \mu^*\big(A^c \cap B^c \cap E\big).\end{aligned}$$

Since $B \in \mathcal{M}(\mu^*)$, $\mu^*(A^c \cap E) = \mu^*(B^c \cap A^c \cap E) + \mu^*(B \cap A^c \cap E)$. Hence,

$$\mu^*\big((A \cup B) \cap E\big) + \mu^*\big((A \cup B)^c \cap E\big) \le \mu^*\big(A \cap E\big) + \mu^*(A^c \cap E) = \mu^*(E),$$

which shows that $A \cup B \in \mathcal{M}(\mu^*)$. By induction, if $A_1, \dots, A_n \in \mathcal{M}(\mu^*)$, we have that $\cup_{i=1}^n A_i \in \mathcal{M}(\mu^*)$. To show (iii) of Definition 4.1, assume first that $\{A_i : i \in \mathbb{N}\} \subset \mathcal{M}(\mu^*)$ is disjoint. Write $A = \cup_{i\in\mathbb{N}} A_i$. We show that $A \in \mathcal{M}(\mu^*)$. We know that for any $n \in \mathbb{N}$, $F_n = \cup_{i=1}^n A_i \in \mathcal{M}(\mu^*)$, i.e., $\mu^*(E) = \mu^*(F_n \cap E) + \mu^*(F_n^c \cap E)$. Using Proposition B.2, we know that $\mu^*(F_n \cap E) = \sum_{i=1}^n \mu^*(A_i \cap E)$. Also, $A^c \subset F_n^c$ ($F_n \subset A$). Hence, $\mu^*(F_n^c \cap E) \ge \mu^*(A^c \cap E)$. Therefore, we obtain, $\mu^*(E) \ge \sum_{i=1}^n \mu^*(A_i \cap E) + \mu^*(A^c \cap E)$. If we let $n \to \infty$, we obtain $\mu^*(E) \ge \sum_{i=1}^\infty \mu^*(A_i \cap E) + \mu^*(A^c \cap E)$. Thus, using Proposition B.2 again, we have that $\mu^*(E) \ge \mu^*(A \cap E) + \mu^*(A^c \cap E)$ and thus $A \in \mathcal{M}(\mu^*)$. Let now $\{B_i : i \in \mathbb{N}\} \subset \mathcal{M}(\mu^*)$, not necessarily disjoint. Write $B = \cup_{i\in\mathbb{N}} B_i$. We want to show that $B \in \mathcal{M}(\mu^*)$. Let $A_1 = B_1$, $A_2 = B_2 \setminus B_1 = B_2 \cap B_1^c$ and so on until we define

$$A_i = B_i \setminus \left(\bigcup_{k=1}^{i-1} B_k \right) = B_i \cap B_1^c \cap \ldots \cap B_{i-1}^c.$$

Then, see the proof of item (vii) in Proposition 5.1, $\{A_i : i \in \mathbb{N}\}$ is disjoint and s.t. for any $n \in \mathbb{N}$, $\cup_{i=1}^n A_i = \cup_{i=1}^n B_i$. In particular, $\cup_{i\in\mathbb{N}} A_i = \cup_{i\in\mathbb{N}} B_i$. Clearly, $\{A_i : i \in \mathbb{N}\} \subset \mathcal{M}(\mu^*)$ and hence, since $\{A_i : i \in \mathbb{N}\}$ is disjoint, $\cup_{i\in\mathbb{N}} A_i \in \mathcal{M}(\mu^*)$ and therefore $\cup_{i\in\mathbb{N}} B_i \in \mathcal{M}(\mu^*)$ as well. This completes the proof of item (I). $\square$

Proof of Proposition 6.7. Using Proposition 6.5, we know that the function

$$\rho^*(A) = \inf_{\xi \in C_{\mathcal{A}}(A)} v_\rho(\xi), \quad A \in \mathcal{P}(\Omega),$$

is an outer measure on $\mathcal{P}(\Omega)$. We also know that $\mathcal{M}(\rho^*)$ is a σ-field and ρ^* restricted to $\mathcal{M}(\rho^*)$ is a measure (cf. Proposition 6.6). As a first step, we show that $\mathcal{M}(\rho^*)$ contains $\mathcal{A}$, i.e., $\mathcal{A} \subset \mathcal{M}(\rho^*)$. Let $A \in \mathcal{A}$. We need to show that for any $E \in \mathcal{P}(\Omega)$,

$$\rho^*(E) \geq \rho^*(A \cap E) + \rho^*(A^c \cap E). \tag{25}$$

It is clear that if $\rho^*(E) = \infty$, (25) is true. Thus consider the case where $\rho^*(E) < \infty$. We apply the same strategy as in the proof of Proposition 6.5 and choose for any $\varepsilon > 0$, a covering $\xi_\varepsilon = \{U_n : n \in \mathbb{N}\} \in C_{\mathcal{A}}(E)$ s.t. $v_\rho(\xi) = \sum_{n\in\mathbb{N}} \rho(U_n) < \rho^*(E) + \varepsilon$. Since $\xi_\varepsilon \in C_{\mathcal{A}}(E)$, we know that for any $n \in \mathbb{N}$, $U_n \in \mathcal{A}$. In particular, since $\mathcal{A}$ is a semiring, $B_n = A \cap U_n \in \mathcal{A}$ for any $n \in \mathbb{N}$. Hence, since $B_n \subset U_n$, we have that for any $n \in \mathbb{N}$, the set $U_n \setminus B_n$ has the form $\cup_{k=1}^{m_n} C_{n_k}$ where $\{C_{n_k} : k = 1, \ldots, m_n\} \subset \mathcal{A}$ is disjoint. Now we notice the following,

1. $U_n = (U_n \setminus B_n) \cup B_n = (\cup_{k=1}^{m_n} C_{n_k}) \cup B_n$, where $\{C_{n_1}, \ldots, C_{n_{m_n}}, B_n\}$ is disjoint;
2. $A \cap E \subset A \cap (\cup_{n\in\mathbb{N}} U_n) = \cup_{n\in\mathbb{N}}(A \cap U_n) = \cup_{n\in\mathbb{N}} B_n$;
3. $U_n \setminus B_n = U_n \cap B_n^c = U_n \cap A^c$;
4. $A^c \cap E \subset A^c \cap (\cup_{n\in\mathbb{N}} U_n) = \cup_{n\in\mathbb{N}}(A^c \cap U_n) = \cup_{n\in\mathbb{N}}(\cup_{k=1}^{m_n} C_{n_k})$.

Therefore, by 1, since ρ is finitely additive on $\mathcal{A}$, for any $n \in \mathbb{N}$, $\sum_{k=1}^{m_k} \rho(C_{n_k}) + \rho(B_n) = \rho((\cup_{k=1}^{m_n} C_{n_k}) \cup B_n) = \rho(U_n)$. Then, using 2 and 4 and the fact that ρ^* is an outer measure it follows that

$$\begin{aligned}
\rho^*(A \cap E) + \rho^*(A^c \cap E) &\leq \rho^*\left(\bigcup_{n\in\mathbb{N}} B_n\right) + \rho^*\left(\bigcup_{n\in\mathbb{N}}\left(\bigcup_{k=1}^{m_n} C_{n_k}\right)\right) \\
&\leq \sum_{n\in\mathbb{N}} \rho^*(B_n) + \sum_{n\in\mathbb{N}}\left(\sum_{k=1}^{m_k} \rho^*(C_{n_k})\right) \\
&\leq \sum_{n\in\mathbb{N}} \rho(B_n) + \sum_{n\in\mathbb{N}}\left(\sum_{k=1}^{m_k} \rho(C_{n_k})\right) \\
&= \sum_{n\in\mathbb{N}}\left(\rho(B_n) + \sum_{k=1}^{m_k} \rho(C_{n_k})\right) \\
&= \sum_{n\in\mathbb{N}} \rho\left(B_n \cup \left(\bigcup_{k=1}^{m_n} C_{n_k}\right)\right) = \sum_{n\in\mathbb{N}} \rho(U_n) < \rho^*(E) + \varepsilon.
\end{aligned}$$

Since $\varepsilon > 0$ was arbitrary, (25) is satisfied. Therefore, $\mathcal{A} \subset \mathcal{M}(\rho^*)$. We show that $\rho^*(A) = \rho(A)$ for any $A \in \mathcal{A}$. Notice that we have already shown this for the special case where $\rho = \ell$ (cf. Proposition 6.4 in Example 6.1). To see the general case, let $A \in \mathcal{A}$ and consider any covering $\xi = \{U_n : n \in \mathbb{N}\} \in C_{\mathcal{A}}(A)$. In particular, $A \subset \cup_{n\in\mathbb{N}} U_n$, and therefore, $A = (\cup_{n\in\mathbb{N}} U_n) \cap A = \cup_{n\in\mathbb{N}}(U_n \cap A)$. Since ρ is countable subadditive on $\mathcal{A}$, it follows that $\rho(A) = \rho(\cup_{n\in\mathbb{N}}(U_n \cap A)) \leq \sum_{n\in\mathbb{N}} \rho(U_n \cap A)$. By Exercise 6.5, ρ is monotone and hence $\rho(A) \leq \sum_{n\in\mathbb{N}} \rho(U_n) = v_\rho(\xi)$. Therefore, it follows that $\rho(A) \leq \inf_{\xi\in C_{\mathcal{A}}(A)} v_\rho(\xi) = \rho^*(A)$, since $\xi \in C_{\mathcal{A}}(A)$ was arbitrary. The other inequality follows immediately, since if $A \in \mathcal{A}$, $\{A\} \in C_{\mathcal{A}}(A)$ and $\rho(A) = v_\rho(\{A\}) \geq \rho^*(A)$. Hence, ρ^* and ρ agree on $\mathcal{A}$. By Proposition 6.6, $\mathcal{M}(\rho^*)$ is a σ-filed. Further, $\mathcal{M}(\rho^*)$ contains $\mathcal{A}$. By definition, $\sigma(\mathcal{A})$ is the smallest σ-field that contains $\mathcal{A}$. Hence, $\sigma(\mathcal{A}) \subset \mathcal{M}(\rho^*)$. Using Proposition 6.6 again, we know that $\rho^*|_{\mathcal{M}(\rho^*)}$ is a measure. Since $\sigma(\mathcal{A})$ is a σ-field and $\sigma(\mathcal{A}) \subset \mathcal{M}(\rho^*)$, it follows that $\rho^*|_{\sigma(\mathcal{A})}$ is a measure as well. Hence, we set $\rho_\uparrow = \rho^*|_{\sigma(\mathcal{A})}$ and obtain a measure on $\sigma(\mathcal{A})$ which is s.t. for any $A \in \mathcal{A}$, $\rho_\uparrow(A) = \rho^*|_{\sigma(\mathcal{A})}(A) = \rho^*(A) = \rho(A)$. $\square$

B.3 π-λ theorem

Definition B.1. *Let $\Omega \neq \emptyset$. A collection $\mathscr{P} \subset \mathcal{P}(\Omega)$ is called a π-system if $A, B \in \mathscr{P}$ implies that $A \cap B \in \mathscr{P}$.*

Definition B.2. *Let $\Omega \neq \emptyset$. A collection $\mathscr{L} \subset \mathcal{P}(\Omega)$ is called a λ-system if*

(i) $\Omega \in \mathscr{L}$;

(ii) $A \in \mathscr{L}$ *implies that* $A^c \in \mathscr{L}$;

(iii) if $\{A_i : i \in \mathbb{N}\} \subset \mathscr{L}$ *s.t.* $\{A_i : i \in \mathbb{N}\}$ *is disjoint, then* $\cup_{i\in\mathbb{N}} A_i \in \mathscr{L}$.

Proposition B.3. *Let $\Omega \neq \emptyset$. If $\mathcal{F}$ is a π-system and a λ-system, then it is a σ-field on Ω.*

Proof. Since $\mathcal{F}$ is a λ-system, it follows that $\Omega \in \mathcal{F}$ and $\mathcal{F}$ is closed under the formation of complements. Hence it remains to show item (iii) of Definition 4.1. Let $\{A_i : i \in \mathbb{N}\} \subset \mathcal{F}$. Given $i \in \mathbb{N}$, we write

$$B_i = A_i \setminus \left(\bigcup_{j=1}^{i-1} A_j \right) = A_i \cap A_1^c \cap \ldots \cap A_{i-1}^c.$$

We already know that $\{B_i : i \in \mathbb{N}\}$ is disjoint (recall the proof of item (vii) of Proposition 5.1). Since $\mathcal{F}$ is a π-system, $B_i \in \mathcal{F}$ for any $i \in \mathbb{N}$ and hence, by item (iii) of Definition B.2, $\cup_{i\in\mathbb{N}} B_i \in \mathcal{F}$. Then, $\mathcal{F}$ is a σ-field, since for any $n \in \mathbb{N}$, $\cup_{i=1}^n B_i = \cup_{i=1}^n A_i$ and hence $\cup_{i\in\mathbb{N}} B_i = \cup_{i\in\mathbb{N}} A_i$. □

The following is known as Dynkin's π-λ theorem.

Proposition B.4. *Let $\Omega \neq \emptyset$. If $\mathscr{P}$ is a π-system and $\mathscr{L}$ is a λ system then, if $\mathscr{P} \subset \mathscr{L}$, it follows that $\sigma(\mathscr{P}) \subset \mathscr{L}$.*

Proof. Suppose that $\mathscr{P} \subset \mathscr{L}$. Define the set

$$\mathscr{C} = \{\mathcal{L} : \mathcal{L} \text{ is a } \lambda\text{-system s.t. } \mathscr{P} \subset \mathcal{L}\}.$$

Then, similar to the definition of the σ-field generated by a family of subsets of Ω, we set

$$\mathscr{L}_0 = \bigcap_{\mathcal{L} \in \mathscr{C}} \mathcal{L}.$$

Again, $\mathscr{L}_0$ is not empty, since $\mathcal{P}(\Omega) \in \mathscr{C}$. Further, $\mathscr{L}_0$ is a λ-system (cf. Exercise 4.5). Upon the assumption that $\mathscr{P} \subset \mathscr{L}$, we obtain $\mathscr{P} \subset \mathscr{L}_0 \subset \mathscr{L}$. If $\mathscr{L}_0$ is a π-system, then $\mathscr{L}_0$ is a σ-field on Ω (Proposition B.3). Hence, since $\sigma(\mathscr{P})$ is the smallest σ-field that contains $\mathscr{P}$, it follows that $\sigma(\mathscr{P}) \subset \mathscr{L}_0 \subset \mathscr{L}$ and we are done. Therefore, we prove that $\mathscr{L}_0$ is a π-system. Given any $A \in \mathcal{P}(\Omega)$ we define

$$\mathscr{L}_A = \{B \in \mathcal{P}(\Omega) : A \cap B \in \mathscr{L}_0\}.$$

Suppose that $A \in \mathscr{P}$. Then $\mathscr{L}_A$ is a λ-system. To see it, notice that $\Omega \in \mathscr{L}_A$, since $A \cap \Omega = A \in \mathscr{L}_0$ (recall that by definition of $\mathscr{L}_0$, $A \in \mathscr{P}$ implies that $A \in \mathcal{L}$ for any $\mathcal{L} \in \mathscr{C}$). Suppose that $B \in \mathscr{L}_A$. Then, $A \cap B \in \mathscr{L}_0$. We also know that $A \cap \Omega \in \mathscr{L}_0$. Hence, since $\mathscr{L}_0$ is a λ-system, it contains $(A\cap\Omega)\cap(A\cap B)^c = (A\cap\Omega)\setminus(A\cap B)$ (notice that $(A\cap\Omega)^c \cap A \cap B = \emptyset$). Then, since $(A \cap \Omega) \setminus (A \cap B) = A \cap (\Omega \setminus B)$, it follows that $\Omega \setminus B = B^c \in \mathscr{L}_A$. Suppose that $\{B_i : i \in \mathbb{N}\} \subset \mathscr{L}_A$ disjoint, then $\{A \cap B_i : i \in \mathbb{N}\} \subset \mathscr{L}_0$, where also $\{A \cap B_i : i \in \mathbb{N}\}$ is disjoint. Thus, since $\mathscr{L}_0$ is a λ-system, $\cup_{i\in\mathbb{N}}(A\cap B_i) = A\cap(\cup_{i\in\mathbb{N}} B_i) \in \mathscr{L}_0$, i.e., $\cup_{i\in\mathbb{N}} B_i \in \mathscr{L}_A$. Therefore, we have shown that if $A \in \mathscr{P}$, then $\mathscr{L}_A$ is a λ-system. Further, if $A \in \mathscr{P}$, then $\mathscr{P} \subset \mathscr{L}_A$, since in this case, for any $B \in \mathscr{P}$, $A \cap B \in \mathscr{P}$ ($\mathscr{P}$ is a π-system) and since $\mathscr{P} \subset \mathscr{L}_0$, it follows that $A \cap B \in \mathscr{L}_0$, i.e., $B \in \mathscr{L}_A$. Then, since $\mathscr{L}_0$ is the smallest λ-system that contains $\mathscr{P}$, $\mathscr{L}_0 \subset \mathscr{L}_A$ if $A \in \mathscr{P}$. This means that if $A \in \mathscr{P}$ and $B \in \mathscr{L}_0$, then $B \in \mathscr{L}_A$ and hence $A \in \mathscr{L}_B$. Notice that $B \in \mathscr{L}_A$ if and only if $A \in \mathscr{L}_B$. Thus, $B \in \mathscr{L}_0$ implies that $\mathscr{P} \subset \mathscr{L}_B$. Then, since $\mathscr{L}_B$ is a λ-system, it follows that $B \in \mathscr{L}_0$ implies that $\mathscr{L}_0 \subset \mathscr{L}_B$. In conclusion, $B \in \mathscr{L}_0$ and $C \in \mathscr{L}_0$ imply that $C \in \mathscr{L}_B$ and hence $B \cap C \in \mathscr{L}_0$. This shows that $\mathscr{L}_0$ is a π-system. $\square$

Proof of Proposition 6.9. Assume that μ_1 is σ-finite on $\mathcal{A}$ (otherwise μ_2 is). With the terminology of Definition B.1, $\mathcal{A}$ is a π-system. Let $B \in \mathcal{A}$ s.t. $\mu_1(B) < \infty$ (this is possible since μ_1 is σ-finite on $\mathcal{A}$). Hence, since by assumption, μ_1 and μ_2 agree on $\mathcal{A}$, $\mu_1(B) = \mu_2(B)$. Define

$$\mathscr{L}_B = \{A \in \sigma(\mathcal{A})\colon \mu_1(B \cap A) = \mu_2(B \cap A)\}.$$

Notice that $\mathcal{A} \subset \mathscr{L}_B$ since if $A \in \mathcal{A}$, $A \cap B \in \mathcal{A}$ ($\mathcal{A}$ is a π-system) and as μ_1 and μ_2 agree on $\mathcal{A}$, $\mu_1(B \cap A) = \mu_2(B \cap A)$, i.e., $A \in \mathscr{L}_B$. We show that $\mathscr{L}_B$ is a λ-system. Clearly, $\Omega \in \mathscr{L}_B$, since $\Omega \in \sigma(\mathcal{A})$ s.t. $\mu_1(B \cap \Omega) = \mu_1(B) = \mu_2(B) = \mu_2(B \cap \Omega)$. If $A \in \mathscr{L}_B$, then $A^c \in \sigma(\mathcal{A})$. Then, since $B \cap A^c = B \setminus (B \cap A)$ and $\mu_1(A \cap B) = \mu_2(A \cap B) < \infty$, we obtain (cf. item (iii) of Proposition 5.1),

$$\mu_1(B \cap A^c) = \mu_1(B \setminus (B \cap A)) = \mu_1(B) - \mu_1(B \cap A) = \mu_2(B) - \mu_2(B \cap A) = \mu_2(B \cap A^c),$$

it follows that $A^c \in \mathscr{L}_B$. If $\{A_i\colon i \in \mathbb{N}\} \subset \mathscr{L}_B$, disjoint, then $\cup_{i\in\mathbb{N}} A_i \in \sigma(\mathcal{A})$ and $\{B \cap A_i\colon i \in \mathbb{N}\}$ is disjoint as well. Then,

$$\begin{aligned}\mu_1(B \cap (\cup_{i\in\mathbb{N}} A_i)) &= \mu_1(\cup_{i\in\mathbb{N}}(B \cap A_i)) \\ &= \sum_{i\in\mathbb{N}} \mu_1(B \cap A_i) = \sum_{i\in\mathbb{N}} \mu_2(B \cap A_i) = \mu_2(B \cap (\cup_{i\in\mathbb{N}} A_i)),\end{aligned}$$

and hence $\cup_{i\in\mathbb{N}} A_i \in \mathscr{L}_B$. Thus, $\mathscr{L}_B$ is a λ-system that contains the π-system $\mathcal{A}$ and with Proposition B.4, we obtain that $\sigma(\mathcal{A}) \subset \mathscr{L}_B$. Therefore, we have shown that $\sigma(\mathcal{A}) \subset \mathscr{L}_B$ for any set $B \in \mathcal{A}$ for which $\mu_1(B) < \infty$. We consider a collection $\{B_i\colon i \in \mathbb{N}\} \subset \mathcal{A}$ s.t. $\cup_{i\in\mathbb{N}} B_i = \Omega$ and for any $i \in \mathbb{N}$, $\mu_1(B_i) < \infty$. In particular, given any $n \in \mathbb{N}$ and $I \subset \{1, \dots, n\}$, since $\mathcal{A}$ is a π-system, $\cap_{i\in I} B_i \in \mathcal{A}$ and $\mu_1(\cap_{i\in I} B_i) < \infty$. Therefore, $\sigma(\mathcal{A}) \subset \mathscr{L}_{\cap_{i\in I} B_i}$. This shows that for any $A \in \sigma(\mathcal{A})$,

$$\mu_1\bigg(\Big(\bigcap_{i\in I} B_i\Big) \cap A\bigg) = \mu_2\bigg(\Big(\bigcap_{i\in I} B_i\Big) \cap A\bigg).$$

Then, upon (Proposition B.1) we obtain that for any $n \in \mathbb{N}$,

$$\mu_1\bigg(\bigcup_{i=1}^{n}(B_i \cap A)\bigg) = \mu_2\bigg(\bigcup_{i=1}^{n}(B_i \cap A)\bigg).$$

Then, we apply item (v) of Proposition 5.1, and conclude that for any $A \in \sigma(\mathcal{A})$,

$$\mu_1(A) = \mu_1\bigg(\bigcup_{i\in\mathbb{N}}(B_i \cap A)\bigg) = \mu_2\bigg(\bigcup_{i\in\mathbb{N}}(B_i \cap A))\bigg) = \mu_2(A).$$

$\square$

B.4 On measurable functions

Proof of Proposition 7.4. We first show that if for any $i = 1, \dots, k$, $f_i\colon \Omega \to \mathbb{R}$ is $\mathcal{F}/\mathfrak{B}(\mathbb{R})$ measurable, then, f is $\mathcal{F}/\mathfrak{B}(\mathbb{R}^k)$ measurable. By Exercise 4.8, we have that $\mathfrak{B}(\mathbb{R}^k) = \sigma(\mathcal{R}'_k)$, where

$$\mathcal{R}'_k = \{(-\infty, x_1] \times \cdots (-\infty, x_k]\colon x = (x_1, \dots, x_k) \in \mathbb{R}^k\}.$$

Hence, by Proposition 7.1, it is sufficient to show that for any $R \in \mathcal{R}'_k$, $f^{-1}(R) \in \mathcal{F}$. Therefore, let $R \in \mathcal{R}'_k$, i.e.,

$$R = (-\infty, x_1] \times \cdots (-\infty, x_k], \quad x = (x_1, \dots, x_k) \in \mathbb{R}^k.$$

Then,

$$f^{-1}(R) = \{\omega \in \Omega \colon f(\omega) \in R\} = \bigcap_{i=1}^{k} \{\omega \in \Omega \colon f_i(\omega) \in (-\infty, x_1]\} = \bigcap_{i=1}^{k} f_i^{-1}((-\infty, x_1]).$$

Then, since for any $i = 1, \dots, k$, $(-\infty, x_1] \in \mathfrak{B}(\mathbb{R})$ and f_i is $\mathcal{F}/\mathfrak{B}(\mathbb{R})$ measurable, it follows that $f_i^{-1}((-\infty, x_1]) \in \mathcal{F}$. Therefore, $f^{-1}(R) \in \mathcal{F}$ as well. For the other direction, assume that f is $\mathcal{F}/\mathfrak{B}(\mathbb{R}^k)$ measurable. Given any $i = 1, \dots, k$, let

$$A_i^n = (-\infty, n] \times \cdots \times \underbrace{(-\infty, x]}_{\text{position } i} \times \cdots \times (-\infty, n],$$

where $x \in \mathbb{R}$ is arbitrary. Then,

$$f^{-1}(A_i^n) = f_1^{-1}((-\infty, n]) \cap \cdots \cap \underbrace{f_i^{-1}((-\infty, x])}_{\text{position } i} \cap \cdots \cap f_k^{-1}((-\infty, n]).$$

Then,

$$\bigcup_{n \in \mathbb{N}} A_i^n = \mathbb{R} \times \cdots \times \underbrace{(-\infty, x]}_{\text{position } i} \times \cdots \times \mathbb{R}.$$

Therefore,

$$\begin{aligned}
\bigcup_{n \in \mathbb{N}} f^{-1}(A_i^n) &= f^{-1}\left(\bigcup_{n \in \mathbb{N}} A_i^n\right) \\
&= f_1^{-1}(\mathbb{R}) \cap \cdots \cap \underbrace{f_i^{-1}((-\infty, x])}_{\text{position } i} \cap \cdots \cap f_k^{-1}(\mathbb{R}) = f_i^{-1}((-\infty, x]).
\end{aligned}$$

Since f is $\mathcal{F}/\mathfrak{B}(\mathbb{R}^k)$ measurable, for any $i = 1, \dots, k$, $\cup_{n \in \mathbb{N}} f^{-1}(A_i^n) \in \mathcal{F}$. Thus, for any $i = 1, \dots, k$ and any $x \in \mathbb{R}$, $f_i^{-1}((-\infty, x]) \in \mathcal{F}$. Then, $\mathfrak{B}(\mathbb{R}) = \sigma(\{(-\infty, x] \colon x \in \mathbb{R}\})$ and hence, by Proposition 7.1, f_i is $\mathcal{F}/\mathfrak{B}(\mathbb{R})$ measurable. This completes the proof. $\square$

Proof of Proposition 7.7. We show by induction (over $n \geq 2$) that there exists $m \in \mathbb{N}$, s.t. f satisfies

$$f(\omega) = \sum_{i=1}^{m} \alpha_i^* \mathbb{1}_{A_i^*}(\omega), \quad \omega \in \Omega,$$

where $\{A_1^*, \dots, A_m^*\} \subset \mathcal{F}$ is disjoint. Let $n = 2$, i.e., $f(\omega) = \alpha_1 \mathbb{1}_{A_1}(\omega) + \alpha_2 \mathbb{1}_{A_2}(\omega)$, $A_1, A_2 \in \mathcal{F}$. We define the sets $A_1^* = A_1 \setminus A_2$, $A_2^* = A_2 \setminus A_1$ and $A_3^* = A_1 \cap A_2$. Further, we set $\alpha_1^* = \alpha_1$, $\alpha_2^* = \alpha_2$ and $\alpha_3^* = \alpha_1 + \alpha_2$. Then, $\{A_1^*, A_2^*, A_3^*\} \subset \mathcal{F}$ is disjoint. Also, for any $\omega \in \Omega$,

$$f(\omega) = \alpha_1^* \mathbb{1}_{A_1^*}(\omega) + \alpha_2^* \mathbb{1}_{A_2^*}(\omega) + \alpha_3^* \mathbb{1}_{A_3^*}(\omega).$$

With regard to the induction step, assume that $f(\omega) = \sum_{i=1}^{n-1} \alpha_i \mathbb{1}_{A_i}(\omega) + \alpha_n \mathbb{1}_{A_n}(\omega)$, $A_i \in \mathcal{F}$, $i = 1, \dots, n$, and the induction hypothesis is that $\{A_1, \dots, A_{n-1}\} \subset \mathcal{F}$ is disjoint. Then, we set

$$A_i^* = A_i \setminus A_n, \quad i = 1, \dots, n-1, \quad A_n^* = A_n \setminus (\cup_{i=1}^{n-1} A_i),$$

and

$$A_{n+i}^* = A_n \cap A_i, \quad i = 1, \dots, n-1.$$

It follows that

$$\{A_1^*, \dots, A_n^*, A_{n+1}^*, \dots, A_{2n-1}^*\} \subset \mathcal{F}$$

is disjoint. We set $\alpha_i^* = \alpha_i$, $i = 1, \dots, n$, and $\alpha_{n+i}^* = \alpha_n + \alpha_i$, $i = 1, \dots, n-1$. It follows that for any $\omega \in \Omega$,

$$f(\omega) = \sum_{i=1}^{2n-1} \alpha_i^* \mathbb{1}_{A_i^*}(\omega).$$

This completes the induction step. Finally, we notice that

$$f(\omega) = \sum_{i=1}^{2n-1} \alpha_i^* \mathbb{1}_{A_i^*}(\omega) + \alpha_{2n}^* \mathbb{1}_{A_{2n}^*}(\omega),$$

with $\alpha_{2n}^* = 0$ and $A_{2n}^* = \Omega \setminus (\cup_{i=1}^{2n-1} A_i^*)$. Thus, $\cup_{i=1}^{2n} A_i^* = \Omega$ and $\{A_i^* : i = 1, \dots, 2n\} \subset \mathcal{F}$ is disjoint. □

Proof of Proposition 7.8. We have that

$$\{\omega \in \Omega \colon f(\omega) < g(\omega)\} = \bigcup_{q \in \mathbb{Q}} \{\omega \in \Omega \colon f(\omega) < q < g(\omega)\}.$$

To see it, take any ω in the left set of the above equation. Then, by Proposition 1.6, there exists $q \in \mathbb{Q}$ s.t. $f(\omega) < q < g(\omega)$. Notice that is is also true if $g(\omega) = \infty$. Clearly, if there exists $q \in \mathbb{Q}$ s.t. $f(\omega) < q < g(\omega)$, then $f(\omega) < g(\omega)$. Then, we have that for any $q \in \mathbb{Q}$,

$$\{\omega \in \Omega \colon f(\omega) < q < g(\omega)\} = \{\omega \in \Omega \colon f(\omega) < q\} \cap \{\omega \in \Omega \colon g(\omega) > q\}.$$

Since f and g are both $\mathcal{F}$ measurable, $\{\omega \in \Omega \colon f(\omega) < q\}$ and $\{\omega \in \Omega \colon g(\omega) > q\}$ are members of $\mathcal{F}$. This shows that $\{\omega \in \Omega \colon f(\omega) < g(\omega)\} \in \mathcal{F}$ since $\mathbb{Q}$ is countable. A similar argument can be used to show that $\{\omega \in \Omega \colon f(\omega) > g(\omega)\} \in \mathcal{F}$. Therefore,

$$\mathcal{F} \ni \{\omega \in \Omega \colon f(\omega) = g(\omega)\} = \big(\{\omega \in \Omega \colon f(\omega) < g(\omega)\} \cup \{\omega \in \Omega \colon f(\omega) > g(\omega)\}\big)^c.$$

□

Proof of Proposition 7.9. Let $x \in \overline{\mathbb{R}}$. We have that

$$\{\omega \in \Omega \colon \sup_{n \in E} f_n(\omega) \le x\} = \cap_{n \in E} \{\omega \in \Omega \colon f_n(\omega) \le x\}.$$

Hence, for any $x \in \overline{\mathbb{R}}$, $\{\omega \in \Omega \colon \sup_{n \in E} f_n \le x\} \in \mathcal{F}$ since we have assumed that f_n is $\mathcal{F}$ measurable for any $n \in E$. We notice that

$$\{\omega \in \Omega \colon \sup_{n \in E} f_n(\omega) < \infty\} = \{\omega \in \Omega \colon \sup_{n \in E} f_n(\omega) \le C\},$$

for some $C \in \mathbb{R}$. Therefore, $\{\omega \in \Omega \colon \sup_{n \in E} f_n(\omega) < \infty\} \in \mathcal{F}$. Hence,

$$\{\omega \in \Omega \colon \sup_{n \in E} f_n(\omega) = \infty\} = \{\omega \in \Omega \colon \sup_{n \in E} f_n(\omega) < \infty\}^c \in \mathcal{F}.$$

Also,

$$\{\omega \in \Omega\colon \sup_{n \in E} f_n(\omega) = -\infty\} = \{\omega \in \Omega\colon \sup_{n \in E} f_n(\omega) \leq -\infty\}$$
$$= \cap_{n\in E}\{\omega \in \Omega\colon f_n(\omega) = -\infty\}.$$

Thus, since for any $n \in E$, f_n is $\mathcal{F}$ measurable, we have that$\{\omega \in \Omega\colon f_n(\omega) = -\infty\} \in \mathcal{F}$. Therefore $\{\omega \in \Omega\colon \sup_{n\in E} f_n(\omega) = -\infty\} \in \mathcal{F}$. As in the solution of Exercise 7.4, we define

$$F = \{\omega \in \Omega\colon \sup_{n\in E} f_n(\omega) \in \mathbb{R}\},$$

and set

$$f^*(\omega) = \sup_{n\in E} f_n(\omega)\mathbb{1}_F(\omega), \quad \omega \in \Omega.$$

We notice that $F \in \mathcal{F}$. Also, since for any $x \in \overline{\mathbb{R}}$, $\{\omega \in \Omega\colon \sup_{n\in E} f_n(\omega) \leq x\} \in \mathcal{F}$, we obtain that

$$\{\omega \in \Omega\colon f^*(\omega) \leq x\} \in \mathcal{F}.$$

Hence, f^* is $\mathcal{F}/\mathfrak{B}(\mathbb{R})$ measurable (cf. Proposition 7.3). Let $A \in \mathfrak{B}(\mathbb{R})$. We obtain

$$\{\omega \in \Omega\colon \sup_{n\in E} f_n(\omega) \in A\} = \big(\{\omega \in \Omega\colon f^*(\omega) \in A\} \cap F\big) \cup (\{\omega \in \Omega\colon \sup_{n\in E} f_n(\omega) \in A\} \cap F^c)$$
$$= \{\omega \in \Omega\colon f^*(\omega) \in A\} \cap F \in \mathcal{F}.$$

This shows that $\omega \mapsto \sup_{n\in E} f_n(\omega)$ is $\mathcal{F}$ measurable. A similar argument shows that $\omega \mapsto \inf_{n\in E} f_n(\omega)$ is $\mathcal{F}$ measurable. With respect to (ii) of Proposition 7.9, we have that

$$\omega \mapsto (\liminf_{n\longrightarrow\infty} f_n)(\omega) = \sup_{n\in\mathbb{N}}(\inf_{k\geq n} f_n)(\omega),$$

and

$$\omega \mapsto (\limsup_{n\longrightarrow\infty} f_n)(\omega) = \inf_{n\in\mathbb{N}}(\sup_{k\geq n} f_n)(\omega),$$

are $\mathcal{F}$ measurable as a consequence of item (i) (cf. Exercise 7.2). Item (iii) is a consequence of Proposition 3.23, we have that

$$\{\omega \in \Omega\colon \lim_{n\to\infty} f_n(\omega) = -\infty\}$$
$$= \{\omega \in \Omega\colon (\liminf_{n\longrightarrow\infty} f_n)(\omega) = -\infty\} \cap \{\omega \in \Omega\colon (\limsup_{n\longrightarrow\infty} f_n)(\omega) = -\infty\}.$$

Therefore, since $\liminf_{n\longrightarrow\infty} f_n$ and $\limsup_{n\longrightarrow\infty} f_n$ are $\mathcal{F}$ measurable, we conclude that

$$\{\omega \in \Omega\colon \lim_{n\to\infty} f_n(\omega) = -\infty\} \in \mathcal{F}.$$

Similarly, we obtain

$$\{\omega \in \Omega\colon \lim_{n\to\infty} f_n(\omega) = \infty\}$$
$$= \{\omega \in \Omega\colon (\liminf_{n\longrightarrow\infty} f_n)(\omega) = \infty\} \cap \{\omega \in \Omega\colon (\limsup_{n\longrightarrow\infty} f_n)(\omega) = \infty\} \in \mathcal{F}.$$

Let $A \in \mathfrak{B}(\mathbb{R})$, we have that

$$\{\omega \in \Omega\colon \lim_{n\to\infty} f_n(\omega) \in A\}$$
$$= \{\omega \in \Omega\colon (\liminf_{n\longrightarrow\infty} f_n)(\omega) = (\limsup_{n\longrightarrow\infty} f_n)(\omega)\} \cap \{\omega \in \Omega\colon (\liminf_{n\longrightarrow\infty} f_n)(\omega) \in A\}.$$

Then, $\{\omega \in \Omega\colon \lim_{n\to\infty} f_n(\omega) \in A\} \in \mathcal{F}$ (cf. Proposition 7.8). Using Propositions 3.21 and 7.8, we obtain

$$\{\omega \in \Omega\colon (f_n(\omega))_{n\in\mathbb{N}} \text{ converges}\} = \{\omega \in \Omega\colon (\liminf_{n\to\infty} f_n)(\omega) = (\limsup_{n\to\infty} f_n)(\omega)\} \in \mathcal{F}.$$

This shows item (iv). Finally, to see item (v), we use again Proposition 3.23 and conclude that

$$\begin{aligned}&\{\omega \in \Omega\colon f_n(\omega) \xrightarrow{n\to\infty} f(\omega)\}\\ &= \{\omega \in \Omega\colon (\liminf_{n\to\infty} f_n)(\omega) = f(\omega)\} \cap \{\omega \in \Omega\colon (\limsup_{n\to\infty} f_n)(\omega) = f(\omega)\}.\end{aligned}$$

Then, since by assumption f is $\mathcal{F}$ measurable, the sets $\{\omega \in \Omega\colon (\liminf_{n\to\infty} f_n)(\omega) = f(\omega)\}$ and $\{\omega \in \Omega\colon (\limsup_{n\to\infty} f_n)(\omega) = f(\omega)\}$ are both elements of $\mathcal{F}$ (cf. Proposition 7.8). This completes the proof of the proposition. □

Proof of Proposition 7.10. The proof is to find such an approximating sequence of simple functions. Given any $n \in \mathbb{N}$, we partition $[0, n)$ as follows:

$$\Big[0, n\Big) = \Big[0, \frac{1}{2^n}\Big) \cup \Big[\frac{1}{2^n}, \frac{2}{2^n}\Big) \cup \dots \cup \Big[\frac{n2^n - 1}{2^n}, \frac{n2^n}{2^n}\Big) = \bigcup_{i=1}^{n2^n} \Big[\frac{i-1}{2^n}, \frac{i}{2^n}\Big).$$

Then, we label,

$$I_{n,i} = \Big[\frac{i-1}{2^n}, \frac{i}{2^n}\Big), \quad i = 1, \dots, n2^n,$$

and $I_n = [n, \infty) \cup \{\infty\}$. Accordingly, we define the sets

$$A_{n,i} = \{\omega \in \Omega\colon f(\omega) \in I_{n,i}\}, \quad i = 1, \dots, n2^n,$$

and $A_n = \{\omega \in \Omega\colon f(\omega) \in I_n\}$. We let

$$f_n(\omega) = \sum_{i=1}^{n2^n} \frac{i-1}{2^n} \mathbb{1}_{A_{n,i}}(\omega) + n\mathbb{1}_{A_n}(\omega).$$

Clearly, for any $\omega \in \Omega$, $f_n(\omega) \le f_{n+1}(\omega)$ and for any $n \in \mathbb{N}$, f_n is $\mathcal{F}$ measurable. Suppose first that $f(\omega) = \infty$. Then, for any $n \in \mathbb{N}$, $f_n(\omega) = n$, i.e., $f_n(\omega) \uparrow f(\omega)$. Assume that $f(\omega) \in [0, \infty)$. Then, there exists $N \in \mathbb{N}$ s.t. $f(\omega) < N$. Hence, there exists precisely one $i \in \{1, \dots, N2^N\}$, s.t. $f(\omega) \in I_{N,i}$, i.e., $(i-1)2^{-N} \le f(\omega) < i2^{-N}$. In particular, $f(\omega) \ge (i-1)2^{-N} = f_N(\omega)$. Therefore,

$$|f(\omega) - f_N(\omega)| = f(\omega) - f_N(\omega) < \frac{i}{2^N} - \frac{i-1}{2^N} = \frac{1}{2^N}.$$

Then, for any $n \ge N$, $f(\omega) < N \le n$. Hence, we repeat the above argument and conclude that for any $n \ge N$, there is precisely one $k \in \{1, \dots, n2^n\}$ s.t. $f(\omega) \in I_{n,k}$ and $|f(\omega) - f_n(\omega)| < 2^{-n}$. Therefore, $\lim_{n\to\infty} f_n(\omega) = f(\omega)$. The proof is complete upon application of Proposition 7.7, i.e., for any $n \in \mathbb{N}$ there exists a standard simple function g_n s.t. for any $\omega \in \Omega$, $f_n(\omega) = g_n(\omega)$. □

Proof of Proposition 7.13. Let $g\colon \mathbb{R}^k \to \mathbb{R}$ be $\mathfrak{B}(\mathbb{R}^k)$ measurable and s.t. $h(\omega) = g(f(\omega))$. Then, h is $\sigma(f)$ measurable (Exercise 7.9). For the other direction, we prove the claim in two steps, first we show it for the case where h is a standard simple function and then we approximate a general h with simple functions. Thus, suppose that h is a simple function in standard form (cf. Definition 7.4 with $\mathcal{F} = \sigma(f)$). Then, there exists $g\colon \mathbb{R}^k \to \mathbb{R}$ which is $\mathfrak{B}(\mathbb{R}^k)$ measurable and s.t. $h(\omega) = g(f(\omega))$ (Exercise 7.10). Let $h\colon \Omega \to \mathbb{R}$ be a general $\sigma(f)$

measurable function. By Proposition 7.11, we find a sequence of $\sigma(f)$ measurable standard simple functions $(h_n)_{n\in\mathbb{N}}$ s.t. $h_n \xrightarrow{n\to\infty} h$. By the previous case, for each $n \in \mathbb{N}$, there exists a $\mathfrak{B}(\mathbb{R}^k)$ measurable function g_n s.t. for any $\omega \in \Omega$, $h_n(\omega) = g_n(f(\omega))$. Consider the set $A = \{x \in \mathbb{R}^k : (g_n(x))_{n\in\mathbb{N}} \text{ converges}\}$. Since for any $n \in \mathbb{N}$, g_n is $\mathfrak{B}(\mathbb{R}^k)$ measurable, it follows that $A \in \mathfrak{B}(\mathbb{R}^k)$ (cf. item (iv) of Proposition 7.9). Define

$$g(x) = \begin{cases} \lim_{n\to\infty} g_n(x), & \text{if } x \in A, \\ 0, & \text{otherwise.} \end{cases}$$

We have that $g(x) = \lim_{n\to\infty} g_n(x)\mathbb{1}_A(x)$ and hence, g is $\mathfrak{B}(\mathbb{R}^k)$ measurable by item (iii) of Proposition 7.9. Let $\omega \in \Omega$, then $f(\omega)$ is s.t.

$$\lim_{n\to\infty} g_n(f(\omega)) = \lim_{n\to\infty} h_n(\omega) = h(\omega).$$

Thus, $f(\omega) \in A$ for any $\omega \in \Omega$. Hence, for any $\omega \in \Omega$,

$$h(\omega) = \lim_{n\to\infty} h_n(\omega) = \lim_{n\to\infty} g_n(f(\omega)) = \lim_{n\to\infty} g_n(f(\omega))\mathbb{1}_A(f(\omega)) = g(f(\omega)).$$

□

B.5 On integration: Part I

Proof of Proposition 8.1. Notice first that we have seen in Example 7.7 that f is $\mathcal{F}$ measurable. Let $\eta = \{B_j : j = 1, \dots, M\} \in Z_0^{\mathcal{F}}$. Since $\eta \in Z_0^{\mathcal{F}}$, we can represent the atoms of ξ using the atoms of η and vice versa, i.e.,

$$A_i = A_i \cap \Omega = \bigcup_{j=1}^{M} (A_i \cap B_j), \quad i = 1, \dots, N,$$

and

$$B_j = \bigcup_{i=1}^{N} (A_i \cap B_j), \quad j = 1, \dots, M.$$

Further, since ξ is disjoint, it follows that for any $j = 1, \dots, M$, $\{A_i \cap B_j : i = 1, \dots, N\}$ is disjoint. Therefore,

$$\mu(B_j) = \sum_{i=1}^{N} \mu(A_i \cap B_j).$$

Further, given any $(j, i) \in \{1, \dots, M\} \times \{1, \dots, N\}$, either $A_i \cap B_j \neq \emptyset$, then

$$\Big(\inf_{\omega\in B_j} f(\omega)\Big)\mu(A_i \cap B_j) \le \Big(\inf_{\omega\in B_j\cap A_i} f(\omega)\Big)\mu(A_i \cap B_j) = \alpha_i\mu(A_i \cap B_j),$$

or $A_i \cap B_j = \emptyset$, then, $(\inf_{\omega\in B_j} f(\omega))\mu(A_i \cap B_j) = 0 = \alpha_i\mu(A_i \cap B_j)$. Hence,

$$\begin{aligned} S_\mu^f(\eta) &= \sum_{j=1}^{M} \Big(\inf_{\omega\in B_j} f(\omega)\Big)\Big(\sum_{i=1}^{N} \mu(A_i \cap B_j)\Big) = \sum_{j=1}^{M}\Big(\sum_{i=1}^{N}\Big(\inf_{\omega\in B_j} f(\omega)\Big)\mu(A_i \cap B_j)\Big) \\ &\le \sum_{j=1}^{M}\Big(\sum_{i=1}^{N} \alpha_i\mu(A_i \cap B_j)\Big) = \sum_{i=1}^{N} \alpha_i\Big(\sum_{j=1}^{M} \mu(A_i \cap B_j)\Big) = \sum_{i=1}^{N} \alpha_i\mu(A_i). \end{aligned}$$

This shows that

$$\int_{\Omega} f(\omega)\mu(d\omega) = \sup_{\eta \in Z_0^{\mathcal{F}}} S_\mu^f(\eta) \leq \sum_{i=1}^{N} \alpha_i \mu(A_i).$$

The reverse inequality follows from the fact that ξ is s.t. $\xi \in Z_0^{\mathcal{F}}$ and $S_\mu^f(\xi) = \sum_{i=1}^N \alpha_i \mu(A_i)$. □

We prove the monotone convergence theorem for nonnegative functions.

Proof of Proposition 8.2. First of all, using the result of the latter exercise, we have that for any $n \in \mathbb{N}$,

$$\int_{\Omega} f_n(\omega)\mu(d\omega) \leq \int_{\Omega} f_{n+1}(\omega)\mu(d\omega)$$

and

$$\sup_{n \in \mathbb{N}} \left(\int_{\Omega} f_n(\omega)\mu(d\omega) \right) \leq \int_{\Omega} f(\omega)\mu(d\omega).$$

Hence, using Proposition 3.8,

$$\lim_{n\to\infty} \left(\int_{\Omega} f_n(\omega)\mu(d\omega) \right) \leq \int_{\Omega} f(\omega)\mu(d\omega).$$

Thus, it remains to show that

$$\lim_{n\to\infty} \left(\int_{\Omega} f_n(\omega)\mu(d\omega) \right) \geq \int_{\Omega} f(\omega)\mu(d\omega).$$

That is, we need to show that for any $\xi = \{A_i \colon i = 1, \dots, N\} \in Z_0^{\mathcal{F}}$,

$$\lim_{n\to\infty} \left(\int_{\Omega} f_n(\omega)\mu(d\omega) \right) \geq S_\mu^f(\xi). \tag{26}$$

Hence, let $\xi = \{A_i \colon i = 1, \dots, N\} \in Z_0^{\mathcal{F}}$. We use the notation

$$S_\mu^f(\xi) = \sum_{i=1}^{N} a_i \mu(A_i), \quad a_i = \inf_{\omega \in A_i} f(\omega), \ i = 1, \dots, N.$$

First we consider the case where ξ is s.t. $S_\mu^f(\xi) < \infty$ and for any $i = 1, \dots, N$, $0 < a_i < \infty$ and $0 < \mu(A_i) < \infty$. Let $\varepsilon > 0$, s.t. $\varepsilon < \min\{a_i \colon i = 1, \dots, N\}$ and

$$\varepsilon < \frac{\delta}{\sum_{i=1}^N \mu(A_i)}, \tag{27}$$

where $\delta > 0$ is arbitrary but nonnegative. Define the sets

$$A_{i,n} = \{\omega \in A_i \colon f_n(\omega) > a_i - \varepsilon\}, \quad i = 1, \dots, N, \ n \in \mathbb{N}.$$

Since $(f_n(\omega))_{n \in \mathbb{N}}$ is increasing, we observe that for any $i = 1, \dots, N$, $A_{i,n} \subset A_{i,n+1}$. Further, for any $i = 1, \dots, N$,

$$\bigcup_{n \in \mathbb{N}} A_{i,n} = A_i.$$

Clearly, if $\omega' \in \cup_{n\in\mathbb{N}} A_{i,n}$, $\omega' \in A_i$. For the other direction, suppose that $\omega' \in A_i$, then, since $f_n(\omega') \uparrow f(\omega')$, there exists $K \in \mathbb{N}$ s.t. for any $n \geq K$,

$$f(\omega') - f_n(\omega') = |f(\omega') - f_n(\omega')| < \varepsilon.$$

Therefore,

$$f_n(\omega') > f(\omega') - \varepsilon \geq \inf_{\omega\in A_i} f(\omega) - \varepsilon.$$

Given any $n \in \mathbb{N}$, consider the partition

$$\xi_n = \big\{A_{i,n} \colon i = 1, \dots, N\big\} \cup \big\{\Omega \setminus (\cup_{i=1}^N A_{i,n})\big\}.$$

We notice that since f_n is $\mathcal{F}$ measurable, $\xi_n \in Z_0^{\mathcal{F}}$. Hence,

$$\begin{aligned}\int_\Omega f_n(\omega)\mu(d\omega) &\geq S_\mu^{f_n}(\xi_n) = \sum_{A\in\xi_n} \Big(\inf_{\omega\in A} f_n(\omega)\Big)\mu(A) \geq \sum_{i=1}^N \Big(\inf_{\omega\in A_{i,n}} f_n(\omega)\Big)\mu(A_{i,n}) \\ &\geq \sum_{i=1}^N \big(a_i - \varepsilon\big)\mu(A_{i,n}).\end{aligned}$$

Then, using item (v) of Proposition 5.1, we obtain that

$$\sum_{i=1}^N \big(a_i - \varepsilon\big)\mu(A_{i,n}) \uparrow \sum_{i=1}^N \big(a_i - \varepsilon\big)\mu(A_i) = S_\mu^f(\xi) - \varepsilon \sum_{i=1}^N \mu(A_i).$$

Therefore,

$$\lim_{n\to\infty} \left(\int_\Omega f_n(\omega)\mu(d\omega)\right) \geq S_\mu^f(\xi) - \varepsilon \sum_{i=1}^N \mu(A_i).$$

Hence, by (27),

$$\lim_{n\to\infty} \left(\int_\Omega f_n(\omega)\mu(d\omega)\right) \geq S_\mu^f(\xi) - \delta.$$

Since $\delta > 0$ was arbitrary, (26) is shown. Consider now the case, where ξ is s.t. $S_\mu^f(\xi) < \infty$. This implies that for any $i = 1, \dots, N$, $a_i\mu(A_i) < \infty$. Clearly, if for any $i = 1, \dots, N$, $a_i\mu(A_i) = 0$, then (26) is true. Thus, suppose that there exists $i_1, \dots, i_{N_0} \in \{1, \dots, N\}$, $N_0 \leq N$, s.t. $a_{i_j}\mu(A_{i_j}) > 0$ for any $j = 1, \dots, N_0$, and $a_i\mu(A_i) = 0$ for any $i \in \{1, \dots, N\} \setminus \{i_1, \dots, i_{N_0}\}$. Hence, since μ is a measure, $a_{i_j} > 0$ and $\mu(A_{i_j}) > 0$ for any $j = 1, \dots, N_0$. Therefore, we obtain

$$S_\mu^f(\xi) = \sum_{j=1}^{N_0} a_{i_j}\mu(A_{i_j}),$$

and proceed towards (26) as in the previous case. Finally, we consider the case where ξ is s.t. $S_\mu^f(\xi) = \infty$. We need to show that

$$\lim_{n\to\infty} \left(\int_\Omega f_n(\omega)\mu(d\omega)\right) = \infty. \tag{28}$$

Since $S_\mu^f(\xi) = \infty$, there exists $j \in \{1, \dots, N\}$ s.t. $a_j\mu(A_j) = \infty$. Hence, $a_j > 0$ and $\mu(A_j) > 0$ and either $a_j = \infty$ or $\mu(A_j) = \infty$. Let $p, q > 0$ s.t. $0 < p < a_j \leq \infty$ and $0 < q < \mu(A_j) \leq \infty$. Set

$$A_{j,n} = \{\omega \in A_j \colon f_n(\omega) > p\}.$$

We have that $A_{j,n} \subset A_{j,n+1}$ and $A_j = \cup_{n\in\mathbb{N}} A_{j,n}$. We notice that $A_j \subset \cup_{n\in\mathbb{N}} A_{j,n}$ follows from the fact that if $\omega \in A_j$, then since $f_n(\omega) \uparrow f(\omega)$, it follows that there exists $K \in \mathbb{N}$ s.t. for any $n \geq K$, $f(\omega) - f_n(\omega) < a_j - p$, i.e., $f_n(\omega) > f(\omega) - a_j + p \geq p$. Therefore, using item (v) of Proposition 5.1 there exists $K_0 \in \mathbb{N}$ s.t. for any $n \geq K_0$, $\mu(A_j) - \mu(A_{j,n}) < \mu(A_j) - q$. That is, $\mu(A_{j,n}) > q$ for any $n \geq K_0$. Then, consider the partition ξ_n composed of the sets $A_{j,n}$ and $A^c_{j,n}$. We have that for any $n \geq K_0$,

$$\int_\Omega f_n(\omega)\mu(d\omega) \geq S^{f_n}_\mu(\xi_n) \geq p\mu(A_{j,n}) > pq.$$

Therefore,

$$\lim_{n\to\infty} \left(\int_\Omega f_n(\omega)\mu(d\omega) \right) \geq pq.$$

Since either $a_j = \infty$ or $\mu(A_j) = \infty$, either p or q can be made arbitrary large and (28) follows. □

We show the linearity propoerty of the integral for nonnegative functions.

Proof of Proposition 8.3. Suppose that $f = \sum_{i=1}^N \alpha_i \mathbb{1}_{A_i}$, $\alpha_i \in [0, \infty)$, $\{A_i : i = 1, \ldots, N\} \subset \mathcal{F}$, $i = 1, \ldots, N$, and $g = \sum_{j=1}^M \beta_j \mathbb{1}_{B_j}$, $\beta_j \in [0, \infty)$, $\{B_j : j = 1, \ldots, M\} \subset \mathcal{F}$, $j = 1, \ldots, M$, i.e., f and g are nonnegative, $\mathcal{F}$ measurable simple functions (cf. Definition 7.3). According to Proposition 7.7, we assume that f and g are already in standard form, i.e., $\cup_{i=1}^N A_i = \cup_{j=1}^M B_j = \Omega$. We write

$$\alpha f = \sum_{i=1}^N \alpha\alpha_i \mathbb{1}_{A_i} = \sum_{i=1}^N \alpha\alpha_i \bigg(\sum_{j=1}^M \mathbb{1}_{A_i \cap B_j} \bigg),$$

and

$$\beta g = \sum_{j=1}^M \beta\beta_i \mathbb{1}_{B_j} = \sum_{j=1}^M \beta\beta_j \bigg(\sum_{i=1}^N \mathbb{1}_{B_j \cap A_i} \bigg).$$

Hence,

$$\begin{aligned}
\alpha f + \beta g &= \sum_{i=1}^N \alpha\alpha_i \bigg(\sum_{j=1}^M \mathbb{1}_{A_i \cap B_j} \bigg) + \sum_{j=1}^M \beta\beta_j \bigg(\sum_{i=1}^N \mathbb{1}_{B_j \cap A_i} \bigg) \\
&= \sum_{i=1}^N \alpha\alpha_i \bigg(\sum_{j=1}^M \mathbb{1}_{A_i \cap B_j} \bigg) + \sum_{i=1}^N \beta\beta_j \bigg(\sum_{j=1}^M \mathbb{1}_{A_i \cap B_j} \bigg) \\
&= \sum_{i=1}^N \bigg(\sum_{j=1}^M \alpha\alpha_i \mathbb{1}_{A_i \cap B_j} \bigg) + \sum_{i=1}^N \bigg(\sum_{j=1}^M \beta\beta_j \mathbb{1}_{A_i \cap B_j} \bigg) \\
&= \sum_{i=1}^N \bigg(\sum_{j=1}^M (\alpha\alpha_i + \beta\beta_j) \mathbb{1}_{A_i \cap B_j} \bigg).
\end{aligned}$$

Assume that $N \geq M$ and set $B_j = \emptyset$ for $j = M+1, \ldots, N$ (if $M \geq N$, we set $A_i = \emptyset$ for $i = N+1, \ldots, M$). Then, we write

$$\alpha f + \beta g = \sum_{i,j=1}^N \gamma_{i,j} \mathbb{1}_{C_{i,j}},$$

where $\gamma_{i,j} = \alpha\alpha_i + \beta\beta_j$ and $C_{i,j} = A_i \cap B_j$, $i, j = 1, \ldots, N$. Therefore, $\alpha f + \beta g$ is a standard simple function. Using Proposition 8.1, we obtain

$$\begin{aligned}\int_\Omega (\alpha f + \beta g)(\omega)\mu(d\omega) &= \sum_{i,j=1}^{N} \gamma_{i,j}\mu(C_{i,j}) \\ &= \alpha \sum_{i=1}^{N} \alpha_i \mu(A_i) + \beta \sum_{j=1}^{M} \beta_j \mu(B_j) \\ &= \alpha \int_\Omega f(\omega)\mu(d\omega) + \beta \int_\Omega g(\omega)\mu(d\omega).\end{aligned}$$

At this point we remark that α or β equal to ∞, would not change the latter result. For the general case, assume that f and g are nonnegative and $\mathcal{F}$ measurable. Using Proposition 7.10, there exists sequences of nonnegative standard simple functions (f_n) and (g_n) s.t. $f_n(\omega) \uparrow f(\omega)$ and $g_n(\omega) \uparrow g(\omega)$, $\omega \in \Omega$. That is, $(\alpha f_n + \beta g_n)(\omega) \uparrow \alpha f + \beta g$. Then, we rely on Proposition 8.2 and conclude that

$$\begin{aligned}\int_\Omega (\alpha f + \beta g)(\omega)\mu(d\omega) &= \lim_{n\to\infty} \left(\int_\Omega (\alpha f_n + \beta g_n)(\omega)\mu(d\omega) \right) \\ &= \alpha \lim_{n\to\infty} \left(\int_\Omega f_n(\omega)\mu(d\omega) \right) + \beta \lim_{n\to\infty} \left(\int_\Omega g_n(\omega)\mu(d\omega) \right) \\ &= \alpha \int_\Omega f(\omega)\mu(d\omega) + \beta \int_\Omega g(\omega)\mu(d\omega).\end{aligned}$$

$\square$

Proof of Proposition 8.5. Let $A_i = \{\omega\colon f(\omega) \geq 1/i\}$, $i \in \mathbb{N}$. Then, clearly $\cup_{i\in\mathbb{N}} A_i = \{\omega\colon f(\omega) > 0\}$. Using item (v) of Proposition 5.1, $\mu(A_n) \uparrow \mu(\{\omega\colon f(\omega) > 0\})$. Thus, since $\mu(\{\omega\colon f(\omega) > 0\}) > 0$, let $\varepsilon > 0$ s.t. $\varepsilon < \mu(\{\omega\colon f(\omega) > 0\})$. We write $\delta = \mu(\{\omega\colon f(\omega) > 0\}) - \varepsilon$. Then, there exists $N \in \mathbb{N}$ s.t. $\mu(\{\omega\colon f(\omega) > 0\}) - \mu(A_N) \leq \varepsilon$, i.e., $\mu(A_N) \geq \mu(\{\omega\colon f(\omega) > 0\}) - \varepsilon = \delta > 0$. Now consider $\xi = \{A_N, A_N^c\}$. Then, $\xi \in Z_0^{\mathcal{F}}$ and $\int_\Omega f(\omega)\mu(d\omega) \geq S_\mu^f(\xi) \geq \frac{1}{N}\mu(A_N) \geq \delta/N > 0$. This completes the proof of (i). Regarding (ii), suppose by contradiction that $\int_\Omega f(\omega)\mu(d\omega) < \infty$ but $\mu(\{\omega\colon f(\omega) = \infty\}) > 0$, i.e., the negation of $f < \infty$ μ a.e. Then, we consider $\xi = \{A, A^c\}$, where $A = \{\omega\colon f(\omega) = \infty\}$. We have that $S_\mu^f(\xi) \geq \infty \cdot \mu(\{\omega\colon f(\omega) = \infty\}) = \infty$, which gives a contradiction. In order to verify (iii), let $G = \{\omega\colon f(\omega) \leq g(\omega)\}$. Let $\xi = \{A_1, \ldots, A_N\} \in Z_0^{\mathcal{F}}$. Then, we consider $\xi^* = \{A_1 \cap G, \ldots, A_N \cap G\} \cup \{G^c\}$ and have that $\xi^* \in Z_0^{\mathcal{F}}$. In particular,

$$S_\mu^f(\xi) \leq S_\mu^f(\xi^*) \leq S_\mu^g(\xi^*) \leq \int_\Omega g(\omega)\mu(d\omega),$$

where the second inequality follows since by assumption on G, for any $A \in \xi$, $\mu(A) = \mu(A \cap G) + \mu(A \cap G^c) = \mu(A \cap G)$. This shows (iii) and in particular (iv) (then also $\int_\Omega g(\omega)\mu(d\omega) \leq \int_\Omega f(\omega)\mu(d\omega)$). $\square$

Proof of Proposition 8.7. We notice that

$$\{\omega\colon f^+(\omega) \leq g^+(\omega)\} \cap \{\omega\colon f^-(\omega) \geq g^-(\omega)\} = \{\omega\colon f(\omega) \leq g(\omega)\}.$$

This follows from the definition of f^+ and f^{-1}, since if ω is s.t. $f(\omega) \leq g(\omega)$, then, $\max\{f(\omega), 0\} \leq \max\{g(\omega), 0\}$ and $\max\{-f(\omega), 0\} \geq \max\{-g(\omega), 0\}$. We obtain

$$\{\omega\colon f^+(\omega) \leq g^+(\omega)\}^c \cup \{\omega\colon f^-(\omega) \geq g^-(\omega)\}^c = \{\omega\colon f(\omega) \leq g(\omega)\}^c$$

Hence, if $f \leq g$ a.e., then, $f^+ \leq g^+$ a.e. and $f^- \geq g^-$ a.e. Therefore, using (iii) in Proposition 8.5, $\int_\Omega f^+(\omega)\mu(d\omega) \leq \int_\Omega g^+(\omega)\mu(d\omega)$ and $\int_\Omega f^-(\omega)\mu(d\omega) \geq \int_\Omega g^-(\omega)\mu(d\omega)$, which makes

$$\begin{aligned}\int_\Omega f(\omega)\mu(d\omega) &= \int_\Omega f^+(\omega)\mu(d\omega) - \int_\Omega f^-(\omega)\mu(d\omega) \\ &\leq \int_\Omega g^+(\omega)\mu(d\omega) - \int_\Omega g^-(\omega)\mu(d\omega) = \int_\Omega g(\omega)\mu(d\omega).\end{aligned}$$

□

We show that the integral for integrable functions is linear.

Proof of Proposition 8.8. Notice first that for any $\omega \in \Omega$,

$$|(\alpha f + \beta g)(\omega)| \leq |\alpha||f(\omega)| + |\beta||g(\omega)|,$$

by the triangular inequality. Thus, since $|\alpha| \int_\Omega |f(\omega)|\mu(d\omega)$ and $|\beta| \int_\Omega |g(\omega)|\mu(d\omega)$ are finite upon the integrability of f and g (cf. Proposition 8.6), $\alpha f + \beta g$ is integrable as well (cf. Exercise 8.1 and Proposition 8.3). Notice first that $\int_\Omega \alpha f(\omega)\mu(d\omega) = \alpha \int_\Omega f(\omega)\mu(d\omega)$. We can see it by considering the cases: $\alpha > 0$, $\alpha < 0$ and $\alpha = 0$. If $\alpha = 0$, then, for any $\omega \in \Omega$, $(\alpha f)^+(\omega) = (\alpha f)^-(\omega) = 0 = \alpha f^-(\omega) = \alpha f^+(\omega)$. If $\alpha > 0$, then,

$$(\alpha f)^+(\omega) = \begin{cases} \alpha f(\omega), & \text{if } \alpha f(\omega) \geq 0, \\ 0, & \text{if } \alpha f(\omega) < 0, \end{cases} = \alpha \begin{cases} f(\omega), & \text{if } f(\omega) \geq 0, \\ 0, & \text{if } f(\omega) < 0, \end{cases} = \alpha f^+(\omega),$$

and

$$(\alpha f)^-(\omega) = \begin{cases} -\alpha f(\omega), & \text{if } -\alpha f(\omega) \geq 0, \\ 0, & \text{if } -\alpha f(\omega) < 0, \end{cases} = \alpha \begin{cases} -f(\omega), & \text{if } -f(\omega) \geq 0, \\ 0, & \text{if } -f(\omega) < 0, \end{cases} = \alpha f^-(\omega),$$

Similarly, if $\alpha < 0$, then $(\alpha f)^+(\omega) = -\alpha f^-(\omega)$ and $(\alpha f)^-(\omega) = -\alpha f^+(\omega)$. Hence, in each case, we use Proposition 8.3 and obtain

$$\begin{aligned}\int_\Omega \alpha f(\omega)\mu(d\omega) &= \int_\Omega (\alpha f)^+(\omega)\mu(d\omega) - \int_\Omega (\alpha f)^-(\omega)\mu(d\omega) \\ &= \alpha \int_\Omega f^+(\omega)\mu(d\omega) - \alpha \int_\Omega f^-(\omega)\mu(d\omega) = \alpha \int_\Omega f(\omega)\mu(d\omega).\end{aligned}$$

Of course the same argument applies for $\int_\Omega \beta g(\omega)\mu(d\omega)$. Hence, if we show that

$$\int_\Omega (\alpha f + \beta g)(\omega)\mu(d\omega) = \int_\Omega \alpha f(\omega)\mu(d\omega) + \int_\Omega \beta g(\omega)\mu(d\omega), \tag{29}$$

we are done. Write $f_* = \alpha f$ and $g_* = \beta g$. We know that $(f_* + g_*)^+ - (f_* + g_*)^- = f_* + g_*$ (cf. Exercise 7.5). Hence, $(f_* + g_*)^+ - (f_* + g_*)^- = f_*^+ - f_*^- + g_*^+ - g_*^-$. This shows that $(f_* + g_*)^+ + f_*^- + g_*^- = (f_* + g_*)^- + f_*^+ + g_*^+$. Hence,

$$\int_\Omega \big((f_* + g_*)^+ + f_*^- + g_*^-\big)(\omega)\mu(d\omega) = \int_\Omega \big((f_* + g_*)^- + f_*^+ + g_*^+\big)(\omega)\mu(d\omega),$$

which gives (cf. Proposition 8.3)

$$\begin{aligned}&\int_\Omega (f_* + g_*)^+(\omega)\mu(d\omega) - \int_\Omega (f_* + g_*)^-(\omega)\mu(d\omega) \\ &= \int_\Omega f_*^+(\omega)\mu(d\omega) - \int_\Omega f_*^-(\omega)\mu(d\omega) + \int_\Omega g_*^+(\omega)\mu(d\omega) - \int_\Omega g_*^-(\omega)\mu(d\omega) \\ &= \int_\Omega f(\omega)\mu(d\omega) + \int_\Omega g(\omega)\mu(d\omega).\end{aligned}$$

This shows (29) and the proof of the proposition is complete. □

We prove Fatou's lemma and Lebesgue's dominated convergence theorem.

Proof of Proposition 8.9. Define $g_n = \inf_{k\geq n} f_k$, then $(g_n)_{n\in\mathbb{N}}$ is a sequence of $\mathcal{F}$ measurable and nonnegative functions s.t. for any $n \in \mathbb{N}$, $\int_\Omega g_n(\omega)\mu(d\omega) \leq \int_\Omega f_n(\omega)\mu(d\omega)$ (cf. Exercise 8.1), Therefore,

$$\liminf_{n\to\infty} \int_\Omega g_n(\omega)\mu(d\omega) \leq \liminf_{n\to\infty} \int_\Omega f_n(\omega)\mu(d\omega).$$

Further, for any $\omega \in \Omega$, $g_n(\omega) \uparrow \liminf_{n\to\infty} f_n(\omega)$ (cf. Proposition A.5). Therefore, using Proposition 8.2, $\int_\Omega g_n(\omega)\mu(d\omega) \uparrow \int_\Omega \liminf_{n\to\infty} f_n(\omega)\mu(d\omega)$. In conclusion, using Proposition 3.23, we obtain

$$\int_\Omega \liminf_{n\to\infty} f_n(\omega)\mu(d\omega) = \liminf_{n\to\infty} \int_\Omega g_n(\omega)\mu(d\omega) \leq \liminf_{n\to\infty} \int_\Omega f_n(\omega)\mu(d\omega).$$

□

Proof of Proposition 8.10. Notice that f is integrable, i.e., $\int_\Omega |f(\omega)|\mu(d\omega) < \infty$. To see it, we notice that (compare also to Proposition 3.25),

$$A = \underbrace{\{\omega\colon \lim_{n\to\infty} f_n(\omega) = f(\omega)\}}_{=A_1} \cap \Bigg(\bigcap_{n\in\mathbb{N}} \underbrace{\{\omega\colon |f_n(\omega)| \leq g(\omega)\}}_{=A_{2,n}} \Bigg) \subset \{\omega\colon f(\omega) \leq g(\omega)\}.$$

Then, upon Proposition 5.1, since $\mu(A_1^c) = 0$ and $\mu(A_{2,n}^c) = 0$ for any $n \in \mathbb{N}$,

$$\mu(A^c) \leq \sum_{n\in\mathbb{N}} \big(\mu(A_1^c) + \mu(A_{2,n}^c)\big) = 0.$$

Thus, $f \leq g$ a.e. and hence f is integrable by Exercise 8.4. Similarly $f^* = \limsup_{n\to\infty} f_n$ and $f_* = \liminf_{n\to\infty} f_n$ are integrable. Then, we notice that $(g + f_n)_{n\in\mathbb{N}}$ and $(g - f_n)_{n\in\mathbb{N}}$ are nonnegative and $\mathcal{F}$ measurable sequences of functions. Given any $\omega \in A$, we rely on Proposition A.4 and obtain

$$(g + f_*)(\omega) = g(\omega) + \liminf_{n\to\infty} f_n(\omega) = \liminf_{n\to\infty} (g + f_n)(\omega),$$

and

$$(g - f^*)(\omega) = \liminf_{n\to\infty} (g - f_n)(\omega).$$

Hence, $g + f_* = \liminf_{n\to\infty}(g + f_n)$ a.e. and $g - f^* = \liminf_{n\to\infty}(g - f_n)$ a.e. Therefore, we use Proposition 8.9 and obtain

$$\begin{aligned}
\int_\Omega g(\omega)\mu(d\omega) + \int_\Omega f_*(\omega)\mu(d\omega) &= \int_\Omega (g + f_*)(\omega)\mu(d\omega) \\
&= \int_\Omega \liminf_{n\to\infty} (g + f_n)(\omega)\mu(d\omega) \\
&\leq \liminf_{n\to\infty} \int_\Omega (g + f_n)(\omega)\mu(d\omega) \\
&= \int_\Omega g(\omega)\mu(d\omega) + \liminf_{n\to\infty} \int_\Omega f_n(\omega)\mu(d\omega),
\end{aligned}$$

where the last equality again follows from Proposition A.4. Similarly, we obtain

$$\begin{aligned}
\int_\Omega g(\omega)\mu(d\omega) - \int_\Omega f^*(\omega)\mu(d\omega) &= \int_\Omega \liminf_{n\to\infty} (g - f_n)(\omega)\mu(d\omega) \\
&\leq \liminf_{n\to\infty} \int_\Omega (g - f_n)(\omega)\mu(d\omega) \\
&= \int_\Omega g(\omega)\mu(d\omega) - \limsup_{n\to\infty} \int_\Omega f_n(\omega)\mu(d\omega).
\end{aligned}$$

This makes (cf. Proposition 3.20)

$$\int_\Omega \liminf_{n\to\infty} f_n(\omega)\mu(d\omega) \le \liminf_{n\to\infty} \int_\Omega f_n(\omega)\mu(d\omega)$$
$$\le \limsup_{n\to\infty} \int_\Omega f_n(\omega)\mu(d\omega) \le \int_\Omega \limsup_{n\to\infty} f_n(\omega)\mu(d\omega).$$

Since $f_n \to f$ a.e. it follows that $\liminf_{n\to\infty} f_n = \limsup_{n\to\infty} f_n = f$ a.e. and hence upon the previous display,

$$\liminf_{n\to\infty} \int_\Omega f_n(\omega)\mu(d\omega) = \limsup_{n\to\infty} \int_\Omega f_n(\omega)\mu(d\omega) = \int_\Omega f(\omega)\mu(d\omega),$$

and therefore (cf. Proposition 3.23),

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

□

References

[1] Patrick Billingsley, *Probability and measure*, 4 ed., Wiley Series in Probability and Statistics, John Wiley & Sons, Inc., Hoboken, NJ, 2012.

[2] Jean-François Le Gall, *Measure theory, probability, and stochastic processes*, Graduate Texts in Mathematics, vol. 295, Springer, Cham, 2022.