Basics of the theory of
MEASURE AND INTEGRAL

M. Papadimitrakis

Contents

1 Measures. **4**
- 1.1 σ-algebras. 4
- 1.2 Measures. 14
- 1.3 Measures from outer measures. 26
- 1.4 Lebesgue measure. 31
- 1.5 Borel measures on topological spaces. 46

2 Measurable functions. **62**
- 2.1 Measurability. 62
- 2.2 The role of null sets. 75
- 2.3 Lusin's Theorem 77

3 Integrals. **83**
- 3.1 Integrals. 83
- 3.2 Lebesgue integral. 106
- 3.3 Lebesgue-Stieltjes integrals. 114
- 3.4 Integrals on Borel measure spaces. 119

4 Product measures. **122**
- 4.1 Product σ-algebra. 122
- 4.2 Product measure. 126
- 4.3 Multiple integrals. 130
- 4.4 Surface measure on $\mathbb{S}^{n-1}$. 137

5 Convergence of functions. **145**
- 5.1 a.e. convergence and uniform a.e. convergence. 145
- 5.2 Convergence in the mean. 147
- 5.3 Convergence in measure. 149
- 5.4 Almost uniform convergence. 152
- 5.5 Relations between various types of convergence. 153

6 Signed measures and complex measures. **157**
- 6.1 Signed measures. 157
- 6.2 The Hahn and Jordan decompositions. 158
- 6.3 Complex measures. 164
- 6.4 Integration. 168
- 6.5 Lebesgue decomposition, Radon-Nikodym derivative. 171
- 6.6 Differentiation. 178
- 6.7 Functions of bounded variation. 187

7 The classical Banach spaces. **189**
7.1 Some facts from functional analysis. 189
7.2 The spaces $B(X)$, $C(X)$, $BC(X)$ and $C_0(X)$. 209
7.3 The spaces $L^p(X, \mathcal{S}, \mu)$ and their duals. 212
7.4 The spaces $M(X)$ and $M_{\mathcal{R}}(X)$. 227
7.5 The spaces $B(X, \mathcal{S})$ and $L^\infty(X, \mathcal{S}, \mu)$ and their duals. 234

Sets, empty set, space, complement, union and intersection (with the index notation and the family notation), set-theoretic difference, symetric difference, countable (finite and infinite) union and intersection, laws of de Morgan, increasing and decreasing sequence of sets (analogy to monotone sequence of numbers and of functions), limsup and liminf (and limit) of a sequence of sets, power set of a set.

Functions, images and inverse images of sets (union, intersection, complement). Sums and products of functions. Convergence and uniform convergence of sequences of functions.

Topology, open sets, closed sets, basic properties (unions, intersections, etc), interior, closure, boundary. Metric spaces, balls. Euclidean spaces. Subspace (or relative) topology. The extended real line, and the extended complex plane. Continuous functions. Diameter of a set and distance between sets in a metric space.

Equivalence and order relations. Equivalence classes. Quotient space. Zorn's Lemma. Axiom of Choice.

Linear algebra.

Chapter 1

Measures.

1.1 σ-algebras.

Definition. *Let X be a set, and $\mathcal{S}$ be a collection of subsets of X. We call $\mathcal{S}$ a* ***σ-algebra*** *of subsets of X if it is non-empty, closed under complements, and closed under countably infinite unions. This means:*
(i) there exists at least one $A \subseteq X$ so that $A \in \mathcal{S}$,
(ii) if $A \in \mathcal{S}$, then $A^c \in \mathcal{S}$,
(iii) if $A_n \in \mathcal{S}$ for all $n \in \mathbb{N}$, then $\bigcup_{n=1}^{+\infty} A_n \in \mathcal{S}$.
The pair $(X, \mathcal{S})$ of a set X and a σ-algebra $\mathcal{S}$ of subsets of X is called a ***measurable space****. The sets $A \in \mathcal{S}$ are called* ***$\mathcal{S}$-measurable*** *subsets of X.*

If there is no danger of confusion, every $A \in \mathcal{S}$ shall be called *measurable subset* of X or just *measurable set*.

Proposition 1.1. *Every σ-algebra of subsets of X contains the sets $\emptyset$ and X, it is closed under finite unions (and, thus, under countable unions), under countable intersections, and under set-theoretic differences.*

Proof. Let $\mathcal{S}$ be any σ-algebra of subsets of X.
Let $A_n \in \mathcal{S}$ for all $n \in \mathbb{N}$. Then $A_n^c \in \mathcal{S}$ for all $n \in \mathbb{N}$. Therefore $\bigcup_{n=1}^{+\infty} A_n^c \in \mathcal{S}$, and hence $(\bigcup_{n=1}^{+\infty} A_n^c)^c \in \mathcal{S}$. Since

$$\bigcap_{n=1}^{+\infty} A_n = (\bigcup_{n=1}^{+\infty} A_n^c)^c,$$

we get that $\bigcap_{n=1}^{+\infty} A_n \in \mathcal{S}$.
Let $A_1, \ldots, A_N \in \mathcal{S}$. We consider $A_n = A_N$ for all $n \in \mathbb{N}$, $n > N$, and we have that $\bigcup_{n=1}^{N} A_n = \bigcup_{n=1}^{+\infty} A_n$. Since $\bigcup_{n=1}^{+\infty} A_n \in \mathcal{S}$, we conclude that $\bigcup_{n=1}^{N} A_n \in \mathcal{S}$.
Similarly, we have that $\bigcap_{n=1}^{N} A_n = \bigcap_{n=1}^{+\infty} A_n$. Since $\bigcap_{n=1}^{+\infty} A_n \in \mathcal{S}$, we get that $\bigcap_{n=1}^{N} A_n \in \mathcal{S}$.
Now let $A \in \mathcal{S}$. Then $A^c \in \mathcal{S}$, and so $\emptyset = A \cap A^c \in \mathcal{S}$ and $X = A \cup A^c \in \mathcal{S}$.
Finally, let $A, B \in \mathcal{S}$. Then $B^c \in \mathcal{S}$, and so $A \setminus B = A \cap B^c \in \mathcal{S}$. □

Here are some simple examples.

Example. The collection $\{\emptyset, X\}$ is a σ-algebra, the smallest possible, of subsets of the set X.

Example. $\mathcal{P}(X)$, the collection of all subsets of X, is a σ-algebra, the largest possible, of subsets of X.

Example. If $E \subseteq X$, then $\{\emptyset, E, E^c, X\}$ is a σ-algebra of subsets of X. In fact, it is the smallest σ-algebra of subsets of X containing E.

Example. Let X be uncountable. The collection $\mathcal{S} = \{A \subseteq X \mid \text{either } A \text{ or } A^c \text{ is countable}\}$ is a σ-algebra of subsets of X. Let us see why.
Firstly, $\emptyset$ is countable, and so $\mathcal{S}$ is non-empty.
If $A \in \mathcal{S}$, then, considering cases, we see that $A^c \in \mathcal{S}$.
Finally, let $A_n \in \mathcal{S}$ for all $n \in \mathbb{N}$. If every A_n is countable, then $\bigcup_{n=1}^{+\infty} A_n$ is also countable, and so $\bigcup_{n=1}^{+\infty} A_n \in \mathcal{S}$. Otherwise, at least one of the A_n^c, say $A_{n_0}^c$, is countable. Since $(\bigcup_{n=1}^{+\infty} A_n)^c \subseteq A_{n_0}^c$, we have that $(\bigcup_{n=1}^{+\infty} A_n)^c$ is also countable, and so again $\bigcup_{n=1}^{+\infty} A_n \in \mathcal{S}$.

The following result is useful.

Lemma 1.1. *Let $\mathcal{S}$ be a σ-algebra of subsets of X. Then for every finite or infinite sequence (A_n) in $\mathcal{S}$ there exists a finite or infinite, respectively, sequence (B_n) in $\mathcal{S}$ such that:*
(i) $B_n \subseteq A_n$ for all n,
(ii) $B_1 \cup B_2 \cup \cdots = A_1 \cup A_2 \cup \cdots$,
(iii) the B_n are pairwise disjoint.

Proof. We consider $B_1 = A_1$, and $B_n = A_n \setminus (A_1 \cup \cdots \cup A_{n-1})$ for all relevant $n \geq 2$. □

Of course we know that a sequence (x_n) of real numbers is called *increasing* or *decreasing* if $x_n \leq x_{n+1}$ for all n or, respectively, if $x_{n+1} \leq x_n$ for all n. Similarly, a sequence (f_n) of real valued functions, with A as their common domain of definition, is called *increasing* or *decreasing* if $f_n \leq f_{n+1}$ on A for all n or, respectively, if $f_{n+1} \leq f_n$ on A for all n. Now, a sequence (A_n) of sets is called *increasing* or *decreasing* if $A_n \subseteq A_{n+1}$ for all n or, respectively, if $A_{n+1} \subseteq A_n$ for all n.

Exercises.

1.1.1. Let $A_n \subseteq X$ for every $n \in \mathbb{N}$. We set

$$\underline{\lim}_{n\to+\infty} A_n = \bigcup_{k=1}^{+\infty} \big(\bigcap_{j=k}^{+\infty} A_j \big), \quad \overline{\lim}_{n\to+\infty} A_n = \bigcap_{k=1}^{+\infty} \big(\bigcup_{j=k}^{+\infty} A_j \big).$$

Only if $\underline{\lim}_{n\to+\infty} A_n = \overline{\lim}_{n\to+\infty} A_n$, we define

$$\lim_{n\to+\infty} A_n = \underline{\lim}_{n\to+\infty} A_n = \overline{\lim}_{n\to+\infty} A_n.$$

Prove the following.
(i) $\underline{\lim}_{n\to+\infty} A_n = \{x \in X \mid x \in A_n \text{ for all large enough } n\}$.
(ii) $\overline{\lim}_{n\to+\infty} A_n = \{x \in X \mid x \in A_n \text{ for infinitely many } n\}$.
(iii) $(\underline{\lim}_{n\to+\infty} A_n)^c = \overline{\lim}_{n\to+\infty} A_n^c$ and $(\overline{\lim}_{n\to+\infty} A_n)^c = \underline{\lim}_{n\to+\infty} A_n^c$.
(iv) $\underline{\lim}_{n\to+\infty} A_n \subseteq \overline{\lim}_{n\to+\infty} A_n$.
(v) If (A_n) is increasing, then $\lim_{n\to+\infty} A_n = \bigcup_{n=1}^{+\infty} A_n$.
(vi) If (A_n) is decreasing, then $\lim_{n\to+\infty} A_n = \bigcap_{n=1}^{+\infty} A_n$.
(vii) If $A_n \subseteq B_n$ for all n, then $\underline{\lim}_{n\to+\infty} A_n \subseteq \underline{\lim}_{n\to+\infty} B_n$ and $\overline{\lim}_{n\to+\infty} A_n \subseteq \overline{\lim}_{n\to+\infty} B_n$.
(viii) If $A_n = B$, if n is even, and $A_n = C$, if n is odd, then $\underline{\lim}_{n\to+\infty} A_n = B \cap C$, and $\overline{\lim}_{n\to+\infty} A_n = B \cup C$.
(ix) If $A_n = B_n \cup C_n$ for all n, then $(\underline{\lim}_{n\to+\infty} B_n) \cup (\underline{\lim}_{n\to+\infty} C_n) \subseteq \underline{\lim}_{n\to+\infty} A_n$ and $\overline{\lim}_{n\to+\infty} A_n = (\overline{\lim}_{n\to+\infty} B_n) \cup (\overline{\lim}_{n\to+\infty} C_n)$.
(x) If $A_n = B_n \cap C_n$ for all n, then $(\underline{\lim}_{n\to+\infty} B_n) \cap (\underline{\lim}_{n\to+\infty} C_n) = \underline{\lim}_{n\to+\infty} A_n$ and $\overline{\lim}_{n\to+\infty} A_n \subseteq (\overline{\lim}_{n\to+\infty} B_n) \cap (\overline{\lim}_{n\to+\infty} C_n)$.

1.1.2. Let $\mathcal{S}_X$ be a σ-algebra of subsets of X, and $f : X \to Y$. Then

$$\mathcal{S}_Y = \{B \subseteq Y \mid f^{-1}(B) \in \mathcal{S}_X\}$$

is called the **push-forward** of $\mathcal{S}_X$ by f on Y. Prove that $\mathcal{S}_Y$ is a σ-algebra of subsets of Y.

1.1.3. Let $\mathcal{S}_Y$ be a σ-algebra of subsets of Y, and $f : X \to Y$. Then

$$\mathcal{S}_X = \{f^{-1}(B) \mid B \in \mathcal{S}_Y\}$$

is called the **pull-back** of $\mathcal{S}_Y$ by f on X. Prove that $\mathcal{S}_X$ is a σ-algebra of subsets of X.

GENERATED σ-ALGEBRAS.

Proposition 1.2. *The intersection of σ-algebras of subsets of X is a σ-algebra of subsets of X.*

Proof. Let **S** be any collection of σ-algebras of subsets of X, and consider $\mathcal{S}_0 = \bigcap\{\mathcal{S} \mid \mathcal{S} \in \mathbf{S}\}$. Since $\emptyset \in \mathcal{S}$ for all $\mathcal{S} \in \mathbf{S}$, we get $\emptyset \in \mathcal{S}_0$, and so $\mathcal{S}_0$ is non-empty.
Let $A \in \mathcal{S}_0$. Then $A \in \mathcal{S}$ for all $\mathcal{S} \in \mathbf{S}$. Since every $\mathcal{S} \in \mathbf{S}$ is a σ-algebra, $A^c \in \mathcal{S}$ for all $\mathcal{S} \in \mathbf{S}$. Therefore, $A^c \in \mathcal{S}_0$.
Let $A_n \in \mathcal{S}_0$ for all $n \in \mathbb{N}$. Then $A_n \in \mathcal{S}$ for all $\mathcal{S} \in \mathbf{S}$ and all $n \in \mathbb{N}$. Since every $\mathcal{S} \in \mathbf{S}$ is a σ-algebra, $\bigcup_{n=1}^{+\infty} A_n \in \mathcal{S}$ for all $\mathcal{S} \in \mathbf{S}$. Thus, $\bigcup_{n=1}^{+\infty} A_n \in \mathcal{S}_0$. ☐

Definition. *Let $\mathcal{C}$ be any collection of subsets of X. The intersection of all σ-algebras $\mathcal{S}$ of subsets of X such that $\mathcal{C} \subseteq \mathcal{S}$ is called the σ-algebra **generated** by $\mathcal{C}$ and we denote it $\mathcal{S}(\mathcal{C})$. I.e.*

$$\mathcal{S}(\mathcal{C}) = \bigcap\{\mathcal{S} \mid \mathcal{S} \text{ is a } \sigma\text{-algebra of subsets of } X \text{ and } \mathcal{C} \subseteq \mathcal{S}\}.$$

Note that there is at least one σ-algebra $\mathcal{S}$ of subsets of X such that $\mathcal{C} \subseteq \mathcal{S}$, namely $\mathcal{S} = \mathcal{P}(X)$. Note also that the term *σ-algebra* used for $\mathcal{S}(\mathcal{C})$ is justified by its definition and Proposition 1.2.

The next straightforward result serves as a tool in many of the following proofs.

Proposition 1.3. *Let $\mathcal{C}$ be any collection of subsets of X. Then $\mathcal{S}(\mathcal{C})$ is the smallest σ-algebra $\mathcal{S}$ of subsets of X such that $\mathcal{C} \subseteq \mathcal{S}$. In other words,*
(i) $\mathcal{S}(\mathcal{C})$ is a σ-algebra of subsets of X,
(ii) $\mathcal{C} \subseteq \mathcal{S}(\mathcal{C})$,
(iii) if $\mathcal{S}$ is any σ-algebra of subsets of X such that $\mathcal{C} \subseteq \mathcal{S}$, then $\mathcal{S}(\mathcal{C}) \subseteq \mathcal{S}$.

Proof. Obvious from the definition of $\mathcal{S}(\mathcal{C})$. ☐

Looking back at two of our examples of σ-algebras, we easily get the following.

Example. Let $E \subseteq X$, and consider $\mathcal{C} = \{E\}$. Then $\mathcal{S}(\mathcal{C}) = \{\emptyset, E, E^c, X\}$.
In fact, $\{\emptyset, E, E^c, X\}$ is a σ-algebra of subsets of X and $\mathcal{C} \subseteq \{\emptyset, E, E^c, X\}$. Moreover, there can be no smaller σ-algebra $\mathcal{S}$ of subsets of X such that $\mathcal{C} \subseteq \mathcal{S}$, since such a σ-algebra $\mathcal{S}$ must necessarily contain $\emptyset, X$ and E^c, besides E.

Example. Let X be an uncountable set, and consider $\mathcal{C} = \{A \subseteq X \mid A \text{ is countable}\}$. Then $\mathcal{S}(\mathcal{C}) = \{A \subseteq X \mid \text{either } A \text{ or } A^c \text{ is countable}\}$.
We know that $\{A \subseteq X \mid \text{either } A \text{ or } A^c \text{ is countable}\}$ is a σ-algebra of subsets of X and, obviously, $\mathcal{C} \subseteq \{A \subseteq X \mid \text{either } A \text{ or } A^c \text{ is countable}\}$. Also, there is no smaller σ-algebra $\mathcal{S}$ of subsets of X such that $\mathcal{C} \subseteq \mathcal{S}$, since any such σ-algebra $\mathcal{S}$ must contain all the complements of countable subsets of X, besides the countable subsets of X.

Exercises.

1.1.4. Let $\mathcal{C}_1, \mathcal{C}_2$ be two collections of subsets of X. If $\mathcal{C}_1 \subseteq \mathcal{C}_2 \subseteq \mathcal{S}(\mathcal{C}_1)$, prove $\mathcal{S}(\mathcal{C}_1) = \mathcal{S}(\mathcal{C}_2)$.

1.1.5. Let $\mathcal{C}_1, \mathcal{C}_2$ be two collections of subsets of X. Prove that $\mathcal{S}(\mathcal{C}_1) = \mathcal{S}(\mathcal{C}_2)$ if and only if $\mathcal{C}_1 \subseteq \mathcal{S}(\mathcal{C}_2)$ and $\mathcal{C}_2 \subseteq \mathcal{S}(\mathcal{C}_1)$.

1.1.6. Let X be a set. In the next three cases find $\mathcal{S}(\mathcal{C})$.
(i) $\mathcal{C} = \emptyset$.
(ii) $\mathcal{C} = \{F \mid E \subseteq F \subseteq X\}$ for some fixed $E \subseteq X$.
(iii) $\mathcal{C} = \{F \mid F$ is a two-point subset of $X\}$.

1.1.7. Let $\mathcal{S}_X$ be a σ-algebra of subsets of X, and $f : X \to Y$, and let $\mathcal{C}_Y$ be a collection of subsets of Y. If $f^{-1}(B) \in \mathcal{S}_X$ for all $B \in \mathcal{C}_Y$, prove that $f^{-1}(B) \in \mathcal{S}_X$ for all $B \in \mathcal{S}(\mathcal{C}_Y)$.
Hint. You may consider the push-forward $\mathcal{S}_Y$ of $\mathcal{S}_X$ by f on Y (see exercise 1.1.2).

1.1.8. Let $\mathcal{C}$ be a collection of subsets of X. Prove that for every $A \in \mathcal{S}(\mathcal{C})$ there is some *countable* subcollection $\mathcal{D}$ of $\mathcal{C}$ so that $A \in \mathcal{S}(\mathcal{D})$.
Hint. Prove that $\bigcup\{\mathcal{S}(\mathcal{D}) \mid \mathcal{D}$ is a countable subcollection of $\mathcal{C}\}$ is a σ-algebra of subsets of X.

ALGEBRAS AND MONOTONE CLASSES.

Definition. *Let $\mathcal{A}$ be a collection of subsets of X. We call $\mathcal{A}$ an* ***algebra*** *of subsets of X if it is non-empty, closed under complements, and closed under unions. This means:*
(i) there exists at least one $A \subseteq X$ so that $A \in \mathcal{A}$,
(ii) if $A \in \mathcal{A}$, then $A^c \in \mathcal{A}$,
(iii) if $A, B \in \mathcal{A}$, then $A \cup B \in \mathcal{A}$.

Proposition 1.4. *Every algebra of subsets of X contains the sets $\emptyset$ and X, it is closed under finite unions, under finite intersections, and under set-theoretic differences.*

Proof. Similar to the proof of Proposition 1.1. $\square$

Example. Every σ-algebra of subsets of X is also an algebra of subsets of X.

Example. If X is an infinite set, then the collection $\mathcal{A} = \{A \subseteq X \mid$ either A or A^c is finite$\}$ is an algebra of subsets of X, but not a σ-algebra of subsets of X.
The proof that $\mathcal{A}$ is an algebra is similar to the proof in the last example of the first subsection. To prove that $\mathcal{A}$ is not a σ-algebra, we consider any countably infinite $A \subseteq X$ so that A^c is infinte. If $A = \{x_1, x_2, \ldots\}$, then the sets $A_n = \{x_1, \ldots, x_n\}$ belong to $\mathcal{A}$ and $\bigcup_{n=1}^{+\infty} A_n = A$, but A does not belong to $\mathcal{A}$.

Definition. *Let $\mathcal{M}$ be a collection of subsets of X. We call $\mathcal{M}$ a* ***monotone class*** *of subsets of X if it is closed under countable increasing unions and under countable decreasing intersections. I.e.*
(i) if $A_n \in \mathcal{M}$ for all $n \in \mathbb{N}$ and (A_n) is increasing, then $\bigcup_{n=1}^{+\infty} A_n \in \mathcal{M}$,
(ii) if $A_n \in \mathcal{M}$ for all $n \in \mathbb{N}$ and (A_n) is decreasing, then $\bigcap_{n=1}^{+\infty} A_n \in \mathcal{M}$.

It is obvious that every σ-algebra is a non-empty monotone class.

Proposition 1.5. *The intersection of monotone classes of subsets of X is a monotone class of subsets of X.*

Proof. Take any collection $\mathbf{M}$ of monotone classes of subsets of X, and $\mathcal{M}_0 = \bigcap\{\mathcal{M} \mid \mathcal{M} \in \mathbf{M}\}$. Let $A_n \in \mathcal{M}_0$ for all $n \in \mathbb{N}$ and (A_n) be increasing. Then $A_n \in \mathcal{M}$ for all $\mathcal{M} \in \mathbf{M}$ and all $n \in \mathbb{N}$. Since every $\mathcal{M} \in \mathbf{M}$ is a monotone class, we have that $\bigcup_{n=1}^{+\infty} A_n \in \mathcal{M}$ for all $\mathcal{M} \in \mathbf{M}$. Thus, $\bigcup_{n=1}^{+\infty} A_n \in \mathcal{M}_0$, and so $\mathcal{M}_0$ is closed under countable increasing unions.
Similarly, let $A_n \in \mathcal{M}_0$ for all $n \in \mathbb{N}$ and (A_n) be decreasing. Then $A_n \in \mathcal{M}$ for all $\mathcal{M} \in \mathbf{M}$ and all $n \in \mathbb{N}$. Since every $\mathcal{M} \in \mathbf{M}$ is a monotone class, we have that $\bigcap_{n=1}^{+\infty} A_n \in \mathcal{M}$ for all $\mathcal{M} \in \mathbf{M}$. Hence, $\bigcap_{n=1}^{+\infty} A_n \in \mathcal{M}_0$, and so $\mathcal{M}_0$ is closed under countable decreasing intersections. $\square$

Definition. *Let $\mathcal{C}$ be any collection of subsets of X. The intersection of all monotone classes $\mathcal{M}$ of subsets of X such that $\mathcal{C} \subseteq \mathcal{M}$ is called the monotone class **generated** by $\mathcal{C}$ and we denote it $\mathcal{M}(\mathcal{C})$. I.e.*

$$\mathcal{M}(\mathcal{C}) = \bigcap\{\mathcal{M} \mid \mathcal{M} \text{ is a monotone class of subsets of } X \text{ and } \mathcal{C} \subseteq \mathcal{M}\}.$$

There is at least one monotone class $\mathcal{M}$ of subsets of X such that $\mathcal{C} \subseteq \mathcal{M}$, namely $\mathcal{M} = \mathcal{P}(X)$. We also note that the term *monotone class* used for $\mathcal{M}(\mathcal{C})$ is justified by its definition and Proposition 1.5.

Proposition 1.6. *Let $\mathcal{C}$ be any collection of subsets of X. Then $\mathcal{M}(\mathcal{C})$ is the smallest monotone class $\mathcal{M}$ of subsets of X such that $\mathcal{C} \subseteq \mathcal{M}$. In other words,*
(i) $\mathcal{M}(\mathcal{C})$ is a monotone class of subsets of X,
(ii) $\mathcal{C} \subseteq \mathcal{M}(\mathcal{C})$,
(iii) if $\mathcal{M}$ is any monotone class of subsets of X such that $\mathcal{C} \subseteq \mathcal{M}$, then $\mathcal{M}(\mathcal{C}) \subseteq \mathcal{M}$.

Proof. Obvious from the definition of $\mathcal{M}(\mathcal{C})$. □

Proposition 1.7. *Let $\mathcal{A}$ be an algebra of subsets of X. Then $\mathcal{M}(\mathcal{A}) = \mathcal{S}(\mathcal{A})$.*

Proof. $\mathcal{S}(\mathcal{A})$ is a σ-algebra and, hence, a monotone class. Since $\mathcal{A} \subseteq \mathcal{S}(\mathcal{A})$, Proposition 1.6 implies $\mathcal{M}(\mathcal{A}) \subseteq \mathcal{S}(\mathcal{A})$.
Now it is enough to prove that $\mathcal{M}(\mathcal{A})$ is a σ-algebra. Since $\mathcal{A} \subseteq \mathcal{M}(\mathcal{A})$, Proposition 1.3 will then immediately imply that $\mathcal{S}(\mathcal{A}) \subseteq \mathcal{M}(\mathcal{A})$, and this will conclude the proof.
$\mathcal{M}(\mathcal{A})$ is non-empty, since $\emptyset \in \mathcal{A} \subseteq \mathcal{M}(\mathcal{A})$.
Now fix any $A \in \mathcal{A}$ and consider the collection

$$\mathcal{M}_A = \{B \subseteq X \mid A \cup B \in \mathcal{M}(\mathcal{A})\}.$$

It is very easy to show that $\mathcal{A} \subseteq \mathcal{M}_A$ and that $\mathcal{M}_A$ is a monotone class of subsets of X.
In fact, if $B \in \mathcal{A}$, then $A \cup B \in \mathcal{A} \subseteq \mathcal{M}(\mathcal{A})$, and so $B \in \mathcal{M}_A$.
Also, let $B_n \in \mathcal{M}_A$ for all $n \in \mathbb{N}$ and (B_n) be increasing. Then $A \cup B_n \in \mathcal{M}(\mathcal{A})$ for all $n \in \mathbb{N}$ and $(A \cup B_n)$ is increasing. Since $\mathcal{M}(\mathcal{A})$ is a monotone class, $\bigcup_{n=1}^{+\infty}(A \cup B_n) \in \mathcal{M}(\mathcal{A})$. Since

$$\bigcup_{n=1}^{+\infty}(A \cup B_n) = A \cup \big(\bigcup_{n=1}^{+\infty} B_n\big),$$

we get $A \cup \big(\bigcup_{n=1}^{+\infty} B_n\big) \in \mathcal{M}(\mathcal{A})$, and so $\bigcup_{n=1}^{+\infty} B_n \in \mathcal{M}_A$. Therefore, $\mathcal{M}_A$ is closed under countable increasing unions.
In the same manner we can prove that $\mathcal{M}_A$ is closed under countable decreasing intersections, and we conclude that it is a monotone class.
Proposition 1.6 implies $\mathcal{M}(\mathcal{A}) \subseteq \mathcal{M}_A$. This means:

$$A \cup B \in \mathcal{M}(\mathcal{A}) \quad \text{for all } A \in \mathcal{A} \text{ and all } B \in \mathcal{M}(\mathcal{A}). \tag{1.1}$$

Now fix any $B \in \mathcal{M}(\mathcal{A})$ and consider $\mathcal{M}_B = \{A \subseteq X \mid A \cup B \in \mathcal{M}(\mathcal{A})\}$ again.
We just proved that $\mathcal{M}_B$ is a monotone class of subsets of X. Moreover, (1.1) implies $\mathcal{A} \subseteq \mathcal{M}_B$.
Again, Proposition 1.6 implies $\mathcal{M}(\mathcal{A}) \subseteq \mathcal{M}_B$, which means:

$$A \cup B \in \mathcal{M}(\mathcal{A}) \quad \text{for all } A \in \mathcal{M}(\mathcal{A}) \text{ and all } B \in \mathcal{M}(\mathcal{A}). \tag{1.2}$$

Now consider the collection

$$\mathcal{M} = \{A \subseteq X \mid A^c \in \mathcal{M}(\mathcal{A})\}.$$

Assume that $A_n \in \mathcal{M}$ for every $n \in \mathbb{N}$ and that (A_n) is increasing. Then $A_n^c \in \mathcal{M}(\mathcal{A})$ for every $n \in \mathbb{N}$ and (A_n^c) is decreasing. Since $\mathcal{M}(\mathcal{A})$ is a monotone class, we get that $\bigcap_{n=1}^{+\infty} A_n^c \in \mathcal{M}(\mathcal{A})$. Since

$$\bigcap_{n=1}^{+\infty} A_n^c = \big(\bigcup_{n=1}^{+\infty} A_n\big)^c,$$

we have that $\big(\bigcup_{n=1}^{+\infty} A_n\big)^c \in \mathcal{M}(\mathcal{A})$ and so $\bigcup_{n=1}^{+\infty} A_n \in \mathcal{M}$. Therefore, $\mathcal{M}$ is closed under countable increasing unions.

In the same manner we can prove that $\mathcal{M}$ is closed under countable decreasing intersections, and we conclude that $\mathcal{M}$ is a monotone class. Moreover, $\mathcal{A} \subseteq \mathcal{M}$ (because, if $A \in \mathcal{A}$, then $A^c \in \mathcal{A}$, and so $A^c \in \mathcal{M}(\mathcal{A})$, and so $A \in \mathcal{M}$). Hence, $\mathcal{M}(\mathcal{A}) \subseteq \mathcal{M}$, which means:

$$A^c \in \mathcal{M}(\mathcal{A}) \quad \text{for all } A \in \mathcal{M}(\mathcal{A}). \tag{1.3}$$

Now (1.2) and (1.3) imply that $\mathcal{M}(\mathcal{A})$ is an algebra of subsets of X.

Finally, let $A_n \in \mathcal{M}(\mathcal{A})$ for all $n \in \mathbb{N}$. We consider $B_n = A_1 \cup \cdots \cup A_n$ for all n. Since $\mathcal{M}(\mathcal{A})$ is an algebra, $B_n \in \mathcal{M}(\mathcal{A})$ for all n. It is clear that (B_n) is increasing, and, since $\mathcal{M}(\mathcal{A})$ is a monotone class, $\bigcup_{n=1}^{+\infty} B_n \in \mathcal{M}(\mathcal{A})$. But

$$\bigcup_{n=1}^{+\infty} A_n = \bigcup_{n=1}^{+\infty} B_n,$$

and so $\bigcup_{n=1}^{+\infty} A_n \in \mathcal{M}(\mathcal{A})$.

Therefore, $\mathcal{M}(\mathcal{A})$ is a σ-algebra. □

Exercises.

1.1.9. Let $\mathcal{A}$ be an algebra of subsets of X. Prove that $\mathcal{A}$ is a σ-algebra if and only if it is closed under countable increasing unions.

1.1.10. Prove that the intersection of algebras of subsets of X is an algebra of subsets of X.

1.1.11. Find $\mathcal{M}(\mathcal{C})$ in the three cases of exercise 1.1.6.

1.1.12. Prove that every *finite* collection of subsets of X is a monotone class of subsets of X.

RESTRICTION OF A σ-ALGEBRA.

Definition. *Let $\mathcal{C}$ be any collection of subsets of X, and $Y \subseteq X$. We define*

$$\mathcal{C}\rceil Y = \{A \cap Y \mid A \in \mathcal{C}\}.$$

This is a collection of subsets of Y, and we call it the ***restriction*** *of $\mathcal{C}$ on Y.*

Proposition 1.8. *Let $\mathcal{S}$ be a σ-algebra of subsets of X, and $Y \subseteq X$. Then $\mathcal{S}\rceil Y$ is a σ-algebra of subsets of Y. If, also, $Y \in \mathcal{S}$, then $\mathcal{S}\rceil Y = \{A \subseteq Y \mid A \in \mathcal{S}\}$.*

Proof. Since $\emptyset \in \mathcal{S}$, we have that $\emptyset = \emptyset \cap Y \in \mathcal{S}\rceil Y$.

Let $B \in \mathcal{S}\rceil Y$. Then $B = A \cap Y$ for some $A \in \mathcal{S}$. Since

$$Y \setminus B = (X \setminus A) \cap Y$$

and $X \setminus A \in \mathcal{S}$, we have that $Y \setminus B \in \mathcal{S}\rceil Y$.

Let $B_n \in \mathcal{S}\rceil Y$ for every $n \in \mathbb{N}$. Then for each n there is $A_n \in \mathcal{S}$ so that $B_n = A_n \cap Y$. Since

$$\bigcup_{n=1}^{+\infty} B_n = \bigcup_{n=1}^{+\infty}(A_n \cap Y) = \big(\bigcup_{n=1}^{+\infty} A_n\big) \cap Y$$

and $\bigcup_{n=1}^{+\infty} A_n \in \mathcal{S}$, we find that $\bigcup_{n=1}^{+\infty} B_n \in \mathcal{S}\rceil Y$.

Therefore, $\mathcal{S}\rceil Y$ is a σ-algebra of subsets of Y.

Now let $Y \in \mathcal{S}$.

If $B \in \mathcal{S}\rceil Y$, then $B = A \cap Y$ for some $A \in \mathcal{S}$, and so $B \subseteq Y$ and $B \in \mathcal{S}$. Therefore, $B \in \{C \subseteq Y \mid C \in \mathcal{S}\}$. Conversely, if $B \in \{C \subseteq Y \mid C \in \mathcal{S}\}$, then $B \subseteq Y$ and $B \in \mathcal{S}$. We set $A = B$, and we get $B = A \cap Y$ and $A \in \mathcal{S}$. Hence, $B \in \mathcal{S}\rceil Y$. □

Proposition 1.9. *Let $\mathcal{C}$ be a collection of subsets of X, and $Y \subseteq X$. If $\mathcal{S}(\mathcal{C}\rceil Y)$ is the σ-algebra of subsets of Y generated by $\mathcal{C}\rceil Y$, then $\mathcal{S}(\mathcal{C}\rceil Y) = \mathcal{S}(\mathcal{C})\rceil Y$.*

Proof. If $B \in \mathcal{C}\rceil Y$, then $B = A \cap Y$ for some $A \in \mathcal{C} \subseteq \mathcal{S}(\mathcal{C})$, and so $B \in \mathcal{S}(\mathcal{C})\rceil Y$. Thus, $\mathcal{C}\rceil Y \subseteq \mathcal{S}(\mathcal{C})\rceil Y$. Proposition 1.8 says that $\mathcal{S}(\mathcal{C})\rceil Y$ is a σ-algebra of subsets of Y, and now Proposition 1.3 implies $\mathcal{S}(\mathcal{C}\rceil Y) \subseteq \mathcal{S}(\mathcal{C})\rceil Y$.
Now we define the following collection of subsets of X:

$$\mathcal{S} = \{A \subseteq X \mid A \cap Y \in \mathcal{S}(\mathcal{C}\rceil Y)\}.$$

We have that $\emptyset \in \mathcal{S}$, because $\emptyset \cap Y = \emptyset \in \mathcal{S}(\mathcal{C}\rceil Y)$.
If $A \in \mathcal{S}$, then $A \cap Y \in \mathcal{S}(\mathcal{C}\rceil Y)$. Then $X \setminus A \in \mathcal{S}$, since

$$(X \setminus A) \cap Y = Y \setminus (A \cap Y) \in \mathcal{S}(\mathcal{C}\rceil Y).$$

If $A_n \in \mathcal{S}$ for all $n \in \mathbb{N}$, then $A_n \cap Y \in \mathcal{S}(\mathcal{C}\rceil Y)$ for all $n \in \mathbb{N}$. This implies that

$$\left(\bigcup_{n=1}^{+\infty} A_n\right) \cap Y = \bigcup_{n=1}^{+\infty}(A_n \cap Y) \in \mathcal{S}(\mathcal{C}\rceil Y),$$

and so $\bigcup_{n=1}^{+\infty} A_n \in \mathcal{S}$.
We conclude that $\mathcal{S}$ is a σ-algebra of subsets of X.
If $A \in \mathcal{C}$, then $A \cap Y \in \mathcal{C}\rceil Y \subseteq \mathcal{S}(\mathcal{C}\rceil Y)$, and so $A \in \mathcal{S}$. Thus, $\mathcal{C} \subseteq \mathcal{S}$, and now Proposition 1.3 implies $\mathcal{S}(\mathcal{C}) \subseteq \mathcal{S}$.
Now, for an arbitrary $B \in \mathcal{S}(\mathcal{C})\rceil Y$, we have that $B = A \cap Y$ for some $A \in \mathcal{S}(\mathcal{C}) \subseteq \mathcal{S}$ and, thus, $B \in \mathcal{S}(\mathcal{C}\rceil Y)$. Hence, $\mathcal{S}(\mathcal{C})\rceil Y \subseteq \mathcal{S}(\mathcal{C}\rceil Y)$. □

Exercises.

1.1.13. Let $Y \subseteq X$, and $\mathcal{A}$ be an algebra of subsets of X. Prove that $\mathcal{A}\rceil Y$ is an algebra of subsets of Y.

BOREL σ-ALGEBRAS.

Definition. *Let X be a topological space, and $\mathcal{T}$ be the topology of X, i.e. the collection of all open subsets of X. The σ-algebra of subsets of X which is generated by $\mathcal{T}$, namely the smallest σ-algebra $\mathcal{S}$ of subsets of X such that $\mathcal{T} \subseteq \mathcal{S}$, is called the **Borel σ-algebra** of X and we denote it $\mathcal{B}_X$. I.e.*

$$\mathcal{B}_X = \mathcal{S}(\mathcal{T}).$$

*The elements of $\mathcal{B}_X$ are called **Borel subsets** of X, and $\mathcal{B}_X$ is also called the σ-algebra of Borel subsets of X.*

If there is no danger of confusion, we shall say *open set* instead of open subset of X and *Borel set* instead of Borel subset of X.

By definition, all open sets are Borel sets and, since $\mathcal{B}_X$ is a σ-algebra, all closed sets (which are the complements of open sets) are also Borel sets. Hence, every countable intersection of open sets and every countable union of closed sets is a Borel set.

If X is a topological space with topology $\mathcal{T}$ and if $Y \subseteq X$, then, as is well-known (and easy to prove), the collection $\mathcal{T}\rceil Y = \{U \cap Y \mid U \in \mathcal{T}\}$ is a topology of Y which is called the **relative topology** or the **subspace topology** of Y.

Proposition 1.10. *Let $Y \subseteq X$. If X is a topological space and Y has the subspace topology, then $\mathcal{B}_Y = \mathcal{B}_X\rceil Y$.*

Proof. An application of Proposition 1.9: $\mathcal{B}_Y = \mathcal{S}(\mathcal{T}\rceil Y) = \mathcal{S}(\mathcal{T})\rceil Y = \mathcal{B}_X\rceil Y$. □

Thus, *the Borel subsets of Y (with the subspace topology of Y) are just the intersections with Y of the Borel subsets of X.*

Examples of topological spaces are the metric spaces. The most familiar metric space is the Euclidean space $\mathbb{R}^n$ with the usual Euclidean metric. Because of the importance of $\mathbb{R}^n$ we shall pay particular attention to $\mathcal{B}_{\mathbb{R}^n}$. Instead of $\mathcal{B}_{\mathbb{R}^n}$ we shall use the simpler symbol $\mathcal{B}_n$:

$$\mathcal{B}_n = \mathcal{B}_{\mathbb{R}^n}.$$

The typical *bounded orthogonal parallelepiped with axis-parallel edges* in $\mathbb{R}^n$ is a set of the form $S = I_1 \times \cdots \times I_n$, where each I_j is a bounded interval in $\mathbb{R}$. The bounded orthogonal parallelepipeds with axis-parallel edges are called *closed* or *open* or *open-closed* or *closed-open* if they are, respectively, of the form $Q = [a_1, b_1] \times \cdots \times [a_n, b_n]$ or $R = (a_1, b_1) \times \cdots \times (a_n, b_n)$ or $P = (a_1, b_1] \times \cdots \times (a_n, b_n]$ or $T = [a_1, b_1) \times \cdots \times [a_n, b_n)$. An orthogonal parallelepiped with axis-parallel edges $S = I_1 \times \cdots \times I_n$ is empty if at least one of the I_j is the empty interval in $\mathbb{R}$.

If we allow at least one of the intervals I_j in $\mathbb{R}$ to be unbounded (and none of them to be empty), then $S = I_1 \times \cdots \times I_n$ is the typical *unbounded orthogonal parallelepiped with axis-parallel edges* in $\mathbb{R}^n$. Again, certain unbounded orthogonal parallelepipeds with axis-parallel edges in $\mathbb{R}^n$ are closed or open or open-closed or closed-open.

Since orthogonal parallelepipeds with axis-parallel edges will play a role in much of the following, we agree to call them, for short, **n-dimensional intervals** or **intervals in $\mathbb{R}^n$**.

The typical open-closed interval in $\mathbb{R}^n$ is of the form $P = (a_1, b_1] \times \cdots \times (a_n, b_n]$, where $-\infty \le a_j \le b_j \le +\infty$ for all j. (Of course, when we write $(a, +\infty]$ we mean $(a, +\infty)$.) The space $\mathbb{R}^n$ is an open-closed interval, as well as any of the half spaces $\{(x_1, \ldots, x_n) \mid x_j \le b_j\}$ and $\{(x_1, \ldots, x_n) \mid a_j < x_j\}$. In fact, every open-closed interval in $\mathbb{R}^n$ is, obviously, the intersection of $2n$ such half-spaces.

Proposition 1.11. *All n-dimensional intervals are Borel sets in $\mathbb{R}^n$.*

Proof. A half-space of the form $\{(x_1, \ldots, x_n) \mid x_j < b_j\}$ or of the form $\{(x_1, \ldots, x_n) \mid x_j \le b_j\}$ is a Borel set in $\mathbb{R}^n$, since it is an open set or a closed set, respectively. Similarly, a half-space of the form $\{(x_1, \ldots, x_n) \mid a_j < x_j\}$ or of the form $\{(x_1, \ldots, x_n) \mid a_j \le x_j\}$ is a Borel set in $\mathbb{R}^n$. Now, an interval S in $\mathbb{R}^n$ is the intersection of $2n$ such half-spaces and so it is a Borel set in $\mathbb{R}^n$. □

Proposition 1.12. *(i) If $\mathcal{C}$ is the collection of all closed or of all open or of all open-closed or of all closed-open bounded intervals or of all bounded intervals in $\mathbb{R}^n$, then $\mathcal{B}_n = \mathcal{S}(\mathcal{C})$.*
(ii) If $\mathcal{C}$ is the collection of all intervals $(a, +\infty)$ in $\mathbb{R}$, then $\mathcal{B}_1 = \mathcal{S}(\mathcal{C})$.

Proof. (i) In all cases, Proposition 1.11 implies $\mathcal{C} \subseteq \mathcal{B}_n$, and so $\mathcal{S}(\mathcal{C}) \subseteq \mathcal{B}_n$.
To show the opposite inclusion we consider any open subset U of $\mathbb{R}^n$. For every $x \in U$ there is a small open ball B_x centered at x which is included in U. Now, considering the case of $\mathcal{C}$ being the collection of all closed bounded intervals, there is a $Q_x = [a_1, b_1] \times \cdots \times [a_n, b_n]$ containing x, which is small enough so that it is included in B_x, and hence in U, and with all $a_1, \ldots, a_n, b_1, \ldots, b_n$ being *rational* numbers. Since $x \in Q_x \subseteq U$ for all $x \in U$, we have that

$$U = \bigcup_{x \in U} Q_x.$$

But the collection of all possible Q_x is countable, and so the general open subset U of $\mathbb{R}^n$ can be written as a countable union of sets in the collection $\mathcal{C}$. Therefore, every open subset U of $\mathbb{R}^n$ belongs to $\mathcal{S}(\mathcal{C})$. Since $\mathcal{S}(\mathcal{C})$ is a σ-algebra of subsets of $\mathbb{R}^n$, and since $\mathcal{B}_n$ is generated by the collection of all open subsets of $\mathbb{R}^n$, we conclude that $\mathcal{B}_n \subseteq \mathcal{S}(\mathcal{C})$.
Of course, the proof of the last inclusion works in the same manner with all other types of intervals.
(ii) Again, we have that $\mathcal{C} \subseteq \mathcal{B}_1$, and so $\mathcal{S}(\mathcal{C}) \subseteq \mathcal{B}_1$.
Moreover, $(a, b] = (a, +\infty) \setminus (b, +\infty) \in \mathcal{S}(\mathcal{C})$ for all $(a, b]$. By (i), the collection of all $(a, b]$ generates $\mathcal{B}_1$. Therefore, $\mathcal{B}_1 \subseteq \mathcal{S}(\mathcal{C})$. □

Proposition 1.13. *The collection*

$$\mathcal{A} = \left\{ \bigcup_{i=1}^{k} P_i \,\middle|\, k \in \mathbb{N},\ P_1, \dots, P_k \text{ are pairwise disjoint open-closed intervals in } \mathbb{R}^n \right\}$$

is an algebra of subsets of $\mathbb{R}^n$*. In particular, the following are true:*
(i) The intersection of two open-closed intervals is an open-closed interval.
(ii) For all open-closed intervals $P, P_1, \dots, P_m$ *there are pairwise disjoint open-closed intervals* $P'_1, \dots, P'_k$ *so that* $P \setminus (P_1 \cup \dots \cup P_m) = P'_1 \cup \dots \cup P'_k$*.*
(iii) For all open-closed intervals $P_1, \dots, P_m$ *there are pairwise disjoint open-closed intervals* $P'_1, \dots, P'_k$ *so that* $P_1 \cup \dots \cup P_m = P'_1 \cup \dots \cup P'_k$*.*

Proof. If $P' = (a'_1, b'_1] \times \dots \times (a'_n, b'_n]$ and $P'' = (a''_1, b''_1] \times \dots \times (a''_n, b''_n]$ are not disjoint, then $a_j < b_j$ for all j, where $a_j = \max\{a'_j, a''_j\}$ and $b_j = \min\{b'_j, b''_j\}$, and then

$$P' \cap P'' = (a_1, b_1] \times \dots \times (a_n, b_n].$$

This proves (i).
If $A' = \bigcup_{i=1}^{k} P'_i$ and $A'' = \bigcup_{j=1}^{l} P''_j$, where the open-closed intervals $P'_1, \dots, P'_k$ are pairwise disjoint and the open-closed intervals $P''_1, \dots, P''_l$ are also pairwise disjoint, then

$$A' \cap A'' = \bigcup_{1 \le i \le k,\ 1 \le j \le l} (P'_i \cap P''_j).$$

The sets $P'_i \cap P''_j$ are pairwise disjoint open-closed intervals, as we have just seen.
Thus, $\mathcal{A}$ is closed under finite intersections.
Consider the open-closed interval $P = (a_1, b_1] \times \dots \times (a_n, b_n]$. It is easy to see that P^c can be written as the union of $2n$ pairwise disjoint open-closed intervals. To express this in a concise way, for every $I = (a, b]$ denote $I^{(l)} = (-\infty, a]$ and $I^{(r)} = (b, +\infty]$ the left and right complementary intervals of I in $\mathbb{R}$ (they may be empty). If we write $P = I_1 \times \dots \times I_n$, then P^c is equal to

$$\begin{gathered}
(I_1^{(l)} \times \mathbb{R} \times \dots \times \mathbb{R}) \ \cup \ (I_1^{(r)} \times \mathbb{R} \times \dots \times \mathbb{R}) \\
\cup\ (I_1 \times I_2^{(l)} \times \mathbb{R} \times \dots \times \mathbb{R}) \ \cup \ (I_1 \times I_2^{(r)} \times \mathbb{R} \times \dots \times \mathbb{R}) \\
\dots\dots\dots\dots \\
\cup\ (I_1 \times \dots \times I_{n-2} \times I_{n-1}^{(l)} \times \mathbb{R}) \ \cup \ (I_1 \times \dots \times I_{n-2} \times I_{n-1}^{(r)} \times \mathbb{R}) \\
\cup\ (I_1 \times \dots \times I_{n-1} \times I_n^{(l)}) \ \cup \ (I_1 \times \dots \times I_{n-1} \times I_n^{(r)}),
\end{gathered}$$

i.e. the union of pairwise disjoint open-closed intervals. Thus, the complement P^c of every open-closed interval P is an element of $\mathcal{A}$.
Now, if $A = \bigcup_{i=1}^{k} P_i$, where the open-closed intervals $P_1, \dots, P_k$ are pairwise disjoint, is any element of $\mathcal{A}$, then $A^c = \bigcap_{i=1}^{k} P_i^c$ is a finite intersection of elements of $\mathcal{A}$. Since $\mathcal{A}$ is closed under finite intersections, we have that $A^c \in \mathcal{A}$, and so $\mathcal{A}$ is closed under complements.
Finally, if $A', A'' \in \mathcal{A}$, then $A' \cup A'' = (A'^c \cap A''^c)^c \in \mathcal{A}$, and so $\mathcal{A}$ is closed under finite unions.
Therefore, $\mathcal{A}$ is an algebra of subsets of $\mathbb{R}^n$, and then (ii) and (iii) are immediate. □

It is convenient for certain purposes, and especially because functions are often infinite valued, to consider $\overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty, -\infty\}$ and $\overline{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$ as topological spaces and define their Borel σ-algebras.

The ϵ-neighborhood of a point $x \in \mathbb{R}$ is, as usual, the interval $(x-\epsilon, x+\epsilon)$. Now we *define* the ϵ-neighborhood of $+\infty$ to be $(\frac{1}{\epsilon}, +\infty]$, and the ϵ-neighborhood of $-\infty$ to be $[-\infty, -\frac{1}{\epsilon})$. We also say that $U \subseteq \overline{\mathbb{R}}$ is an **open** subset of $\overline{\mathbb{R}}$ if every point of U has an ϵ-neighborhood (the ϵ depending on the point) which is included in U. It is trivial to see (justifying the term *open*) that the collection of all open subsets of $\overline{\mathbb{R}}$ is a topology of $\overline{\mathbb{R}}$, namely that it contains the sets $\emptyset$ and $\overline{\mathbb{R}}$ and that it is closed under arbitrary unions and under finite intersections. It is obvious that a set $U \subseteq \mathbb{R}$ is an open subset of $\overline{\mathbb{R}}$ if and only if it is an open subset of $\mathbb{R}$. In particular, $\mathbb{R}$ itself is an open subset of $\overline{\mathbb{R}}$. It is also obvious that, if a set $U \subseteq \overline{\mathbb{R}}$ is an open subset of $\overline{\mathbb{R}}$, then $U \cap \mathbb{R}$ is an open subset of $\mathbb{R}$. Therefore, the topology of $\mathbb{R}$ coincides with its subspace topology as a subset of $\overline{\mathbb{R}}$.

The next result says, in particular, that *we may construct the general Borel subset of* $\overline{\mathbb{R}}$ *by taking the general Borel subset of* $\mathbb{R}$ *and adjoining none or any one or both of the points* $+\infty$, $-\infty$ *to it.*

Proposition 1.14. *(i)* $\mathcal{B}_1 = \overline{\mathcal{B}}_1 \rceil \mathbb{R}$, *where we denote* $\overline{\mathcal{B}}_1$ *the Borel* σ*-algebra of* $\overline{\mathbb{R}}$.
(ii) $\overline{\mathcal{B}}_1 = \big\{A, A \cup \{+\infty\}, A \cup \{-\infty\}, A \cup \{+\infty, -\infty\} \,|\, A \in \mathcal{B}_1\big\}$.
(iii) If $\mathcal{C}$ *is the collection containing* $\{+\infty\}$ *or* $\{-\infty\}$ *and all closed or all open or all open-closed or all closed-open or all bounded intervals in* $\mathbb{R}$, *then* $\overline{\mathcal{B}}_1 = \mathcal{S}(\mathcal{C})$.
(iv) If $\mathcal{C}$ *is the collection of all intervals* $(a, +\infty]$ *in* $\overline{\mathbb{R}}$, *then* $\overline{\mathcal{B}}_1 = \mathcal{S}(\mathcal{C})$.

Proof. (i) Immediate from Proposition 1.10.
(ii) $\mathbb{R}$ is open in $\overline{\mathbb{R}}$, and so $\mathbb{R} \in \overline{\mathcal{B}}_1$. Now (i) and the last statement in Proposition 1.8 imply that

$$\mathcal{B}_1 = \{A \subseteq \mathbb{R} \,|\, A \in \overline{\mathcal{B}}_1\}.$$

Therefore, if $A \in \mathcal{B}_1$, then $A \in \overline{\mathcal{B}}_1$. Also, $[-\infty, +\infty)$ is open in $\overline{\mathbb{R}}$, and so $\{+\infty\} \in \overline{\mathcal{B}}_1$. Similarly, $\{-\infty\} \in \overline{\mathcal{B}}_1$ and $\{+\infty, -\infty\} \in \overline{\mathcal{B}}_1$, and we conclude that

$$\big\{A, A \cup \{+\infty\}, A \cup \{-\infty\}, A \cup \{+\infty, -\infty\} \,|\, A \in \mathcal{B}_1\big\} \subseteq \overline{\mathcal{B}}_1.$$

Conversely, let $B \in \overline{\mathcal{B}}_1$ and consider $A = B \cap \mathbb{R} \in \mathcal{B}_1$. Then $B = A$ or $B = A \cup \{+\infty\}$ or $B = A \cup \{-\infty\}$ or $B = A \cup \{+\infty, -\infty\}$, and we conclude that

$$\overline{\mathcal{B}}_1 \subseteq \big\{A, A \cup \{+\infty\}, A \cup \{-\infty\}, A \cup \{+\infty, -\infty\} \,|\, A \in \mathcal{B}_1\big\}.$$

(iii) Let $\mathcal{C} = \big\{\{+\infty\}, (a, b] \,|\, -\infty < a \leq b < +\infty\big\}$.
From all the above we get that $\mathcal{C} \subseteq \overline{\mathcal{B}}_1$, and so $\mathcal{S}(\mathcal{C}) \subseteq \overline{\mathcal{B}}_1$.
If $A \in \mathcal{B}_1$, then Proposition 1.12 implies $A \in \mathcal{S}(\mathcal{C})$. In particular, $\mathbb{R} \in \mathcal{S}(\mathcal{C})$, and so $(-\infty, +\infty] = \mathbb{R} \cup \{+\infty\} \in \mathcal{S}(\mathcal{C})$. Therefore, $\{-\infty\} = \overline{\mathbb{R}} \setminus (-\infty, +\infty] \in \mathcal{S}(\mathcal{C})$, and $\{+\infty, -\infty\} = \{+\infty\} \cup \{-\infty\} \in \mathcal{S}(\mathcal{C})$. From all these and from (ii) we conclude that

$$\overline{\mathcal{B}}_1 = \big\{A, A \cup \{+\infty\}, A \cup \{-\infty\}, A \cup \{+\infty, -\infty\} \,|\, A \in \mathcal{B}_1\big\} \subseteq \mathcal{S}(\mathcal{C}).$$

The proof is similar for all other choices of $\mathcal{C}$.
(iv) We have that $\mathcal{C} \subseteq \overline{\mathcal{B}}_1$, and so $\mathcal{S}(\mathcal{C}) \subseteq \overline{\mathcal{B}}_1$.
Now, $\{+\infty\} = \bigcap_{n=1}^{+\infty}(n, +\infty] \in \mathcal{S}(\mathcal{C})$. Also $(a, b] = (a, +\infty] \setminus (b, +\infty] \in \mathcal{S}(\mathcal{C})$ for all $(a, b]$. By (iii), the collection containing $\{+\infty\}$ and all $(a, b]$ generates $\overline{\mathcal{B}}_1$. Therefore, $\overline{\mathcal{B}}_1 \subseteq \mathcal{S}(\mathcal{C})$. $\square$

We now turn to the case of $\overline{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$.

The ϵ-neighborhood of a point $x = (x_1, x_2) = x_1 + ix_2 \in \mathbb{C}$ is, as usual, the open disc $B(x; \epsilon) = \{y = (y_1, y_2) \in \mathbb{C} \,|\, |y - x| < \epsilon\}$, where $|y - x| = ((y_1 - x_1)^2 + (y_2 - x_2)^2)^{1/2}$. We *define* the ϵ-neighborhood of ∞ to be the set $\big\{y \in \mathbb{C} \,\big|\, |y| > \frac{1}{\epsilon}\big\} \cup \{\infty\}$, i.e. the complement of a closed disc centered at 0 (together with the point ∞). We say that $U \subseteq \overline{\mathbb{C}}$ is an **open** subset of $\overline{\mathbb{C}}$ if every point of U has an ϵ-neighborhood (the ϵ depending on the point) which is included in U. The collection of all open subsets of $\overline{\mathbb{C}}$ contains $\emptyset$ and $\overline{\mathbb{C}}$ and it is closed under arbitrary unions and under finite intersections, thus forming a topology of $\overline{\mathbb{C}}$. It is clear that $U \subseteq \mathbb{C}$ is an open subset of $\overline{\mathbb{C}}$ if and only if it is an open subset of $\mathbb{C}$. In particular, $\mathbb{C}$ itself is an open subset of $\overline{\mathbb{C}}$. Also, if $U \subseteq \overline{\mathbb{C}}$ is an open subset of $\overline{\mathbb{C}}$, then $U \cap \mathbb{C}$ is an open subset of $\mathbb{C}$. Therefore, the topology of $\mathbb{C}$ coincides with its subspace topology as a subset of $\overline{\mathbb{C}}$.

As in the case of $\overline{\mathbb{R}}$, *we may construct the general Borel subset of* $\overline{\mathbb{C}}$ *by taking the general Borel subset of* $\mathbb{C}$ *and at most adjoining the point* ∞ *to it.*

Proposition 1.15. *(i)* $\mathcal{B}_2 = \overline{\mathcal{B}}_2 \rceil \mathbb{C}$*, where we denote* $\overline{\mathcal{B}}_2$ *the Borel* σ*-algebra of* $\overline{\mathbb{C}}$*.*
(ii) $\overline{\mathcal{B}}_2 = \big\{A, A \cup \{\infty\} \,|\, A \in \mathcal{B}_2\big\}$*.*
(iii) If $\mathcal{C}$ *is the collection of all closed or all open or all open-closed or all closed-open or all bounded intervals in* $\mathbb{C} = \mathbb{R}^2$*, then* $\overline{\mathcal{B}}_2 = \mathcal{S}(\mathcal{C})$*.*

Proof. The proof is very similar to (and slightly simpler than) the proof of Proposition 1.14. □

Exercises.

1.1.14. Let $Y \subseteq X$. If $\mathcal{T}$ is a topology of X, prove that $\mathcal{T}\rceil Y$ is a topology of Y.

1.1.15. Let X be a topological space, and $\mathcal{F}$ be the collection of all closed subsets of X. Prove that $\mathcal{B}_X = \mathcal{S}(\mathcal{F})$.

1.1.16. If X, Y are two topological spaces and $f : X \to Y$ is continuous, prove that $f^{-1}(B)$ is a Borel subset of X for every Borel subset B of Y.
Hint. Exercise 1.1.7 may help.

1.1.17. If Y is a Borel subset of the topological space X, prove that $\mathcal{B}_Y = \{A \subseteq Y \,|\, A \in \mathcal{B}_X\}$.

1.1.18. (i) Let $\mathcal{C}$ be the collection of all half-spaces in $\mathbb{R}^n$ of the form $\{(x_1, \ldots, x_n) \,|\, a_j < x_j\}$, where $j = 1, \ldots, n$ and $a_j \in \mathbb{R}$. Prove that $\mathcal{B}_n = \mathcal{S}(\mathcal{C})$.
(ii) Let $\mathcal{C}$ be the collection of all open balls $B(x; r)$ or of all closed balls $\overline{B}(x; r)$, where $x \in \mathbb{R}^n$ and $r > 0$. Prove that $\mathcal{B}_n = \mathcal{S}(\mathcal{C})$.

1.1.19. Let $\mathcal{C}$ be the collection of all open discs $B(x; r)$ or of all closed discs $\overline{B}(x; r)$, where $x \in \mathbb{C}$ and $r > 0$. Prove that $\overline{\mathcal{B}}_2 = \mathcal{S}(\mathcal{C})$.

1.1.20. Let X be a metric space. Prove that every closed subset of X is a countable intersection of open subsets of X, and that every open subset of X is a countable union of closed subsets of X.
Hint. If F is a closed subset of X, consider the sets $U_n = \{x \in X \,|\, d(x, y) < \frac{1}{n}$ for some $y \in F\}$, where $n \in \mathbb{N}$ and d is the metric of X.

1.1.21. Let X be a topological space, Y be a metric space, and $f : X \to Y$. Prove that the set $\{x \in X \,|\, f$ is continuous at $x\}$ is a countable intersection of open subsets of X.
Hint. Consider the sets $U_n = \bigcup_{y \in Y} \mathrm{int}\big(f^{-1}\big(B\big(y; \frac{1}{n}\big)\big)\big)$ for $n \in \mathbb{N}$, where $\mathrm{int}(A)$ is the interior of $A \subseteq X$, and $B(y; r)$ is the open ball in Y with center $y \in Y$ and radius $r > 0$.

1.1.22. Let X be a topological space, Y be a metric space, and $f_k : X \to Y$ for $k \in \mathbb{N}$. Assume that Y is complete and separable, and that every f_k is continuous on X. Prove that $\{x \in \mathbb{R}^n \,|\, (f_k(x))$ converges$\}$ is a countable intersection of countable unions of closed subsets of X.
Hint. Consider the sets $U_n = \bigcup_{y \in A} \bigcup_{k=1}^{+\infty} \bigcap_{j=k}^{+\infty} f_j^{-1}\big(\overline{B}\big(y; \frac{1}{n}\big)\big)$ for $n \in \mathbb{N}$, where A is a countable set which is dense in Y, and $\overline{B}(y; r)$ is the closed ball in Y with center $y \in Y$ and radius $r > 0$.

1.2 Measures.

Definition. *Let* $(X, \mathcal{S})$ *be a measurable space. A function* $\mu : \mathcal{S} \to [0, +\infty]$ *is called a* ***measure*** *on* $(X, \mathcal{S})$ *if*
(i) $\mu(\emptyset) = 0$*,*
(ii) $\mu\big(\bigcup_{n=1}^{+\infty} A_n\big) = \sum_{n=1}^{+\infty} \mu(A_n)$ *for all sequences* (A_n) *of pairwise disjoint elements of* $\mathcal{S}$*.*
The triple $(X, \mathcal{S}, \mu)$ *of a set* X*, a* σ*-algebra* $\mathcal{S}$ *of subsets of* X *and a measure* μ *on* $(X, \mathcal{S})$ *is called a* ***measure space****.*

If there is no danger of confusion, we shall say that μ is a measure on $\mathcal{S}$ or a measure on X. Recall that each $A \in \mathcal{S}$ is called a measurable set. Now, the quantity $\mu(A)$ is called the **μ-measure** of A or, if there is no danger of confusion, just the *measure* of A.

Note that the values of a measure are non-negative real numbers or $+\infty$.

Property (ii) of a measure is called **σ-additivity**. Sometimes a measure is also called **σ-additive measure** to distinguish from a **finitely additive measure** μ which, by definition, satisfies $\mu(\emptyset) = 0$ and $\mu\big(\bigcup_{n=1}^{N} A_n\big) = \sum_{n=1}^{N} \mu(A_n)$ for all $N \in \mathbb{N}$ and all pairwise disjoint $A_1, \ldots, A_N \in \mathcal{S}$.

In fact, it is easy to see that a (σ-additive) measure on a σ-algebra is finitely additive. Indeed, if $A_1, \ldots, A_N \in \mathcal{S}$ are pairwise disjoint, we take $A_n = \emptyset$ (and hence $\mu(A_n) = 0$) for $n > N$, and then

$$\mu\big(\bigcup_{n=1}^{N} A_n\big) = \mu\big(\bigcup_{n=1}^{+\infty} A_n\big) = \sum_{n=1}^{+\infty} \mu(A_n) = \sum_{n=1}^{N} \mu(A_n).$$

Example. The simplest measure on a measurable space $(X, \mathcal{S})$ is the *zero measure* which is denoted 0 and it is defined by $0(A) = 0$ for every $A \in \mathcal{S}$.

Example. Let X be an uncountable set and consider $\mathcal{S} = \{A \subseteq X \mid \text{either } A \text{ or } A^c \text{ is countable}\}$. We define: $\mu(A) = 0$, if A is countable, and $\mu(A) = 1$, if A^c is countable.
Then it is clear that $\mu(\emptyset) = 0$, and let $A_1, A_2, \ldots \in \mathcal{S}$ be pairwise disjoint. If all A_n are countable, then $\bigcup_{n=1}^{+\infty} A_n$ is also countable, and we get

$$\mu\big(\bigcup_{n=1}^{+\infty} A_n\big) = 0 = \sum_{n=1}^{+\infty} \mu(A_n).$$

Now, assume that one of the A_n, say A_{n_0}, is uncountable. Then for all $n \neq n_0$ we have $A_n \subseteq A_{n_0}^c$, and so A_n is countable. Therefore, $\mu(A_{n_0}) = 1$, and $\mu(A_n) = 0$ for all $n \neq n_0$. Moreover, $(\bigcup_{n=1}^{+\infty} A_n)^c \subseteq A_{n_0}^c$, and so $(\bigcup_{n=1}^{+\infty} A_n)^c$ is countable. Thus,

$$\mu\big(\bigcup_{n=1}^{+\infty} A_n\big) = 1 = \sum_{n=1}^{+\infty} \mu(A_n).$$

Therefore, μ is a measure on X.

Example. We consider the measurable space $(X, \mathcal{P}(X))$, and we define $\sharp : \mathcal{P}(X) \to [0, +\infty]$ in the following manner. We set $\sharp(A) = \text{card}(A)$, i.e. the cardinality of A, if A is a finite subset of X. We also set $\sharp(A) = +\infty$ if A is an infinite subset of X.
Clearly, $\sharp(\emptyset) = \text{card}(\emptyset) = 0$. Now let $A_1, A_2, \ldots$ be pairwise disjoint subsets of X. If at most finitely many of the A_n are non-empty and those which are non-empty are finite, then $\bigcup_{n=1}^{+\infty} A_n$ is also finite, and

$$\sharp\big(\bigcup_{n=1}^{+\infty} A_n\big) = \text{card}\,\big(\bigcup_{n=1}^{+\infty} A_n\big) = \sum_{n=1}^{+\infty} \text{card}(A_n) = \sum_{n=1}^{+\infty} \sharp(A_n).$$

If *either* at most finitely many of the A_n are non-empty and at least one of those which are non-empty is infinite *or* if infinitely many of the A_n are non-empty, then $\bigcup_{n=1}^{+\infty} A_n$ is infinite, and

$$\sharp\big(\bigcup_{n=1}^{+\infty} A_n\big) = +\infty = \sum_{n=1}^{+\infty} \sharp(A_n).$$

Therefore, $\sharp$ is a measure on $(X, \mathcal{P}(X))$, and it is called the **counting measure** on X.

Example. Again, we consider the measurable space $(X, \mathcal{P}(X))$ and a particular $x_0 \in X$, and we define $\delta_{x_0} : \mathcal{P}(X) \to [0, +\infty]$ as follows. We set $\delta_{x_0}(A) = 1$, if $x_0 \in A$, and $\delta_{x_0}(A) = 0$, if $x_0 \notin A$.
Of course, $\delta_{x_0}(\emptyset) = 0$. Let $A_1, A_2, \ldots$ be pairwise disjoint subsets of X. If $x_0 \notin A_n$ for every n, then $x_0 \notin \bigcup_{n=1}^{+\infty} A_n$, and so

$$\delta_{x_0}\big(\bigcup_{n=1}^{+\infty} A_n\big) = 0 = \sum_{n=1}^{+\infty} \delta_{x_0}(A_n).$$

If $x_0 \in A_n$ for some n, then this n is unique, and also $x_0 \in \bigcup_{n=1}^{+\infty} A_n$. Hence,

$$\delta_{x_0}\big(\bigcup_{n=1}^{+\infty} A_n\big) = 1 = \sum_{n=1}^{+\infty} \delta_{x_0}(A_n).$$

Therefore, δ_{x_0} is a measure on $(X, \mathcal{P}(X))$, and it is called the **Dirac measure** at x_0 or the **Dirac mass** at x_0.

Proposition 1.16. *Let $(X, \mathcal{S}, \mu)$ be a measure space.*
(i) If $A, B \in \mathcal{S}$ and $A \subseteq B$, then $\mu(A) \leq \mu(B)$.
(ii) If $A, B \in \mathcal{S}$, $A \subseteq B$ and $\mu(A) < +\infty$, then $\mu(B \setminus A) = \mu(B) - \mu(A)$.
(iii) If $A_1, A_2, \ldots \in \mathcal{S}$, then $\mu\big(\bigcup_{n=1}^{+\infty} A_n\big) \leq \sum_{n=1}^{+\infty} \mu(A_n)$.
(iv) If $A_1, A_2, \ldots \in \mathcal{S}$ and (A_n) is increasing, then $\mu\big(\bigcup_{n=1}^{+\infty} A_n\big) = \lim_{n\to+\infty} \mu(A_n)$.
(v) If $A_1, A_2, \ldots \in \mathcal{S}$, $\mu(A_N) < +\infty$ for some N and (A_n) is decreasing, then $\mu\big(\bigcap_{n=1}^{+\infty} A_n\big) = \lim_{n\to+\infty} \mu(A_n)$.

Proof. (i) We write $B = A \cup (B \setminus A)$. By finite additivity of μ,

$$\mu(B) = \mu(A) + \mu(B \setminus A) \geq \mu(A).$$

(ii) From both sides of $\mu(B) = \mu(A) + \mu(B \setminus A)$ we subtract $\mu(A)$.
(iii) Due to Lemma 1.1, there are $B_1, B_2, \ldots \in \mathcal{S}$ which are pairwise disjoint, and satisfy $B_n \subseteq A_n$ for all n, and $\bigcup_{n=1}^{+\infty} B_n = \bigcup_{n=1}^{+\infty} A_n$. By σ-additivity of μ and (i), we get

$$\mu\big(\bigcup_{n=1}^{+\infty} A_n\big) = \mu\big(\bigcup_{n=1}^{+\infty} B_n\big) = \sum_{n=1}^{+\infty} \mu(B_n) \leq \sum_{n=1}^{+\infty} \mu(A_n).$$

(iv) We have that

$$\bigcup_{n=1}^{+\infty} A_n = A_1 \cup \big(\bigcup_{k=1}^{+\infty}(A_{k+1} \setminus A_k)\big),$$

where all sets whose union is taken in the right side are pairwise disjoint. Applying σ-additivity (and finite additivity),

$$\begin{aligned}\mu\big(\bigcup_{n=1}^{+\infty} A_n\big) &= \mu(A_1) + \sum_{k=1}^{+\infty} \mu(A_{k+1} \setminus A_k) = \lim_{n\to+\infty}\big(\mu(A_1) + \sum_{k=1}^{n-1} \mu(A_{k+1} \setminus A_k)\big)\\ &= \lim_{n\to+\infty} \mu\big(A_1 \cup \bigcup_{k=1}^{n-1}(A_{k+1} \setminus A_k)\big) = \lim_{n\to+\infty} \mu(A_n).\end{aligned}$$

(v) We write $A = \bigcap_{n=1}^{+\infty} A_n$. Then $(A_N \setminus A_n)$ is increasing and $\bigcup_{n=1}^{+\infty}(A_N \setminus A_n) = A_N \setminus A$. So from (iv) we get

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

Now, $\mu(A_N) < +\infty$ implies $\mu(A_n) < +\infty$ for all $n \geq N$ and $\mu(A) < +\infty$. From (ii) we get

$$\lim_{n\to+\infty}(\mu(A_N) - \mu(A_n)) = \mu(A_N) - \mu(A)$$

and, since $\mu(A_N) < +\infty$, we find $\lim_{n\to+\infty} \mu(A_n) = \mu(A)$. □

Property (i) of a measure is called **monotonicity**, property (iii) is called **σ-subadditivity**, property (iv) is called **continuity from below**, and property (v) is called **continuity from above**.

Definition. *Let $(X, \mathcal{S}, \mu)$ be a measure space.*
*(i) μ is called **finite** if $\mu(X) < +\infty$.*
*(ii) μ is called **σ-finite** if there exist $X_1, X_2, \ldots \in \mathcal{S}$ so that $\bigcup_{n=1}^{+\infty} X_n = X$ and $\mu(X_n) < +\infty$ for all n.*
*(iii) μ is called **semifinite** if for every $E \in \mathcal{S}$ with $\mu(E) = +\infty$ there is an $F \in \mathcal{S}$ so that $F \subseteq E$ and $0 < \mu(F) < +\infty$.*
*(iv) We say that $E \in \mathcal{S}$ is **of finite μ-measure** if $\mu(E) < +\infty$.*
*(v) We say that $E \in \mathcal{S}$ is **of σ-finite μ-measure** if there exist $E_1, E_2, \ldots \in \mathcal{S}$ so that $E \subseteq \bigcup_{n=1}^{+\infty} E_n$ and $\mu(E_n) < +\infty$ for all n.*

If there is no danger of confusion, we may say that E is *of finite measure* or *of σ-finite measure*.
Some observations related to the last definition are immediate.

1. If μ is finite, then all sets in $\mathcal{S}$ are of finite measure. More generally, if $E \in \mathcal{S}$ is of finite

measure, then all subsets of E which belong to $\mathcal{S}$ are of finite measure.

2. If μ is σ-finite, then all sets in $\mathcal{S}$ are of σ-finite measure. More generally, if $E \in \mathcal{S}$ is of σ-finite measure, then all subsets of E which belong to $\mathcal{S}$ are of σ-finite measure.

3. The collection of sets of finite measure is closed under finite unions.

4. The collection of sets of σ-finite measure is closed under countable unions.

5. If μ is finite, then it is also σ-finite.

Lemma 1.2. *Let $(X, \mathcal{S}, \mu)$ be a measure space. If μ is σ-finite, then*
(i) there exist pairwise disjoint $X_1, X_2, \ldots \in \mathcal{S}$ so that $\bigcup_{n=1}^{+\infty} X_n = X$ and $\mu(X_n) < +\infty$ for every n.
(ii) there exist $X_1, X_2, \ldots \in \mathcal{S}$ so that (X_n) is increasing and $\bigcup_{n=1}^{+\infty} X_n = X$ and $\mu(X_n) < +\infty$ for every n.

Proof. By definition, there exist $X_1', X_2', \ldots \in \mathcal{S}$ so that $\bigcup_{n=1}^{+\infty} X_n' = X$ and $\mu(X_n') < +\infty$ for every n.
(i) Due to Lemma 1.1 there are pairwise disjoint $X_1, X_2, \ldots \in \mathcal{S}$ so that $\bigcup_{n=1}^{+\infty} X_n = X$ and $X_n \subseteq X_n'$ for every n. From the last inclusion we get $\mu(X_n) \leq \mu(X_n') < +\infty$ for every n.
(ii) We take the successive unions $X_1 = X_1'$ and $X_n = X_1' \cup \cdots \cup X_n'$ for $n \geq 2$. Then, clearly, (X_n) is increasing and $\bigcup_{n=1}^{+\infty} X_n = X$. Moreover, $\mu(X_1) = \mu(X_1') < +\infty$ and also

$$\mu(X_n) \leq \mu(X_1') + \cdots + \mu(X_n') < +\infty$$

for $n \geq 2$. □

Proposition 1.17. *Let $(X, \mathcal{S}, \mu)$ be a measure space. If μ is σ-finite, then it is semifinite.*

Proof. By Lemma 1.2, there are $X_1, X_2, \ldots \in \mathcal{S}$ so that (X_n) is increasing and $\bigcup_{n=1}^{+\infty} X_n = X$ and $\mu(X_n) < +\infty$ for all n.
Let $E \in \mathcal{S}$ have $\mu(E) = +\infty$. Since $(E \cap X_n)$ is increasing and $\bigcup_{n=1}^{+\infty}(E \cap X_n) = E$, we get

$$\lim_{n\to+\infty} \mu(E \cap X_n) = \mu(E) = +\infty.$$

Hence, $\mu(E \cap X_{n_0}) > 0$ for some n_0. Also, $E \cap X_{n_0} \subseteq E$ and $\mu(E \cap X_{n_0}) \leq \mu(X_{n_0}) < +\infty$. □

Definition. *Let $(X, \mathcal{S}, \mu)$ be a measure space. $E \in \mathcal{S}$ is called **μ-null** if $\mu(E) = 0$.*

If there is no danger of confusion, we shall say that E is *null* instead of μ-null.
The following is trivial but basic.

Proposition 1.18. *Let $(X, \mathcal{S}, \mu)$ be a measure space.*
(i) If $E \in \mathcal{S}$ is null, then every subset of E which belongs to $\mathcal{S}$ is also null.
(ii) If $E_1, E_2, \ldots \in \mathcal{S}$ are all null, then $\bigcup_{n=1}^{+\infty} E_n$ is null.

Proof. The proof is based on the monotonicity and on the σ-subadditivity of μ. □

LINEAR COMBINATIONS OF MEASURES.

Proposition 1.19. *Let μ, ν be measures on the measurable space $(X, \mathcal{S})$ and $\lambda \in [0, +\infty)$.*
(i) We define the function $\mu + \nu : \mathcal{S} \to [0, +\infty]$ by

$$(\mu + \nu)(E) = \mu(E) + \nu(E) \quad \text{for all } E \in \mathcal{S}.$$

Then $\mu + \nu$ is a measure on $(X, \mathcal{S})$.
(ii) We define the function $\lambda\mu : \mathcal{S} \to [0, +\infty]$ by

$$(\lambda\mu)(E) = \lambda\, \mu(E) \quad \text{for all } E \in \mathcal{S}$$

(where we follow the convention: $0\,(+\infty) = 0$ whenever $\lambda = 0$ and $\mu(E) = +\infty$). Then $\lambda\mu$ is a measure on $(X, \mathcal{S})$.

Proof. (i) We have $(\mu+\nu)(\emptyset)=\mu(\emptyset)+\nu(\emptyset)=0+0=0$.
If $A_1,A_2,\ldots\in\mathcal{S}$ are pairwise disjoint, then

$$\begin{aligned}(\mu+\nu)\big(\bigcup_{n=1}^{+\infty}A_n\big)&=\mu\big(\bigcup_{n=1}^{+\infty}A_n\big)+\nu\big(\bigcup_{n=1}^{+\infty}A_n\big)=\sum_{n=1}^{+\infty}\mu(A_n)+\sum_{n=1}^{+\infty}\nu(A_n)\\&=\sum_{n=1}^{+\infty}(\mu(A_n)+\nu(A_n))=\sum_{n=1}^{+\infty}(\mu+\nu)(A_n).\end{aligned}$$

Hence, $\mu+\nu$ is a measure on $(X,\mathcal{S})$.
(ii) We have $(\lambda\mu)(\emptyset)=\lambda\,\mu(\emptyset)=\lambda\,0=0$.
If $A_1,A_2,\ldots\in\mathcal{S}$ are pairwise disjoint, then

$$(\lambda\mu)\big(\bigcup_{n=1}^{+\infty}A_n\big)=\lambda\,\mu\big(\bigcup_{n=1}^{+\infty}A_n\big)=\lambda\sum_{n=1}^{+\infty}\mu(A_n)=\sum_{n=1}^{+\infty}\lambda\,\mu(A_n)=\sum_{n=1}^{+\infty}(\lambda\mu)(A_n).$$

Hence, $\lambda\mu$ is a measure on $(X,\mathcal{S})$. $\square$

Definition. *Let μ,ν be measures on the measurable space $(X,\mathcal{S})$ and $\lambda\in[0,+\infty)$. The measures $\mu+\nu$ and $\lambda\mu$ on $(X,\mathcal{S})$ which are defined in Proposition 1.19 are called **sum** of μ and ν and **product** of μ by λ.*

Thus, we may define more general **non-negative linear combinations**

$$\lambda_1\mu_1+\cdots+\lambda_n\mu_n$$

of measures.

Exercises.

1.2.1. Let X be uncountable and define $\mu(E)=0$, if $E\subseteq X$ is countable, and $\mu(E)=+\infty$, if $E\subseteq X$ is uncountable. Prove that μ is a measure on $(X,\mathcal{P}(X))$ which is not semifinite.

1.2.2. Let X be infinite and define $\mu(E)=0$, if $E\subseteq X$ is finite, and $\mu(E)=+\infty$, if $E\subseteq X$ is infinite. Prove that μ is a finitely additive measure on $(X,\mathcal{P}(X))$ which is not a measure.

1.2.3. Let μ be a finitely additive measure on the measurable space $(X,\mathcal{S})$.
(i) Prove that μ is a measure if and only if it is continuous from below.
(ii) If $\mu(X)<+\infty$, prove that μ is a measure if and only if it is continuous from above.

1.2.4. Let $(X,\mathcal{S},\mu)$ be a measure space. If $A\in\mathcal{S}$, $B\subseteq X$ and $\mu(A\triangle B)=0$, prove that $B\in\mathcal{S}$ and $\mu(B)=\mu(A)$.

1.2.5. Let $(X,\mathcal{S},\mu)$ be a measure space and $A_1,A_2,\ldots\in\mathcal{S}$. See exercise 1.1.1, and prove that:
(i) $\mu\big(\underline{\lim}_{n\to+\infty}A_n\big)\leq\underline{\lim}_{n\to+\infty}\mu(A_n)$,
(ii) $\overline{\lim}_{n\to+\infty}\mu(A_n)\leq\mu\big(\overline{\lim}_{n\to+\infty}A_n\big)$, if $\mu\big(\bigcup_{n=1}^{+\infty}A_n\big)<+\infty$,
(iii) $\mu\big(\overline{\lim}_{n\to+\infty}A_n\big)=0$, if $\sum_{n=1}^{+\infty}\mu(A_n)<+\infty$.

1.2.6. Let μ be a semifinite measure on the measurable space $(X,\mathcal{S})$. Prove that for every $E\in\mathcal{S}$ with $\mu(E)=+\infty$ and every $M>0$ there is an $F\in\mathcal{S}$ so that $F\subseteq E$ and $M<\mu(F)<+\infty$.
Hint. Consider the $\sup\{\mu(F)\,|\,F\in\mathcal{S},F\subseteq E,\mu(F)<+\infty\}$.

1.2.7. Let $(X,\mathcal{S},\mu)$ be a measure space and $E\in\mathcal{S}$ be of σ-finite measure. If $\{D_i\}_{i\in I}$ is a collection of pairwise disjoint sets in $\mathcal{S}$, prove that the set $\{i\in I\,|\,\mu(E\cap D_i)>0\}$ is countable.
Hint. If $\mu(E)<+\infty$ and $n\in\mathbb{N}$, prove that the set $\big\{i\in I\,\big|\,\mu(E\cap D_i)\geq\frac{1}{n}\big\}$ is finite.

1.2.8. Let (μ_n) be an increasing sequence of measures on the measurable space $(X,\mathcal{S})$. We define $\mu(E)=\lim_{n\to+\infty}\mu_n(E)$ for all $E\in\mathcal{S}$. Prove that μ is a measure on $(X,\mathcal{S})$.

1.2.9. Let $(X, \mathcal{S}, \mu)$ be a measure space. Prove that for all n and $A_1, \dots, A_n \in \mathcal{S}$ we have

$$\begin{aligned}\mu\big(\textstyle\bigcup_{j=1}^n A_j\big) + \sum_{k \text{ even}} \big(\sum_{1\leq i_1<\cdots<i_k\leq n} \mu(A_{i_1} \cap \cdots \cap A_{i_k})\big)\\ = \textstyle\sum_{k \text{ odd}} \big(\sum_{1\leq i_1<\cdots<i_k\leq n} \mu(A_{i_1} \cap \cdots \cap A_{i_k})\big).\end{aligned}$$

This is called **inclusion-exclusion formula**.

1.2.10. Let $(X, \mathcal{S}_X, \mu_X)$ be a measure space and $f : X \to Y$. We consider the push-forward of $\mathcal{S}_X$ by f on Y i.e. the σ-algebra $\mathcal{S}_Y = \{B \subseteq Y \mid f^{-1}(B) \in \mathcal{S}_X\}$ (see exercise 1.1.2). We define

$$\mu_Y(B) = \mu_X(f^{-1}(B)), \quad B \in \mathcal{S}_Y.$$

Prove that μ_Y is a measure on $(Y, \mathcal{S}_Y)$. It is called the **push-forward** of μ_X by f on Y.

1.2.11. Let $(Y, \mathcal{S}_Y, \mu_Y)$ be a measure space and $f : X \to Y$ be one-to-one on X and onto Y. We consider the pull-back of $\mathcal{S}_Y$ by f on X i.e. the σ-algebra $\mathcal{S}_X = \{f^{-1}(B) \mid B \in \mathcal{S}_Y\}$ (see exercise 1.1.3). We define

$$\mu_X(A) = \mu_Y(f(A)), \quad A \in \mathcal{S}_X.$$

Prove that μ_X is a measure on $(X, \mathcal{S}_X)$. It is called the **pull-back** of μ_Y by f on X.

1.2.12. Let $(X, \mathcal{S}, \mu)$ be a measure space.
(i) If $A, B \in \mathcal{S}$ and $\mu(A \triangle B) = 0$, prove that $\mu(A) = \mu(B)$.
(ii) Define $A \sim B$ if $A, B \in \mathcal{S}$ and $\mu(A \triangle B) = 0$. Prove that $\sim$ is an equivalence relation on $\mathcal{S}$.
For the rest we assume that $\mu(X) < +\infty$, and we define $\overline{d}(A, B) = \mu(A \triangle B)$ for all $A, B \in \mathcal{S}$.
(iii) Prove that $\overline{d}$ is a pseudometric on $\mathcal{S}$. This means: $0 \leq \overline{d}(A, B) < +\infty$, $\overline{d}(A, B) = \overline{d}(B, A)$ and $\overline{d}(A, C) \leq \overline{d}(A, B) + \overline{d}(B, C)$ for all $A, B, C \in \mathcal{S}$.
(iv) On the set $\mathcal{S}/_\sim$ of all equivalence classes we define $d([A], [B]) = \overline{d}(A, B) = \mu(A \triangle B)$ for all $[A], [B] \in \mathcal{S}/_\sim$. Prove that $d([A], [B])$ is well defined and that d is a metric on $\mathcal{S}/_\sim$.

1.2.13. Let $\mathcal{A}$ be an algebra of subsets of X. If
(i) $\mu(\emptyset) = 0$,
(ii) $\mu\big(\bigcup_{j=1}^{+\infty} A_j\big) = \sum_{j=1}^{+\infty} \mu(A_j)$ for all pairwise disjoint $A_1, A_2, \ldots \in \mathcal{A}$ with $\bigcup_{j=1}^{+\infty} A_j \in \mathcal{A}$,
then we say that $\mu : \mathcal{A} \to [0, +\infty]$ is a **measure** on $(X, \mathcal{A})$.
Prove that if μ is a measure on $(X, \mathcal{A})$, where $\mathcal{A}$ is an algebra of subsets of X, then μ is finitely additive, monotone, σ-subadditive, continuous from below and continuous from above (provided that, every time a countable union or countable intersection of elements of $\mathcal{A}$ appears, we assume that this is also an element of $\mathcal{A}$).

1.2.14. Let $\big((X_n, \mathcal{S}_n, \mu_n)\big)$ be a sequence of measure spaces, where the X_n are pairwise disjoint. We define $X = \bigcup_{n=1}^{+\infty} X_n$, $\mathcal{S} = \{E \subseteq X \mid E \cap X_n \in \mathcal{S}_n \text{ for all } n \in \mathbb{N}\}$ and $\mu(E) = \sum_{n=1}^{+\infty} \mu_n(E \cap X_n)$ for all $E \in \mathcal{S}$.
(i) Prove that $(X, \mathcal{S}, \mu)$ is a measure space. It is called the **direct sum** of $\big((X_n, \mathcal{S}_n, \mu_n)\big)$ and it is denoted $\bigoplus_{n=1}^{+\infty}(X_n, \mathcal{S}_n, \mu_n)$.
(ii) Prove that μ is σ-finite if and only if μ_n is σ-finite for all $n \in \mathbb{N}$.

POINT-MASS DISTRIBUTIONS.

Before introducing a particular class of measures we shall define *sums of non-negative terms over general sets of indices*. We shall follow the standard practice of using both notations $a(i)$ and a_i for the values of a function a on a set I of indices.

Definition. *Let I be a non-empty set of indices, and $a : I \to [0, +\infty]$. We define* ***the sum of the values of a*** *by*

$$\sum_{i\in I} a_i = \sup\Big\{ \sum_{i\in F} a_i \,\Big|\, F \text{ is a non-empty finite subset of } I\Big\}.$$

If $I = \emptyset$, we define $\sum_{i\in I} a_i = 0$.

Of course, if F is a non-empty finite set, then

$$\sum_{i\in F} a_i = \sum_{k=1}^{N} a_{i_k},$$

where $F = \{a_{i_1}, \ldots, a_{i_N}\}$ is an arbitrary enumeration of F.

We first make sure that this definition extends a simpler situation.

Proposition 1.20. *If I is countable and $I = \{i_1, i_2, \ldots\}$ is an arbitrary enumeration of I, then $\sum_{i\in I} a_i = \sum_{k=1}^{+\infty} a_{i_k}$ for all $a : I \to [0, +\infty]$.*

Proof. For arbitrary N we consider the finite subset $F = \{i_1, \ldots, i_N\}$ of I. Then, by the definition of $\sum_{i\in I} a_i$, we have

$$\sum_{k=1}^{N} a_{i_k} = \sum_{i\in F} a_i \le \sum_{i\in I} a_i.$$

Since N is arbitrary, we get $\sum_{k=1}^{+\infty} a_{i_k} \le \sum_{i\in I} a_i$.
Now for an arbitrary non-empty finite $F \subseteq I$ we consider the indices of the elements of F provided by the enumeration $I = \{i_1, i_2, \ldots\}$ and we take the largest, say N, of them. Of course, this implies $F \subseteq \{i_1, i_2, \ldots, i_N\}$. Therefore

$$\sum_{i\in F} a_i \le \sum_{k=1}^{N} a_{i_k} \le \sum_{k=1}^{+\infty} a_{i_k}.$$

Since F is arbitrary, we find, by the definition of $\sum_{i\in I} a_i$, that $\sum_{i\in I} a_i \le \sum_{k=1}^{+\infty} a_{i_k}$. □

Proposition 1.21. *Let $a : I \to [0, +\infty]$. If $\sum_{i\in I} a_i < +\infty$, then $a_i < +\infty$ for all $i \in I$ and the set $\{i \in I \mid a_i > 0\}$ is countable.*

Proof. Let $\sum_{i\in I} a_i < +\infty$.
We take any $i_0 \in I$. Considering the finite set $F = \{i_0\}$, we see that

$$a_{i_0} = \sum_{i\in F} a_i \le \sum_{i\in I} a_i < +\infty.$$

Now, for arbitrary $n \in \mathbb{N}$, we consider the set

$$I_n = \Big\{ i \in I \,\Big|\, a_i \ge \tfrac{1}{n}\Big\}.$$

If F is an arbitrary finite subset of I_n, then

$$\tfrac{1}{n}\ \mathrm{card}(F) \le \sum_{i\in F} a_i \le \sum_{i\in I} a_i.$$

Hence, the cardinality of the arbitrary finite subset of I_n is not larger than the number $n \sum_{i\in I} a_i$, and so I_n is finite. But we have that

$$\{i \in I \mid a_i > 0\} = \bigcup_{n=1}^{+\infty} I_n,$$

and so $\{i \in I \mid a_i > 0\}$ is countable. □

Proposition 1.22. *(i) If $a, b : I \to [0, +\infty]$ and $a_i \le b_i$ for all $i \in I$, then $\sum_{i\in I} a_i \le \sum_{i\in I} b_i$.*
(ii) If $a : I \to [0, +\infty]$ and $J \subseteq I$, then $\sum_{i\in J} a_i \le \sum_{i\in I} a_i$.
(iii) If $a : I \to [0, +\infty]$ and $J = \{i \in I \mid a_i > 0\}$, then $\sum_{i\in I} a_i = \sum_{i\in J} a_i$.

Proof. (i) For arbitrary finite $F \subseteq I$ we have

$$\sum_{i\in F} a_i \le \sum_{i\in F} b_i \le \sum_{i\in I} b_i.$$

Taking the supremum over the finite subsets F of I, we find $\sum_{i\in I} a_i \le \sum_{i\in I} b_i$.
(ii) For arbitrary finite $F \subseteq J$ we have that $F \subseteq I$, and so

$$\sum_{i\in F} a_i \le \sum_{i\in I} a_i.$$

Taking the supremum over the finite subsets F of J, we get $\sum_{i\in J} a_i \le \sum_{i\in I} a_i$.
(iii) Since $J \subseteq I$, (ii) implies that $\sum_{i\in J} a_i \le \sum_{i\in I} a_i$.
For an arbitrary finite $F \subseteq I$ we write $F_1 = \{i \in F \,|\, a_i > 0\}$ and $F_2 = \{i \in F \,|\, a_i = 0\}$. Then $F_1 \cup F_2 = F$ and $F_1 \cap F_2 = \emptyset$, and also $F_1 \subseteq J$. Hence,

$$\sum_{i\in F} a_i = \sum_{i\in F_1} a_i + \sum_{i\in F_2} a_i = \sum_{i\in F_1} a_i \le \sum_{i\in J} a_i.$$

Taking the supremum over the finite subsets F of I, we get $\sum_{i\in I} a_i \le \sum_{i\in J} a_i$. ☐

Proposition 1.23. *Let $I = \bigcup_{k\in K} J_k$, where K is non-empty and the J_k are non-empty and pairwise disjoint. Then for every $a : I \to [0, +\infty]$ we have $\sum_{i\in I} a_i = \sum_{k\in K} \big(\sum_{i\in J_k} a_i\big)$.*

Proof. We take an arbitrary finite $F \subseteq I$ and we consider the finite sets $F_k = F \cap J_k$. We observe that the set

$$L = \{k \in K \,|\, F_k \neq \emptyset\}$$

is a finite subset of K. Then, using trivial properties of sums over finite sets of indices, we find

$$\sum_{i\in F} a_i = \sum_{k\in L} \big(\sum_{i\in F_k} a_i\big).$$

The definitions of $\sum_{i\in J_k}$ and of $\sum_{k\in K}$ imply that

$$\sum_{i\in F} a_i \le \sum_{k\in L} \big(\sum_{i\in J_k} a_i\big) \le \sum_{k\in K} \big(\sum_{i\in J_k} a_i\big).$$

Taking the supremum over the finite subsets F of I we find $\sum_{i\in I} a_i \le \sum_{k\in K} \big(\sum_{i\in J_k} a_i\big)$.
Now we take an arbitrary finite $L \subseteq K$, and an arbitrary finite $F_k \subseteq J_k$ for each $k \in L$. Then $\sum_{k\in L} \big(\sum_{i\in F_k} a_i\big)$ is, clearly, a sum (without repetitions) over the finite subset $\bigcup_{k\in L} F_k$ of I. Hence

$$\sum_{k\in L} \big(\sum_{i\in F_k} a_i\big) \le \sum_{i\in I} a_i.$$

Taking the supremum over the finite subsets F_k of J_k for each $k \in L$, one at a time, we get that

$$\sum_{k\in L} \big(\sum_{i\in J_k} a_i\big) \le \sum_{i\in I} a_i.$$

Taking the supremum over the finite subsets L of K, we get $\sum_{k\in K} \big(\sum_{i\in J_k} a_i\big) \le \sum_{i\in I} a_i$. ☐

After this short investigation of the general summation notion we define a class of measures.

Proposition 1.24. *Let X be non-empty and $a : X \to [0, +\infty]$. We define $\mu : \mathcal{P}(X) \to [0, +\infty]$ by*

$$\mu(E) = \sum_{x\in E} a_x \quad \textit{for all } E \subseteq X.$$

Then μ is a measure on $(X, \mathcal{P}(X))$.

Proof. It is obvious that $\mu(\emptyset) = \sum_{x\in\emptyset} a_x = 0$.
If $E_1, E_2, \dots$ are pairwise disjoint and $E = \bigcup_{n=1}^{+\infty} E_n$, we have

$$\mu(E) = \sum_{x\in E} a_x = \sum_{n\in\mathbb{N}} \big(\sum_{x\in E_n} a_x\big) = \sum_{n\in\mathbb{N}} \mu(E_n) = \sum_{n=1}^{+\infty} \mu(E_n),$$

applying Propositions 1.20 and 1.23. ☐

Definition. *The measure on $(X, \mathcal{P}(X))$ defined in Proposition 1.24 is called the **point-mass distribution** on X induced by the function a. The value a_x is called the **point-mass** at x.*

Example. Consider the function which assigns point-mass $a_x = 1$ at every $x \in X$. It is easy to see that the induced point-mass distribution is $\sharp$, i.e. the counting measure on X.

Example. Take a particular $x_0 \in X$ and the function which assigns point-mass $a_{x_0} = 1$ at x_0 and point-mass $a_x = 0$ at all other points of X. Then the induced point-mass distribution is δ_{x_0}, i.e. the Dirac measure at x_0.

Exercises.

1.2.15. Let X be non-empty and consider a finite $A \subseteq X$. If $a : X \to [0, +\infty)$ satisfies $a_x = 0$ for all $x \notin A$, prove that the point-mass distribution μ on X induced by a can be written as a non-negative linear combination of Dirac measures: $\mu = \sum_{x \in A} a_x \delta_x$.

1.2.16. Let I be a set of indices, $a, b : I \to [0, +\infty]$ and $\kappa \in [0, +\infty)$.
(i) Prove that $\sum_{i \in I} a_i = 0$ if and only if $a_i = 0$ for all $i \in I$.
(ii) Prove that $\sum_{i \in I} \kappa a_i = \kappa \sum_{i \in I} a_i$ (consider $0\,(+\infty) = 0$).
(iii) Prove that $\sum_{i \in I}(a_i + b_i) = \sum_{i \in I} a_i + \sum_{i \in I} b_i$.

1.2.17. Let I, J be two sets of indices and consider any $a : I \times J \to [0, +\infty]$. Using Proposition 1.23, prove that

$$\sum_{i \in I} \big(\sum_{j \in J} a_{i,j} \big) = \sum_{(i,j) \in I \times J} a_{i,j} = \sum_{j \in J} \big(\sum_{i \in I} a_{i,j} \big).$$

Recognize as a special case of this the result of exercise 1.2.16 (iii).

1.2.18. Let X be non-empty and consider the point-mass distribution μ defined by the function $a : X \to [0, +\infty]$. Prove that
(i) μ is semifinite if and only if $a_x < +\infty$ for every $x \in X$,
(ii) μ is σ-finite if and only if $a_x < +\infty$ for every $x \in X$ and the set $\{x \in X \mid a_x > 0\}$ is countable.

1.2.19. (i) Let X be any non-empty countable set. Prove that every measure μ on $(X, \mathcal{P}(X))$ is a point-mass distribution.
(ii) Consider the measure in exercise 1.2.1. Prove that this measure is not a point-mass distribution.

1.2.20. A generalization of exercise 1.2.14.
Let $\{(X_i, \mathcal{S}_i, \mu_i) \mid i \in I\}$ be a collection of measure spaces, where the X_i are pairwise disjoint. We define $X = \bigcup_{i \in I} X_i$, $\mathcal{S} = \{E \subseteq X \mid E \cap X_i \in \mathcal{S}_i \text{ for all } i \in I\}$ and $\mu(E) = \sum_{i \in I} \mu_i(E \cap X_i)$ for all $E \in \mathcal{S}$.
(i) Prove that $(X, \mathcal{S}, \mu)$ is a measure space. It is called the **direct sum** of $\{(X_i, \mathcal{S}_i, \mu_i) \mid i \in I\}$ and it is denoted $\bigoplus_{i \in I}(X_i, \mathcal{S}_i, \mu_i)$.
(ii) Prove that μ is σ-finite if and only if the set $J = \{i \in I \mid \mu_i \neq 0\}$ is countable and μ_i is σ-finite for all $i \in J$.

COMPLETE MEASURES.

Proposition 1.18 says that a subset of a μ-null set is also μ-null, *provided* that the subset is contained in the σ-algebra on which the measure μ is defined.

Definition. *Let $(X, \mathcal{S}, \mu)$ be a measure space. Suppose that for every $E \in \mathcal{S}$ with $\mu(E) = 0$ and every $F \subseteq E$ it is implied that $F \in \mathcal{S}$ (and hence $\mu(F) = 0$). Then μ is called **complete** and $(X, \mathcal{S}, \mu)$ is called a **complete measure space**.*

Thus, a measure μ is complete if the σ-algebra on which it is defined contains all subsets of all μ-null sets.

Definition. *If $(X, \mathcal{S}_1, \mu_1)$ and $(X, \mathcal{S}_2, \mu_2)$ are two measure spaces on the same set X, we say that $(X, \mathcal{S}_2, \mu_2)$ is an* ***extension*** *of $(X, \mathcal{S}_1, \mu_1)$ if $\mathcal{S}_1 \subseteq \mathcal{S}_2$ and $\mu_1(E) = \mu_2(E)$ for all $E \in \mathcal{S}_1$.*

Theorem 1.1. *Let $(X, \mathcal{S}, \mu)$ be a measure space. Then there is a unique smallest complete extension $(X, \overline{\mathcal{S}}, \overline{\mu})$ of $(X, \mathcal{S}, \mu)$. In other words, there is a unique measure space $(X, \overline{\mathcal{S}}, \overline{\mu})$ so that*
(i) $(X, \overline{\mathcal{S}}, \overline{\mu})$ is an extension of $(X, \mathcal{S}, \mu)$,
(ii) $(X, \overline{\mathcal{S}}, \overline{\mu})$ is complete,
(iii) if $(X, \overline{\overline{\mathcal{S}}}, \overline{\overline{\mu}})$ is another complete extension of $(X, \mathcal{S}, \mu)$, then it is an extension also of $(X, \overline{\mathcal{S}}, \overline{\mu})$.

Proof. We shall first construct $(X, \overline{\mathcal{S}}, \overline{\mu})$. We define

$$\overline{\mathcal{S}} = \{A \cup F \mid A \in \mathcal{S} \text{ and } F \subseteq E \text{ for some } E \in \mathcal{S} \text{ with } \mu(E) = 0\}.$$

We shall prove that $\overline{\mathcal{S}}$ is a σ-algebra.
We write $\emptyset = \emptyset \cup \emptyset$, where the first $\emptyset$ belongs to $\mathcal{S}$ and the second $\emptyset$ is a subset of $\emptyset \in \mathcal{S}$ with $\mu(\emptyset) = 0$. Therefore, $\emptyset \in \overline{\mathcal{S}}$.
Let $B \in \overline{\mathcal{S}}$. Then $B = A \cup F$, where $A \in \mathcal{S}$ and $F \subseteq E$ for some $E \in \mathcal{S}$ with $\mu(E) = 0$. We then write $B^c = A_1 \cup F_1$, where $A_1 = (A \cup E)^c$ and $F_1 = E \setminus (A \cup F)$. Then $A_1 \in \mathcal{S}$ and $F_1 \subseteq E$. Hence, $B^c \in \overline{\mathcal{S}}$.
Let $B_1, B_2, \ldots \in \overline{\mathcal{S}}$. Then for every n we have $B_n = A_n \cup F_n$, where $A_n \in \mathcal{S}$ and $F_n \subseteq E_n$ for some $E_n \in \mathcal{S}$ with $\mu(E_n) = 0$. Now

$$\bigcup_{n=1}^{+\infty} B_n = \big(\bigcup_{n=1}^{+\infty} A_n\big) \cup \big(\bigcup_{n=1}^{+\infty} F_n\big),$$

where $\bigcup_{n=1}^{+\infty} A_n \in \mathcal{S}$ and $\bigcup_{n=1}^{+\infty} F_n \subseteq \bigcup_{n=1}^{+\infty} E_n \in \mathcal{S}$ with

$$\mu\big(\bigcup_{n=1}^{+\infty} E_n\big) \leq \sum_{n=1}^{+\infty} \mu(E_n) = 0$$

and hence $\mu(\bigcup_{n=1}^{+\infty} E_n) = 0$. Thus, $\bigcup_{n=1}^{+\infty} B_n \in \overline{\mathcal{S}}$.
Now, we construct $\overline{\mu}$.
For every $B \in \overline{\mathcal{S}}$ we write $B = A \cup F$, where $A \in \mathcal{S}$ and $F \subseteq E$ for some $E \in \mathcal{S}$ with $\mu(E) = 0$, and we define

$$\overline{\mu}(B) = \mu(A).$$

To prove that $\overline{\mu}(B)$ is well defined, we assume that $B = A' \cup F'$, where $A' \in \mathcal{S}$ and $F' \subseteq E'$ for some $E' \in \mathcal{S}$ with $\mu(E') = 0$, and we shall prove that $\mu(A) = \mu(A')$. Since $A \subseteq B \subseteq A' \cup E'$, we have

$$\mu(A) \leq \mu(A') + \mu(E') = \mu(A')$$

and, symmetrically, $\mu(A') \leq \mu(A)$.
To prove that $\overline{\mu}$ is a measure on $(X, \overline{\mathcal{S}})$, write $\emptyset = \emptyset \cup \emptyset$ as above, and get $\overline{\mu}(\emptyset) = \mu(\emptyset) = 0$.
Let also $B_1, B_2, \ldots \in \overline{\mathcal{S}}$ be pairwise disjoint. Then $B_n = A_n \cup F_n$, where $A_n \in \mathcal{S}$ and $F_n \subseteq E_n$ for some $E_n \in \mathcal{S}$ with $\mu(E_n) = 0$. Observe that the A_n are pairwise disjoint. Then

$$\bigcup_{n=1}^{+\infty} B_n = \big(\bigcup_{n=1}^{+\infty} A_n\big) \cup \big(\bigcup_{n=1}^{+\infty} F_n\big)$$

and $\bigcup_{n=1}^{+\infty} F_n \subseteq \bigcup_{n=1}^{+\infty} E_n \in \mathcal{S}$ with

$$\mu\big(\bigcup_{n=1}^{+\infty} E_n\big) \leq \sum_{n=1}^{+\infty} \mu(E_n) = 0,$$

and hence $\mu\big(\bigcup_{n=1}^{+\infty} E_n\big) = 0$. Therefore,

$$\overline{\mu}\big(\bigcup_{n=1}^{+\infty} B_n\big) = \mu\big(\bigcup_{n=1}^{+\infty} A_n\big) = \sum_{n=1}^{+\infty} \mu(A_n) = \sum_{n=1}^{+\infty} \overline{\mu}(B_n).$$

We now prove that $\overline{\mu}$ is complete. Let $B \in \overline{\mathcal{S}}$ with $\overline{\mu}(B) = 0$ and let $B' \subseteq B$. Write $B = A \cup F$, where $A \in \mathcal{S}$ and $F \subseteq E$ for some $E \in \mathcal{S}$ with $\mu(E) = 0$, and so $\mu(A) = \overline{\mu}(B) = 0$. Then write $B' = \emptyset \cup B'$, with $\emptyset \in \mathcal{S}$ and $B' \subseteq E'$, where $E' = A \cup E \in \mathcal{S}$ and $\mu(E') \leq \mu(A) + \mu(E) = 0$. Hence, $B' \in \overline{\mathcal{S}}$.
To prove that $(X, \overline{\mathcal{S}}, \overline{\mu})$ is an extension of $(X, \mathcal{S}, \mu)$, we take any $A \in \mathcal{S}$ and we write $A = A \cup \emptyset$, where $\emptyset \subseteq \emptyset \in \mathcal{S}$ with $\mu(\emptyset) = 0$. This implies that $A \in \overline{\mathcal{S}}$ and $\overline{\mu}(A) = \mu(A)$.
Now suppose that $(X, \overline{\overline{\mathcal{S}}}, \overline{\overline{\mu}})$ is another complete extension of $(X, \mathcal{S}, \mu)$. Take any $B \in \overline{\mathcal{S}}$, and then $B = A \cup F$, where $A \in \mathcal{S}$ and $F \subseteq E$ for some $E \in \mathcal{S}$ with $\mu(E) = 0$. But then $A, E \in \overline{\overline{\mathcal{S}}}$ and $\overline{\overline{\mu}}(E) = \mu(E) = 0$. Since $\overline{\overline{\mu}}$ is complete, we get that also $F \in \overline{\overline{\mathcal{S}}}$ and hence $B = A \cup F \in \overline{\overline{\mathcal{S}}}$. Moreover,

$$\overline{\overline{\mu}}(A) \leq \overline{\overline{\mu}}(B) \leq \overline{\overline{\mu}}(A) + \overline{\overline{\mu}}(F) = \overline{\overline{\mu}}(A),$$

which implies

$$\overline{\overline{\mu}}(B) = \overline{\overline{\mu}}(A) = \mu(A) = \overline{\mu}(B).$$

It only remains to prove the uniqueness of a smallest complete extension of $(X, \mathcal{S}, \mu)$. This is obvious, since two *smallest* complete extensions of $(X, \mathcal{S}, \mu)$ must be extensions of each other and, hence, identical. $\square$

Definition. *If $(X, \mathcal{S}, \mu)$ is a measure space, then its smallest complete extension is called the* ***completion*** *of $(X, \mathcal{S}, \mu)$.*

Exercises.

1.2.21. Let $(X, \mathcal{S}, \mu)$ be a measure space. We say that $E \subseteq X$ **belongs locally** to $\mathcal{S}$ if $E \cap A \in \mathcal{S}$ for all $A \in \mathcal{S}$ with $\mu(A) < +\infty$. We define $\widetilde{\mathcal{S}} = \{E \subseteq X \mid E \text{ belongs locally to } \mathcal{S}\}$.
(i) Prove that $\mathcal{S} \subseteq \widetilde{\mathcal{S}}$ and that $\widetilde{\mathcal{S}}$ is a σ-algebra. If $\mathcal{S} = \widetilde{\mathcal{S}}$, then $(X, \mathcal{S}, \mu)$ is called **saturated**.
(ii) If μ is σ-finite, prove that $(X, \mathcal{S}, \mu)$ is saturated.
We define $\widetilde{\mu}(E) = \mu(E)$, if $E \in \mathcal{S}$, and $\widetilde{\mu}(E) = +\infty$, if $E \in \widetilde{\mathcal{S}} \setminus \mathcal{S}$.
(iii) Prove that $\widetilde{\mu}$ is a measure on $(X, \widetilde{\mathcal{S}})$, and, hence, $(X, \widetilde{\mathcal{S}}, \widetilde{\mu})$ is an extension of $(X, \mathcal{S}, \mu)$.
(iv) If $(X, \mathcal{S}, \mu)$ is complete, prove that $(X, \widetilde{\mathcal{S}}, \widetilde{\mu})$ is also complete.
(v) Prove that $(X, \widetilde{\mathcal{S}}, \widetilde{\mu})$ is a saturated measure space.
$(X, \widetilde{\mathcal{S}}, \widetilde{\mu})$ is called the **saturation** of $(X, \mathcal{S}, \mu)$.

RESTRICTION OF A MEASURE.

Let $(X, \mathcal{S}, \mu)$ be a measure space and $Y \in \mathcal{S}$. We recall that the restriction $\mathcal{S}\rceil Y$ of the σ-algebra $\mathcal{S}$ of subsets of X on Y is $\mathcal{S}\rceil Y = \{A \subseteq Y \mid A \in \mathcal{S}\}$.

Proposition 1.25. *Let $(X, \mathcal{S}, \mu)$ be a measure space, $Y \in \mathcal{S}$ and $\mathcal{S}\rceil Y = \{A \subseteq Y \mid A \in \mathcal{S}\}$. We define $\mu\rceil Y : \mathcal{S}\rceil Y \to [0, +\infty]$ by*

$$(\mu\rceil Y)(A) = \mu(A), \quad A \in \mathcal{S}\rceil Y \ (\textit{i.e. } A \in \mathcal{S},\ A \subseteq Y).$$

Then $\mu\rceil Y$ is a measure on $(Y, \mathcal{S}\rceil Y)$.

Proof. Exercise. $\square$

Definition. *Let $(X, \mathcal{S}, \mu)$ be a measure space and $Y \in \mathcal{S}$. The measure $\mu\rceil Y$ on $(Y, \mathcal{S}\rceil Y)$ of Proposition 1.25 is called* ***the restriction of*** *μ* ***on*** *$(Y, \mathcal{S}\rceil Y)$.*

Informally speaking, we may say that $\mu\rceil Y$ is the same as μ but applied only to the measurable subsets of Y.
There is a second kind of restriction of a measure.

Proposition 1.26. *Let $(X, \mathcal{S}, \mu)$ be a measure space and $Y \in \mathcal{S}$. We define $\mu_Y : \mathcal{S} \to [0, +\infty]$ by*

$$\mu_Y(A) = \mu(A \cap Y), \quad A \in \mathcal{S}.$$

Then μ_Y is a measure on $(X, \mathcal{S})$ with the properties:
(i) $\mu_Y(A) = \mu(A)$ if $A \in \mathcal{S}$, $A \subseteq Y$,
(ii) $\mu_Y(A) = 0$ if $A \in \mathcal{S}$, $A \subseteq Y^c$.

Proof. We have $\mu_Y(\emptyset) = \mu(\emptyset \cap Y) = \mu(\emptyset) = 0$.
If $A_1, A_2, \ldots \in \mathcal{S}$ are pairwise disjoint, then

$$\begin{aligned}\mu_Y\big(\textstyle\bigcup_{j=1}^{+\infty} A_j\big) &= \mu\big(\big(\textstyle\bigcup_{j=1}^{+\infty} A_j\big) \cap Y\big) = \mu\big(\textstyle\bigcup_{j=1}^{+\infty}(A_j \cap Y)\big) \\ &= \textstyle\sum_{j=1}^{+\infty} \mu(A_j \cap Y) = \sum_{j=1}^{+\infty} \mu_Y(A_j).\end{aligned}$$

Therefore, μ_Y is a measure on $(X, \mathcal{S})$ and its two properties are trivial to prove. □

Definition. *Let $(X, \mathcal{S}, \mu)$ be a measure space and $Y \in \mathcal{S}$. The measure μ_Y on $(X, \mathcal{S})$ of Proposition 1.26 is called **the Y-restriction of μ on** $(X, \mathcal{S})$.*

Informally speaking, we may describe the relation between the two restrictions of μ as follows. The restriction μ_Y assigns value 0 to all sets in $\mathcal{S}$ which are included in the complement of Y while the restriction $\mu\rceil Y$ simply ignores all those sets. Both restrictions μ_Y and $\mu\rceil Y$ assign the same values (the same to the values that μ assigns) to all sets in $\mathcal{S}$ which are included in Y.

UNIQUENESS OF MEASURES.

The next result is very useful when we want to prove that two measures are equal on a σ-algebra $\mathcal{S}$. It says that it is enough to prove that they are equal on an *algebra* which generates $\mathcal{S}$, provided that an extra assumption of σ-finiteness of the two measures on the algebra is satisfied.

Proposition 1.27. *Let $\mathcal{A}$ be an algebra of subsets of X and let μ, ν be two measures on $(X, \mathcal{S}(\mathcal{A}))$. Suppose there are $A_1, A_2, \ldots \in \mathcal{A}$ so that (A_n) is increasing, $\bigcup_{n=1}^{+\infty} A_n = X$ and $\mu(A_n) < +\infty$ and $\nu(A_n) < +\infty$ for all n.*
If μ, ν are equal on $\mathcal{A}$, then they are equal also on $\mathcal{S}(\mathcal{A})$.

Proof. (a) We assume that $\mu(X) < +\infty$ and $\nu(X) < +\infty$.
We define the collection

$$\mathcal{M} = \{E \in \mathcal{S}(\mathcal{A}) \mid \mu(E) = \nu(E)\}.$$

It is easy to see that $\mathcal{M}$ is a monotone class. Indeed, let $E_1, E_2, \ldots \in \mathcal{M}$ and (E_n) be increasing and $\bigcup_{n=1}^{+\infty} E_n = E$. By continuity of measures from below, we get

$$\lim_{n\to+\infty} \mu(E_n) = \mu(E), \quad \lim_{n\to+\infty} \nu(E_n) = \nu(E).$$

Since $\mu(E_n) = \nu(E_n)$ for all n, we find $\mu(E) = \nu(E)$, and so $E \in \mathcal{M}$. Now, we just repeat the same argument, assuming that (E_n) is decreasing and $\bigcap_{n=1}^{+\infty} E_n = E$, and using the continuity of measures from above and the assumption $\mu(X) < +\infty$ and $\nu(X) < +\infty$.
Since $\mathcal{M}$ is a monotone class including $\mathcal{A}$, Proposition 1.4 implies that $\mathcal{M}(\mathcal{A}) \subseteq \mathcal{M}$. Now, Proposition 1.5 implies that $\mathcal{S}(\mathcal{A}) \subseteq \mathcal{M}$, and so $\mu(E) = \nu(E)$ for all $E \in \mathcal{S}(\mathcal{A})$.
(b) We consider the general case: we do not assume that $\mu(X) < +\infty$ or that $\nu(X) < +\infty$.
For each n, we consider the A_n-restrictions of μ, ν on $(X, \mathcal{S}(\mathcal{A}))$. Namely,

$$\mu_{A_n}(E) = \mu(E \cap A_n), \;\; \nu_{A_n}(E) = \nu(E \cap A_n) \quad \text{for all } E \in \mathcal{S}(\mathcal{A}).$$

Then all μ_{A_n} and ν_{A_n} are finite measures on $(X, \mathcal{S}(\mathcal{A}))$, since $\mu_{A_n}(X) = \mu(A_n) < +\infty$ and $\nu_{A_n}(X) = \nu(A_n) < +\infty$.
If $A \in \mathcal{A}$, then $A \cap A_n \in \mathcal{A}$ for all n, and so

$$\mu_{A_n}(A) = \mu(A \cap A_n) = \nu(A \cap A_n) = \nu_{A_n}(A).$$

Now, by the result of (a) we get that μ_{A_n} and ν_{A_n} are equal on $\mathcal{S}(\mathcal{A})$. I.e.

$$\mu(E \cap A_n) = \mu_{A_n}(E) = \nu_{A_n}(E) = \nu(E \cap A_n), \quad E \in \mathcal{S}(\mathcal{A}).$$

Now let $E \in \mathcal{S}(\mathcal{A})$. Then $(E \cap A_n)$ is increasing and $\bigcup_{n=1}^{+\infty}(E \cap A_n) = E$. Continuity of μ and ν from below implies

$$\lim_{n\to+\infty} \mu(E \cap A_n) = \mu(E), \quad \lim_{n\to+\infty} \nu(E \cap A_n) = \nu(E).$$

Since $\mu(E \cap A_n) = \nu(E \cap A_n)$ for every n, we get $\mu(E) = \nu(E)$.
Thus, μ, ν are equal on $\mathcal{S}(\mathcal{A})$. $\square$

1.3 Measures from outer measures.

Definition. *Let X be a set. A function $\mu^* : \mathcal{P}(X) \to [0, +\infty]$ is called an* ***outer measure*** *on X if*
(i) $\mu^*(\emptyset) = 0$,
(ii) $\mu^*(E) \le \mu^*(F)$ if $E \subseteq F \subseteq X$,
(iii) $\mu^*\big(\bigcup_{n=1}^{+\infty} E_n\big) \le \sum_{n=1}^{+\infty} \mu^*(E_n)$ for all sequences (E_n) of subsets of X.

Note that, if μ^* is an outer measure on X, then $\mu^*(E)$ is defined for *all* subsets E of X. Property (ii) of an outer measure is called **monotonicity**, and property (iii) is called **σ-subadditivity**. It is easy to see that an outer measure is also finitely subadditive: taking $E_n = \emptyset$ for $n > N$, we get

$$\mu^*\big(\bigcup_{n=1}^{N} E_n\big) = \mu^*\big(\bigcup_{n=1}^{+\infty} E_n\big) \le \sum_{n=1}^{+\infty} \mu^*(E_n) = \sum_{n=1}^{N} \mu^*(E_n).$$

Exercises.

1.3.1. Let μ^*, μ_1^*, μ_2^* be outer measures on X and $\kappa \in [0, +\infty)$. Prove that $\kappa\mu^*, \mu_1^* + \mu_2^*$ and $\max\{\mu_1^*, \mu_2^*\}$ are outer measures on X, where these are defined by the formulas $(\kappa\mu^*)(E) = \kappa\, \mu^*(E)$ (consider $0\,(+\infty) = 0$), $(\mu_1^* + \mu_2^*)(E) = \mu_1^*(E) + \mu_2^*(E)$ and $\max\{\mu_1^*, \mu_2^*\}(E) = \max\{\mu_1^*(E), \mu_2^*(E)\}$ for all $E \subseteq X$.

1.3.2. Let (μ_n^*) be a sequence of outer measures on X. Define $\mu^*(E) = \sup_n \mu_n^*(E)$ for all $E \subseteq X$. Prove that μ^* is an outer measure on X.

1.3.3. For every $E \subseteq \mathbb{N}$ define $\lambda(E) = \overline{\lim}_{n\to+\infty} \frac{\mathrm{card}(E\cap\{1,2,\ldots,n\})}{n}$. Prove that λ is not an outer measure on $\mathbb{N}$.

CONSTRUCTION OF OUTER MEASURES.

Proposition 1.28. *Let $\mathcal{C}$ be any collection of subsets of X so that $\emptyset \in \mathcal{C}$, and let $\tau : \mathcal{C} \to [0, +\infty]$ satisfy $\tau(\emptyset) = 0$. We define*

$$\mu^*(E) = \inf\big\{ \sum_{j=1}^{+\infty} \tau(C_j) \,\big|\, C_1, C_2, \ldots \in \mathcal{C} \text{ so that } E \subseteq \bigcup_{j=1}^{+\infty} C_j \big\}$$

for all $E \subseteq X$, where we agree that $\inf \emptyset = +\infty$. *Then μ^* is an outer measure on X.*

It is clear that, if there is at least one countable covering of E with elements of $\mathcal{C}$, then the set $\{\sum_{j=1}^{+\infty} \tau(C_j) \,|\, C_1, C_2, \ldots \in \mathcal{C}$ so that $E \subseteq \bigcup_{j=1}^{+\infty} C_j\}$ is non-empty. If there is no countable covering of E with elements of $\mathcal{C}$, then this set is empty, and so $\mu^*(E) = \inf \emptyset = +\infty$.

Proof. The inclusion $\emptyset \subseteq \bigcup_{j=1}^{+\infty} \emptyset$ implies

$$\mu^*(\emptyset) \le \sum_{j=1}^{+\infty} \tau(\emptyset) = 0,$$

and so $\mu^*(\emptyset) = 0$.
Now, let $A \subseteq B \subseteq X$. If there is no countable covering of B by elements of $\mathcal{C}$, then $\mu^*(B) = +\infty$ and the inequality $\mu^*(A) \le \mu^*(B)$ is obviously true. Otherwise, we take an arbitrary covering $B \subseteq \bigcup_{j=1}^{+\infty} C_j$ with $C_1, C_2, \ldots \in \mathcal{C}$. Then we also have $A \subseteq \bigcup_{j=1}^{+\infty} C_j$ and, by the definition of $\mu^*(A)$, we get

$$\mu^*(A) \le \sum_{j=1}^{+\infty} \tau(C_j).$$

Taking the infimum of the right side, we find $\mu^*(A) \le \mu^*(B)$.
Finally, let us prove

$$\mu^*\big(\bigcup_{n=1}^{+\infty} A_n\big) \le \sum_{n=1}^{+\infty} \mu^*(A_n) \tag{1.4}$$

for all $A_1, A_2, \ldots \subseteq X$. If the right side of (1.4) is equal to $+\infty$, the inequality is clear. So we may assume that the right side of (1.4) is $< +\infty$. Then $\mu^*(A_n) < +\infty$ for all n. Now we take an arbitrary $\epsilon > 0$. By the definition of $\mu^*(A_n)$, there exist $C_{n,1}, C_{n,2}, \ldots \in \mathcal{C}$ so that $A_n \subseteq \bigcup_{j=1}^{+\infty} C_{n,j}$ and

$$\sum_{j=1}^{+\infty} \tau(C_{n,j}) < \mu^*(A_n) + \tfrac{\epsilon}{2^n}.$$

Then

$$\bigcup_{n=1}^{+\infty} A_n \subseteq \bigcup_{(n,j)\in\mathbb{N}\times\mathbb{N}} C_{n,j},$$

and so, using an arbitrary enumeration of $\mathbb{N} \times \mathbb{N}$ and Proposition 1.20, we get by the definition of $\mu^*\big(\bigcup_{n=1}^{+\infty} A_n\big)$ that

$$\mu^*\big(\bigcup_{n=1}^{+\infty} A_n\big) \le \sum_{(n,j)\in\mathbb{N}\times\mathbb{N}} \tau(C_{n,j}).$$

Proposition 1.23 implies

$$\mu^*\big(\bigcup_{n=1}^{+\infty} A_n\big) \le \sum_{n=1}^{+\infty} \big(\sum_{j=1}^{+\infty} \tau(C_{n,j})\big) < \sum_{n=1}^{+\infty} \big(\mu^*(A_n) + \tfrac{\epsilon}{2^n}\big) = \sum_{n=1}^{+\infty} \mu^*(A_n) + \epsilon.$$

Since ϵ is arbitrary, this proves (1.4). □

CONSTRUCTION OF A MEASURE FROM AN OUTER MEASURE.

We shall see now how a measure is constructed from an outer measure.

Definition. *Let μ^* be an outer measure on X. We say that the set $A \subseteq X$ is **μ^*-measurable** if*

$$\mu^*(E \cap A) + \mu^*(E \cap A^c) = \mu^*(E) \quad \text{for all } E \subseteq X.$$

We denote $\mathcal{S}_{\mu^*}$ the collection of all μ^*-measurable subsets of X.

Thus, a set A is μ^*-measurable if and only if it decomposes every subset E of X into two disjoint pieces, namely $E \cap A$ and $E \cap A^c$, the outer measures of which add to give the outer measure of the subset.

Observe that $E = (E \cap A) \cup (E \cap A^c)$, and so $\mu^*(E) \le \mu^*(E \cap A) + \mu^*(E \cap A^c)$ due to the subadditivity of μ^*. Therefore, in order to check the validity of the equality in the definition, it is enough to check the inequality

$$\mu^*(E \cap A) + \mu^*(E \cap A^c) \le \mu^*(E).$$

Furthermore, it is enough to check this last inequality whenever $\mu^*(E) < +\infty$.

Caratheodory's Theorem. *If μ^* is an outer measure on X, then $\mathcal{S}_{\mu^*}$ is a σ-algebra of subsets of X. If we denote μ the restriction of μ^* on $\mathcal{S}_{\mu^*}$, then $(X, \mathcal{S}_{\mu^*}, \mu)$ is a complete measure space.*

Proof. We have

$$\mu^*(E \cap \emptyset) + \mu^*(E \cap \emptyset^c) = \mu^*(\emptyset) + \mu^*(E) = \mu^*(E)$$

for all $E \subseteq X$, and so $\emptyset \in \mathcal{S}_{\mu^*}$.
Let $A \in \mathcal{S}_{\mu^*}$. Then

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

for all $E \subseteq X$. Therefore, $A^c \in \mathcal{S}_{\mu^*}$, and so $\mathcal{S}_{\mu^*}$ is closed under complements.
Now let $A, B \in \mathcal{S}_{\mu^*}$ and $E \subseteq X$. For the first inequality below we use the subadditivity of μ^*, for the second equality we use μ^*-measurability of B, and for the last equality we use the μ^*-measurability of A:

$$\begin{aligned}\mu^*(E \cap (A \cup B)) + \mu^*(E \cap (A \cup B)^c) &= \mu^*(E \cap (A \cup B)) + \mu^*(E \cap (A^c \cap B^c)) \\ &\leq \mu^*(E \cap (A \cap B^c)) + \mu^*(E \cap (B \cap A^c)) \\ &\quad + \mu^*(E \cap (A \cap B)) + \mu^*(E \cap (A^c \cap B^c)) \\ &= \mu^*(E \cap A) + \mu^*(E \cap A^c) = \mu^*(E).\end{aligned}$$

Thus, $A \cup B \in \mathcal{S}_{\mu^*}$, and by induction we get that $\mathcal{S}_{\mu^*}$ is closed under finite unions. Since it is also closed under complements, $\mathcal{S}_{\mu^*}$ is an algebra of subsets of X, and so it is also closed under finite intersections and under set-theoretic differences.
Let $A, B \in \mathcal{S}_{\mu^*}$ and $A \cap B = \emptyset$. Then for all $E \subseteq X$ we have

$$\mu^*(E \cap (A \cup B)) = \mu^*([E \cap (A \cup B)] \cap A) + \mu^*([E \cap (A \cup B)] \cap A^c) = \mu^*(E \cap A) + \mu^*(E \cap B).$$

By an obvious induction we find that, if $A_1, \dots, A_N \in \mathcal{S}_{\mu^*}$ are pairwise disjoint and $E \subseteq X$ is arbitrary, then

$$\mu^*(E \cap (A_1 \cup \dots \cup A_N)) = \mu^*(E \cap A_1) + \dots + \mu^*(E \cap A_N).$$

Now, if $A_1, A_2, \ldots \in \mathcal{S}_{\mu^*}$ are pairwise disjoint and $E \subseteq X$ is arbitrary, then for all N we have

$$\mu^*(E \cap A_1) + \dots + \mu^*(E \cap A_N) = \mu^*(E \cap (A_1 \cup \dots \cup A_N)) \leq \mu^*\big(E \cap \big(\bigcup_{n=1}^{+\infty} A_n\big)\big)$$

by the monotonicity of μ^*. Hence

$$\sum_{n=1}^{+\infty} \mu^*(E \cap A_n) \leq \mu^*\big(E \cap \big(\bigcup_{n=1}^{+\infty} A_n\big)\big).$$

The opposite inequality is immediate by the σ-subadditivity of μ^*:

$$\mu^*\big(E \cap \big(\bigcup_{n=1}^{+\infty} A_n\big)\big) = \mu^*\big(\bigcup_{n=1}^{+\infty}(E \cap A_n)\big) \leq \sum_{n=1}^{+\infty} \mu^*(E \cap A_n).$$

We conclude with the equality

$$\sum_{n=1}^{+\infty} \mu^*(E \cap A_n) = \mu^*\big(E \cap \big(\bigcup_{n=1}^{+\infty} A_n\big)\big) \tag{1.5}$$

for all pairwise disjoint $A_1, A_2, \ldots \in \mathcal{S}_{\mu^*}$ and all $E \subseteq X$.
If $A_1, A_2, \ldots \in \mathcal{S}_{\mu^*}$ are pairwise disjoint and $E \subseteq X$ is arbitrary, then, since $\mathcal{S}_{\mu^*}$ is closed under finite unions, $\bigcup_{n=1}^{N} A_n \in \mathcal{S}_{\mu^*}$ for all N. Hence

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

where we used the finite version of (1.5) for the first term and the monotonicity of μ^* for the second.
Since N is arbitrary,

$$\begin{aligned}\mu^*(E) &\geq \sum_{n=1}^{+\infty} \mu^*(E \cap A_n) + \mu^*\big(E \cap \big(\bigcup_{n=1}^{+\infty} A_n\big)^c\big) \\ &= \mu^*\big(E \cap \big(\bigcup_{n=1}^{+\infty} A_n\big)\big) + \mu^*\big(E \cap \big(\bigcup_{n=1}^{+\infty} A_n\big)^c\big)\end{aligned}$$

by (1.5). Therefore, $\bigcup_{n=1}^{+\infty} A_n \in \mathcal{S}_{\mu^*}$.
If $A_1, A_2, \ldots \in \mathcal{S}_{\mu^*}$ are not necessarily pairwise disjoint, then in the spirit of Lemma 1.1 we write $B_1 = A_1$ and $B_n = A_n \setminus (A_1 \cup \cdots \cup A_{n-1})$ for all $n \geq 2$. Since $\mathcal{S}_{\mu^*}$ is an algebra, all B_n belong to $\mathcal{S}_{\mu^*}$ and they are pairwise disjoint. Hence $\bigcup_{n=1}^{+\infty} A_n = \bigcup_{n=1}^{+\infty} B_n \in \mathcal{S}_{\mu^*}$. We conclude that $\mathcal{S}_{\mu^*}$ is a σ-algebra.
We now define $\mu : \mathcal{S}_{\mu^*} \to [0, +\infty]$ as the restriction of μ^*, i.e.

$$\mu(A) = \mu^*(A) \quad \text{for all } A \in \mathcal{S}_{\mu^*}.$$

Using $E = X$ in (1.5), we get that for all pairwise disjoint $A_1, A_2, \ldots \in \mathcal{S}_{\mu^*}$,

$$\textstyle\sum_{n=1}^{+\infty} \mu(A_n) = \sum_{n=1}^{+\infty} \mu^*(A_n) = \mu^*\big(\bigcup_{n=1}^{+\infty} A_n\big) = \mu\big(\bigcup_{n=1}^{+\infty} A_n\big).$$

Since $\mu(\emptyset) = \mu^*(\emptyset) = 0$, we see that $(X, \mathcal{S}_{\mu^*}, \mu)$ is a measure space.
Finally, let $A \in \mathcal{S}_{\mu^*}$ with $\mu(A) = 0$ and $B \subseteq A$. Then

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

and so

$$\mu^*(E \cap B) + \mu^*(E \cap B^c) \leq \mu^*(B) + \mu^*(E) = \mu^*(E)$$

for all $E \subseteq X$. Therefore, $B \in \mathcal{S}_{\mu^*}$, and so μ is complete. $\square$

As a by-product of the proof of Caratheodory's Theorem we get the useful

Proposition 1.29. *Let μ^* be an outer measure on X.*
(i) If $B \subseteq X$ and $\mu^*(B) = 0$, then B is μ^*-measurable.
(ii) We have $\sum_{n=1}^{+\infty} \mu^*(E \cap A_n) = \mu^*\big(E \cap \big(\bigcup_{n=1}^{+\infty} A_n\big)\big)$ for all pairwise disjoint μ^*-measurable $A_1, A_2, \ldots$ and all $E \subseteq X$.

Proof. The proof of (i) is in the last part of the proof of the Theorem of Caratheodory, and (ii) is just (1.5). $\square$

Thus, every outer measure μ^* produces a specific σ-algebra, the elements of which are the μ^*-measurable sets, and a measure μ, which is the same as μ^* but applied only on the μ^*-measurable sets and not on all subsets of the whole space X. If there is no danger of confusion, we shall call the μ^*-measurable sets just *measurable sets* (and keep in mind that they are defined by a specific procedure starting from the outer measure μ^*).

The most widely used method of producing measures is based on the Theorem of Caratheodory and it is the one we just described: one starts with an outer measure μ^* on X and produces the measure space $(X, \mathcal{S}_{\mu^*}, \mu)$. There is another method of producing measures, the so-called *Daniell method* which we shall describe later.

Exercises.

1.3.4. Let μ^* be an outer measure on X and $Y \subseteq X$.
(i) Define $(\mu^* \rceil Y)(E) = \mu^*(E)$ for all $E \subseteq Y$, and prove that $\mu^* \rceil Y$ is an outer measure on Y.
(ii) Define $\mu^*_Y(E) = \mu^*(E \cap Y)$ for all $E \subseteq X$, and prove that μ^*_Y is an outer measure on X. Moreover, prove that Y is μ^*_Y-measurable.

1.3.5. Let $X \neq \emptyset$. We define: $\mu^*(E) = 0$, if $E = \emptyset$, and $\mu^*(E) = 1$, if $\emptyset \neq E \subseteq X$. Prove that μ^* is an outer measure on X, and that $\emptyset$ and X are the only μ^*-measurable subsets of X.

1.3.6. Let μ^* be an outer measure on X. If $A_1, A_2, \ldots \in \mathcal{S}_{\mu^*}$ and (A_n) is increasing, prove that $\lim_{n\to+\infty} \mu^*(E \cap A_n) = \mu^*\big(E \cap \big(\bigcup_{n=1}^{+\infty} A_n\big)\big)$ for every $E \subseteq X$.
Hint. Use Proposition 1.29.

1.3.7. Let μ^* be an outer measure on X and μ be the induced measure (the restriction of μ^*) on $\mathcal{S}_{\mu^*}$. If $E, G \subseteq X$ we say that G is a μ^***-measurable cover** of E if: $E \subseteq G$, $G \in \mathcal{S}_{\mu^*}$, and for all $A \in \mathcal{S}_{\mu^*}$ for which $A \subseteq G \setminus E$ we have $\mu(A) = 0$.
(i) If G_1, G_2 are μ^*-measurable covers of E, prove that $\mu(G_1 \triangle G_2) = 0$ and hence $\mu(G_1) = \mu(G_2)$.
(ii) Suppose $E \subseteq G$, $G \in \mathcal{S}_{\mu^*}$ and $\mu^*(E) = \mu(G)$. If $\mu^*(E) < +\infty$, prove that G is a μ^*-measurable cover of E.

1.3.8. We say $E \subseteq \mathbb{R}$ has a **condensation point at infinity** if E has uncountably many points outside every bounded interval.
For any $E \subseteq \mathbb{R}$ define: $\mu^*(E) = 0$, if E is countable, $\mu^*(E) = 1$, if E is uncountable and does not have a condensation point at infinity, and $\mu^*(E) = +\infty$, if E has a condensation point at infinity. Prove that μ^* is an outer measure on $\mathbb{R}$, and that $A \subseteq \mathbb{R}$ is μ^*-measurable if and only if either A or A^c is countable. Does every $E \subseteq \mathbb{R}$ have a μ^*-measurable cover? (See exercise 1.3.7).

1.3.9. Consider the collection $\mathcal{C}$ of subsets of $\mathbb{N}$ which only contains $\emptyset$ and all the two-point subsets of $\mathbb{N}$. Define: $\tau(C) = 0$, if $C = \emptyset$, and $\tau(C) = 2$, if $C \in \mathcal{C}$, $C \neq \emptyset$. Calculate $\mu^*(E)$ for all $E \subseteq \mathbb{N}$, where μ^* is the outer measure defined as in Proposition 1.28. Prove that $\emptyset$ and $\mathbb{N}$ are the only μ^*-measurable subsets of $\mathbb{N}$.

1.3.10. *Extension of a measure, I.*
Let $(X, \mathcal{S}_0, \mu_0)$ be a measure space. Define

$$\mu^*(E) = \inf\big\{ \textstyle\sum_{j=1}^{+\infty} \mu_0(A_j) \,\big|\, A_1, A_2, \ldots \in \mathcal{S}_0 \text{ so that } E \subseteq \bigcup_{j=1}^{+\infty} A_j \big\}$$

for every $E \subseteq X$. Proposition 1.28 implies that μ^* is an outer measure on X. We say that μ^* is **induced** by the measure μ_0.
(i) Prove that $\mu^*(E) = \min\{\mu_0(A) \,|\, A \in \mathcal{S}_0, E \subseteq A\}$.
(ii) If $(X, \mathcal{S}_{\mu^*}, \mu)$ is the complete measure space resulting from μ^* by Caratheodory's Theorem (i.e. μ is the restriction of μ^* on $\mathcal{S}_{\mu^*}$), prove that $(X, \mathcal{S}_{\mu^*}, \mu)$ is an extension of $(X, \mathcal{S}_0, \mu_0)$.
(iii) Assume that $E \subseteq X$, and $A_1, A_2, \ldots \in \mathcal{S}_0$, and $E \subseteq \bigcup_{j=1}^{+\infty} A_j$, and $\mu(A_j) < +\infty$ for all j. Prove that $E \in \mathcal{S}_{\mu^*}$ if and only if there is some $A \in \mathcal{S}_0$ so that $E \subseteq A$ and $\mu^*(A \setminus E) = 0$.
(iv) If μ is σ-finite, prove that $(X, \mathcal{S}_{\mu^*}, \mu)$ is the completion of $(X, \mathcal{S}_0, \mu_0)$.
(v) Let X be an uncountable set, $\mathcal{S}_0 = \{A \subseteq X \mid \text{either } A \text{ or } A^c \text{ is countable}\}$ and $\mu_0(A) = \sharp(A)$ for every $A \in \mathcal{S}_0$. Prove that $(X, \mathcal{S}_0, \mu_0)$ is a complete measure space and that $\mathcal{S}_{\mu^*} = \mathcal{P}(X)$. Thus, the result of (iv) does not hold in general.
(vi) See exercise 1.2.21 and prove that $(X, \mathcal{S}_{\mu^*}, \mu)$ is always the saturation of the completion of $(X, \mathcal{S}_0, \mu_0)$.

1.3.11. *Extension of a measure, II.*
Let $\mathcal{A}_0$ be an algebra of subsets of X, and μ_0 be a measure on $(X, \mathcal{A}_0)$ (see exercise 1.2.13). Let

$$\mu^*(E) = \inf\big\{ \textstyle\sum_{j=1}^{+\infty} \mu_0(A_j) \,\big|\, A_1, A_2, \ldots \in \mathcal{A}_0 \text{ so that } E \subseteq \bigcup_{j=1}^{+\infty} A_j \big\}$$

for all $E \subseteq X$. Proposition 1.28 implies that μ^* is an outer measure on X. We say that μ^* is **induced** by the measure μ_0.
(i) Prove that $\mu^*(A) = \mu_0(A)$ for every $A \in \mathcal{A}_0$.
(ii) Prove that every $A \in \mathcal{A}_0$ is μ^*-measurable, and so $\mathcal{S}(\mathcal{A}_0) \subseteq \mathcal{S}_{\mu^*}$.
Thus, if $(X, \mathcal{S}_{\mu^*}, \mu)$ is the complete measure space resulting from μ^* by Caratheodory's Theorem (i.e. μ is the restriction of μ^* on $\mathcal{S}_{\mu^*}$), then $(X, \mathcal{S}_{\mu^*}, \mu)$ is an extention of $(X, \mathcal{S}(\mathcal{A}_0), \mu)$, and this is an extention of $(X, \mathcal{A}_0, \mu_0)$.
(iii) If $(X, \mathcal{S}(\mathcal{A}_0), \nu)$ is another measure space which is an extension of $(X, \mathcal{A}_0, \mu_0)$, prove that

$\nu(E) \leq \mu(E)$ for all $E \in \mathcal{S}(\mathcal{A}_0)$, and that $\nu(E) = \mu(E)$ for all $E \in \mathcal{S}(\mathcal{A}_0)$ with $\mu(E) < +\infty$.
(iv) If the original $(X, \mathcal{A}_0, \mu_0)$ is σ-finite, prove that μ is the unique measure on $(X, \mathcal{S}(\mathcal{A}_0))$ which is an extension of μ_0 on $(X, \mathcal{A}_0)$.

1.3.12. Let μ^* be an outer measure on X. We say that μ^* is a **regular outer measure** if for every $E \subseteq X$ there is $A \in \mathcal{S}_{\mu^*}$ so that $E \subseteq A$ and $\mu^*(E) = \mu(A)$ (where μ is the usual restriction of μ^* on $\mathcal{S}_{\mu^*}$).
(i) Prove that μ^* is a regular outer measure if and only if μ^* is induced by some measure on some algebra of subsets of X (as described in exercise 1.3.11).
(ii) Consider the outer measure μ^* in exercise 1.3.8. Is μ^* a regular outer measure?

1.4 Lebesgue measure.

VOLUME OF INTERVALS.

We consider the quantity $\mathrm{vol}_n(S)$, the n-dimensional **volume** of S, defined for any bounded interval $S = I_1 \times \cdots \times I_n$ in $\mathbb{R}^n$ by

$$\mathrm{vol}_n(S) = \mathrm{length}(I_1) \cdots \mathrm{length}(I_n).$$

Clearly, $\mathrm{vol}_n(S) < +\infty$ for every bounded interval S. Moreover, if $S = I_1 \times \cdots \times I_n$, then $\mathrm{vol}_n(S) = 0$ if and only if at least one of the I_j is an one-point interval or the empty interval. Note, also, that, if $n = 1$, then the one-dimensional volume of a bounded interval in $\mathbb{R}$ is just its length.

Proposition 1.30 summarizes some geometrically obvious properties of volumes of bounded intervals.

Proposition 1.30. *(i) We consider $P = (a_1, b_1] \times \cdots \times (a_n, b_n]$ and, for each $k = 1, \ldots, n$, we take $a_k = c_k^{(0)} < c_k^{(1)} < \cdots < c_k^{(m_k)} = b_k$. We consider $P_{i_1,\ldots,i_n} = (c_1^{(i_1-1)}, c_1^{(i_1)}] \times \cdots \times (c_n^{(i_n-1)}, c_n^{(i_n)}]$ for $1 \leq i_1 \leq m_1, \ldots, 1 \leq i_n \leq m_n$, and we say that the intervals $P_{i_1,\ldots,i_n}$ result from P by subdivision of its edges. Then $\mathrm{vol}_n(P) = \sum_{1 \leq i_1 \leq m_1, \ldots, 1 \leq i_n \leq m_n} \mathrm{vol}_n(P_{i_1,\ldots,i_n})$.*
(ii) Assume that $P, P_1, \ldots, P_l$ are bounded open-closed intervals, that $P_1, \ldots, P_l$ are pairwise disjoint and that $P = \bigcup_{j=1}^{l} P_j$. Then $\mathrm{vol}_n(P) = \sum_{j=1}^{l} \mathrm{vol}_n(P_j)$.
(iii) Assume that $P, P_1, \ldots, P_l$ are bounded open-closed intervals, that $P_1, \ldots, P_l$ are pairwise disjoint and that $\bigcup_{j=1}^{l} P_j \subseteq P$. Then $\sum_{j=1}^{l} \mathrm{vol}_n(P_j) \leq \mathrm{vol}_n(P)$.
(iv) Assume that $P, P_1, \ldots, P_l$ are bounded open-closed intervals and that $P \subseteq \bigcup_{j=1}^{l} P_j$. Then $\mathrm{vol}_n(P) \leq \sum_{j=1}^{l} \mathrm{vol}_n(P_j)$.
(v) Assume that Q is a bounded closed interval, that $R_1, \ldots, R_l$ are bounded open intervals and that $Q \subseteq \bigcup_{j=1}^{l} R_j$. Then $\mathrm{vol}_n(Q) \leq \sum_{j=1}^{l} \mathrm{vol}_n(R_j)$.

Proof. (i) For the second equality in the following calculation we use the distributive property of multiplication of sums:

$$\begin{aligned}
&\textstyle\sum_{1 \leq i_1 \leq m_1, \ldots, 1 \leq i_n \leq m_n} \mathrm{vol}_n(P_{i_1,\ldots,i_n}) \\
&\qquad = \textstyle\sum_{1 \leq i_1 \leq m_1, \ldots, 1 \leq i_n \leq m_n} (c_1^{(i_1)} - c_1^{(i_1-1)}) \cdots (c_n^{(i_n)} - c_n^{(i_n-1)}) \\
&\qquad = \textstyle\sum_{1 \leq i_1 \leq m_1} (c_1^{(i_1)} - c_1^{(i_1-1)}) \cdots \sum_{1 \leq i_n \leq m_n} (c_n^{(i_n)} - c_n^{(i_n-1)}) \\
&\qquad = (b_1 - a_1) \cdots (b_n - a_n) = \mathrm{vol}_n(P).
\end{aligned}$$

(ii) Let $P = (a_1, b_1] \times \cdots \times (a_n, b_n]$ and $P_j = (a_1^{(j)}, b_1^{(j)}] \times \cdots \times (a_n^{(j)}, b_n^{(j)}]$ for $j = 1, \ldots, l$. For every $k = 1, \ldots, n$ we set

$$\{c_k^{(0)}, \ldots, c_k^{(m_k)}\} = \{a_k^{(1)}, \ldots, a_k^{(l)}, b_k^{(1)}, \ldots, b_k^{(l)}\},$$

so that $a_k = c_k^{(0)} < c_k^{(1)} < \cdots < c_k^{(m_k)} = b_k$. This simply means that we *rename* the numbers $a_k^{(1)}, \ldots, a_k^{(l)}, b_k^{(1)}, \ldots, b_k^{(l)}$ in increasing order and so that there are no repetitions. Of course, the smallest of these numbers is a_k and the largest is b_k, otherwise the $P_1, \ldots, P_l$ would not cover P. It is obvious that every interval $(a_k^{(j)}, b_k^{(j)}]$ is the union of some successive among the intervals $(c_k^{(0)}, c_k^{(1)}], \ldots, (c_k^{(m_k-1)}, c_k^{(m_k)}]$.
We now set $P_{i_1,\ldots,i_n} = (c_1^{(i_1-1)}, c_1^{(i_1)}] \times \cdots \times (c_n^{(i_n-1)}, c_n^{(i_n)}]$ for $1 \leq i_1 \leq m_1, \ldots, 1 \leq i_n \leq m_n$. It is clear that the intervals $P_{i_1,\ldots,i_n}$ result from P by subdivision of its edges. It is also (almost) clear that the intervals among the $P_{i_1,\ldots,i_n}$ which belong to a P_j result from it by subdivision of its edges, and that every $P_{i_1,\ldots,i_n}$ is included in exactly one from $P_1, \ldots, P_l$ (because the $P_1, \ldots, P_l$ are disjoint and cover P).
Now, using (i) for the first and third equality, and grouping together the intervals $P_{i_1,\ldots,i_n}$ which are included in the same P_j for the second equality, we find

$$\begin{aligned}\mathrm{vol}_n(P) &= \textstyle\sum_{1\leq i_1\leq m_1,\ldots,1\leq i_n\leq m_n} \mathrm{vol}_n(P_{i_1,\ldots,i_n}) = \sum_{j=1}^{l} \big(\sum_{P_{i_1,\ldots,i_n}\subseteq P_j} \mathrm{vol}_n(P_{i_1,\ldots,i_n})\big)\\ &= \textstyle\sum_{j=1}^{l} \mathrm{vol}_n(P_j).\end{aligned}$$

(iii) We know from Proposition 1.13 that $P \setminus (P_1 \cup \cdots \cup P_l) = P_1' \cup \cdots \cup P_k'$ for some pairwise disjoint bounded open-closed intervals $P_1', \ldots, P_k'$. Then $P = P_1 \cup \cdots \cup P_l \cup P_1' \cup \cdots \cup P_k'$, and so (ii) implies

$$\mathrm{vol}_n(P) = \textstyle\sum_{j=1}^{l} \mathrm{vol}_n(P_j) + \sum_{i=1}^{k} \mathrm{vol}_n(P_i') \geq \sum_{j=1}^{l} \mathrm{vol}_n(P_j).$$

(iv) We first write $P = P_1' \cup \cdots \cup P_l'$ where $P_j' = P_j \cap P$ are open-closed intervals included in P. We then write

$$P = P_1' \cup (P_2' \setminus P_1') \cup \cdots \cup \big(P_l' \setminus (P_1' \cup \cdots \cup P_{l-1}')\big),$$

where each of these l pairwise disjoint sets can, by Proposition 1.13, be written as a finite union of pairwise disjoint bounded open-closed intervals:

$$P_1' = P_1', \quad P_j' \setminus (P_1' \cup \cdots \cup P_{j-1}') = P_1^{(j)} \cup \cdots \cup P_{m_j}^{(j)} \ \text{ for } 2 \leq j \leq l.$$

Now, using (ii) for the equality and (iii) for the two inequalities, we get

$$\begin{aligned}\mathrm{vol}_n(P) &= \mathrm{vol}_n(P_1') + \textstyle\sum_{j=2}^{l} \big(\sum_{m=1}^{m_j} \mathrm{vol}_n(P_m^{(j)})\big)\\ &\leq \mathrm{vol}_n(P_1') + \textstyle\sum_{j=2}^{l} \mathrm{vol}_n(P_j') \leq \sum_{j=1}^{l} \mathrm{vol}_n(P_j).\end{aligned}$$

(v) Let P and P_j be the open-closed intervals with the same edges as Q and, respectively, R_j. Then $P \subseteq Q \subseteq R_1 \cup \cdots \cup R_l \subseteq P_1 \cup \cdots \cup P_l$ and we get

$$\mathrm{vol}_n(Q) = \mathrm{vol}_n(P) \leq \textstyle\sum_{j=1}^{l} \mathrm{vol}_n(P_j) = \sum_{j=1}^{l} \mathrm{vol}_n(R_j)$$

using (iv). $\square$

LEBESGUE MEASURE.

Now we consider the collection $\mathcal{C}$ of all bounded open intervals in $\mathbb{R}^n$ and the $\tau : \mathcal{C} \to [0, +\infty]$ defined by $\tau(R) = \mathrm{vol}_n(R) = (b_1 - a_1) \cdots (b_n - a_n)$ for every $R = (a_1, b_1) \times \cdots \times (a_n, b_n) \in \mathcal{C}$.

If we define

$$m_n^*(E) = \inf\big\{ \textstyle\sum_{j=1}^{+\infty} \mathrm{vol}_n(R_j) \,\big|\, R_1, R_2, \ldots \in \mathcal{C} \text{ so that } E \subseteq \bigcup_{j=1}^{+\infty} R_j \big\}$$

for all $E \subseteq \mathbb{R}^n$, then Proposition 1.28 implies that m_n^* is an outer measure on $\mathbb{R}^n$.

We observe that, since $\mathbb{R}^n = \bigcup_{k=1}^{+\infty} R_k$, where $R_k = (-k, k) \times \cdots \times (-k, k)$, there is a countable covering by elements of $\mathcal{C}$ for every $E \subseteq \mathbb{R}^n$.

Now Caratheodory's Theorem implies that the collection $\mathcal{S}_{m_n^*}$ of m_n^*-measurable sets is a σ-algebra of subsets of $\mathbb{R}^n$, and, if m_n is defined as the restriction of m_n^* on $\mathcal{S}_{m_n^*}$, then m_n is a complete measure on $(X, \mathcal{S}_{m_n^*})$. Now we simplify the notation and instead of $\mathcal{S}_{m_n^*}$ we write $\mathcal{L}_n$:

$$\mathcal{L}_n = \mathcal{S}_{m_n^*}.$$

So $\mathcal{L}_n$ is the σ-algebra of m_n^*-measurable subsets of $\mathbb{R}^n$, and m_n is a complete measure on $(X, \mathcal{L}_n)$.

Definition. *$\mathcal{L}_n$ is called the σ-algebra of **Lebesgue subsets** of $\mathbb{R}^n$, m_n^* is called the **Lebesgue outer measure** on $\mathbb{R}^n$, and m_n is called the **Lebesgue measure** on $\mathbb{R}^n$.*

We shall also say that m_n^* is the *n-dimensional Lebesgue outer measure* and that m_n is the *n-dimensional Lebesgue measure*. If there is no danger of confusion, we shall say *Lebesgue set* instead of Lebesgue subset of $\mathbb{R}^n$.

Our aim now is to study properties of Lebesgue sets and especially their relation with the Borel sets or even more special sets in $\mathbb{R}^n$, like open sets or closed sets or unions of intervals.

Proposition 1.31. *(i) Every bounded interval S in $\mathbb{R}^n$ is a Lebesgue set, and $m_n(S) = \mathrm{vol}_n(S)$.*
(ii) Every countable subset A of $\mathbb{R}^n$ is a Lebesgue set and $m_n(A) = 0$.

Proof. (i) Let $Q = [a_1, b_1] \times \cdots \times [a_n, b_n]$ and $R = (a_1 - \epsilon, b_1 + \epsilon) \times \cdots \times (a_n - \epsilon, b_n + \epsilon)$. Then $Q \subseteq R$, and by the definition of m_n^* we get

$$m_n^*(Q) \leq \mathrm{vol}_n(R) = (b_1 - a_1 + 2\epsilon) \cdots (b_n - a_n + 2\epsilon).$$

Since $\epsilon > 0$ is arbitrary, we find $m_n^*(Q) \leq \mathrm{vol}_n(Q)$.
Now we take any covering, $Q \subseteq \bigcup_{j=1}^{+\infty} R_j$, of Q by bounded open intervals R_j. Since Q is compact, there is l so that $Q \subseteq \bigcup_{j=1}^{l} R_j$, and then Proposition 1.30 implies

$$\mathrm{vol}_n(Q) \leq \textstyle\sum_{j=1}^{l} \mathrm{vol}_n(R_j) \leq \sum_{j=1}^{+\infty} \mathrm{vol}_n(R_j).$$

Taking the infimum of the right side, we get $\mathrm{vol}_n(Q) \leq m_n^*(Q)$, and so

$$m_n^*(Q) = \mathrm{vol}_n(Q). \tag{1.6}$$

Let S be a bounded interval and $a_1, b_1, \ldots, a_n, b_n$ be the end-points of its edges.
If $a_j < b_j$ for all j, then $Q' \subseteq S \subseteq Q''$, where $Q' = [a_1 + \epsilon, b_1 - \epsilon] \times \cdots \times [a_n + \epsilon, b_n - \epsilon]$ and $Q'' = [a_1 - \epsilon, b_1 + \epsilon] \times \cdots \times [a_n - \epsilon, b_n + \epsilon]$ for small $\epsilon > 0$. Then $m_n^*(Q') \leq m_n^*(S) \leq m_n^*(Q'')$, which, due to (1.6), becomes

$$(b_1 - a_1 - 2\epsilon) \cdots (b_n - a_n - 2\epsilon) \leq m_n^*(S) \leq (b_1 - a_1 + 2\epsilon) \cdots (b_n - a_n + 2\epsilon).$$

Since $\epsilon > 0$ is arbitrarily small, we find

$$m_n^*(S) = \mathrm{vol}_n(S). \tag{1.7}$$

If $a_j = b_j$ for at least one j, then of course $\mathrm{vol}_n(S) = 0$. Moreover, we have $S \subseteq Q''$, where $Q'' = [a_1 - \epsilon, b_1 + \epsilon] \times \cdots \times [a_n - \epsilon, b_n + \epsilon]$, as before. Then $m_n^*(S) \leq m_n^*(Q'')$, which, due to (1.6) again, becomes

$$m_n^*(S) \leq (b_1 - a_1 + 2\epsilon) \cdots (b_n - a_n + 2\epsilon).$$

Since $\epsilon > 0$ is arbitrarily small, we find $m_n^*(S) \leq \mathrm{vol}_n(S)$. And, since $\mathrm{vol}_n(S) = 0$, we get (1.7) again. Therefore, (1.7) holds for every bounded interval S.
Consider a bounded open-closed interval P and a bounded open interval R. Take the open-closed interval P_R with the same edges as R. Then (1.7) implies

$$m_n^*(R \cap P) \leq m_n^*(P_R \cap P) = \mathrm{vol}_n(P_R \cap P) \tag{1.8}$$

and we also have

$$m_n^*(R \cap P^c) \le m_n^*(P_R \cap P^c). \tag{1.9}$$

Now Proposition 1.13 implies $P_R \cap P^c = P_R \setminus P = P_1' \cup \cdots \cup P_k'$ for some pairwise disjoint bounded open-closed intervals $P_1', \ldots, P_k'$. From (1.9) and (1.7) we get

$$m_n^*(R \cap P^c) \le \sum_{i=1}^k m_n^*(P_i') = \sum_{i=1}^k \mathrm{vol}_n(P_i'). \tag{1.10}$$

And now from (1.8) and (1.10) and from Proposition 1.30 we get

$$m_n^*(R \cap P) + m_n^*(R \cap P^c) \le \mathrm{vol}_n(P_R \cap P) + \sum_{i=1}^k \mathrm{vol}_n(P_i') = \mathrm{vol}_n(P_R) = \mathrm{vol}_n(R).$$

We have just proved that

$$m_n^*(R \cap P) + m_n^*(R \cap P^c) \le \mathrm{vol}_n(R). \tag{1.11}$$

Now consider any bounded open-closed interval P and any $E \subseteq \mathbb{R}^n$ with $m_n^*(E) < +\infty$. Take, for arbitrary $\epsilon > 0$, a covering $E \subseteq \bigcup_{j=1}^{+\infty} R_j$ of E by bounded open intervals R_j so that

$$\sum_{j=1}^{+\infty} \mathrm{vol}_n(R_j) < m_n^*(E) + \epsilon.$$

Using the σ-subadditivity of m_n^* and (1.11), we get

$$\begin{aligned} m_n^*(E \cap P) + m_n^*(E \cap P^c) &\le \sum_{j=1}^{+\infty} m_n^*(R_j \cap P) + \sum_{j=1}^{+\infty} m_n^*(R_j \cap P^c) \\ &= \sum_{j=1}^{+\infty} (m_n^*(R_j \cap P) + m_n^*(R_j \cap P^c)) \\ &\le \sum_{j=1}^{+\infty} \mathrm{vol}_n(R_j) < m_n^*(E) + \epsilon. \end{aligned}$$

This implies $m_n^*(E \cap P) + m_n^*(E \cap P^c) \le m_n^*(E)$, and so P is a Lebesgue set.
If T is any bounded interval at least one of whose edges is a single point, then $m_n^*(T) = \mathrm{vol}_n(T) = 0$, and so, by Proposition 1.29, T is a Lebesgue set. Now, any bounded interval S differs from the open-closed interval P, which has the same edges as S, by finitely many (at most $2n$) T's, and so S is also a Lebesgue set. Moreover, $m_n(S) = m_n^*(S) = \mathrm{vol}_n(S)$.
(ii) If $x \in \mathbb{R}^n$, then $\{x\}$ is a degenerate interval, and so $m_n(\{x\}) = \mathrm{vol}_n(\{x\}) = 0$. Now, if $A = \{x_1, x_2, \ldots\} \subseteq \mathbb{R}^n$ is an infinite countable set, then $A = \bigcup_{k=1}^{+\infty} \{x_k\}$ is a Lebesgue set, and

$$m_n(A) = \sum_{k=1}^{+\infty} m_n(\{x_k\}) = 0.$$

Of course, the same is true if A is finite. ☐

Proposition 1.32. *Lebesgue measure is σ-finite but not finite.*

Proof. $\mathbb{R}^n = \bigcup_{k=1}^{+\infty} Q_k$, where $Q_k = [-k, k] \times \cdots \times [-k, k]$ and $m_n(Q_k) = \mathrm{vol}_n(Q_k) < +\infty$ for all k.
On the other hand, $m_n(\mathbb{R}^n) \ge m_n(Q_k) = (2k)^n$ for all k, and so $m_n(\mathbb{R}^n) = +\infty$. ☐

LEBESGUE MEASURE AND BOREL SETS.

Proposition 1.33. *All Borel sets are Lebesgue sets, i.e. $\mathcal{B}_n \subseteq \mathcal{L}_n$.*

Proof. Proposition 1.31 says that, if $\mathcal{C}$ is the collection of all bounded intervals in $\mathbb{R}^n$, then $\mathcal{C} \subseteq \mathcal{L}_n$. But then $\mathcal{B}_n = \mathcal{S}(\mathcal{C}) \subseteq \mathcal{L}_n$. ☐

Proposition 1.34. *Let $E \subseteq \mathbb{R}^n$. Then*
(i) E is a Lebesgue set if and only if there is a set A, which is a countable intersection of open sets, such that $E \subseteq A$ and $m_n^*(A \setminus E) = 0$.
(ii) E is a Lebesgue set if and only if there is a set B, which is a countable union of compact sets, such that $B \subseteq E$ and $m_n^*(E \setminus B) = 0$.

Proof. (i) Assume that there is a set A, a countable intersection of open sets, such that $E \subseteq A$ and $m_n^*(A \setminus E) = 0$. Then $A \in \mathcal{B}_n$, and so $A \in \mathcal{L}_n$. Also, by Proposition 1.29, $A \setminus E \in \mathcal{L}_n$. Hence, $E = A \setminus (A \setminus E) \in \mathcal{L}_n$.
For the converse, consider, by means of Proposition 1.32, $Y_1, Y_2, \ldots \in \mathcal{L}_n$ so that $\mathbb{R}^n = \bigcup_{k=1}^{+\infty} Y_k$ and $m_n(Y_k) < +\infty$ for all k. Define $E_k = E \cap Y_k$ and then get $E = \bigcup_{k=1}^{+\infty} E_k$ and $m_n(E_k) < +\infty$ for all k.
For all k and arbitrary $l \in \mathbb{N}$ we consider a covering $E_k \subseteq \bigcup_{j=1}^{+\infty} R_j^{(k,l)}$ by bounded open intervals $R_j^{(k,l)}$ so that

$$\textstyle\sum_{j=1}^{+\infty} \mathrm{vol}_n(R_j^{(k,l)}) < m_n(E_k) + \frac{1}{l2^k}.$$

The set $U^{(k,l)} = \bigcup_{j=1}^{+\infty} R_j^{(k,l)}$ is open, and we have that $E_k \subseteq U^{(k,l)}$ and

$$m_n(U^{(k,l)}) \leq \textstyle\sum_{j=1}^{+\infty} m_n(R_j^{(k,l)}) = \sum_{j=1}^{+\infty} \mathrm{vol}_n(R_j^{(k,l)}) < m_n(E_k) + \frac{1}{l2^k},$$

from which we get

$$m_n(U^{(k,l)} \setminus E_k) < \tfrac{1}{l2^k}.$$

Now, the set $U^{(l)} = \bigcup_{k=1}^{+\infty} U^{(k,l)}$ is open, with $E \subseteq U^{(l)}$ and $U^{(l)} \setminus E \subseteq \bigcup_{k=1}^{+\infty}(U^{(k,l)} \setminus E_k)$, from which we get

$$m_n(U^{(l)} \setminus E) \leq \textstyle\sum_{k=1}^{+\infty} m_n(U^{(k,l)} \setminus E_k) < \sum_{k=1}^{+\infty} \frac{1}{l2^k} = \frac{1}{l}.$$

Finally, we define $A = \bigcap_{l=1}^{+\infty} U^{(l)}$. Then $E \subseteq A$ and

$$m_n(A \setminus E) \leq m_n(U^{(l)} \setminus E) < \tfrac{1}{l}$$

for all l, and so $m_n(A \setminus E) = 0$.
(ii) Assume that B is a countable union of compact sets so that $B \subseteq E$ and $m_n^*(E \setminus B) = 0$. Then $B \in \mathcal{B}_n$, and so $B \in \mathcal{L}_n$. Also, by Proposition 1.29, $E \setminus B \in \mathcal{L}_n$. Thus, $E = B \cup (E \setminus B) \in \mathcal{L}_n$.
Now take $E \in \mathcal{L}_n$. Then $E^c \in \mathcal{L}_n$ and by (i) there is a set A, a countable intersection of open sets, so that $E^c \subseteq A$ and $m_n(A \setminus E^c) = 0$.
We set $B = A^c$, a countable union of closed sets, and we get $m_n(E \setminus B) = m_n(A \setminus E^c) = 0$. Now, let $B = \bigcup_{j=1}^{+\infty} F_j$, where each F_j is closed. We then write $F_j = \bigcup_{k=1}^{+\infty} F_{j,k}$, where $F_{j,k} = F_j \cap ([-k,k] \times \cdots \times [-k,k])$ is a compact set. This proves that B is a countable union of compact sets: $B = \bigcup_{(j,k) \in \mathbb{N} \times \mathbb{N}} F_{j,k}$. $\square$

Proposition 1.34 says that *every Lebesgue set is, except from a null set, equal to a Borel set.*

Proposition 1.35. *(i) m_n is the only measure on $(\mathbb{R}^n, \mathcal{B}_n)$ satisfying $m_n(S) = \mathrm{vol}_n(S)$ for every bounded interval S.*
(ii) $(\mathbb{R}^n, \mathcal{L}_n, m_n)$ is the completion of $(\mathbb{R}^n, \mathcal{B}_n, m_n)$.

Proof. (i) Let μ be a measure on $(\mathbb{R}^n, \mathcal{B}_n)$ with $\mu(S) = \mathrm{vol}_n(S)$, and hence $\mu(S) = m_n(S)$, for all bounded intervals S. If $S = I_1 \times \cdots \times I_n$ is an unbounded interval, we take any increasing sequence (S_k) of bounded intervals (for example, $S_k = S \cap ([-k,k] \times \cdots \times [-k,k])$) so that $\bigcup_{k=1}^{+\infty} S_k = S$, and we get that

$$\mu(S) = \lim_{k \to +\infty} \mu(S_k) = \lim_{k \to +\infty} m_n(S_k) = m_n(S).$$

Therefore,

$$\mu\big(\textstyle\bigcup_{j=1}^{m} P_j\big) = \sum_{j=1}^{m} \mu(P_j) = \sum_{j=1}^{m} m_n(P_j) = m_n\big(\bigcup_{j=1}^{m} P_j\big)$$

for all pairwise disjoint open-closed intervals $P_1, \ldots, P_m$. So the measures μ and m_n are equal on the algebra $\mathcal{A} = \big\{ \bigcup_{j=1}^{m} P_j \,\big|\, m \in \mathbb{N}, P_1, \ldots, P_m \text{ pairwise disjoint open-closed intervals in } \mathbb{R}^n \big\}$. By Proposition 1.27, the two measures are equal also on $\mathcal{S}(\mathcal{A}) = \mathcal{B}_n$.

(ii) Let $(\mathbb{R}^n, \overline{\mathcal{B}_n}, \overline{m_n})$ be the completion of $(\mathbb{R}^n, \mathcal{B}_n, m_n)$.
By Proposition 1.33, $(\mathbb{R}^n, \mathcal{L}_n, m_n)$ is a complete extension of $(\mathbb{R}^n, \mathcal{B}_n, m_n)$. Hence, $\overline{\mathcal{B}_n} \subseteq \mathcal{L}_n$ and $\overline{m_n}(E) = m_n(E)$ for every $E \in \overline{\mathcal{B}_n}$.
Now take any $E \in \mathcal{L}_n$. Proposition 1.34 implies that there is a Borel set B so that $B \subseteq E$ and $m_n(E \setminus B) = 0$. Once more, Proposition 1.34 imples that there is a Borel set A so that $E \setminus B \subseteq A$ and $m_n(A \setminus (E \setminus B)) = 0$. Then

$$m_n(A) = m_n(A \setminus (E \setminus B)) + m_n(E \setminus B) = 0.$$

Since $A \in \mathcal{B}_n$, we have $A \in \overline{\mathcal{B}_n}$. Now, since $E \setminus B \subseteq A$ and $\overline{m_n}(A) = m_n(A) = 0$, and since $(\mathbb{R}^n, \overline{\mathcal{B}_n}, \overline{m_n})$ is complete, we get that $E \setminus B \in \overline{\mathcal{B}_n}$. We also have $B \in \mathcal{B}_n$, and so $B \in \overline{\mathcal{B}_n}$. Hence, $E = B \cup (E \setminus B) \in \overline{\mathcal{B}_n}$.
Therefore, $\mathcal{L}_n \subseteq \overline{\mathcal{B}_n}$, and the proof is complete. □

Proposition 1.36. *Let $m_n^*(E) < +\infty$. Then $E \in \mathcal{L}_n$ if and only if for any $\epsilon > 0$ there are pairwise disjoint bounded intervals $S_1, \ldots, S_l$ (of any kind we like) so that $m_n(E \triangle (S_1 \cup \cdots \cup S_l)) < \epsilon$.*

Proof. Let $E \in \mathcal{L}_n$ and $m_n(E) < +\infty$. We consider a covering $E \subseteq \bigcup_{j=1}^{+\infty} R_j'$ by bounded open intervals R_j' such that

$$\textstyle\sum_{j=1}^{+\infty} \mathrm{vol}_n(R_j') < m_n(E) + \frac{\epsilon}{4}.$$

Now we consider the bounded open-closed interval P_j' which has the same edges as R_j', and then we have $E \subseteq \bigcup_{j=1}^{+\infty} P_j'$ and

$$\textstyle\sum_{j=1}^{+\infty} \mathrm{vol}_n(P_j') < m_n(E) + \frac{\epsilon}{4}.$$

We take m so that $\sum_{j=m+1}^{+\infty} \mathrm{vol}_n(P_j') < \frac{\epsilon}{4}$, and we observe the inclusions

$$E \setminus \big(\textstyle\bigcup_{j=1}^{m} P_j'\big) \subseteq \bigcup_{j=m+1}^{+\infty} P_j', \quad \big(\bigcup_{j=1}^{m} P_j'\big) \setminus E \subseteq \big(\bigcup_{j=1}^{+\infty} P_j'\big) \setminus E.$$

Thus,

$$m_n\big(E \setminus \big(\textstyle\bigcup_{j=1}^{m} P_j'\big)\big) \leq \sum_{j=m+1}^{+\infty} \mathrm{vol}_n(P_j') < \frac{\epsilon}{4},$$
$$m_n\big(\big(\textstyle\bigcup_{j=1}^{m} P_j'\big) \setminus E\big) \leq m_n\big(\bigcup_{j=1}^{+\infty} P_j'\big) - m_n(E) < \frac{\epsilon}{4}.$$

Adding, we find

$$m_n\big(E \triangle \big(\textstyle\bigcup_{j=1}^{m} P_j'\big)\big) < \frac{\epsilon}{2}.$$

Proposition 1.13 implies that there are pairwise disjoint bounded open-closed intervals $P_1, \ldots, P_l$ so that $\bigcup_{j=1}^{m} P_j' = \bigcup_{k=1}^{l} P_k$, and so

$$m_n\big(E \triangle \big(\textstyle\bigcup_{k=1}^{l} P_k\big)\big) < \frac{\epsilon}{2}.$$

Using a technique which appeared in the proof of Proposition 1.31, for each P_k we can find an interval S_k (of any kind we like) so that $S_k \subseteq P_k$ and

$$m_n(P_k \setminus S_k) = m_n(P_k) - m_n(S_k) < \frac{\epsilon}{2l}.$$

Then the $S_1, \ldots, S_l$ are pairwise disjoint, and $\bigcup_{k=1}^{l} S_k \subseteq \bigcup_{k=1}^{l} P_k$. Moreover,

$$\big(\textstyle\bigcup_{k=1}^{l} P_k\big) \setminus \big(\bigcup_{k=1}^{l} S_k\big) \subseteq \bigcup_{k=1}^{l} (P_k \setminus S_k),$$

and so

$$m_n\big(\big(\textstyle\bigcup_{k=1}^{l} P_k\big) \setminus \big(\bigcup_{k=1}^{l} S_k\big)\big) \leq m_n\big(\bigcup_{k=1}^{l} (P_k \setminus S_k)\big) = \sum_{k=1}^{l} m_n(P_k \setminus S_k) < \sum_{k=1}^{l} \frac{\epsilon}{2l} = \frac{\epsilon}{2}.$$

Since

$$E \triangle \big(\textstyle\bigcup_{k=1}^{l} S_k\big) \subseteq \big(E \triangle \big(\bigcup_{k=1}^{l} P_k\big)\big) \cup \big(\big(\bigcup_{k=1}^{l} P_k\big) \setminus \big(\bigcup_{k=1}^{l} S_k\big)\big),$$

we finally get

$$m_n\big(E\triangle\big(\bigcup_{k=1}^{l} S_k\big)\big) < \tfrac{\epsilon}{2}+\tfrac{\epsilon}{2}=\epsilon.$$

Conversely, assume that for every $k\in\mathbb{N}$ there is a set B_k, a finite union of pairwise disjoint bounded intervals, so that

$$m_n^*(E\triangle B_k) < \tfrac{1}{2^k}.$$

We consider the set

$$F=\bigcap_{m=1}^{+\infty}\big(\bigcup_{k=m}^{+\infty} B_k\big),$$

and since $B_k\in\mathcal{L}_n$ for all k, we have that $F\in\mathcal{L}_n$. Then

$$F\setminus E=\bigcap_{m=1}^{+\infty}\big(\bigcup_{k=m}^{+\infty}(B_k\setminus E)\big)\subseteq\bigcup_{k=m}^{+\infty}(B_k\setminus E)$$

for every m, and so

$$m_n^*(F\setminus E)\le\sum_{k=m}^{+\infty} m_n^*(B_k\setminus E)\le\sum_{k=m}^{+\infty} m_n^*(B_k\triangle E)\le\sum_{k=m}^{+\infty}\tfrac{1}{2^k}=\tfrac{1}{2^{m-1}}$$

for every m. Hence $m_n^*(F\setminus E)=0$, which implies that $F\setminus E\in\mathcal{L}_n$.
Also,

$$E\setminus F=\bigcup_{m=1}^{+\infty}\big(\bigcap_{k=m}^{+\infty}(E\setminus B_k)\big)=\bigcup_{m=M}^{+\infty}\big(\bigcap_{k=m}^{+\infty}(E\setminus B_k)\big)\subseteq\bigcup_{m=M}^{+\infty}(E\setminus B_m)$$

for every M, and so

$$m_n^*(E\setminus F)\le\sum_{m=M}^{+\infty} m_n^*(E\setminus B_m)\le\sum_{m=M}^{+\infty} m_n^*(E\triangle B_m)\le\sum_{m=M}^{+\infty}\tfrac{1}{2^m}=\tfrac{1}{2^{M-1}}$$

for every M. Hence $m_n^*(E\setminus F)=0$, which implies that $E\setminus F\in\mathcal{L}_n$.
Now, since $E=(E\setminus F)\cup(E\cap F)=(E\setminus F)\cup\big(F\setminus(F\setminus E)\big)$, we get that $E\in\mathcal{L}_n$. $\square$

Exercises.

1.4.1. If $A\in\mathcal{L}_n$ and A is bounded, prove that $m_n(A)<+\infty$. Give an example of an $A\in\mathcal{L}_n$ which is not bounded but has $m_n(A)<+\infty$.

1.4.2. Let $A=\mathbb{Q}\cap[0,1]$. If $R_1,\dots,R_m$ are open intervals so that $A\subseteq\bigcup_{j=1}^{m}R_j$, prove that $1\le\sum_{j=1}^{m}\mathrm{vol}_1(R_j)$. Discuss the contrast to $m_1^*(A)=0$.

1.4.3. Let $E\subseteq\mathbb{R}^n$ with $m_n^*(E)>0$, and $0\le\alpha<1$. Prove that there is a non-empty bounded open interval R so that $m_n^*(E\cap R)\ge\alpha\ \mathrm{vol}_n(R)$.

1.4.4. Let $E\subseteq\mathbb{R}^n$ be a Lebesgue set, and $\delta>0$. If $m_n(E\cap R)\ge\delta\ \mathrm{vol}_n(R)$ for all bounded open intervals R, prove that $m_n(E^c)=0$.
Hint. Use the result of exercise 1.4.3.

LEBESGUE MEASURE AND SIMPLE TRANSFORMATIONS.

Some of the simplest and most important transformations of $\mathbb{R}^n$ are the translations and the linear transformations.

Every $z\in\mathbb{R}^n$ defines the **translation by** z, namely the function $\tau_z:\mathbb{R}^n\to\mathbb{R}^n$ given by

$$\tau_z(x)=x+z,\quad x\in\mathbb{R}^n.$$

Then τ_z is an one-to-one transformation of $\mathbb{R}^n$ onto $\mathbb{R}^n$ and its inverse transformation is τ_{-z}. For every $E\subseteq\mathbb{R}^n$ we define

$$E+z=\{x+z\,|\,x\in E\}=\tau_z(E).$$

If S is any bounded interval in $\mathbb{R}^n$, then any translation transforms it onto another interval (of the same type) with the same volume. In fact, if $a_1,b_1,\dots,a_n,b_n$ are the end-points of the edges of S, then $S + z$ has $a_1 + z_1, b_1 + z_1, \dots, a_n + z_n, b_n + z_n$ as end-points of its edges, where $z = (z_1, \dots, z_n)$. Therefore,

$$\begin{aligned}\mathrm{vol}_n(S+z) &= \big((b_1+z_1)-(a_1+z_1)\big)\cdots\big((b_n+z_n)-(a_n+z_n)\big)\\ &= (b_1-a_1)\cdots(b_n-a_n) = \mathrm{vol}_n(S).\end{aligned}$$

So we may say that the volume of intervals in $\mathbb{R}^n$ is invariant under translations. We shall see that the same is true for the Lebesgue measure of Lebesgue sets in $\mathbb{R}^n$.

Proposition 1.37. *(i) $\mathcal{L}_n$ is invariant under translations: $A + z \in \mathcal{L}_n$ for every $A \in \mathcal{L}_n$ and $z \in \mathbb{R}^n$.*
(ii) m_n is invariant under translations:

$$m_n(A+z) = m_n(A)$$

for every $A \in \mathcal{L}_n$ and $z \in \mathbb{R}^n$.

Proof. Let $E \subseteq \mathbb{R}^n$ and $z \in \mathbb{R}^n$. Then for all coverings $E \subseteq \bigcup_{j=1}^{+\infty} R_j$ by bounded open intervals R_j we get $E + z \subseteq \bigcup_{j=1}^{+\infty}(R_j + z)$. Therefore,

$$m_n^*(E+z) \le \textstyle\sum_{j=1}^{+\infty} \mathrm{vol}_n(R_j+z) = \sum_{j=1}^{+\infty} \mathrm{vol}_n(R_j).$$

Taking the infimum of the right side, we find that $m_n^*(E+z) \le m_n^*(E)$. Now, applying this to $E + z$ translated by $-z$, we get

$$m_n^*(E) = m_n^*((E+z)-z) \le m_n^*(E+z).$$

Hence, $m_n^*(E+z) = m_n^*(E)$ for all $E \subseteq \mathbb{R}^n$ and $z \in \mathbb{R}^n$.
Suppose now that $A \in \mathcal{L}_n$ and $E \subseteq \mathbb{R}^n$. We have

$$\begin{aligned}&m_n^*(E\cap(A+z)) + m_n^*(E\cap(A+z)^c)\\ &\qquad= m_n^*\big([(E-z)\cap A]+z\big) + m_n^*\big([(E-z)\cap A^c]+z\big)\\ &\qquad= m_n^*\big((E-z)\cap A\big) + m_n^*\big((E-z)\cap A^c\big) = m_n^*(E-z) = m_n^*(E).\end{aligned}$$

Therefore, $A + z \in \mathcal{L}_n$ and $m_n(A+z) = m_n^*(A+z) = m_n^*(A) = m_n(A)$. □

As is well known, a **linear transformation** of $\mathbb{R}^n$ is a function $T : \mathbb{R}^n \to \mathbb{R}^n$ such that

$$T(x+y) = T(x) + T(y), \quad T(\kappa x) = \kappa T(x)$$

for all $x, y \in \mathbb{R}^n$ and $\kappa \in \mathbb{R}$.

Every linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ has a *determinant*, $\det(T) \in \mathbb{R}$. The linear transfomation $T : \mathbb{R}^n \to \mathbb{R}^n$ is one-to-one if and only if it is onto if and only if $\det(T) \neq 0$. Moreover, if $\det(T) \neq 0$, then $T^{-1} : \mathbb{R}^n \to \mathbb{R}^n$ is also a linear transformation and $\det(T^{-1}) = (\det(T))^{-1}$. Finally, if T, T_1, T_2 are linear transformations of $\mathbb{R}^n$ and $T = T_1 \circ T_2$, then $\det(T) = \det(T_1)\det(T_2)$. All these are standard results of Linear Algebra.

Proposition 1.38. *Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a linear transformation. If $A \in \mathcal{L}_n$, then $T(A) \in \mathcal{L}_n$ and*

$$m_n(T(A)) = |\det(T)|\, m_n(A).$$

If $\det(T) = 0$ and $m_n(A) = +\infty$, we interpret the right side as $0\,(+\infty) = 0$.

Proof. At first we assume that $\det(T) \neq 0$.
If T has the form

$$T(x_1, x_2, \dots, x_n) = (\lambda x_1, x_2, \dots, x_n) \tag{1.12}$$

for a certain $\lambda \in \mathbb{R} \setminus \{0\}$, then $\det(T) = \lambda$. Also, if $R = (a_1, b_1) \times \cdots \times (a_n, b_n)$, then, depending on whether $\lambda > 0$ or $\lambda < 0$, we have, respectively,

$$T(R) = (\lambda a_1, \lambda b_1) \times (a_2, b_2) \times \cdots \times (a_n, b_n) \text{ or } (\lambda b_1, \lambda a_1) \times (a_2, b_2) \times \cdots \times (a_n, b_n),$$

Thus, $T(R)$ is an interval and $m_n(T(R)) = |\lambda|\, m_n(R) = |\det(T)|\, m_n(R)$.
If T has the form

$$T(x_1, x_2, \dots, x_{i-1}, x_i, x_{i+1}, \dots, x_n) = (x_i, x_2, \dots, x_{i-1}, x_1, x_{i+1}, \dots, x_n) \tag{1.13}$$

for a certain $i \neq 1$, then $\det(T) = -1$. Also, if $R = (a_1, b_1) \times \cdots \times (a_n, b_n)$ as before, then

$$T(R) = (a_i, b_i) \times (a_2, b_2) \times \cdots \times (a_{i-1}, b_{i-1}) \times (a_1, b_1) \times (a_{i+1}, b_{i+1}) \times \cdots \times (a_n, b_n).$$

Thus, $T(R)$ is an interval and, again, $m_n(T(R)) = m_n(R) = |\det(T)|\, m_n(R)$.
Now, let T have the form

$$T(x_1, \dots, x_{i-1}, x_i, x_{i+1}, \dots, x_n) = (x_1, \dots, x_{i-1}, x_i + x_1, x_{i+1}, \dots, x_n) \tag{1.14}$$

for a certain $i \neq 1$. Then $\det(T) = 1$. Now it is more convenient to work with an interval of the form

$$S = [a_1, b_1] \times \cdots \times [a_{i-1}, b_{i-1}] \times (a_i, b_i] \times [a_{i+1}, b_{i+1}] \times \cdots \times [a_n, b_n]. \tag{1.15}$$

Then $T(S)$ is not an interval any more. In fact,

$$T(S) = \{(y_1, \dots, y_n) \,|\, y_j \in [a_j, b_j] \text{ for } j \neq i,\ y_i - y_1 \in (a_i, b_i]\}.$$

We also define the following three auxilliary sets:

$$\begin{aligned}
L &= [a_1, b_1] \times \cdots \times [a_{i-1}, b_{i-1}] \times [a_i + a_1, b_i + a_1) \times [a_{i+1}, b_{i+1}] \times \cdots \times [a_n, b_n], \\
M &= \{(y_1, \dots, y_n) \,|\, y_j \in [a_j, b_j] \text{ for } j \neq i,\ a_i + a_1 \leq y_i \leq a_i + y_1\}, \\
N &= \{(y_1, \dots, y_n) \,|\, y_j \in [a_j, b_j] \text{ for } j \neq i,\ b_i + a_1 \leq y_i \leq b_i + y_1\}.
\end{aligned}$$

It is easy to see that

$$T(S) \cap M = \emptyset, \quad L \cap N = \emptyset, \quad T(S) \cup M = L \cup N, \quad N = M + z,$$

where $z = (0, \dots, 0, b_i - a_i, 0, \dots, 0)$. Moreover, L is an interval and so it is a Borel set. It is easy to see that M, N are closed sets and so they are also Borel sets. Now, from $T(S) = (L \cup N) \setminus M$ we get that $T(S)$ is also a Borel set. Then we have

$$m_n(T(S)) + m_n(M) = m_n(T(P) \cup M) = m_n(L \cup N) = m_n(L) + m_n(N)$$

and

$$m_n(M) = m_n(M + z) = m_n(N).$$

Hence, L, S being intervals,

$$m_n(T(S)) = m_n(L) = \mathrm{vol}_n(L) = \mathrm{vol}_n(S) = m_n(S) = |\det(T)|\, m_n(S).$$

Now, if $R = (a_1, b_1) \times \cdots \times (a_n, b_n)$ is any bounded open interval, we take the corresponding interval S of the form (1.15) with the same endpoints as R. Then $R \subseteq S$ and $m_n(R) = m_n(S)$, and we get $T(R) \subseteq T(S)$ and

$$m_n^*(T(R)) \leq m_n(T(S)) = |\det(T)|\, m_n(S) = |\det(T)|\, m_n(R).$$

We have shown that for every linear transformation T of the above three types (1.12), (1.13), (1.14) we have

$$m_n^*(T(R)) \leq |\det(T)|\, m_n(R)$$

for every bounded open interval R. (In the first two cases, it was obvious that $T(R)$ was an interval, and so $T(R) \in \mathcal{B}_n$. In the third case, with a little more work, we can also show that $T(R) \in \mathcal{B}_n$ and that the equality $m_n(T(R)) = |\det(T)|\, m_n(R)$ holds, but we do not really need this for the rest of the proof.)

Let, again, T be any linear transformation of one of the above three types. Take any $E \subseteq \mathbb{R}^n$ and consider an arbitrary covering $E \subseteq \bigcup_{j=1}^{+\infty} R_j$ by bounded open intervals R_j. Then $T(E) \subseteq \bigcup_{j=1}^{+\infty} T(R_j)$, and so

$$m_n^*(T(E)) \leq \sum_{j=1}^{+\infty} m_n^*(T(R_j)) \leq |\det(T)| \sum_{j=1}^{+\infty} m_n(R_j) = |\det(T)| \sum_{j=1}^{+\infty} \mathrm{vol}_n(R_j).$$

Taking the infimum over all such coverings, we conclude that

$$m_n^*(T(E)) \leq |\det(T)|\, m_n^*(E).$$

If T is any linear transformation with $\det(T) \neq 0$, then, by a well-known result of Linear Algebra, there are linear transformations $T_1, \dots, T_N$, where each is of one of the above three types so that $T = T_1 \circ \cdots \circ T_N$. Applying the last result repeatedly, we find

$$m_n^*(T(E)) \leq |\det(T_1)| \cdots |\det(T_N)|\, m_n^*(E)| = |\det(T)|\, m_n^*(E)$$

for every $E \subseteq \mathbb{R}^n$. If in this inequality we use the set $T(E)$ in the place of E and T^{-1} in the place of T, we get

$$m_n^*(E) \leq |\det(T^{-1})|\, m_n^*(T(E)) = |\det(T)|^{-1} m_n^*(T(E)).$$

Combining the last two inequalities, we conclude that

$$m_n^*(T(E)) = |\det(T)|\, m_n^*(E)$$

for every linear transformation T with $\det(T) \neq 0$ and every $E \subseteq \mathbb{R}^n$.

Now let $A \in \mathcal{L}_n$. For all $E \subseteq \mathbb{R}^n$ we get

$$\begin{aligned} m_n^*(E \cap T(A)) + m_n^*(E \cap (T(A))^c) &= m_n^*\big(T(T^{-1}(E) \cap A)\big) + m_n^*\big(T(T^{-1}(E) \cap A^c)\big) \\ &= |\det(T)|\, \big(m_n^*(T^{-1}(E) \cap A) + m_n^*(T^{-1}(E) \cap A^c)\big) \\ &= |\det(T)|\, m_n^*(T^{-1}(E)) = m_n^*(E). \end{aligned}$$

Thus, $T(A) \in \mathcal{L}_n$. Moreover,

$$m_n(T(A)) = m_n^*(T(A)) = |\det(T)|\, m_n^*(A) = |\det(T)|\, m_n(A).$$

If $\det(T) = 0$, then $V = T(\mathbb{R}^n)$ is a linear subspace of $\mathbb{R}^n$ with $\dim(V) \leq n-1$. We shall prove that $m_n(V) = 0$ and, since $T(A) \subseteq V$, from the completeness of m_n we shall conclude that $T(A) \in \mathcal{L}_n$ and

$$m_n(T(A)) = 0 = |\det(T)|\, m_n(A)$$

for every $A \in \mathcal{L}_n$.

We consider any basis $\{f_1, \dots, f_m\}$ of V with $m = \dim(V) \leq n-1$, and we complete it to a basis $\{f_1, \dots, f_m, f_{m+1}, \dots, f_n\}$ of $\mathbb{R}^n$. We consider the linear transformation $S : \mathbb{R}^n \to \mathbb{R}^n$ given by

$$S(x_1 f_1 + \cdots + x_n f_n) = (x_1, \dots, x_n).$$

Then S is one-to-one, and so $\det(S) \neq 0$. Moreover,

$$S(V) = \{(x_1, \dots, x_m, 0, \dots, 0) \,|\, x_1, \dots, x_m \in \mathbb{R}\}.$$

We have $S(V) = \bigcup_{k=1}^{+\infty} Q_k$, where $Q_k = [-k,k] \times \cdots \times [-k,k] \times \{0\} \times \cdots \times \{0\}$ (the first m factors of Q_k are equal to $[-k,k]$). Each Q_k is a closed interval in $\mathbb{R}^n$ with $m_n(Q_k) = 0$. Therefore, $m_n(S(V)) = 0$, and this implies $m_n(V) = |\det(S)|^{-1} m_n(S(V)) = 0$. $\square$

Two special examples of linear transformations of $\mathbb{R}^n$ are the dilations and the reflection.

Every $\lambda > 0$ defines the **dilation** $l_\lambda : \mathbb{R}^n \to \mathbb{R}^n$, given by

$$l_\lambda(x) = \lambda x, \quad x \in \mathbb{R}^n.$$

Then l_λ is an one-to-one linear transformation with $\det(l_\lambda) = \lambda^n$. The inverse linear transformation of l_λ is $l_{1/\lambda}$. For every $E \subseteq \mathbb{R}^n$ we define

$$\lambda E = \{\lambda x \,|\, x \in E\} = l_\lambda(E)$$

and we have

$$m_n(\lambda A) = \lambda^n m_n(A)$$

for all $A \in \mathcal{L}_n$.

Another linear transformation is $r : \mathbb{R}^n \to \mathbb{R}^n$, **reflection** through 0, defined by

$$r(x) = -x, \quad x \in \mathbb{R}^n.$$

Reflection r is one-to-one with $\det(r) = (-1)^n$, and it is the inverse of itself. We define

$$-E = \{-x \,|\, x \in E\} = r(E)$$

for all $E \subseteq \mathbb{R}^n$ and we have

$$m_n(-A) = m_n(A)$$

for all $A \in \mathcal{L}_n$.

If $b, b_1, \ldots, b_n \in \mathbb{R}^n$, then the set

$$M = \{b + \kappa_1 b_1 + \cdots + \kappa_n b_n \,|\, 0 \le \kappa_1 \le 1, \ldots, 0 \le \kappa_n \le 1\}$$

is the typical bounded *closed parallelepiped* in $\mathbb{R}^n$. One of the vertices of M is b, and then $b_1, \ldots, b_n$ (interpreted as vectors) are the edges of M which start from b. For such an M we define the linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ by

$$T(x) = T(x_1, \ldots, x_n) = x_1 b_1 + \cdots + x_n b_n, \quad x = (x_1, \ldots, x_n) \in \mathbb{R}^n.$$

We also consider the translation τ_b and the unit cube $Q_0 = [0,1]^n = [0,1] \times \cdots \times [0,1]$ in $\mathbb{R}^n$. We observe that $M = \tau_b\big(T(Q_0)\big)$, and now Propositions 1.37 amd 1.38 imply that M is a Lebesgue set and

$$m_n(M) = m_n\big(T(Q_0)\big) = |\det(T)|\, m_n(Q_0) = |\det(T)|.$$

The columns of the matrix of T with respect to the standard basis $\{e_1, \ldots, e_n\}$ of $\mathbb{R}^n$ are the vectors $T(e_1) = b_1, \ldots, T(e_n) = b_n$. We conclude with the rule: *the Lebesgue measure of a bounded closed parallelepiped is equal to the absolute value of the determinant of the matrix having as columns the sides of the parallelepiped starting from one of its vertices.* Of course, it is easy to see that the same is true for any bounded parallelepiped.

A *hyperplane* of $\mathbb{R}^n$ is a set of the form $V + z$, where $z \in \mathbb{R}^n$ and V is a linear subspace of $\mathbb{R}^n$ with $\dim(V) = n - 1$.

Proposition 1.39. *If A is included in a hyperplane of $\mathbb{R}^n$, then A is a Lebesgue set and $m_n(A) = 0$.*

Proof. If V is a linear subspace of $\mathbb{R}^n$ with $\dim(V) = n - 1$, then there is a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ so that $V = T(\mathbb{R}^n)$ and $\det(T) = 0$. Now, Proposition 1.38 implies that V is a Lebesgue set and, as we saw in the proof of Proposition 1.38, we have $m_n(V) = 0$. Then Proposition 1.37 says that $V + z$ is a Lebesgue set and $m_n(V + z) = m_n(V) = 0$.
Now, if $A \subseteq V + z$, then by the completeness of Lebesgue measure we have that A is a Lebesgue set and $m_n(A) = 0$. $\square$

Exercises.

1.4.5. Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be an **isometric linear transformation**. This means that T is a linear transformation satisfying $|T(x) - T(y)| = |x - y|$ for every $x, y \in \mathbb{R}^n$ or, equivalently, $TT^* = T^*T = I$, where T^* is the adjoint of T and I is the identity transformation.
Prove that $m_n(T(A)) = m_n(A)$ for every $A \in \mathcal{L}_n$.

1.4.6. A parallelepiped in $\mathbb{R}^n$ is called **degenerate** if it is included in a hyperplane of $\mathbb{R}^n$.
Prove that a parallelepiped M is degenerate if and only if $m_n(M) = 0$.

1.4.7. State in a formal way and prove the rule "volume $=$ base area $\times$ height" for parallelepipeds in $\mathbb{R}^n$.

1.4.8. Prove that m_n is the only measure μ on $(\mathbb{R}^n, \mathcal{B}_n)$ which is invariant under translations (i.e. $\mu(A + z) = \mu(A)$ for all $A \in \mathcal{B}_n$ and all $z \in \mathbb{R}^n$), and which satisfies $\mu(Q_0) = 1$, where $Q_0 = [0, 1] \times \cdots \times [0, 1]$.
Hint. For every $m \in \mathbb{N}$ and for all cubes of the form $Q = \left[x_1, x_1 + \frac{1}{m}\right] \times \cdots \times \left[x_n, x_n + \frac{1}{m}\right]$, prove that $\mu(Q) = \left(\frac{1}{m}\right)^n$.

1.4.9. Let $E \subseteq \mathbb{R}^n$ be a Lebesgue set with $m_n(E) > 0$. Prove that the **difference set** of E, i.e. the set $D(E) = \{x - y \,|\, x, y \in E\}$, includes some open interval in $\mathbb{R}^n$ which is centered at 0.
Hint. Take $\alpha = \frac{2}{3}$. Then exercise 1.4.3 says that there is a non-empty bounded open interval $R = I_1 \times \cdots \times I_n$ so that $m_n(E \cap R) \geq \alpha \ \mathrm{vol}_n(R)$. Consider the open interval $R' = J_1 \times \cdots \times J_n$, where J_k is the open interval in $\mathbb{R}$ which is centered at 0 and with $\mathrm{length}(J_k) = 2(1 - \alpha^{1/n})\,\mathrm{length}(I_k)$. Prove that $E \cap (E + z) \cap R \neq \emptyset$ for all $z \in R'$.

1.4.10. Let $E \subseteq \mathbb{R}^n$ be a Lebesgue set, and A be a dense subset of $\mathbb{R}^n$. If $m_n(E \triangle (E + z)) = 0$ for all $z \in A$, prove that $m_n(E) = 0$ or $m_n(E^c) = 0$.

THE CANTOR SET AND THE CANTOR FUNCTION.

If $x \in \mathbb{R}^n$, then $\{x\}$ is a degenerate interval, and so $m_n(\{x\}) = \mathrm{vol}_n(\{x\}) = 0$. In fact, every countable set in $\mathbb{R}^n$ has Lebesgue measure zero: if $A = \{x_1, x_2, \ldots\}$, then

$$m_n(A) = \sum_{k=1}^{+\infty} m_n(\{x_k\}) = 0.$$

The aim of this subsection is to construct an *uncountable* set in $\mathbb{R}$ whose one-dimensional Lebesgue measure is zero.

We start with the interval $I_0 = [0, 1]$, we then take $I_1 = [0, \frac{1}{3}] \cup [\frac{2}{3}, 1]$, we continue with $I_2 = [0, \frac{1}{9}] \cup [\frac{2}{9}, \frac{1}{3}] \cup [\frac{2}{3}, \frac{7}{9}] \cup [\frac{8}{9}, 1]$ and so on: at every stage we divide each of the intervals which we get at the previous stage into three subintervals of equal length and we keep only the two closed subintervals on the sides.

We thus construct a decreasing sequence (I_k) of closed sets so that every I_k consists of 2^k closed intervals all of which have the same length $\frac{1}{3^k}$. We define

$$C = \bigcap_{k=1}^{+\infty} I_k$$

and call it the **Cantor set**.

C is a compact subset of $[0, 1]$ with $m_1(C) = 0$. To see this we observe that for every k we have

$$0 \leq m_1(C) \leq m_1(I_k) = 2^k \frac{1}{3^k}$$

and that $\lim_{k \to +\infty} \left(\frac{2}{3}\right)^k = 0$.

We shall prove, by contradiction, that C is uncountable: let us assume that $C = \{x_1, x_2, \ldots\}$. We shall now describe an inductive process of picking one of the subintervals constituting each I_k.

It is obvious that every x_k belongs to I_k, since it belongs to C. At the first step we choose the interval $I^{(1)}$ to be the subinterval of I_1 which does *not* contain x_1. Now, $I^{(1)}$ includes two subintervals of I_2, and at the second step we choose the interval $I^{(2)}$ to be whichever of these two subintervals of $I^{(1)}$ does *not* contain x_2. (If both do not contain x_2, we just take the left one.) And we continue inductively: if we have already chosen $I^{(k-1)}$ from the subintervals of I_{k-1}, then this includes two subintervals of I_k. We choose as $I^{(k)}$ whichever of these two subintervals of $I^{(k-1)}$ does *not* contain x_k. (If both do not contain x_k, we just take the left one.)

This produces a sequence $(I^{(k)})$ of closed intervals with the following properties:
(i) $I^{(k)} \subseteq I_k$ for all k,
(ii) $I^{(k)} \subseteq I^{(k-1)}$ for all k,
(iii) $\text{length}(I^{(k)}) = \frac{1}{3^k}$, and so $\lim_{k\to+\infty} \text{length}(I^{(k)}) = 0$,
(iv) $x_k \notin I^{(k)}$ for all k.
From (ii) and (iii) we conclude that the intersection of all $I^{(k)}$ contains a single point: $\bigcap_{k=1}^{+\infty} I^{(k)} = \{x_0\}$ for some x_0. From (i) we see that $x_0 \in I_k$ for all k, and so $x_0 \in C$. Therefore, $x_0 = x_k$ for some $k \in \mathbb{N}$. But then $x_0 \in I^{(k)}$ and, by (iv), *the same* point x_k does not belong to $I^{(k)}$.

We arrived at a contradiction, and we conclude that C is uncountable.

Now, for each $k \in \mathbb{N}$ we shall define a function $f_k : [0,1] \to [0,1]$ as follows. We observe that the set $[0,1] \setminus I_k$ consists of $2^k - 1$ open intervals, and we denote these intervals $J_1^{(k)}, \dots, J_{2^k-1}^{(k)}$, going from left to right:

$$[0,1] \setminus I_k = J_1^{(k)} \cup \cdots \cup J_{2^k-1}^{(k)}.$$

We define $f_k(0) = 0$, $f_k(1) = 1$, then we define f_k to be constant $\frac{j}{2^k}$ on $J_j^{(k)}$ for $j = 1, \dots, 2^k - 1$, and, finally, we define f_k to be linear on each of the 2^k subintervals of I_k in such a way that f_k is continuous on $[0,1]$. The resulting function f_k is strictly increasing on each of the 2^k subintervals of I_k, and constant on each of the $2^k - 1$ subintervals of $[0,1] \setminus I_k$.

We observe that the subintervals of $[0,1] \setminus I_{k-1}$ are also subintervals of $[0,1] \setminus I_k$ and that $f_{k-1} = f_k$ on each of them. Moreover, on each of the subintervals of I_{k-1} the functions f_{k-1}, f_k are increasing, they coincide at the endpoints and the difference of their common values at the endpoints is $\frac{1}{2^{k-1}}$. Therefore, we get that $|f_k - f_{k-1}| \le \frac{1}{2^{k-1}}$ on each of the subintervals of I_{k-1} and, hence, $|f_k - f_{k-1}| \le \frac{1}{2^{k-1}}$ everywhere on $[0,1]$ for all $k \ge 2$. This implies that the series of functions $f_1 + \sum_{k=2}^{+\infty}(f_k - f_{k-1})$ converges to a function, say f, uniformly on $[0,1]$:

$$f_1 + \sum_{k=2}^{+\infty}(f_k - f_{k-1}) = f \quad \text{uniformly on } [0,1].$$

The k-th partial sum of the series is $f_1 + (f_2 - f_1) + \cdots + (f_k - f_{k-1}) = f_k$, and so

$$\lim_{k\to+\infty} f_k = f \quad \text{uniformly on } [0,1].$$

Since $f_k(0) = 0$ and $f_k(1) = 1$ for all k, we have that $f(0) = 0$ and $f(1) = 1$. Moreover, f is increasing on $[0,1]$ since it is the limit of increasing functions on $[0,1]$. Furthermore, all f_k are continuous on $[0,1]$ and from uniform convergence we conclude that f is continuous on $[0,1]$.

The function f_k was defined to be constant $\frac{j}{2^k}$ on $J_j^{(k)}$ for all $j = 1, \dots, 2^k - 1$. But we observe that for *all* $m \ge k$ we have $f_m = f_k$ on each $J_j^{(k)}$. Therefore, f is constant $\frac{j}{2^k}$ on $J_j^{(k)}$ for all k and all $j = 1, \dots, 2^k - 1$.

The function f is called the **Cantor function**. We restate its main properties:
(i) f is increasing and continuous on $[0,1]$.
(ii) $f(0) = 0$ and $f(1) = 1$, and f is constant on each of the subintervals of $[0,1] \setminus C$. More precisely, for every $k \ge 1$ the function f is constant $\frac{j}{2^k}$ on $J_j^{(k)}$ for all $j = 1, \dots, 2^k - 1$.

It is standard to extend the Cantor function on $\mathbb{R}$ by defining $f = 0$ on $(-\infty, 0)$ and $f = 1$ on $(1, +\infty)$. Thus, f is continuous and increasing on $\mathbb{R}$.

Exercises.

1.4.11. *An example of an m_1-null uncountable set which is dense in an interval.*
Let $\mathbb{Q} \cap [0,1] = \{x_1, x_2, \ldots\}$. Take $U(\epsilon) = \bigcup_{j=1}^{+\infty} \left(x_j - \frac{\epsilon}{2^j}, x_j + \frac{\epsilon}{2^j}\right)$ and $A = \bigcap_{n=1}^{+\infty} U\left(\frac{1}{n}\right)$.
(i) Prove that $m_1(U(\epsilon)) \leq 2\epsilon$.
(ii) If $\epsilon < \frac{1}{2}$, prove that $[0,1]$ is not a subset of $U(\epsilon)$.
(iii) Prove that $A \subseteq [0,1]$ and $m_1(A) = 0$.
(iv) Prove that $\mathbb{Q} \cap [0,1] \subseteq A$, and that A is uncountable.

1.4.12. Prove that the Cantor set is perfect: it is closed and has no isolated points.

1.4.13. (i) Prove that for every sequence (a_k) in $\{0,1,2\}$ the series $\sum_{k=1}^{+\infty} \frac{a_k}{3^k}$ converges and its sum is a number in $[0,1]$.
Conversely, prove that for every number x in $[0,1]$ there is a sequence (a_k) in $\{0,1,2\}$ so that $x = \sum_{k=1}^{+\infty} \frac{a_k}{3^k}$. Then we say that $0.a_1a_2\ldots$ is a **ternary expansion** of x and that $a_1, a_2, \ldots$ are the **ternary digits** of this expansion.
(ii) If $x \in [0,1]$ is of the form $x = \frac{m}{3^N}$, where $m \equiv 1 (\operatorname{mod} 3)$ and $N \in \mathbb{N}$, prove that x has exactly two ternary expansions: one of the form $0.a_1 \ldots a_{N-1}1000\ldots$ and another of the form $0.a_1 \ldots a_{N-1}0222\ldots$.
If $x \in [0,1]$ is either irrational or of the form $x = \frac{m}{3^N}$, where $m \equiv 0 (\operatorname{mod} 3)$ or $m \equiv 2 (\operatorname{mod} 3)$ and $N \in \mathbb{N}$, prove that x has exactly one ternary expansion which is not of either one of the above forms.
(iii) Let C be the Cantor set. If $x \in [0,1]$, prove that $x \in C$ if and only if x has at least one ternary expansion containing *no* ternary digit 1.

1.4.14. *More Cantor sets.*
(a) We take an arbitrary sequence (ϵ_k) so that $0 < \epsilon_k < \frac{1}{2}$ for all k. We split $I_0 = [0,1]$ into the three intervals $\left[0, \frac{1}{2} - \epsilon_1\right], \left(\frac{1}{2} - \epsilon_1, \frac{1}{2} + \epsilon_1\right), \left[\frac{1}{2} + \epsilon_1, 1\right]$, and we form I_1 as the union of the two closed intervals. Inductively, if we have already constructed I_{k-1} as a union of certain closed intervals, we split each of these intervals into three subintervals of which the two side ones are closed and their proportion to the original is $\frac{1}{2} - \epsilon_k$. Then we denote I_k the union of the new intervals. Clearly, I_k consists of 2^k disjoint closed intervals.
We set $K = \bigcap_{k=1}^{+\infty} I_k$.
Observe that, if $\epsilon_k = \frac{1}{6}$ for every k, then $K = C$, i.e. the usual Cantor set.
(i) Prove that K is compact, has no isolated points, includes no open interval, and is uncountable.
(ii) Prove that $m_1(K) = \lim_{k\to+\infty}(1 - 2\epsilon_1) \cdots (1 - 2\epsilon_k)$.
(iii) Taking $0 < \epsilon < 1$, and $\epsilon_k = \frac{\epsilon}{3^k}$ for all k, prove that $m_1(K) > 1 - \epsilon$.
Hint. $(1 - a_1) \cdots (1 - a_k) > 1 - (a_1 + \cdots + a_k)$ for all k and all $a_1, \ldots, a_k \in (0,1]$.
(iv) Prove that $m_1(K) > 0$ if and only if $\sum_{k=1}^{+\infty} \epsilon_k < +\infty$.
Hint. Use the inequality in the hint for (iii) and also that $1 - a \leq e^{-a}$ for all a.
(b) We can produce Cantor sets in $\mathbb{R}^n$. Using the sequence (ϵ_k) and the sequence (I_k) of closed subsets of $[0,1]$ in part (a), we consider the cartesian products $I_k^n = I_k \times \cdots \times I_k$. Then $I_0^n = [0,1]^n$ is the closed unit cube in $\mathbb{R}^n$, and every I_k^n is the union of 2^{kn} closed cubes. Each of the $2^{(k-1)n}$ cubes of I_{k-1}^n contains 2^n cubes of I_k^n. Now, if K is the set of part (a), then the cartesian product $K^n = K \times \cdots \times K$ is the intersection of the I_k^n, i.e. $K^n = \bigcap_{k=1}^{+\infty} I_k^n$.
Adjusting (i)-(iv) of part (a), prove that K^n is compact, has no isolated points, includes no open interval, is uncountable, has $m_n(K^n) = \lim_{k\to+\infty}((1-2\epsilon_1) \cdots (1-2\epsilon_k))^n$, and that $m_n(K^n) > 0$ if and only if $\sum_{k=1}^{+\infty} \epsilon_k < +\infty$.

A NON-LEBESGUE SET IN $\mathbb{R}$.

For any $x, y \in \mathbb{R}$ we write
$$x \sim y$$

if $x - y \in \mathbb{Q}$. It is easy to see that $\sim$ is an equivalence relation. Indeed, $x \sim x$, because $x - x = 0 \in \mathbb{Q}$. Also, if $x \sim y$, then $x - y \in \mathbb{Q}$, and then $y - x = -(x - y) \in \mathbb{Q}$, and so $y \sim x$. Finally, if $x \sim y$ and $y \sim z$, then $x-y \in \mathbb{Q}$ and $y-z \in \mathbb{Q}$, and then $x-z = (x-y)+(y-z) \in \mathbb{Q}$, and then $x \sim z$.

It is easy to see that every equivalence class of $\sim$ has non-empty intersection with $[0, 1]$. Indeed, let ξ be any equivalence class of $\sim$, and let $x \in \xi$. Since $\mathbb{Q}$ is dense in $\mathbb{R}$, there exists $r \in \mathbb{Q} \cap [-x, -x + 1]$. Now we consider $y = x + r$, and then $y \in [0, 1]$. Also $y \sim x$, and so $y \in \xi$.

Now, using the Axiom of Choice, we form a set N containing exactly one element from the intersection of each equivalence class of $\sim$ with $[0, 1]$.

Obviously, $N \subseteq [0, 1]$. Our aim is to prove that N is not a Lebesgue set in $\mathbb{R}$.

We form the set

$$A = \bigcup_{r\in\mathbb{Q}\cap[-1,1]}(N + r).$$

We shall need three properties of A.

(i) *If* $r_1, r_2 \in \mathbb{Q} \cap [-1, 1]$ *and* $r_1 \neq r_2$*, then* $(N + r_1) \cap (N + r_2) = \emptyset$*.*

Indeed, if $x \in (N + r_1) \cap (N + r_2)$, then $x - r_1, x - r_2 \in N$. But $x \sim x - r_1$ and $x \sim x - r_2$, and so N contains two different elements from the equivalence class of $\sim$ which contains x.

(ii) $A \subseteq [-1, 2]$.

This is clear, since $N \subseteq [0, 1]$ implies $N + r \subseteq [-1, 2]$ for every $r \in \mathbb{Q} \cap [-1, 1]$.

(iii) $[0, 1] \subseteq A$.

Indeed, let $x \in [0, 1]$ and let us consider the equivalence class ξ of $\sim$ which contains x. Then N contains exactly one element $\overline{x}$ from $\xi \cap [0, 1]$. Then $\overline{x} \in N$ and $x - \overline{x} \in \mathbb{Q}$. We consider $r = x - \overline{x}$, and then $r \in \mathbb{Q} \cap [-1, 1]$. Hence, $x = \overline{x} + r \in N + r$ for some $r \in \mathbb{Q} \cap [-1, 1]$, and so $x \in A$.

Now let us suppose that N is a Lebesgue set in $\mathbb{R}$. By (i) and by the invariance of m_1 under translations, we get that

$$m_1(A) = \sum_{r\in\mathbb{Q}\cap[-1,1]} m_1(N + r) = \sum_{r\in\mathbb{Q}\cap[-1,1]} m_1(N).$$

If $m_1(N) > 0$, then $m_1(A) = +\infty$, contradicting (ii). If $m_1(N) = 0$, then $m_1(A) = 0$, contradicting (iii).

Therefore, N is not a Lebesgue set in $\mathbb{R}$.

Exercises.

1.4.15. *Another construction of a non-Lebesgue set in* $\mathbb{R}$*.*

Consider the equivalence relation $\sim$ which we used in this section, and let L be a set containing exactly one element from each of the equivalence classes of $\sim$.

(i) Prove that $\mathbb{R} = \bigcup_{r\in\mathbb{Q}}(L + r)$, and that the sets $L + r$ are pairwise disjoint.

(ii) Prove that the difference set of L (see exercise 1.4.9) contains no rational number $\neq 0$.

(iii) Using the result of exercise 1.4.9, prove that L is not a Lebesgue set in $\mathbb{R}$.

1.4.16. *Non-Lebesgue sets in* $\mathbb{R}$ *are everywhere, I.*

We shall prove that every $E \subseteq \mathbb{R}$ with $m_1^*(E) > 0$ includes at least one non-Lebesgue set in $\mathbb{R}$.

(i) Consider the non-Lebesgue set $N \subseteq [0, 1]$ which was constructed in this section, and prove that, if $B \subseteq N$ is a Lebesgue set, then $m_1(B) = 0$. Therefore, if $M \subseteq N$ has $m_1^*(M) > 0$, then M is a non-Lebesgue set in $\mathbb{R}$.

(ii) Consider an arbitrary $E \subseteq \mathbb{R}$ with $m_1^*(E) > 0$, and $\alpha = 1 - m_1^*(N)$. Then $0 \leq \alpha < 1$. Exercise 1.4.3 implies that there is a bounded interval (a, b) so that $m_1^*(E \cap (a, b)) \geq \alpha\,(b - a)$. Now, the set $N' = (b - a)N + a$ is included in $[a, b]$, and has $m_1^*(N') = (1 - \alpha)(b - a)$. If $M' \subseteq N'$ has $m_1^*(M') > 0$, then M' is not a Lebesgue set in $\mathbb{R}$.

(iii) Prove that $E \cap N'$ is not a Lebesgue set in $\mathbb{R}$.

1.4.17. *Non-Lebesgue sets in $\mathbb{R}$ are everywhere, II.*

Consider $E \subseteq \mathbb{R}$ with $m_1^*(E) > 0$.

(i) Consider the set L in exercise 1.4.15. Then $E = \bigcup_{r\in\mathbb{Q}}(E \cap (L + r))$. Prove that the difference set (exercise 1.4.9) of each $E \cap (L + r)$ contains no rational number $\neq 0$.

(ii) Use the result of exercise 1.4.9, and prove that, for at least one $r \in \mathbb{Q}$, the set $E \cap (L + r)$ is not a Lebesgue set in $\mathbb{R}$.

1.4.18. *Not all Lebesgue sets in $\mathbb{R}$ are Borel sets, and not all continuous functions map Lebesgue sets onto Lebesgue sets.*

Let $f : [0, 1] \to [0, 1]$ be the Cantor function. We define $g(x) = f(x) + x$ for $x \in [0, 1]$.

(i) Prove that g is continuous, strictly increasing, one-to-one, and onto $[0, 2]$. Its inverse function $g^{-1} : [0, 2] \to [0, 1]$ is also continuous, strictly increasing, one-to-one, and onto $[0, 1]$.

(ii) Prove that the set $A = g([0, 1] \setminus C)$, where C is the Cantor set, is an open set in $\mathbb{R}$, with $m_1(A) = 1$. Therefore, the set $E = g(C)$ is a closed set in $\mathbb{R}$, with $m_1(E) = 1$.

(iii) Exercises 1.4.16 and 1.4.17 provide us with non-Lebesgue sets $M \subseteq E$. For any such set M, consider the set $K = g^{-1}(M) \subseteq C$. Prove that K is a Lebesgue set in $\mathbb{R}$.

(iv) Using exercise 1.1.7, prove that K is not a Borel set.

(v) g maps K onto M.

1.5 Borel measures on topological spaces.

LEBESGUE-STIELTJES MEASURES ON $\mathbb{R}$.

Lemma 1.3. *If $-\infty \le a < b \le +\infty$ and $F : (a, b) \to \mathbb{R}$ is increasing, then*

(i) $F(x+) = \inf\{F(y) \mid x < y\}$ if $x \in [a, b)$,

(ii) $F(x-) = \sup\{F(y) \mid y < x\}$ if $x \in (a, b]$,

(iii) $F(x-) \le F(x) \le F(x+) \le F(y) \le F(z-) \le F(z) \le F(z+)$ if $a < x < y < z < b$,

(iv) $F(x+) = \lim_{y\to x+} F(y\pm)$ if $x \in [a, b)$,

(v) $F(x-) = \lim_{y\to x-} F(y\pm)$ if $x \in (a, b]$.

Proof. Exercise. □

We consider a_0, b_0 with $-\infty \le a_0 < b_0 \le +\infty$ and an increasing function $F : (a_0, b_0) \to \mathbb{R}$. We define a *non-negative* function τ acting on bounded subintervals of (a_0, b_0), as follows:

$$\tau((a, b)) = F(b-) - F(a+), \qquad \tau([a, b]) = F(b+) - F(a-),$$
$$\tau((a, b]) = F(b+) - F(a+), \qquad \tau([a, b)) = F(b-) - F(a-).$$

The mnemonic rule is: if the end-point is included in the interval, then we approach it from outside of the interval, while, if the end-point is not included in the interval, then we approach it from inside of the interval.

We use the collection of all bounded open subintervals of (a_0, b_0) and the function τ to define, as an application of Proposition 1.28, the following outer measure on (a_0, b_0):

$$\mu_F^*(E) = \inf\big\{ \textstyle\sum_{j=1}^{+\infty} \tau((a_j, b_j)) \,\big|\, (a_j, b_j) \subseteq (a_0, b_0) \text{ for all } j \text{ so that } E \subseteq \bigcup_{j=1}^{+\infty}(a_j, b_j)\big\}$$

for every $E \subseteq (a_0, b_0)$. Caratheodory's Theorem implies that the collection of μ_F^*-measurable sets is a σ-algebra of subsets of (a_0, b_0). As we know, this σ-algebra is denoted $\mathcal{S}_{\mu_F^*}$, but we shall simplify the notation using the symbol $\mathcal{S}_F$. The restriction of μ_F^* on the σ-algebra of μ_F^*-measurable sets, i.e. $\mathcal{S}_F$, is denoted μ_F. Thus, we get the measure space

$$((a_0, b_0), \mathcal{S}_F, \mu_F)$$

which, by Caratheodory's Theorem, is complete.

Definition. *The measure μ_F is called the **Lebesgue-Stieltjes measure** induced by the (increasing) function $F : (a_0, b_0) \to \mathbb{R}$.*

If $F(x) = x$ for all $x \in \mathbb{R}$, then $\tau(S) = \text{vol}_1(S)$ for all bounded intervals S and, in this special case, μ_F coincides with the 1-dimensional Lebesgue measure m_1 on $\mathbb{R}$. Thus, Lebesgue-Stieltjes measure is a generalization of Lebesgue measure.

Following the same procedure as with Lebesgue measure, we shall investigate the relation between the σ-algebra $\mathcal{S}_F$ and the Borel subsets of (a_0, b_0). Proposition 1.40 is analogous to Proposition 1.30.

Proposition 1.40. *(i) Let $P = (a, b] \subseteq (a_0, b_0)$ and $a = c^{(0)} < c^{(1)} < \cdots < c^{(m)} = b$. If $P_i = (c^{(i-1)}, c^{(i)}]$, then $\tau(P) = \sum_{i=1}^m \tau(P_i)$.*
(ii) Assume that $P, P_1, \dots, P_l$ are bounded open-closed subintervals of (a_0, b_0), that $P_1, \dots, P_l$ are pairwise disjoint and that $P = \bigcup_{j=1}^l P_j$. Then $\tau(P) = \sum_{j=1}^l \tau(P_j)$.
(iii) Assume that $P, P_1, \dots, P_l$ are bounded open-closed subintervals of (a_0, b_0), that $P_1, \dots, P_l$ are pairwise disjoint and that $\bigcup_{j=1}^l P_j \subseteq P$. Then $\sum_{j=1}^l \tau(P_j) \leq \tau(P)$.
(iv) Assume that $P, P_1, \dots, P_l$ are bounded open-closed subintervals of (a_0, b_0) and that $P \subseteq \bigcup_{j=1}^l P_j$. Then $\tau(P) \leq \sum_{j=1}^l \tau(P_j)$.
(v) Assume that Q is a bounded closed interval, that $R_1, \dots, R_l$ are bounded open subintervals of (a_0, b_0) and that $Q \subseteq \bigcup_{j=1}^l R_j$. Then $\tau(Q) \leq \sum_{j=1}^l \tau(R_j)$.

Proof. (i) We have a telescoping sum:

$$\textstyle\sum_{i=1}^m \tau(P_i) = \sum_{i=1}^m (F(c^{(i)}+) - F(c^{(i-1)}+)) = F(b+) - F(a+) = \tau((a, b]).$$

(ii) Exactly one of $P_1, \dots, P_l$ has the same right end-point as P. We rename and call it P_l. Then exactly one of $P_1, \dots, P_{l-1}$ has right end-point coinciding with the left end-point of P_l. We rename and call it P_{l-1}. We continue until the left end-point of the last remaining subinterval, which we shall rename P_1, coincides with the left end-point of P. Then the result is clear from (i).
(iii) We know that $P \setminus (P_1 \cup \cdots \cup P_l) = P_1' \cup \cdots \cup P_k'$ for some pairwise disjoint open-closed intervals $P_1', \dots, P_k'$. Then $P = \big(\bigcup_{j=1}^l P_j\big) \cup \big(\bigcup_{i=1}^k P_i'\big)$, and from (ii) we get

$$\textstyle\tau(P) = \sum_{j=1}^l \tau(P_j) + \sum_{i=1}^k \tau(P_i') \geq \sum_{j=1}^l \tau(P_j).$$

(iv) We write $P = P_1' \cup \cdots \cup P_l'$, where $P_j' = P_j \cap P$ are open-closed intervals included in P. Then we write

$$P = P_1' \cup (P_2' \setminus P_1') \cup \cdots \cup \big(P_l' \setminus (P_1' \cup \cdots \cup P_{l-1}')\big).$$

Each of these l pairwise disjoint sets can be written as a finite union of pairwise disjoint open-closed intervals:

$$P_1' = P_1', \quad P_j' \setminus (P_1' \cup \cdots \cup P_{j-1}') = P_1^{(j)} \cup \cdots \cup P_{m_j}^{(j)} \text{ for } 2 \leq j \leq l.$$

Now, using (ii) for the equality and (iii) for the two inequalities, we get

$$\textstyle\tau(P) = \tau(P_1') + \sum_{j=2}^l \big(\sum_{m=1}^{m_j} \tau(P_m^{(j)})\big) \leq \tau(P_1') + \sum_{j=2}^l \tau(P_j') \leq \sum_{j=1}^l \tau(P_j).$$

(v) Let $Q = [a, b]$ and $R_j = (a_j, b_j)$. For small $\epsilon > 0$ we define $P_\epsilon = (a - \epsilon, b]$ and $P_{j,\epsilon} = (a_j, b_j - \epsilon]$. It is easy to see that $P_\epsilon \subseteq P_{1,\epsilon} \cup \cdots \cup P_{l,\epsilon}$ if $\epsilon > 0$ is small enough. Now, (iv) implies that

$$\textstyle F(b+) - F((a - \epsilon)+) \leq \sum_{j=1}^l \big(F((b_j - \epsilon)+) - F(a_j+)\big)$$

for small $\epsilon > 0$. We take the limit as $\epsilon \to 0+$, and we get

$$\textstyle\tau(Q) = F(b+) - F(a-) \leq \sum_{j=1}^l \big(F(b_j-) - F(a_j+)\big) = \sum_{j=1}^l \tau(R_j)$$

using Lemma 1.3. $\square$

Proposition 1.41 corresponds to Proposition 1.31.

Proposition 1.41. *Every bounded subinterval S of (a_0, b_0) is μ_F^*-measurable and $\mu_F(S) = \tau(S)$.*

Proof. Let $Q = [a, b] \subseteq (a_0, b_0)$.
Then

$$\mu_F^*(Q) \le \tau((a - \epsilon, b + \epsilon)) = F((b + \epsilon)-) - F((a - \epsilon)+)$$

for all small enough $\epsilon > 0$. Taking the limit as $\epsilon \to 0+$ and using Lemma 1.3, we get

$$\mu_F^*(Q) \le F(b+) - F(a-) = \tau(Q).$$

For every covering $Q \subseteq \bigcup_{j=1}^{+\infty} R_j$ by bounded open subintervals R_j of (a_0, b_0), there is (by compactness) l so that $Q \subseteq \bigcup_{j=1}^{l} R_j$. Proposition 1.40 implies

$$\tau(Q) \le \textstyle\sum_{j=1}^{l} \tau(R_j) \le \sum_{j=1}^{+\infty} \tau(R_j).$$

Hence $\tau(Q) \le \mu_F^*(Q)$, and we conclude that

$$\mu_F^*(Q) = \tau(Q)$$

for all closed intervals $Q \subseteq (a_0, b_0)$.
If $P = (a, b] \subseteq (a_0, b_0)$, then

$$\mu_F^*(P) \le \tau((a, b + \epsilon)) = F((b + \epsilon)-) - F(a+)$$

for all small enough $\epsilon > 0$. We take the limit as $\epsilon \to 0+$, and we get

$$\mu_F^*(P) \le F(b+) - F(a+) = \tau(P).$$

If $R = (a, b) \subseteq (a_0, b_0)$, then

$$\mu_F^*(R) \le \tau((a, b)) = \tau(R).$$

Now let $P = (a, b]$ and $R = (c, d)$ be included in (a_0, b_0). We take $P_R = (c, d - \epsilon]$, and we write

$$\begin{aligned}\mu_F^*(R \cap P) &= \mu_F^*\big((P_R \cap P) \cup ((d - \epsilon, d) \cap P)\big) \le \mu_F^*(P_R \cap P) + \mu_F^*((d - \epsilon, d)) \\ &\le \tau(P_R \cap P) + F(d-) - F((d - \epsilon)+)\end{aligned}$$

by the previous results. The same inequalities, with P^c instead of P, give

$$\mu_F^*(R \cap P^c) \le \mu_F^*(P_R \cap P^c) + F(d-) - F((d - \epsilon)+).$$

We sum the last two inequalities, and we find

$$\mu_F^*(R \cap P) + \mu_F^*(R \cap P^c) \le \tau(P_R \cap P) + \mu_F^*(P_R \cap P^c) + 2(F(d-) - F((d - \epsilon)+)).$$

Now, we have $P_R \cap P^c = P_1 \cup \cdots \cup P_l$ for pairwise disjoint open-closed intervals P_j, and we get

$$\begin{aligned}\tau(P_R \cap P) + \mu_F^*(P_R \cap P^c) &\le \tau(P_R \cap P) + \textstyle\sum_{j=1}^{l} \mu_F^*(P_j) \\ &\le \tau(P_R \cap P) + \textstyle\sum_{j=1}^{l} \tau(P_j) = \tau(P_R)\end{aligned}$$

by our first results and Lemma 1.3. Therefore,

$$\begin{aligned}\mu_F^*(R \cap P) + \mu_F^*(R \cap P^c) &\le \tau(P_R) + 2(F(d-) - F((d - \epsilon)+)) \\ &= F((d - \epsilon)+) - F(c+) + 2(F(d-) - F((d - \epsilon)+)).\end{aligned}$$

Taking limit as $\epsilon \to 0+$, we find

$$\mu_F^*(R \cap P) + \mu_F^*(R \cap P^c) \le F(d-) - F(c+) = \tau(R).$$

We proved that
$$\mu_F^*(R\cap P)+\mu_F^*(R\cap P^c)\leq \tau(R)$$
for all bounded open intervals R and bounded open-closed intervals P included in (a_0,b_0).
Now, we consider an arbitrary $E\subseteq (a_0,b_0)$ with $\mu_F^*(E)<+\infty$. We take a covering $E\subseteq \bigcup_{j=1}^{+\infty}R_j$ by bounded open subintervals R_j of (a_0,b_0) so that
$$\sum_{j=1}^{+\infty}\tau(R_j)<\mu_F^*(E)+\epsilon.$$
By σ-subadditivity of μ_F^* and by the last result we find
$$\mu_F^*(E\cap P)+\mu_F^*(E\cap P^c)\leq \sum_{j=1}^{+\infty}\big(\mu_F^*(R_j\cap P)+\mu_F^*(R_j\cap P^c)\big)\leq \sum_{j=1}^{+\infty}\tau(R_j)<\mu_F^*(E)+\epsilon.$$
Taking limit as $\epsilon\to 0+$, we find
$$\mu_F^*(E\cap P)+\mu_F^*(E\cap P^c)\leq \mu_F^*(E),$$
concluding that $P\in\mathcal{S}_F$.
If $Q=[a,b]\subseteq (a_0,b_0)$, we take any increasing (a_k) in (a_0,b_0) so that $\lim_{k\to+\infty}a_k=a$ and then $Q=\bigcap_{k=1}^{+\infty}(a_k,b]\in\mathcal{S}_F$. Moreover, by our first result,
$$\mu_F(Q)=\mu_F^*(Q)=\tau(Q).$$
If $P=(a,b]\subseteq (a_0,b_0)$, we take any decreasing (a_k) in $(a,b]$ so that $\lim_{k\to+\infty}a_k=a$, and we get that
$$\mu_F(P)=\lim_{k\to+\infty}\mu_F([a_k,b])=\lim_{k\to+\infty}(F(b+)-F(a_k-))=F(b+)-F(a+)=\tau(P).$$
If $T=[a,b)\subseteq (a_0,b_0)$, we take any increasing (b_k) in $[a,b)$ so that $\lim_{k\to+\infty}b_k=b$, and we get that $T=\bigcup_{k=1}^{+\infty}[a,b_k]\in\mathcal{S}_F$. Moreover,
$$\mu_F(T)=\lim_{k\to+\infty}\mu_F([a,b_k])=\lim_{k\to+\infty}(F(b_k+)-F(a-))=F(b-)-F(a-)=\tau(T).$$
Finally, if $R=(a,b)\subseteq (a_0,b_0)$, we take any decreasing (a_k) and any increasing (b_k) in (a,b) so that $\lim_{k\to+\infty}a_k=a$, $\lim_{k\to+\infty}b_k=b$ and $a_1\leq b_1$. Then $R=\bigcup_{k=1}^{+\infty}[a_k,b_k]\in\mathcal{S}_F$. Moreover,
$$\mu_F(R)=\lim_{k\to+\infty}\mu_F([a_k,b_k])=\lim_{k\to+\infty}(F(b_k+)-F(a_k-))=F(b-)-F(a+)=\tau(R).$$
We have thus proved that $\mu_F(S)=\tau(S)$ for every bounded interval $S\subseteq (a_0,b_0)$. $\square$

Proposition 1.42 corresponds to Proposition 1.32.

Proposition 1.42. *μ_F is σ-finite. Moreover, μ_F is finite if and only if F is bounded.*

Proof. We take any decreasing (a_k) and any increasing (b_k) in (a_0,b_0) so that $\lim_{k\to+\infty}a_k=a_0$, $\lim_{k\to+\infty}b_k=b_0$. Then $(a_0,b_0)=\bigcup_{k=1}^{+\infty}[a_k,b_k]$ and $\mu_F([a_k,b_k])=F(b_k+)-F(a_k-)<+\infty$ for all k. Hence, μ_F is σ-finite.
We know that $\mu_F((a_0,b_0))=F(b_0-)-F(a_0+)$. Therefore, if μ_F is finite, then $-\infty<F(a_0+)$ and $F(b_0-)<+\infty$. Since all values of F lie in the bounded interval $[F(a_0+),F(b_0-)]$, we get that F is bounded. Conversely, if F is bounded, then $F(a_0+)$ and $F(b_0-)$ are finite, and so $\mu_F((a_0,b_0))<+\infty$. $\square$

It is easy to prove that *the collection of all subintervals of (a_0,b_0) generates the σ-algebra of all Borel subsets of (a_0,b_0)*. Indeed, let $\mathcal{C}$ be the collection of all intervals in $\mathbb{R}$ and $\mathcal{F}$ be the collection of all subintervals of (a_0,b_0). It is clear that $\mathcal{F}=\mathcal{C}\rceil(a_0,b_0)$, and then Propositions 1.9 and 1.10 imply that $\mathcal{B}_{(a_0,b_0)}=\mathcal{B}_1\rceil(a_0,b_0)=\mathcal{S}(\mathcal{C})\rceil(a_0,b_0)=\mathcal{S}(\mathcal{F})$.

Proposition 1.43 corresponds to Proposition 1.33.

Proposition 1.43. *All Borel subsets of* (a_0, b_0) *belong to* $\mathcal{S}_F$.

Proof. Proposition 1.41 implies that the collection $\mathcal{F}$ of all subintervals of (a_0, b_0) is included in $\mathcal{S}_F$. By the discussion of the previous paragraph, we conclude that $\mathcal{B}_{(a_0,b_0)} = \mathcal{S}(\mathcal{F}) \subseteq \mathcal{S}_F$. ☐

Proposition 1.44 corresponds to Proposition 1.34.

Proposition 1.44. *Let* $E \subseteq (a_0, b_0)$. *Then*
(i) $E \in \mathcal{S}_F$ *if and only if there is an* $A \subseteq (a_0, b_0)$, *which is a countable intersection of open sets, such that* $E \subseteq A$ *and* $\mu_F^*(A \setminus E) = 0$.
(ii) $E \in \mathcal{S}_F$ *if and only if there is a* B, *which is a countable union of compact sets, such that* $B \subseteq E$ *and* $\mu_F^*(E \setminus B) = 0$.

Proof. The proof is exactly the same as the proof of the similar Proposition 1.34. Only the obvious changes have to be made: m_n changes to μ_F, and m_n^* to μ_F^*, $\mathbb{R}^n$ changes to (a_0, b_0), vol_n changes to τ, and $\mathcal{L}_n$ changes to $\mathcal{S}_F$. ☐

Therefore, *every set in* $\mathcal{S}_F$ *is, except from a* μ_F*-null set, equal to a Borel set.*
Proposition 1.45 corresponds to Proposition 1.35.

Proposition 1.45. *(i)* μ_F *is the only measure on* $\big((a_0, b_0), \mathcal{B}_{(a_0,b_0)}\big)$ *satisfying* $\mu_F(S) = \tau(S)$ *for all bounded intervals* $S \subseteq (a_0, b_0)$.
(ii) $\big((a_0, b_0), \mathcal{S}_F, \mu_F\big)$ *is the completion of* $\big((a_0, b_0), \mathcal{B}_{(a_0,b_0)}, \mu_F\big)$.

Proof. The proof is similar to the proof of Proposition 1.35. Only some obvious notational modifications are needed. ☐

It should be observed that the Lebesgue-Stieltjes measure of a set $\{x\}$, consisting of a single point $x \in (a_0, b_0)$, is equal to

$$\mu_F(\{x\}) = F(x+) - F(x-),$$

i.e. to the jump of F at x. In other words, the measure of a one-point set is positive if and only if F is discontinuous there. Also, observe that the measure of an *open* subinterval of (a_0, b_0) is 0 if and only if F is constant on this interval.

It is very common in practice to consider the increasing function F with the extra property of being *continuous from the right*. In this case the measure of an open-closed interval takes the simpler form

$$\mu_F((a, b]) = F(b) - F(a).$$

Proposition 1.46 shows that this is not a serious restriction.

Proposition 1.46. *Given any increasing function on* (a_0, b_0) *there is another increasing function which is continuous from the right so that the Lebesgue-Stieltjes measures induced by the two functions are equal.*

Proof. Given any increasing $F : (a_0, b_0) \to \mathbb{R}$, we define $F_0 : (a_0, b_0) \to \mathbb{R}$ by $F_0(x) = F(x+)$ for all $x \in (a_0, b_0)$. It is immediate from Lemma 1.3 that F_0 is increasing, that F_0 is continuous from the right, i.e. $F_0(x+) = F_0(x)$ for all x, and that $F_0(x+) = F(x+), F_0(x-) = F(x-)$ for all x. Now, F_0 and F induce the same Lebesgue-Stieltjes measure on (a_0, b_0), simply because the corresponding functions $\tau(S)$ (from which the constructions of the measures μ_{F_0}, μ_F start) assign the same values to every interval $S \subseteq (a_0, b_0)$. ☐

The functions F_0 and F of Proposition 1.46 have the same jump at every x and, in particular, they have the same continuity points.

Example. We consider the Cantor function $f : \mathbb{R} \to [0, 1]$ which is increasing, continuous and bounded. We call μ_f the **Cantor measure** on $(-\infty, +\infty)$.
Since f is continuous, we have that $\mu_f(\{x\}) = 0$ for every x. We recall that f is constant on every subinterval of $[0, 1] \setminus C$, where C is the Cantor set, and that f is constant 0 on $(-\infty, 0]$ and constant 1 on $[1, +\infty)$. Therefore, $\mu_f((-\infty, 0]) = \mu_f([1, +\infty)) = 0$, and $\mu_f(J_m) = 0$ for each of the subintervals $J_1, J_2, \dots$ of $[0, 1] \setminus C$. Since $f(0) = 0$ and $f(1) = 1$, we get

$$\mu_f((-\infty, +\infty)) = \mu_f([0, 1]) = f(1) - f(0) = 1.$$

Moreover,

$$\mu_f(C) = \mu_f([0, 1]) - \textstyle\sum_{m=1}^{+\infty} \mu_f(J_m) = 1 - \sum_{m=1}^{+\infty} 0 = 1.$$

Since $\mu_f(C) = \mu_f((-\infty, +\infty)) = 1$, we get that $\mu_f(A) = 0$ for every Borel set A in $\mathbb{R}$ with $A \cap C = \emptyset$.
Finally, since the difference of the values of f at the endpoints of each of the 2^k subintervals of I_k (look at the construction of C) is equal to $\frac{1}{2^k}$, we have that $\mu_f(I) = \frac{1}{2^k}$ for each of these subintervals I of I_k.

BOREL MEASURES ON TOPOLOGICAL SPACES.

Definition. *Let X be a topological space and $(X, \mathcal{S}, \mu)$ be a measure space. The measure μ is called a **Borel measure** on X if $\mathcal{B}_X \subseteq \mathcal{S}$, i.e. if all Borel subsets of X are measurable.*

Observe that, for μ to be a Borel measure, it is enough that all open sets are measurable. This is because $\mathcal{B}_X$ is generated by the collection of all open sets.

Example. Lebesgue measure m_n on $\mathbb{R}^n$ is a Borel measure.

Example. Every Lebesgue-Stieltjes measure μ_F on any interval (a_0, b_0) is a Borel measure.
It is easy to see that $\mu_F(K) < +\infty$ for every compact $K \subseteq (a_0, b_0)$. Indeed, we have $K \subseteq [a, b] \subseteq (a_0, b_0)$ for some a, b, and so

$$\mu_F(K) \subseteq \mu_F([a, b]) = F(b+) - F(a-) < +\infty.$$

In fact, Proposition 1.47 says that Lebesgue-Stieltjes measures are the only Borel measures μ on an interval (a_0, b_0) with the property that $\mu([a, b]) < +\infty$ for every $[a, b] \subseteq (a_0, b_0)$.

Proposition 1.47. *Let $-\infty \le a_0 < b_0 \le +\infty$ and $c_0 \in (a_0, b_0)$. Also let μ be a Borel measure on (a_0, b_0) so that $\mu([a, b]) < +\infty$ for every $[a, b] \subseteq (a_0, b_0)$. Then there is a unique function $F : (a_0, b_0) \to \mathbb{R}$, which is increasing and continuous from the right, so that $\mu = \mu_F$ on $\mathcal{B}_{(a_0,b_0)}$ and $F(c_0) = 0$. For any other function $G : (a_0, b_0) \to \mathbb{R}$, which is increasing and continuous from the right, we have: $\mu = \mu_G$ if and only if G differs from F by a constant.*

Proof. We define $F(x) = \mu((c_0, x])$, if $c_0 \le x < b_0$, and $F(x) = -\mu((x, c_0])$, if $a_0 < x < c_0$.
F is real valued, and it is clear, by the monotonicity of μ, that F is increasing.
We take any decreasing (x_n) so that $\lim_{n\to+\infty} x_n = x$. If $c_0 \le x$, by continuity of μ from above, we get

$$\lim_{n\to+\infty} F(x_n) = \lim_{n\to+\infty} \mu((c_0, x_n]) = \mu((c_0, x]) = F(x).$$

Also, if $x < c_0$, then $x_n < c_0$ for large n, and, by continuity of μ from below, we get

$$\lim_{n\to+\infty} F(x_n) = -\lim_{n\to+\infty} \mu((x_n, c_0]) = -\mu((x, c_0]) = F(x).$$

Therefore, F is continuous from the right at every x.
Now we have that

$$\mu_F((a, b]) = F(b) - F(a) = \mu((a, b]),$$

where the second equality becomes clear by considering cases: $a < b < c_0$, $a < c_0 \le b$ and $c_0 \le a < b$. We easily get the same result, namely $\mu_F(S) = \mu(S)$, for all other types of intervals S, and then Proposition 1.45 implies that $\mu_F = \mu$ on $\mathcal{B}_{(a_0,b_0)}$.
Let $G : (a_0, b_0) \to \mathbb{R}$ be increasing and continuous from the right, and let $\mu_G = \mu (= \mu_F)$ on $\mathcal{B}_{(a_0,b_0)}$. Then we have that

$$G(x) - G(c_0) = \mu_G((c_0, x]) = \mu_F((c_0, x]) = F(x) - F(c_0)$$

for all $x \ge c_0$. Similarly,

$$G(c_0) - G(x) = \mu_G((x, c_0]) = \mu_F((x, c_0]) = F(c_0) - F(x)$$

for all $x < c_0$. Thus, F, G differ by a constant: $G - F = G(c_0) - F(c_0)$ on (a_0, b_0). Moreover, if $F(c_0) = 0 = G(c_0)$, then F, G are equal on (a_0, b_0). ☐

If the Borel measure μ of Proposition 1.47 satisfies $\mu((a_0, c_0]) < +\infty$, then we may make a different choice for F than the one we made in the proof of Proposition 1.47. We add the constant $\mu((a_0, c_0])$ to the function F in the proof, and we get the function

$$F(x) = \mu((a_0, x]), \quad x \in (a_0, b_0).$$

This last function is called the **cumulative distribution function** of μ.

A central notion related to Borel measures is the notion of regularity, and this is because of the need to relate the general Borel set (a somewhat obscure object) to appropriate open or closed sets.

We recall that a topological space X is called **Hausdorff** if for every $x_1, x_2 \in X$ with $x_1 \ne x_2$ there are *disjoint* open neighborhoods V_{x_1}, V_{x_2} of x_1, x_2, respectively. We know that every compact subset of a Hausdorff topological space is closed and, hence, a Borel set.

Let E be a Borel subset of a Hausdorff topological space X and μ be a Borel measure on X. It is clear that $\mu(K) \le \mu(E) \le \mu(U)$ for all compact K and open U with $K \subseteq E \subseteq U$. Hence,

$$\sup\{\mu(K) \,|\, K \text{ compact}, \ K \subseteq E\} \le \mu(E) \le \inf\{\mu(U) \,|\, U \text{ open}, \ E \subseteq U\}.$$

Definition. *Let X be a Hausdorff topological space and μ be a Borel measure on X. Then μ is called **regular** if the following are true for every Borel subset E of X:*
(i) $\mu(E) = \inf\{\mu(U) \,|\, U \text{ open}, \ E \subseteq U\}$,
(ii) $\mu(E) = \sup\{\mu(K) \,|\, K \text{ compact}, \ K \subseteq E\}$.

In other words, μ is regular if the measure of every Borel set can be approximated from above by the measures of larger open sets and from below by the measures of smaller compact sets.

In the proof of Proposition 1.48 we shall use the Euclidean norm $\|\cdot\|_2$ on $\mathbb{R}^n$, defined by

$$\|x\|_2 = (x_1^2 + \cdots + x_n^2)^{1/2} \quad \text{for all } x = (x_1, \ldots, x_n).$$

We also recall the Euclidean open balls in $\mathbb{R}^n$: the open ball with center $x \in \mathbb{R}^n$ and radius $r > 0$ is

$$B(x; r) = \{y \in \mathbb{R}^n \,|\, \|y - x\|_2 < r\}.$$

Proposition 1.48. *Let O be any open subset of $\mathbb{R}^n$. Then there is an increasing sequence (K_m) of compact sets so that $\bigcup_{m=1}^{+\infty} K_m = O$.*

Proof. We consider the sets

$$K_m = \left\{x \in O \,\middle|\, \|x\|_2 \le m \text{ and } \|y - x\|_2 \ge \tfrac{1}{m} \text{ for all } y \notin O\right\}.$$

The set K_m is bounded, since $\|x\|_2 \le m$ for all $x \in K_m$.
Let (x_j) be a sequence in K_m converging to some x in $\mathbb{R}^n$. From $\|x_j\|_2 \le m$ for all j, we get

$\|x\|_2 \leq m$. Also, from $\|y - x_j\|_2 \geq \frac{1}{m}$ for all j and for all $y \notin O$, we get $\|y - x\|_2 \geq \frac{1}{m}$ for all $y \notin O$. Thus, $x \in K_m$, and so K_m is closed.
Therefore, K_m is compact.
It is clear that $K_m \subseteq K_{m+1} \subseteq O$ for all m.
Now we take any $x \in O$, and an $\epsilon > 0$ such that $B(x;\epsilon) \subseteq O$. We also consider any $m \in \mathbb{N}$ such that $m \geq \max\left\{\|x\|_2, \frac{1}{\epsilon}\right\}$, and then it is trivial to see that $x \in K_m$. Thus, $\bigcup_{m=1}^{+\infty} K_m = O$. □

Theorem 1.2. *Let X be a Hausdorff topological space and μ be a Borel measure on X. We assume that for every open set O there is an increasing sequence of compact subsets of O which cover O, and that there is an increasing sequence of open sets with finite μ-measure which cover X. Then:*
(i) $\mu(K) < +\infty$ for every compact set K.
(ii) For every Borel set E and every $\epsilon > 0$ there is an open set U and a closed set F so that $F \subseteq E \subseteq U$ and $\mu(U \setminus F) < \epsilon$. If also $\mu(E) < +\infty$, then F can be taken compact.
(iii) For every Borel set E there is a set A, which is a countable intersection of open sets, and a set B, which is a countable union of compact sets, so that $B \subseteq E \subseteq A$ and $\mu(A \setminus B) = 0$.
(iv) μ is regular.

Proof. There is an increasing sequence (G_m) of open sets so that $\mu(G_m) < +\infty$ for every m and $\bigcup_{m=1}^{+\infty} G_m = X$.
Now, let K be compact. Since $K \subseteq \bigcup_{m=1}^{+\infty} G_m$, there is M so that $K \subseteq \bigcup_{m=1}^{M} G_m$. Then

$$\mu(K) \leq \textstyle\sum_{m=1}^{M} \mu(G_m) < +\infty,$$

and we have proved (i).
(a) Let $\mu(X) < +\infty$.
We consider the collection $\mathcal{S}$ of all Borel sets E with the property expressed in (ii), namely, that for every $\epsilon > 0$ there is an open U and a closed F so that $F \subseteq E \subseteq U$ and $\mu(U \setminus F) < \epsilon$.
We take any open O, and any $\epsilon > 0$. By assumption there is an increasing sequence (K_m) of compact sets so that $\bigcup_{m=1}^{+\infty} K_m = O$. Then $(O \setminus K_m)$ is decreasing and $\bigcap_{m=1}^{+\infty}(O \setminus K_m) = \emptyset$. Since $\mu(O \setminus K_1) \leq \mu(X) < +\infty$, continuity of μ from above implies that

$$\lim_{m\to+\infty} \mu(O \setminus K_m) = 0.$$

Hence, there is some m so that $\mu(O \setminus K_m) < \epsilon$. Considering $U = O$ and $F = K_m$, we get $F \subseteq O \subseteq U$ and $\mu(U \setminus F) < \epsilon$. Thus, all open sets belong to $\mathcal{S}$.
If $E \in \mathcal{S}$ and $\epsilon > 0$ is arbitrary, there is an open U and a closed F so that $F \subseteq E \subseteq U$ and $\mu(U \setminus F) < \epsilon$. Then F^c is open, U^c is closed, $U^c \subseteq E^c \subseteq F^c$ and

$$\mu(F^c \setminus U^c) = \mu(U \setminus F) < \epsilon.$$

This implies that $E^c \in \mathcal{S}$.
Now, we take $E_1, E_2, \ldots \in \mathcal{S}$ and $E = \bigcup_{j=1}^{+\infty} E_j$. If $\epsilon > 0$, for each E_j there is an open U_j and a closed F_j' so that $F_j' \subseteq E_j \subseteq U_j$ and $\mu(U_j \setminus F_j') < \frac{\epsilon}{2^j}$. We consider $B = \bigcup_{j=1}^{+\infty} F_j'$ and the open set $U = \bigcup_{j=1}^{+\infty} U_j$, and then $B \subseteq E \subseteq U$. Then $U \setminus B \subseteq \bigcup_{j=1}^{+\infty}(U_j \setminus F_j')$, and so

$$\mu(U \setminus B) \leq \textstyle\sum_{j=1}^{+\infty} \mu(U_j \setminus F_j') < \sum_{j=1}^{+\infty} \frac{\epsilon}{2^j} = \epsilon.$$

Since B is not necessarily closed, we consider the closed sets $F_j = F_1' \cup \cdots \cup F_j'$. Then (F_j) is increasing and $\bigcup_{j=1}^{+\infty} F_j = B$, and so $(U \setminus F_j)$ is decreasing and $\bigcap_{j=1}^{+\infty}(U \setminus F_j) = U \setminus B$. Since $\mu(U \setminus F_1) \leq \mu(X) < +\infty$, continuity of μ from below gives

$$\lim_{j\to+\infty} \mu(U \setminus F_j) = \mu(U \setminus B).$$

Thus, there is some j so that $\mu(U \setminus F_j) < \epsilon$. The inclusion $F_j \subseteq E \subseteq U$ is clearly true. We conclude that $E \in \mathcal{S}$.
Therefore, $\mathcal{S}$ is a σ-algebra. Since $\mathcal{S}$ contains all open sets, we have that $\mathcal{B}_X \subseteq \mathcal{S}$, and we have finished the proof of the first statement of (ii) in the special case $\mu(X) < +\infty$.
(b) Let E be a Borel set such that there is some open G with $E \subseteq G$ and $\mu(G) < +\infty$.
We consider the G-restriction μ_G of μ, which is defined by $\mu_G(A) = \mu(A \cap G)$ for all Borel sets A. Clearly, $\mu_G(X) = \mu(G) < +\infty$.
By the result of (a), for any $\epsilon > 0$ there is an open U' and a closed F so that $F \subseteq E \subseteq U'$ and $\mu_G(U' \setminus F) < \epsilon$. We consider the open set $U = U' \cap G$. Since $E \subseteq G$, we get $F \subseteq E \subseteq U \subseteq G$ and $\mu(U \setminus F) = \mu_G(U \setminus F) < \epsilon$.
Therefore, the first statement of (ii) is now proved with no restriction on $\mu(X)$ but only for Borel sets which are included in open sets of finite μ-measure.
(c) Now, we consider the general case, and the sequence (G_m) of open sets as in the beginning of the proof.
For any Borel set E we consider the Borel sets

$$E_1 = E \cap G_1, \quad E_m = E \cap (G_m \setminus G_{m-1}) \text{ for all } m \geq 2,$$

and we have that $E = \bigcup_{m=1}^{+\infty} E_m$. Since $E_m \subseteq G_m$, (b) implies that for each m and every $\epsilon > 0$ there is an open U_m and a closed F_m so that $F_m \subseteq E_m \subseteq U_m$ and $\mu(U_m \setminus F_m) < \frac{\epsilon}{2^m}$. Now we consider the sets

$$U = \bigcup_{m=1}^{+\infty} U_m, \quad F = \bigcup_{m=1}^{+\infty} F_m.$$

Then U is clearly open, and it is easy to see that F is closed. Indeed, let $x \in F^c = \bigcap_{m=1}^{+\infty} F_m^c$. Then $x \in G_M$ for some large M. Also, $x \in \bigcap_{m=1}^{M} F_m^c$ and $\bigcap_{m=1}^{M} F_m^c$ is open. Hence, there is an open neighborhood U_x of x which is included in $G_M \cap \bigcap_{m=1}^{M} F_m^c$. Since G_M is included in $\bigcap_{m=M+1}^{+\infty} F_m^c$, we get that U_x is included in $\bigcap_{m=1}^{+\infty} F_m^c = F^c$. Therefore, F^c is open.
Finally, $F \subseteq E \subseteq U$, and, as in the proof of (a), we have $U \setminus F \subseteq \bigcup_{m=1}^{+\infty} (U_m \setminus F_m)$, and so

$$\mu(U \setminus F) \leq \sum_{m=1}^{+\infty} \mu(U_m \setminus F_m) < \sum_{m=1}^{+\infty} \frac{\epsilon}{2^m} = \epsilon.$$

This concludes the proof of the first statement of (ii).
(d) Let $\mu(E) < +\infty$. There are an open U and a closed F so that $F \subseteq E \subseteq U$ and $\mu(U \setminus F) < \frac{\epsilon}{2}$. By assumption, there is an increasing sequence (K_m) of compact sets so that $\bigcup_{m=1}^{+\infty} K_m = X$. Then the sets $F_m = F \cap K_m$ are compact, the (F_m) is increasing and $\bigcup_{m=1}^{+\infty} F_m = F$. Hence, $(E \setminus F_m)$ is decreasing and $\bigcap_{m=1}^{+\infty} (E \setminus F_m) = E \setminus F$ and $\mu(E \setminus F_1) \leq \mu(E) < +\infty$. By continuity of μ from above,

$$\lim_{m \to +\infty} \mu(E \setminus F_m) = \mu(E \setminus F) < \frac{\epsilon}{2}.$$

Hence, there is m so that $\mu(E \setminus F_m) < \frac{\epsilon}{2}$, and so

$$\mu(U \setminus F_m) = \mu(U \setminus E) + \mu(E \setminus F_m) < \epsilon.$$

This proves the second statement of (ii).
(e) Let E be a Borel set. We take open U_j and closed F_j so that $F_j \subseteq E \subseteq U_j$ and $\mu(U_j \setminus F_j) < \frac{1}{j}$. We define $A = \bigcap_{j=1}^{+\infty} U_j$ and $B = \bigcup_{j=1}^{+\infty} F_j$, and then $B \subseteq E \subseteq A$. Now, for all j we have

$$\mu(A \setminus B) \leq \mu(U_j \setminus F_j) < \frac{1}{j},$$

and so $\mu(A \setminus B) = 0$. We consider the compact sets K_m of part (d), and we define the compact sets $K_{j,m} = F_j \cap K_m$. Then $B = \bigcup_{(j,m) \in \mathbb{N} \times \mathbb{N}} K_{j,m}$, and we conclude the proof of (iii).
(f) If $\mu(E) = +\infty$, then we have that $\mu(U) = +\infty$ for all open U such that $E \subseteq U$, and so we get $\mu(E) = \inf\{\mu(U) \,|\, U \text{ open}, \ E \subseteq U\}$.

If $\mu(E) < +\infty$, then, from (i), for every $\epsilon > 0$ there is an open U so that $E \subseteq U$ and $\mu(U \setminus E) < \epsilon$. This implies

$$\mu(U) = \mu(E) + \mu(U \setminus E) < \mu(E) + \epsilon,$$

and so again $\mu(E) = \inf\{\mu(U) \,|\, U \text{ open}, \ E \subseteq U\}$.
Finally, from (iii), there is some $B = \bigcup_{m=1}^{+\infty} H'_m$, where all H'_m are compact, so that $B \subseteq E$ and $\mu(E \setminus B) = 0$. Hence,

$$\mu(B) = \mu(B) + \mu(E \setminus B) = \mu(E).$$

We take the compact sets $H_m = H'_1 \cup \cdots \cup H'_m$, and then (H_m) is increasing and $\bigcup_{m=1}^{+\infty} H_m = B$. Then

$$\lim_{m \to +\infty} \mu(H_m) = \mu(B) = \mu(E),$$

and so $\sup\{\mu(K) \,|\, K \text{ compact}, \ K \subseteq E\} = \mu(E)$. □

Example. Let us consider the Euclidean space $\mathbb{R}^n$ with any Borel measure μ on $\mathbb{R}^n$ such that $\mu(B(0;m)) < +\infty$ for every $m \in \mathbb{N}$.
Then Proposition 1.48 implies that $\mathbb{R}^n$ and μ satisfy the assumptions of Theorem 1.2, and so, in particular, μ is regular.
A special case of this is the Lebesgue measure m_n on $\mathbb{R}^n$.

Example. Let (a_0, b_0) be an interval in $\mathbb{R}$ and μ be a Borel measure on (a_0, b_0) so that $\mu([a,b]) < +\infty$ for every $[a,b] \subseteq (a_0, b_0)$.
It is easy to see, by means of Proposition 1.48, that the assumptions of Theorem 1.2 are satisfied, and so μ is regular. On the other hand, since Proposition 1.47 implies that μ is a Lebesgue-Stieltjes measure, this result (the regularity of μ) is also easily implied by Proposition 1.44.

Exercises.

1.5.1. If $-\infty < x_1 < x_2 < \cdots < x_N < +\infty$ and $0 < \lambda_1, \dots, \lambda_N < +\infty$, then find (and draw) the cumulative distribution function of $\mu = \sum_{k=1}^{N} \lambda_k \delta_{x_k}$.

1.5.2. Let μ be a Borel measure on $\mathbb{R}$ so that $\mu(K) < +\infty$ for every compact $K \subseteq \mathbb{R}$ and so that $\mu((-\infty, 0]) < +\infty$. Prove that there is a unique $F : \mathbb{R} \to \mathbb{R}$, which is increasing and continuous from the right, so that $\mu = \mu_F$ and $\lim_{x \to -\infty} F(x) = 0$. Which is this function?

1.5.3. If μ, ν are regular Borel measures on the Hausdorff topological space X and $\lambda \in [0, +\infty)$, prove that $\lambda\mu$ and $\mu + \nu$ are regular Borel measures on X.

1.5.4. Let μ be a Borel measure on the topological space X. A point $x \in X$ is called a **support point** for μ if $\mu(U_x) > 0$ for every open neighborhood U_x of x.
The set

$$\text{supp}(\mu) = \{x \in X \,|\, x \text{ is a support point for } \mu\}$$

is called the **support** of μ.
(i) Prove that $\text{supp}(\mu)$ is a closed set.
(ii) If X is Hausdorff, prove that $\mu(K) = 0$ for all compact sets $K \subseteq (\text{supp}(\mu))^c$.
(iii) If X is Hausdorff and μ is regular, prove that $\mu\big((\text{supp}(\mu))^c\big) = 0$, and that $(\text{supp}(\mu))^c$ is the largest open set which is μ-null.

1.5.5. If f is the Cantor function, prove that the support (exercise 1.5.4) of μ_f is the Cantor set C.

1.5.6. Let $F : \mathbb{R} \to \mathbb{R}$ be any increasing function. Prove that the complement of the support (exercise 1.5.4) of the measure μ_F is the union of all *open* intervals on each of which F is constant.

1.5.7. Let $a : \mathbb{R} \to [0, +\infty]$ induce the point-mass distribution μ on $(\mathbb{R}, \mathcal{P}(\mathbb{R}))$. Then μ is a Borel measure on $\mathbb{R}$.
(i) Prove that $\mu(K) < +\infty$ for every compact $K \subseteq \mathbb{R}$ if and only if $\sum_{-R\leq x\leq R} a_x < +\infty$ for every $R > 0$.
(ii) In particular, prove that, if $\mu(K) < +\infty$ for every compact $K \subseteq \mathbb{R}$, then $\{x \in \mathbb{R} \,|\, a_x > 0\}$ is countable.
(iii) If $\mu(K) < +\infty$ for every compact $K \subseteq \mathbb{R}$, find (in terms of the function a) an increasing, continuous from the right $F : \mathbb{R} \to \mathbb{R}$ so that $\mu = \mu_F$ on $\mathcal{B}_1$. Describe the sets E such that $\mu_F^*(E) = 0$ and find the σ-algebra $\mathcal{S}_F$ of all μ_F^*-measurable sets. Is $\mathcal{S}_F = \mathcal{P}(\mathbb{R})$?

1.5.8. Let μ be a σ-finite regular Borel measure on the Hausdorff topological space X and Y be a Borel subset of X. Prove that both restrictions, $\mu\rceil Y$ and μ_Y, are regular Borel measures.

1.5.9. Let μ be a regular Borel measure on the Hausdorff topological space X so that $\mu(\{x\}) = 0$ for all $x \in X$. A measure satisfying this last property is called **continuous.** Prove that for every Borel set A with $0 < \mu(A) < +\infty$ and every $t \in (0, \mu(A))$ there is some Borel set B so that $B \subseteq A$ and $\mu(B) = t$.

1.5.10. Let X be a separable and complete metric space and let μ be a Borel measure on X such that $\mu(X) = 1$. Prove that there is a B, which is a countable union of compact sets, so that $\mu(B) = 1$.

1.5.11. Let $\mathcal{T} = \{\emptyset, X\}$ be the trivial topology on the non-empty set X. Prove that every subset of X is compact, while the only Borel sets in X are $\emptyset$ and X.

1.5.12. Let X be a Hausdorff topological space and μ be a measure on $(X, \mathcal{B}_X)$ which satisfy the assumptions of Theorem 1.2. Let Y be an *open* or *closed* subset of X with its subspace topology and let $\mu\rceil Y$ be the restriction of μ on $(Y, \mathcal{B}_Y)$. Prove that Y and $\mu\rceil Y$ also satisfy the assumptions of Theorem 1.2.

METRIC OUTER MEASURES.

Let (X, d) be a metric space. As usual, we denote $B(x; r)$ the open ball in X with center $x \in X$ and radius $r > 0$, i.e.

$$B(x; r) = \{y \in X \,|\, d(y, x) < r\}.$$

We recall that, if E, F are non-empty subsets of X, the quantity

$$d(E, F) = \inf\{d(x, y) \,|\, x \in E, y \in F\}$$

is the **distance** between E and F.

Definition. *Let (X, d) be a metric space and μ^* be an outer measure on X. We say that μ^* is a* ***metric outer measure*** *if*

$$\mu^*(E \cup F) = \mu^*(E) + \mu^*(F)$$

for every non-empty sets $E, F \subseteq X$ with $d(E, F) > 0$.

Proposition 1.49. *Let (X, d) be a metric space and μ^* be an outer measure on X. Then, the measure μ which is induced by μ^* on $(X, \mathcal{S}_{\mu^*})$ is a Borel measure (i.e. all Borel subsets of X are μ^*-measurable) if and only if μ^* is a metric outer measure.*

Proof. We assume that all Borel sets are μ^*-measurable, and we take arbitrary non-empty $E, F \subseteq X$ with $d(E, F) > 0$. We consider $r = d(E, F)$ and the open set $U = \bigcup_{x\in E} B(x; r)$. It is clear that $E \subseteq U$ and $F \cap U = \emptyset$. Since U is μ^*-measurable, we have

$$\mu^*(E \cup F) = \mu^*\big((E \cup F) \cap U\big) + \mu^*\big((E \cup F) \cap U^c\big) = \mu^*(E) + \mu^*(F).$$

Therefore, μ^* is a metric outer measure on X.
Now let μ^* be a metric outer measure. We consider any open set $U \subseteq X$. If A is a non-empty subset of U, we define

$$A_n = \left\{x \in A \,\middle|\, d(x,y) \geq \tfrac{1}{n} \text{ for every } y \notin U\right\}.$$

It is obvious that (A_n) is increasing. If $x \in A \subseteq U$, there is $r > 0$ so that $B(x;r) \subseteq U$. Now, if we take $n \in \mathbb{N}$ so that $\frac{1}{n} \leq r$, then $x \in A_n$. Therefore, $\bigcup_{n=1}^{+\infty} A_n = A$.
Now, we define $B_1 = A_1$ and $B_n = A_n \setminus A_{n-1}$ for all $n \geq 2$, and we have that the sets $B_1, B_2, \dots$ are pairwise disjoint and that $\bigcup_{n=1}^{+\infty} B_n = A$.
If $x \in A_n$ and $z \in B_{n+2}$, then $z \notin A_{n+1}$, and so there is $y \notin U$ so that $d(y,z) < \frac{1}{n+1}$. Then

$$d(x,z) \geq d(x,y) - d(y,z) > \tfrac{1}{n} - \tfrac{1}{n+1} = \tfrac{1}{n(n+1)}.$$

Therefore, $d(A_n, B_{n+2}) \geq \frac{1}{n(n+1)} > 0$ for every n. Since $A_{n+2} \supseteq A_n \cup B_{n+2}$, we find

$$\mu^*(A_{n+2}) \geq \mu^*(A_n \cup B_{n+2}) = \mu^*(A_n) + \mu^*(B_{n+2}).$$

By induction, we get

$$\mu^*(B_1) + \mu^*(B_3) + \cdots + \mu^*(B_{2k-1}) \leq \mu^*(A_{2k-1}),$$

$$\mu^*(B_2) + \mu^*(B_4) + \cdots + \mu^*(B_{2k}) \leq \mu^*(A_{2k}).$$

If $\sum_{k=1}^{+\infty} \mu^*(B_{2k-1}) = +\infty$ then $\lim_{k\to+\infty} \mu^*(A_{2k-1}) = +\infty$ and, if $\sum_{k=1}^{+\infty} \mu^*(B_{2k}) = +\infty$, then $\lim_{k\to+\infty} \mu^*(A_{2k}) = +\infty$. Since the sequence $(\mu^*(A_n))$ is increasing, in both cases we get $\lim_{n\to+\infty} \mu^*(A_n) = +\infty$. Since $\mu^*(A_n) \leq \mu^*(A)$ for all n, we get

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

If $\sum_{k=1}^{+\infty} \mu^*(B_{2k-1}) < +\infty$ and $\sum_{k=1}^{+\infty} \mu^*(B_{2k}) < +\infty$, then $\sum_{k=1}^{+\infty} \mu^*(B_k) < +\infty$. So for every $\epsilon > 0$ there is n so that $\sum_{k=n+1}^{+\infty} \mu^*(B_k) < \epsilon$. Now, from $A = A_n \cup \left(\bigcup_{k=n+1}^{+\infty} B_k\right)$ we get

$$\mu^*(A) \leq \mu^*(A_n) + \textstyle\sum_{k=n+1}^{+\infty} \mu^*(B_k) \leq \mu^*(A_n) + \epsilon.$$

This implies that

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

Therefore, in any case, $\lim_{n\to+\infty} \mu^*(A_n) = \mu^*(A)$ for all $A \subseteq U$.
Now, we take an arbitrary $E \subseteq X$, and we consider $A = E \cap U$ and the corresponding sets A_n. Since $E \cap U^c \subseteq U^c$, we have that

$$d(A_n, E \cap U^c) \geq d(A_n, U^c) \geq \tfrac{1}{n} > 0$$

by the definition of A_n. Therefore,

$$\mu^*(E) \geq \mu^*\big(A_n \cup (E \cap U^c)\big) = \mu^*(A_n) + \mu^*(E \cap U^c)$$

for all n. Taking the limit as $n \to +\infty$, we find

$$\mu^*(E) \geq \mu^*(E \cap U) + \mu^*(E \cap U^c).$$

Thus, every open set U is μ^*-measurable, and so every Borel set is μ^*-measurable. $\square$

HAUSDORFF MEASURE, HAUSDORFF DIMENSION.

Let (X, d) be a metric space. The *diameter* of a non-empty set $E \subseteq X$ is defined by

$$\text{diam}(E) = \sup\{d(x, y) \mid x, y \in E\}$$

and the diameter of $\emptyset$ is defined by $\text{diam}(\emptyset) = 0$.

If $\text{cl}(E)$ is the closure of $E \subseteq X$, then it is easy to see that $\text{diam}(\text{cl}(E)) = \text{diam}(E)$.

We take an arbitrary $\delta > 0$, and we consider the collection $\mathcal{C}_\delta$ of all subsets of X of diameter not larger than δ. We fix α with $0 < \alpha < +\infty$, and we consider the function $\tau_{\alpha,\delta} : \mathcal{C}_\delta \to [0, +\infty]$ defined by $\tau_{\alpha,\delta}(E) = (\text{diam}(E))^\alpha$ for every $E \in \mathcal{C}_\delta$. We are now ready to apply Proposition 1.28, and for any $E \subseteq X$ we define

$$h^*_{\alpha,\delta}(E) = \inf\big\{ \textstyle\sum_{j=1}^{+\infty} (\text{diam}(E_j))^\alpha \,\big|\, \text{diam}(E_j) \le \delta \text{ for all } j \text{ and } E \subseteq \bigcup_{j=1}^{+\infty} E_j \big\}.$$

We have that $h^*_{\alpha,\delta}$ is an outer measure on X, and we further define

$$h^*_\alpha(E) = \sup_{\delta>0} h^*_{\alpha,\delta}(E), \quad E \subseteq X.$$

We observe that, if $0 < \delta_1 < \delta_2$, then the set whose infimum is $h^*_{\alpha,\delta_1}(E)$ is included in the set whose infimum is $h^*_{\alpha,\delta_2}(E)$. Therefore, $h^*_{\alpha,\delta_2}(E) \le h^*_{\alpha,\delta_1}(E)$, and so

$$h^*_\alpha(E) = \lim_{\delta\to 0+} h^*_{\alpha,\delta}(E), \quad E \subseteq X.$$

Proposition 1.50. *Let (X, d) be a metric space and $0 < \alpha < +\infty$. Then h^*_α is a metric outer measure on X.*

Proof. We have $h^*_\alpha(\emptyset) = \sup_{\delta>0} h^*_{\alpha,\delta}(\emptyset) = 0$, since $h^*_{\alpha,\delta}$ is an outer measure for every $\delta > 0$.
If $E \subseteq F \subseteq X$, then for every $\delta > 0$ we have

$$h^*_{\alpha,\delta}(E) \le h^*_{\alpha,\delta}(F) \le h^*_\alpha(F).$$

Taking the supremum of the left side, we find $h^*_\alpha(E) \le h^*_\alpha(F)$.
If $E = \bigcup_{j=1}^{+\infty} E_j$, then for every $\delta > 0$ we have

$$h^*_{\alpha,\delta}(E) \le \textstyle\sum_{j=1}^{+\infty} h^*_{\alpha,\delta}(E_j) \le \sum_{j=1}^{+\infty} h^*_\alpha(E_j),$$

and, taking the supremum of the left side, we find $h^*_\alpha(E) \le \sum_{j=1}^{+\infty} h^*_\alpha(E_j)$.
Therefore, h^*_α is an outer measure on X.
Now, we consider any $E, F \subseteq X$ with $d(E, F) > 0$.
If $h^*_\alpha(E\cup F) = +\infty$, then $h^*_\alpha(E\cup F) \le h^*_\alpha(E) + h^*_\alpha(F)$ implies $h^*_\alpha(E\cup F) = h^*_\alpha(E) + h^*_\alpha(F)$.
Now, we assume that $h^*_\alpha(E \cup F) < +\infty$, and so $h^*_{\alpha,\delta}(E \cup F) < +\infty$ for every $\delta > 0$. We take arbitrary δ so that

$$0 < \delta < d(E, F)$$

and an arbitrary covering

$$E \cup F \subseteq \textstyle\bigcup_{j=1}^{+\infty} A_j$$

with $\text{diam}(A_j) \le \delta$ for every j. It is obvious that each A_j intersects *at most one* of the E and F. We define $B_j = A_j$, if A_j intersects E, and $B_j = \emptyset$, otherwise. Similarly, we define $C_j = A_j$, if A_j intersects F, and $C_j = \emptyset$, otherwise. Then

$$E \subseteq \textstyle\bigcup_{j=1}^{+\infty} B_j, \quad F \subseteq \bigcup_{j=1}^{+\infty} C_j,$$

and so

$$h^*_{\alpha,\delta}(E) \le \textstyle\sum_{j=1}^{+\infty} (\text{diam}(B_j))^\alpha, \quad h^*_{\alpha,\delta}(F) \le \sum_{j=1}^{+\infty} (\text{diam}(C_j))^\alpha.$$

Adding, we find

$$h^*_{\alpha,\delta}(E) + h^*_{\alpha,\delta}(F) \le \textstyle\sum_{j=1}^{+\infty} (\text{diam}(A_j))^\alpha,$$

and, taking the infimum of the right side, $h^*_{\alpha,\delta}(E)+h^*_{\alpha,\delta}(F) \le h^*_{\alpha,\delta}(E\cup F)$. Now, taking the limit as $\delta \to 0+$, we get $h^*_\alpha(E) + h^*_\alpha(F) \le h^*_\alpha(E \cup F)$. Finally, since $h^*_\alpha(E \cup F) \le h^*_\alpha(E) + h^*_\alpha(F)$, we conclude that $h^*_\alpha(E \cup F) = h^*_\alpha(E) + h^*_\alpha(F)$. $\square$

Definition. *Let (X, d) be a metric space and $0 < \alpha < +\infty$. We call h^*_α the **α-dimensional Hausdorff outer measure** on X, and the measure h_α on $(X, \mathcal{S}_{h^*_\alpha})$ is called the **α-dimensional Hausdorff measure** on X.*

Proposition 1.51. *If (X, d) is a metric space and $0 < \alpha < +\infty$, then h_α is a Borel measure on X. Namely, $\mathcal{B}_X \subseteq \mathcal{S}_{h^*_\alpha}$.*

Proof. Immediate, by Proposition 1.49 and 1.50. $\square$

Lemma 1.4. *Let (X, d) be a metric space, E be a Borel set in X, and $0 < \alpha_1 < \alpha_2 < +\infty$. If $h_{\alpha_1}(E) < +\infty$, then $h_{\alpha_2}(E) = 0$.*

Proof. Since $h^*_{\alpha_1}(E) = h_{\alpha_1}(E) < +\infty$, we have that $h^*_{\alpha_1,\delta}(E) < +\infty$ for every $\delta > 0$. We fix such a $\delta > 0$, and we consider a covering $E \subseteq \bigcup_{j=1}^{+\infty} A_j$ by subsets of X with $\mathrm{diam}(A_j) \leq \delta$ for all j so that

$$\textstyle\sum_{j=1}^{+\infty} (\mathrm{diam}(A_j))^{\alpha_1} < h^*_{\alpha_1,\delta}(E) + 1 \leq h^*_{\alpha_1}(E) + 1.$$

Then

$$h^*_{\alpha_2,\delta}(E) \leq \textstyle\sum_{j=1}^{+\infty} (\mathrm{diam}(A_j))^{\alpha_2} \leq \delta^{\alpha_2-\alpha_1} \sum_{j=1}^{+\infty} (\mathrm{diam}(A_j))^{\alpha_1} \leq (h^*_{\alpha_1}(E) + 1)\delta^{\alpha_2-\alpha_1}.$$

Taking the limit as $\delta \to 0+$, we find $h^*_{\alpha_2}(E) = 0$, and so $h_{\alpha_2}(E) = 0$. $\square$

Proposition 1.52. *Let (X, d) be a metric space and E be a Borel set in X. Then there is an $\alpha_0 \in [0, +\infty]$ (depending on E) such that: $h_\alpha(E) = +\infty$ for every $\alpha \in (0, \alpha_0)$, and $h_\alpha(E) = 0$ for every $\alpha \in (\alpha_0, +\infty)$.*

Proof. We consider various cases.
If $h_\alpha(E) = 0$ for every $\alpha > 0$, then it is enough to define $\alpha_0 = 0$.
If $h_\alpha(E) = +\infty$ for every $\alpha > 0$, then it is enough to define $\alpha_0 = +\infty$.
Otherwise, there are α_1 and α_2 in $(0, +\infty)$ so that $0 < h_{\alpha_1}(E)$ and $h_{\alpha_2}(E) < +\infty$. In this case, Lemma 1.4 implies $\alpha_1 \leq \alpha_2$, and $h_\alpha(E) = +\infty$ for every $\alpha \in (0, \alpha_1)$, and $h_\alpha(E) = 0$ for every $\alpha \in (\alpha_2, +\infty)$. Now, we consider

$$\alpha_0 = \sup\{\alpha \in (0, +\infty) \,|\, h_\alpha(E) = +\infty\}.$$

Then $\alpha_0 \in [\alpha_1, \alpha_2]$. Again, Lemma 1.4 implies $h_\alpha(E) = +\infty$ for every $\alpha \in (0, \alpha_0)$, and $h_\alpha(E) = 0$ for every $\alpha \in (\alpha_0, +\infty)$. $\square$

Definition. *If E is any Borel subset of a metric space (X, d), the a_0 of Proposition 1.52 is called the **Hausdorff dimension** of E, and it is denoted*

$$\mathrm{dim}_h(E).$$

In other words, we have $0 \leq \mathrm{dim}_h(E) \leq +\infty$, and $h_\alpha(E) = +\infty$ for $0 \leq \alpha < \mathrm{dim}_h(E)$, and $h_\alpha(E) = 0$ for $\alpha > \mathrm{dim}_h(E)$. If $\alpha = \mathrm{dim}_h(E)$, then $h_\alpha(E)$ can take any value in $[0, +\infty]$.

Proposition 1.53. *For every Borel set E in the Euclidean space $\mathbb{R}^n$ we have $\mathrm{dim}_h(E) \leq n$. Moreover, there is a positive constant c_n, depending only on n, so that $h_n(E) = c_n m_n(E)$. Therefore, if $m_n(E) > 0$, then $\mathrm{dim}_h(E) = n$.*

Proof. Consider an arbitrary $\alpha > n$ and any bounded Borel set E. We take a closed cube Q large enough so that $E \subseteq Q$. By subdividing each of the edges of Q into N intervals of the same length we can subdivide Q into N^n closed cubes Q_j, $j = 1, \ldots, N^n$, of the same Lebesgue measure. If the side length of Q is l, then the diameter of Q is $\sqrt{n}\, l$ and the diameter of each Q_j is $\frac{\sqrt{n}\, l}{N}$. Now, E is covered by the union of all Q_j, and so

$$h^*_{\alpha,\sqrt{n}\,l/N}(E) \leq \textstyle\sum_{j=1}^{N^n} \left(\frac{\sqrt{n}\,l}{N}\right)^\alpha = \frac{(\sqrt{n}\,l)^\alpha}{N^{\alpha-n}}. \tag{1.16}$$

Therefore,

$$h_\alpha(E) = \lim_{N\to+\infty} h^*_{\alpha,\sqrt{n}\,l/N}(E) = 0.$$

Now, if E is not bounded, we can write it as $E = \bigcup_{k=1}^{+\infty} E_k$, where all E_k are bounded Borel sets. Hence, $h_\alpha(E) = 0$ again.
Since $h_\alpha(E) = 0$ for all $\alpha > n$, we get that $\dim_h(E) \le n$.
Now, we consider the closed cube

$$Q_0 = [0,1] \times \cdots \times [0,1] = [0,1]^n.$$

Let $\delta > 0$, and let $Q_0 \subseteq \bigcup_{j=1}^{+\infty} E_j$ with $\mathrm{diam}(E_j) \le \delta$ for all j. Then each E_j is contained in a closed ball B_j of radius $\mathrm{diam}(E_j)$. Also, the closed ball B_j is contained in a closed cube Q_j of side-length $2\,\mathrm{diam}(E_j)$. Therefore, $Q_0 \subseteq \bigcup_{j=1}^{+\infty} Q_j$, and so

$$1 = m_n(Q_0) \le \sum_{j=1}^{+\infty} m_n(Q_j) = \sum_{j=1}^{+\infty} (2\,\mathrm{diam}(E_j))^n.$$

Thus, $\frac{1}{2^n} \le \sum_{j=1}^{+\infty} (\mathrm{diam}(E_j))^n$. Taking the infimum of the left side, we get

$$\tfrac{1}{2^n} \le h^*_{n,\delta}(Q_0) \le h^*_n(Q_0) = h_n(Q_0).$$

On the other hand, we may repeat the argument at the beginning of this proof with the closed cube $Q = Q_0$, which has side length $l = 1$, and with $\alpha = n$. Then (1.16) becomes

$$h^*_{n,\sqrt{n}/N}(Q_0) \le \sum_{j=1}^{N^n} \left(\tfrac{\sqrt{n}}{N}\right)^n = n^{n/2}.$$

Finally,

$$h_n(Q_0) = \lim_{N\to+\infty} h^*_{n,\sqrt{n}/N}(Q_0) \le n^{n/2}.$$

We conclude that $0 < h_n(Q_0) < +\infty$.
Now it is easy to show (exactly as with the Lebesgue measure m_n) that for all Borel sets A, all $z \in \mathbb{R}^n$ and all $\lambda > 0$ we have

$$h_n(A + z) = h_n(A), \quad h_n(\lambda A) = \lambda^n h_n(A).$$

This implies that $h_n(Q) = l^n h_n(Q_0)$ for every closed cube Q, where l is the side length of Q. But we also have that $m_n(Q) = l^n m_n(Q_0)$, and so

$$h_n(Q) = c_n m_n(Q)$$

for every closed cube Q, where $c_n = \frac{h_n(Q_0)}{m_n(Q_0)}$. We may easily extend this result, i.e. $h_n(S) = c_n m_n(S)$, to all bounded intervals S with rational vertices (indeed, such an interval can be decomposed into pairwise disjoint cubes), and then to all bounded intervals. Now, Proposition 1.35 implies that the Borel measures h_n and $c_n m_n$ are equal. □

Example. We shall calculate the Hausdorff dimension of the Cantor set $C \subseteq [0,1]$.
From Proposition 1.53 we know that $0 \le \dim_h(C) \le 1$.
We consider the sets I_k which are involved in the construction of $C = \bigcap_{k=1}^{+\infty} I_k$. Each I_k consists of 2^k closed intervals of length $\frac{1}{3^k}$ and, since $C \subseteq I_k$, we get

$$h^*_{\alpha,1/3^k}(C) \le 2^k \left(\tfrac{1}{3^k}\right)^\alpha = \left(\tfrac{2}{3^\alpha}\right)^k.$$

If $\alpha > \frac{\log 2}{\log 3}$, then $\frac{2}{3^\alpha} < 1$, and we get

$$h_\alpha(C) = \lim_{k\to+\infty} h^*_{\alpha,1/3^k}(C) = 0 \quad \text{for } \alpha > \tfrac{\log 2}{\log 3}.$$

Therefore, $\dim_h(C) \leq \frac{\log 2}{\log 3}$.
Now, we consider $\alpha = \frac{\log 2}{\log 3}$, and we take any δ with $0 < \delta < \frac{1}{3}$, and any covering $C \subseteq \bigcup_{j=1}^{+\infty} E_j$ with $\mathrm{diam}(E_j) \leq \delta$ for all j. Considering the closure $\mathrm{cl}(E_j)$ of each E_j, we have that $C \subseteq \bigcup_{j=1}^{+\infty} \mathrm{cl}(E_j)$ and $\mathrm{diam}(\mathrm{cl}(E_j)) = \mathrm{diam}(E_j) \leq \delta$. Therefore, without loss of generality, we may assume that every E_j is a Borel set.
We also consider the Cantor measure μ_f, where f is the Cantor function.
Now, assume that $E_j \cap C \neq \emptyset$ and $\mathrm{diam}(E_j) > 0$. Then there is exactly one $k \in \mathbb{N}$ so that

$$\tfrac{1}{3^{k+1}} \leq \mathrm{diam}(E_j) < \tfrac{1}{3^k}.$$

Then $E_j \cap I_k \neq \emptyset$ and, since $\mathrm{diam}(E_j) < \frac{1}{3^k}$, we have that E_j intersects *exactly one*, say I, of the 2^k subintervals of I_k. Hence,

$$\mu_f(E_j) = \mu_f(E_j \cap I) \leq \mu_f(I) = \tfrac{1}{2^k} = \tfrac{2}{2^{k+1}} = \tfrac{2}{3^{(k+1)\alpha}} \leq 2(\mathrm{diam}(E_j))^\alpha.$$

Next, assume that $E_j \cap C \neq \emptyset$ and $\mathrm{diam}(E_j) = 0$. Then E_j contains only one point, and so

$$\mu_f(E_j) = 0 = 2(\mathrm{diam}(E_j))^\alpha.$$

Finally, if $E_j \cap C = \emptyset$, then

$$\mu_f(E_j) = 0 \leq 2(\mathrm{diam}(E_j))^\alpha.$$

In any case we have $\mu_f(E_j) \leq 2(\mathrm{diam}(E_j))^\alpha$ for all j. Therefore,

$$1 = \mu_f(C) \leq \textstyle\sum_{j=1}^{+\infty} \mu_f(E_j) \leq 2 \sum_{j=1}^{+\infty} (\mathrm{diam}(E_j))^\alpha.$$

Taking the infimum of the right side, we get that $\frac{1}{2} \leq h^*_{\alpha,\delta}(C)$ for all δ with $0 < \delta < \frac{1}{3}$, and so

$$\tfrac{1}{2} \leq \lim_{\delta \to 0+} h^*_{\alpha,\delta}(C) = h_\alpha(C) \quad \text{for } \alpha = \tfrac{\log 2}{\log 3}.$$

Hence, $\dim_h(C) \geq \frac{\log 2}{\log 3}$, and we conclude that $\dim_h(C) = \frac{\log 2}{\log 3}$.

Exercises.

1.5.13. Let K be the set constructed in part (a) of exercise 1.4.14 using $\epsilon_k = \epsilon$ for all k, where $0 < \epsilon < \frac{1}{2}$. Prove that $\dim_h(K) = (\log 2)/(\log \frac{2}{1-2\epsilon})$. Thus, by varying ϵ in the interval $(0, \frac{1}{2})$ we get Borel sets in $\mathbb{R}$ whose Hausdorff dimensions cover the whole range between 0 and 1.
Find a Borel set K in $\mathbb{R}$ with $\dim_h(K) = 0$.

Chapter 2

Measurable functions.

2.1 Measurability.

Definition. *Let $(X, \mathcal{S}_X)$ and $(Y, \mathcal{S}_Y)$ be measurable spaces and $f : X \to Y$. We say that f is $(\mathcal{S}_X, \mathcal{S}_Y)$-**measurable** if $f^{-1}(E) \in \mathcal{S}_X$ for all $E \in \mathcal{S}_Y$.*
*If the second space $(Y, \mathcal{S}_Y)$ is $\mathbb{R}$ or $\mathbb{C}$ or $\overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ or $\mathbb{R}^n$ with the corresponding σ-algebra of Borel sets, then we just say that f is $\mathcal{S}_X$-**measurable**. If, moreover, the first space $(X, \mathcal{S}_X)$ is $\mathbb{R}^n$ with the σ-algebra of Borel sets or the σ-algebra of Lebesgue sets, then we just say that f is **Borel measurable** or **Lebesgue measurable**, respectively. And, if $(X, \mathcal{S}_X)$ is a topological space with the σ-algebra of Borel sets, then we just say that f is **Borel measurable**.*

In the general case and if there is no danger of confusion, we may just say that f is *measurable*.

If $f : X \to \mathbb{R}$, then it is also true that $f : X \to \overline{\mathbb{R}}$. Thus, according to the definition we have given, there might be a conflict between the two meanings of $\mathcal{S}_X$-measurability of f. But, actually, there is no such conflict. Indeed, suppose that f is $(\mathcal{S}_X, \mathcal{B}_1)$-measurable. If $E \in \overline{\mathcal{B}}_1$, then $E \cap \mathbb{R} \in \mathcal{B}_1$, and so $f^{-1}(E) = f^{-1}(E \cap \mathbb{R}) \in \mathcal{S}_X$. Hence, f is $(\mathcal{S}_X, \overline{\mathcal{B}}_1)$-measurable. Conversely, let f be $(\mathcal{S}_X, \overline{\mathcal{B}}_1)$-measurable. If $E \in \mathcal{B}_1$, then $E \in \overline{\mathcal{B}}_1$, and so $f^{-1}(E) \in \mathcal{S}_X$. Hence, f is $(\mathcal{S}_X, \mathcal{B}_1)$-measurable.

The same question arises when $f : X \to \mathbb{C}$, since it is then also true that $f : X \to \overline{\mathbb{C}}$. Exactly as before we may prove that f is $(\mathcal{S}_X, \mathcal{B}_2)$-measurable if and only if it is $(\mathcal{S}_X, \overline{\mathcal{B}}_2)$-measurable, and so there is no conflict in the meaning of $\mathcal{S}_X$-measurability of f.

Example. Any constant function is measurable.
Indeed, let $(X, \mathcal{S}_X)$ and $(Y, \mathcal{S}_Y)$ be measurable spaces and $f(x) = y_0 \in Y$ for all $x \in X$. We take an arbitrary $E \in \mathcal{S}_Y$. If $y_0 \in E$, then $f^{-1}(E) = X \in \mathcal{S}_X$. If $y_0 \notin E$, then $f^{-1}(E) = \emptyset \in \mathcal{S}_X$.

Proposition 2.1. *Let $(X, \mathcal{S}_X)$ and $(Y, \mathcal{S}_Y)$ be measurable spaces and $f : X \to Y$. Suppose that $\mathcal{C}_Y$ is a collection of subsets of Y so that $\mathcal{S}(\mathcal{C}_Y) = \mathcal{S}_Y$. If $f^{-1}(E) \in \mathcal{S}_X$ for all $E \in \mathcal{C}_Y$, then f is $(\mathcal{S}_X, \mathcal{S}_Y)$-measurable.*

Proof. We consider the collection

$$\mathcal{S}'_Y = \{E \subseteq Y \mid f^{-1}(E) \in \mathcal{S}_X\}$$

of subsets of Y. (According to exercise 1.1.2, this is the push-forward of $\mathcal{S}_X$.)
Since $f^{-1}(\emptyset) = \emptyset \in \mathcal{S}_X$, we get that $\emptyset \in \mathcal{S}'_Y$.
Let $E \in \mathcal{S}'_Y$. Then $f^{-1}(E) \in \mathcal{S}_X$, and so $f^{-1}(E^c) = (f^{-1}(E))^c \in \mathcal{S}_X$. Hence, $E^c \in \mathcal{S}'_Y$.
Let $E_j \in \mathcal{S}'_Y$ for all j. Then $f^{-1}(E_j) \in \mathcal{S}_X$ for all j, and so $f^{-1}\big(\bigcup_{j=1}^{+\infty} E_j\big) = \bigcup_{j=1}^{+\infty} f^{-1}(E_j) \in \mathcal{S}_X$. Hence, $\bigcup_{j=1}^{+\infty} E_j \in \mathcal{S}'_Y$.
Therefore, $\mathcal{S}'_Y$ is a σ-algebra of subsets of Y.
Since, by hypothesis, $\mathcal{C}_Y \subseteq \mathcal{S}'_Y$, we get that $\mathcal{S}_Y = \mathcal{S}(\mathcal{C}_Y) \subseteq \mathcal{S}'_Y$. This concludes the proof. $\square$

Proposition 2.2. *Let X, Y be topological spaces and $f : X \to Y$ be continuous on X. Then f is $(\mathcal{B}_X, \mathcal{B}_Y)$-measurable.*

Proof. Let $\mathcal{T}_Y$ be the topology of Y, i.e. the collection of all open subsets of Y. By continuity of f, for all $E \in \mathcal{T}_Y$ we have that $f^{-1}(E)$ is an open subset of X, and so $f^{-1}(E) \in \mathcal{B}_X$. Since $\mathcal{S}(\mathcal{T}_Y) = \mathcal{B}_Y$, Proposition 2.1 implies that f is $(\mathcal{B}_X, \mathcal{B}_Y)$-measurable. □

COMPOSITION.

Proposition 2.3. *Let $(X, \mathcal{S}_X)$, $(Y, \mathcal{S}_Y)$, $(Z, \mathcal{S}_Z)$ be measurable spaces and let $f : X \to Y$ and $g : Y \to Z$. If f is $(\mathcal{S}_X, \mathcal{S}_Y)$-measurable and g is $(\mathcal{S}_Y, \mathcal{S}_Z)$-measurable, then $g \circ f : X \to Z$ is $(\mathcal{S}_X, \mathcal{S}_Z)$-measurable.*

Proof. For all $E \in \mathcal{S}_Z$ we have $g^{-1}(E) \in \mathcal{S}_Y$, and so $(g \circ f)^{-1}(E) = f^{-1}\big(g^{-1}(E)\big) \in \mathcal{S}_X$. □

Hence, *composition of measurable functions is measurable*.

MEASURABILITY AND SIMPLE TRANSFORMATIONS OF $\mathbb{R}^n$.

We recall that the function $\tau_z : \mathbb{R}^n \to \mathbb{R}^n$ given by $\tau_z(x) = x + z$ for all $x \in \mathbb{R}^n$ is called *translation* by z. The inverse of τ_z is τ_{-z} given by $\tau_{-z}(x) = x - z$. If $A \subseteq \mathbb{R}^n$, then

$$\tau_z : A \to \tau_z(A), \quad \tau_{-z} : \tau_z(A) \to A.$$

There is a corresponding *translation* by z of a function $f : A \to Y$, where $A \subseteq \mathbb{R}^n$. This is the function

$$\tau_z(f) = f \circ \tau_{-z} : \tau_z(A) \to Y$$

given by

$$\tau_z(f)(x) = f(\tau_{-z}(x)) = f(x - z), \quad x \in \tau_z(A) = A + z.$$

We note that *the domain of definition of the translation of f by z is the translation of the domain of definition of f by z.*

Proposition 2.4 says that *the translation of a measurable function is a measurable function*.

Proposition 2.4. *Let $(Y, \mathcal{S}_Y)$ be a measure space and $A \in \mathcal{L}_n$. If $f : A \to Y$ is $(\mathcal{L}_n \rceil A, \mathcal{S}_Y)$-measurable, then $\tau_z(f) : \tau_z(A) \to Y$ is $(\mathcal{L}_n \rceil \tau_z(A), \mathcal{S}_Y)$-measurable.*

Proof. If we prove that $\tau_{-z} : \tau_z(A) \to A$ is $(\mathcal{L}_n \rceil \tau_z(A), \mathcal{L}_n \rceil A)$-measurable, then, in view of Proposition 2.3, the proof will be complete.
So let $E \in \mathcal{L}_n \rceil A$, i.e. $E \subseteq A$ and $E \in \mathcal{L}_n$. Then

$$(\tau_{-z})^{-1}(E) = \tau_z(E) \subseteq \tau_z(A), \quad (\tau_{-z})^{-1}(E) = \tau_z(E) \in \mathcal{L}_n,$$

where the second relation is implied by Proposition 1.37. Hence $(\tau_{-z})^{-1}(E) \in \mathcal{L}_n \rceil \tau_z(A)$. □

Now we consider any linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ with $\det(T) \neq 0$, so that the inverse linear transformation $T^{-1} : \mathbb{R}^n \to \mathbb{R}^n$ is also defined. If $A \subseteq \mathbb{R}^n$, then

$$T : A \to T(A), \quad T^{-1} : T(A) \to A.$$

There is a corresponding linear transformation of a function $f : A \to Y$, where $A \subseteq \mathbb{R}^n$. This is the function

$$T(f) = f \circ T^{-1} : T(A) \to Y$$

given by

$$T(f)(x) = f(T^{-1}(x)), \quad x \in T(A).$$

Proposition 2.5 says that *the linear transformation of a measurable function is a measurable function*.

Proposition 2.5. *Let $(Y, \mathcal{S}_Y)$ be a measure space and $A \in \mathcal{L}_n$. If $f : A \to Y$ is $(\mathcal{L}_n \rceil A, \mathcal{S}_Y)$-measurable, then $T(f) : T(A) \to Y$ is $(\mathcal{L}_n \rceil T(A), \mathcal{S}_Y)$-measurable.*

Proof. As in the proof of Proposition 2.4, if we prove that $T^{-1} : T(A) \to A$ is $(\mathcal{L}_n \rceil T(A), \mathcal{L}_n \rceil A)$-measurable, then Proposition 2.3 will conclude the proof.
So let $E \in \mathcal{L}_n \rceil A$, i.e. $E \subseteq A$ and $E \in \mathcal{L}_n$. Then

$$(T^{-1})^{-1}(E) = T(E) \subseteq T(A), \quad (T^{-1})^{-1}(E) = T(E) \in \mathcal{L}_n,$$

where the second relation is implied by Proposition 1.38. Hence $(T^{-1})^{-1}(E) \in \mathcal{L}_n \rceil T(A)$. $\square$

As a special case of an invertible linear transformation we consider the function $l_\lambda : \mathbb{R}^n \to \mathbb{R}^n$, i.e. the *dilation* by $\lambda > 0$, given by $l_\lambda(x) = \lambda x$ for all $x \in \mathbb{R}^n$. The inverse of l_λ is $l_{1/\lambda}$.

The corresponding transformation of the function $f : A \to Y$, where $A \subseteq \mathbb{R}^n$, is the function $l_\lambda(f) = f \circ l_{1/\lambda} : l_\lambda(A) \to Y$ given by

$$l_\lambda(f)(x) = f(l_{1/\lambda}(x)) = f\big(\tfrac{x}{\lambda}\big), \quad x \in l_\lambda(A) = \lambda A.$$

The function $l_\lambda(f)$ is called *dilation* of f by λ.

Another special case of an invertible linear transformation is the function $r : \mathbb{R}^n \to \mathbb{R}^n$, i.e. the *reflection*, given by $r(x) = -x$ for all $x \in \mathbb{R}^n$. The inverse of r is itself.

The corresponding *reflection* of the function $f : A \to Y$, where $A \subseteq \mathbb{R}^n$, is the function $r(f) = f \circ r : r(A) \to Y$ given by

$$r(f)(x) = f(r(x)) = f(-x), \quad x \in r(A) = -A.$$

RESTRICTION AND GLUING.

If $f : X \to Y$ and $A \subseteq X$ is non-empty, then the function $f \rceil A : A \to Y$, defined by

$$(f \rceil A)(x) = f(x) \quad \text{for all } x \in A,$$

is the usual **restriction** of f on A.

Proposition 2.6. *Let $(X, \mathcal{S}_X)$, $(Y, \mathcal{S}_Y)$ be measurable spaces and $f : X \to Y$ be $(\mathcal{S}_X, \mathcal{S}_Y)$-measurable. If $A \in \mathcal{S}_X$ is non-empty, then $f \rceil A$ is $(\mathcal{S}_X \rceil A, \mathcal{S}_Y)$-measurable.*

Proof. Let $E \in \mathcal{S}_Y$. Then

$$\begin{aligned}(f \rceil A)^{-1}(E) &= \{x \in A \,|\, (f \rceil A)(x) \in E\} = \{x \in A \,|\, f(x) \in E\} = \{x \in X \,|\, f(x) \in E\} \cap A \\ &= f^{-1}(E) \cap A.\end{aligned}$$

Now, since $f^{-1}(E) \in \mathcal{S}_X$, we get that $(f \rceil A)^{-1}(E) \in \mathcal{S}_X \rceil A$. $\square$

We may say that *measurability of a function on the whole space implies its measurability on every (measurable) subset of the space.*

Proposition 2.7. *Let $(X, \mathcal{S}_X)$, $(Y, \mathcal{S}_Y)$ be measurable spaces and $f : X \to Y$. Let the (finitely many or infinitely many) non-empty $A_1, A_2, \ldots \in \mathcal{S}_X$ be pairwise disjoint and $A_1 \cup A_2 \cup \cdots = X$. If $f \rceil A_j$ is $(\mathcal{S}_X \rceil A_j, \mathcal{S}_Y)$-measurable for all j, then f is $(\mathcal{S}_X, \mathcal{S}_Y)$-measurable.*

Proof. Let $E \in \mathcal{S}_Y$. Then $f^{-1}(E) \cap A_j = (f \rceil A_j)^{-1}(E) \in \mathcal{S}_X \rceil A_j$ for all j. This implies that $f^{-1}(E) \cap A_j \in \mathcal{S}_X$ for all j, and so $f^{-1}(E) = (f^{-1}(E) \cap A_1) \cup (f^{-1}(E) \cap A_2) \cup \cdots \in \mathcal{S}_X$. Therefore, f is $(\mathcal{S}_X, \mathcal{S}_Y)$-measurable. $\square$

Thus, *measurability of a function separately on complementary (measurable) subsets of the space implies its measurability on the whole space*.

There are two operations on measurable functions that are taken care of by Propositions 2.6 and 2.7. One is the restriction of a function $f : X \to Y$ on some non-empty $A \subseteq X$ and the other is the **gluing** of functions $f\rceil A_j : A_j \to Y$ to form a single $f : X \to Y$, whenever the countably many A_j are non-empty, pairwise disjoint and cover X. The rules are: *restriction of a measurable function on a measurable set is measurable*, and *gluing of measurable functions defined on measurable subsets results to a measurable function*.

Example. Let X, Y be topological spaces, $f : X \to Y$, and $A_1, A_2, \ldots \in \mathcal{B}_X$ be pairwise disjoint and $A_1 \cup A_2 \cup \cdots = X$. Let also every $f\rceil A_j : A_j \to Y$ be continuous on A_j.
By Proposition 2.2, each $f\rceil A_j : A_j \to Y$ is $(\mathcal{B}_{A_j}, \mathcal{S}_Y)$-measurable. Since $\mathcal{B}_{A_j} = \mathcal{B}_X\rceil A_j$, we have that each $f\rceil A_j : A_j \to Y$ is $(\mathcal{B}_X\rceil A_j, \mathcal{S}_Y)$-measurable. Therefore, f is $(\mathcal{B}_X, \mathcal{B}_Y)$-measurable.
Loosely speaking, *if a function is piecewise continuous, then it is Borel measurable*.

FUNCTIONS WITH ARITHMETICAL VALUES.

Proposition 2.8. *Let $(X, \mathcal{S})$ be a measurable space and $f : X \to \mathbb{R}^n$. Let, for each $j = 1, \ldots, n$, $f_j : X \to \mathbb{R}$ denote the j-th component function of f. Namely, $f(x) = (f_1(x), \ldots, f_n(x))$ for all $x \in X$. Then f is $\mathcal{S}$-measurable if and only if every f_j is $\mathcal{S}$-measurable.*

Proof. Let f be $\mathcal{S}$-measurable.
Let I be any interval in $\mathbb{R}$. We consider the interval $S = \mathbb{R} \times \cdots \times \mathbb{R} \times I \times \mathbb{R} \times \cdots \times \mathbb{R}$ in $\mathbb{R}^n$, where I is its j-th factor. Then

$$f_j^{-1}(I) = \{x \in X \mid f_j(x) \in I\} = \{x \in X \mid f(x) \in S\} = f^{-1}(S).$$

Since $S \in \mathcal{B}_n$, we get $f^{-1}(S) \in \mathcal{S}$ and so $f_j^{-1}(I) \in \mathcal{S}$. Since the collection of all I generates $\mathcal{B}_1$, Proposition 2.1 implies that f_j is $\mathcal{S}$-measurable.
Now let every f_j be $\mathcal{S}$-measurable.
Let $S = I_1 \times \cdots \times I_n$ be any interval in $\mathbb{R}^n$. Now

$$f^{-1}(S) = \{x \in X \mid f(x) \in S\} = \textstyle\bigcap_{j=1}^n \{x \in X \mid f_j(x) \in I_j\} = \bigcap_{j=1}^n f_j^{-1}(I_j).$$

Since $f_j^{-1}(I_j) \in \mathcal{S}$ for all j, we get $f^{-1}(S) \in \mathcal{S}$. The collection of all intervals S generates $\mathcal{B}_n$, and Proposition 2.1, again, implies that f is $\mathcal{S}$- measurable. $\square$

Loosely speaking, *measurability of a vector function is equivalent to measurability of all its component functions*.

The next two results give simple criteria for measurability of real or complex valued functions.

Proposition 2.9. *Let $(X, \mathcal{S})$ be a measurable space and $f : X \to \mathbb{R}$. Then f is $\mathcal{S}$-measurable if and only if $f^{-1}((a, +\infty)) \in \mathcal{S}$ for all $a \in \mathbb{R}$.*

Proof. Since $(a, +\infty) \in \mathcal{B}_1$, one direction is trivial. The other direction is a corollary of Proposition 2.1, since, by Proposition 1.12, the collection of all intervals $(a, +\infty)$ generates $\mathcal{B}_1$. $\square$

Of course, in the statement of Proposition 2.9 one may replace the intervals $(a, +\infty)$ by the intervals $[a, +\infty)$ or $(-\infty, b)$ or $(-\infty, b]$.

If $f : X \to \mathbb{C}$, then the functions $\mathrm{Re}(f) : X \to \mathbb{R}$ and $\mathrm{Im}(f) : X \to \mathbb{R}$ are defined by

$$\mathrm{Re}(f)(x) = \mathrm{Re}(f(x)), \quad \mathrm{Im}(f)(x) = \mathrm{Im}(f(x)) \quad \text{for all } x \in X$$

and they are called the **real part** and the **imaginary part** of f, respectively.

Proposition 2.10. *Let $(X, \mathcal{S})$ be a measurable space and $f : X \to \mathbb{C}$. Then f is $\mathcal{S}$-measurable if and only if both* $\mathrm{Re}(f) : X \to \mathbb{R}$ *and* $\mathrm{Im}(f) : X \to \mathbb{R}$ *are $\mathcal{S}$-measurable.*

Proof. An immediate application of Proposition 2.8. □

The next two results investigate extended-real and extended-complex valued functions.

Proposition 2.11. *Let $(X, \mathcal{S})$ be a measurable space and $f : X \to \overline{\mathbb{R}}$. The following are equivalent.*
(i) f is $\mathcal{S}$-measurable.
(ii) $f^{-1}(\{+\infty\}) \in \mathcal{S}$, $f^{-1}(\mathbb{R}) \in \mathcal{S}$, and, if $A = f^{-1}(\mathbb{R})$ is non-empty, the function $f\rceil A : A \to \mathbb{R}$ is $\mathcal{S}\rceil A$-measurable.
(iii) $f^{-1}((a, +\infty]) \in \mathcal{S}$ for all $a \in \mathbb{R}$.

Proof. Using Proposition 2.6, we easily see that (i) implies (ii).
Now, we assume (ii). We consider the sets $B = f^{-1}(\{+\infty\})$ and $C = f^{-1}(\{-\infty\}) = (A \cup B)^c$. Both B and C belong to $\mathcal{S}$, and the restrictions $f\rceil B = +\infty$ and $f\rceil C = -\infty$ are constants, and so they are, respectively, $\mathcal{S}\rceil B$-measurable and $\mathcal{S}\rceil C$-measurable. Then Proposition 2.7 implies that f is $\mathcal{S}$-measurable, and so (ii) implies (i).
It is clear that (i) implies (iii).
Now, we assume (iii). Proposition 1.14 says that the collection of all $(a, +\infty]$ generates $\overline{\mathcal{B}}_1$. Then Proposition 2.1 implies that f is $\mathcal{S}$-measurable, and so (iii) implies (i). □

Proposition 2.12. *Let $(X, \mathcal{S})$ be a measurable space and $f : X \to \overline{\mathbb{C}}$. The following are equivalent.*
(i) f is $\mathcal{S}$-measurable.
(ii) $f^{-1}(\mathbb{C}) \in \mathcal{S}$, and, if $A = f^{-1}(\mathbb{C})$ is non-empty, the function $f\rceil A : A \to \mathbb{C}$ is $\mathcal{S}\rceil A$-measurable.

Proof. Using Proposition 2.6, we easily see that (i) implies (ii).
We assume (ii), and we consider the set $B = f^{-1}(\{\infty\}) = (f^{-1}(\mathbb{C}))^c$.
Then $B \in \mathcal{S}$, and the restriction $f\rceil B = \infty$ is constant, and so it is $\mathcal{S}\rceil B$-measurable. Then Proposition 2.7 implies that f is $\mathcal{S}$-measurable, and so (ii) implies (i). □

Exercises.

2.1.1. Let $(X, \mathcal{S})$ be a measurable space and $f : X \to \overline{\mathbb{R}}$. Prove that f is $\mathcal{S}$-measurable if $f^{-1}((a, +\infty]) \in \mathcal{S}$ for all $a \in \mathbb{Q}$.

2.1.2. Prove that every monotone $f : \mathbb{R} \to \mathbb{R}$ is Borel measurable.

2.1.3. Let $(X, \mathcal{S})$ be a measurable space and assume that the collection $\{E_\lambda\}_{\lambda \in \mathbb{R}}$ of subsets of X which belong to $\mathcal{S}$ has the properties:
(i) $E_\lambda \subseteq E_\kappa$ for all λ, κ with $\lambda \leq \kappa$,
(ii) $\bigcup_{\lambda \in \mathbb{R}} E_\lambda = X$, $\bigcap_{\lambda \in \mathbb{R}} E_\lambda = \emptyset$,
(iii) $\bigcap_{\kappa,\, \kappa > \lambda} E_\kappa = E_\lambda$ for all $\lambda \in \mathbb{R}$.
Consider the function $f : X \to \mathbb{R}$ defined by $f(x) = \inf\{\lambda \in \mathbb{R} \,|\, x \in E_\lambda\}$. Prove that f is $\mathcal{S}$-measurable and that $E_\lambda = \{x \in X \mid f(x) \leq \lambda\}$ for every $\lambda \in \mathbb{R}$.
How will the result change if we drop any of the assumptions in (ii) and (iii)?

SUM AND PRODUCT.

The next result is that *sums and products of real or complex valued measurable functions are measurable functions*.

Proposition 2.13. *Let $(X, \mathcal{S})$ be a measurable space and $f, g : X \to \mathbb{R}$ or $\mathbb{C}$ be $\mathcal{S}$-measurable. Then $f + g, fg$ are $\mathcal{S}$-measurable.*

Proof. (a) In the case $f, g : X \to \mathbb{R}$, we consider $H : X \to \mathbb{R}^2$ defined by $H(x) = (f(x), g(x))$ for all $x \in X$. Proposition 2.8 implies that H is $\mathcal{S}$-measurable. Now, we consider $\phi, \psi : \mathbb{R}^2 \to \mathbb{R}$ defined by

$$\phi(y, z) = y + z, \quad \psi(y, z) = yz.$$

Since ϕ, ψ are continuous, Proposition 2.2 implies that they are Borel measurable. Therefore, by Proposition 2.3, $\phi \circ H, \psi \circ H : X \to \mathbb{R}$ are $\mathcal{S}$-measurable. But,

$$\phi \circ H = f + g, \quad \psi \circ H = fg.$$

(b) In the case $f, g : X \to \mathbb{C}$, we consider $\mathrm{Re}(f), \mathrm{Im}(f), \mathrm{Re}(g), \mathrm{Im}(g) : X \to \mathbb{R}$, which, by Proposition 2.10, are all $\mathcal{S}$-measurable. Then part (a) implies that

$$\mathrm{Re}(f + g) = \mathrm{Re}(f) + \mathrm{Re}(g), \quad \mathrm{Im}(f + g) = \mathrm{Im}(f) + \mathrm{Im}(g),$$

$$\mathrm{Re}(fg) = \mathrm{Re}(f)\,\mathrm{Re}(g) - \mathrm{Im}(f)\,\mathrm{Im}(g), \quad \mathrm{Im}(fg) = \mathrm{Re}(f)\,\mathrm{Im}(g) + \mathrm{Im}(f)\,\mathrm{Re}(g)$$

are all $\mathcal{S}$-measurable. By Proposition 2.10 again, $f + g, fg$ are $\mathcal{S}$-measurable. $\square$

If we want to extend the previous results to functions with infinite values, we must be more careful.

The sums $(+\infty) + (-\infty), (-\infty) + (+\infty)$ are not defined in $\overline{\mathbb{R}}$ and neither is $\infty + \infty$ defined in $\overline{\mathbb{C}}$. Hence, when we add $f, g : X \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ we must agree on how to treat the summation on, respectively, the set

$$B = \{x \in X \mid f(x) = +\infty,\, g(x) = -\infty \text{ or } f(x) = -\infty,\, g(x) = +\infty\}$$

or the set

$$B = \{x \in X \mid f(x) = \infty,\, g(x) = \infty\}.$$

There are two standard ways to do this. One is to ignore the bad set and consider $f + g$ defined on $A = X \setminus B$ on which it is naturally defined. The other way is to choose some appropriate h defined on B and define $f + g = h$ on B. The usual choice for h is some constant, e.g. $h = 0$.

Proposition 2.14. *Let $(X, \mathcal{S})$ be a measurable space and $f, g : X \to \overline{\mathbb{R}}$ be $\mathcal{S}$-measurable. Then the set $B = \{x \in X \mid f(x) = +\infty,\, g(x) = -\infty$ or $f(x) = -\infty,\, g(x) = +\infty\}$ belongs to $\mathcal{S}$.*
(i) If $A = X \setminus B$, then the function $f + g : A \to \overline{\mathbb{R}}$ is $\mathcal{S}\rceil A$-measurable.
(ii) Let $h : B \to \overline{\mathbb{R}}$ be $\mathcal{S}\rceil B$-measurable. We define $(f + g)(x) = f(x) + g(x)$, if $x \in A$, and $(f + g)(x) = h(x)$, if $x \in B$. Then $f + g : X \to \overline{\mathbb{R}}$ is $\mathcal{S}$-measurable.
Similar results hold if $f, g : X \to \overline{\mathbb{C}}$ and $B = \{x \in X \mid f(x) = \infty,\, g(x) = \infty\}$.

Proof. Let $f, g : X \to \overline{\mathbb{R}}$ be $\mathcal{S}$-measurable.
We have

$$B = \big(f^{-1}(\{+\infty\}) \cap g^{-1}(\{-\infty\})\big) \cup \big(f^{-1}(\{-\infty\}) \cap g^{-1}(\{+\infty\})\big),$$

and so $B \in \mathcal{S}$.
(i) We consider the sets:

$$C = \{x \in X \mid f(x), g(x) \in \mathbb{R}\}.$$

$$D_1 = \{x \in X \mid f(x) = +\infty,\, g(x) \neq -\infty \text{ or } f(x) \neq -\infty,\, g(x) = +\infty\},$$

$$D_2 = \{x \in X \mid f(x) = -\infty,\, g(x) \neq +\infty \text{ or } f(x) \neq +\infty,\, g(x) = -\infty\},$$

It is clear that $C, D_1, D_2 \in \mathcal{S}$, that $A = C \cup D_1 \cup D_2$, and that the three sets are pairwise disjoint. The restriction of $f + g$ on C is the sum of the real valued $f\rceil C, g\rceil C$. By Proposition 2.6, both $f\rceil C, g\rceil C$ are $\mathcal{S}\rceil C$-measurable. Now, since

$$(f + g)\rceil C = (f\rceil C) + (g\rceil C),$$

Proposition 2.13 implies that $(f+g)\rceil C$ is $\mathcal{S}\rceil C$-measurable. The restriction $(f+g)\rceil D_1 = +\infty$, is $\mathcal{S}\rceil D_1$-measurable. Also, the restriction $(f+g)\rceil D_2 = -\infty$ is $\mathcal{S}\rceil D_2$-measurable. Finally, Proposition 2.7 implies that $f+g : A \to \overline{\mathbb{R}}$ is $\mathcal{S}\rceil A$-measurable.

(ii) This is immediate after the result of (i) and Proposition 2.7.

The case $f, g : X \to \overline{\mathbb{C}}$ is similar, if not simpler. □

Thus, *there is always a measurable sum of measurable functions*.

For multiplication we make the following

Convention: $(\pm\infty)\, 0 = 0\, (\pm\infty) = 0$ *in* $\overline{\mathbb{R}}$ *and* $\infty\, 0 = 0\, \infty = 0$ *in* $\overline{\mathbb{C}}$.

Thus, multiplication is always defined and we may say that *the product of measurable functions is measurable*.

Proposition 2.15. *Let* $(X, \mathcal{S})$ *be a measurable space and* $f, g : X \to \overline{\mathbb{R}}$ *or* $\overline{\mathbb{C}}$ *be* $\mathcal{S}$*-measurable. Then the function* fg *is* $\mathcal{S}$*-measurable.*

Proof. Let $f, g : X \to \overline{\mathbb{R}}$ be $\mathcal{S}$-measurable.

We consider the sets

$$\begin{aligned}
A &= \{x \in X \mid f(x), g(x) \in \mathbb{R}\},\\
C_1 &= \{x \in X \mid f(x) = +\infty,\, g(x) > 0 \text{ or } f(x) = -\infty,\, g(x) < 0\\
&\qquad \text{or } f(x) > 0,\, g(x) = +\infty \text{ or } f(x) < 0,\, g(x) = -\infty\},\\
C_2 &= \{x \in X \mid f(x) = -\infty,\, g(x) > 0 \text{ or } f(x) = +\infty,\, g(x) < 0\\
&\qquad \text{or } f(x) > 0,\, g(x) = -\infty \text{ or } f(x) < 0,\, g(x) = +\infty\},\\
D &= \{x \in X \mid f(x) = \pm\infty,\, g(x) = 0 \text{ or } f(x) = 0,\, g(x) = \pm\infty\}.
\end{aligned}$$

These four sets are pairwise disjoint, their union is X and they all belong to $\mathcal{S}$.

Now, we have

$$(fg)\rceil A = (f\rceil A)(g\rceil A).$$

By Proposition 2.6, $f\rceil A, g\rceil A$ are $\mathcal{S}\rceil A$-measurable, and then Proposition 2.13 implies that $(fg)\rceil A$ is $\mathcal{S}\rceil A$-measurable. The restriction $(fg)\rceil C_1 = +\infty$ is $\mathcal{S}\rceil C_1$-measurable. Similarly, the restriction $(fg)\rceil C_2 = -\infty$ is $\mathcal{S}\rceil C_2$-measurable. Finally, $(fg)\rceil D = 0$ is $\mathcal{S}\rceil D$-measurable.

Now, Proposition 2.7 implies that fg is $\mathcal{S}$-measurable.

If $f, g : X \to \overline{\mathbb{C}}$, the proof is similar and slightly simpler. □

ABSOLUTE VALUE AND SIGNUM.

The action of the absolute value on infinities is: $|+\infty| = |-\infty| = +\infty$ and $|\infty| = +\infty$.

Proposition 2.16. *Let* $(X, \mathcal{S})$ *be a measurable space and* $f : X \to \overline{\mathbb{R}}$ *or* $\overline{\mathbb{C}}$ *be* $\mathcal{S}$*-measurable. Then the function* $|f| : X \to [0, +\infty]$ *is* $\mathcal{S}$*-measurable.*

Proof. Let $f : X \to \overline{\mathbb{R}}$. The function $|\cdot| : \overline{\mathbb{R}} \to [0, +\infty]$ is continuous, and so it is Borel measurable. Therefore, $|f|$, the composition of $|\cdot|$ and f, is $\mathcal{S}$-measurable.

The same proof applies in the case $f : X \to \overline{\mathbb{C}}$. □

Definition. *For every* $z \in \overline{\mathbb{C}}$ *we define:* $\operatorname{sign}(z) = \frac{z}{|z|}$*, if* $z \neq 0$ *and* $z \neq \infty$*, and* $\operatorname{sign}(0) = 0$*, and* $\operatorname{sign}(\infty) = \infty$.

If we denote $\mathbb{C}^* = \overline{\mathbb{C}} \setminus \{0, \infty\}$, then the restriction $\operatorname{sign}\rceil\mathbb{C}^* : \mathbb{C}^* \to \overline{\mathbb{C}}$ is continuous. The restriction $\operatorname{sign}\rceil\{0\}$ is constant 0 and the restriction $\operatorname{sign}\rceil\{\infty\}$ is constant ∞. Now, Proposition 2.7 implies that $\operatorname{sign} : \overline{\mathbb{C}} \to \overline{\mathbb{C}}$ is Borel measurable.

All this applies in the same way to the well-known function $\operatorname{sign} : \overline{\mathbb{R}} \to \overline{\mathbb{R}}$ defined by: $\operatorname{sign}(x) = 1$, if $0 < x \leq +\infty$, and $\operatorname{sign}(x) = -1$, if $-\infty \leq x < 0$, and $\operatorname{sign}(0) = 0$. Hence,

$\text{sign} : \overline{\mathbb{R}} \to \overline{\mathbb{R}}$ is Borel measurable.
For all $z \in \overline{\mathbb{C}}$ we may write

$$z = |z| \, \text{sign}(z)$$

and this is called the **polar decomposition** of z.

Proposition 2.17. *Let $(X, \mathcal{S})$ be a measurable space and $f : X \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ be $\mathcal{S}$-measurable. Then the function* $\text{sign}(f)$*, defined by* $\text{sign}(f)(x) = \text{sign}(f(x))$ *for all $x \in X$, is $\mathcal{S}$-measurable.*

Proof. If $f : X \to \overline{\mathbb{R}}$, then $\text{sign}(f)$ is the composition of $\text{sign} : \overline{\mathbb{R}} \to \overline{\mathbb{R}}$ and f. Hence, the result is clear by Proposition 2.3. The same argument applies if $f : X \to \overline{\mathbb{C}}$. $\square$

Exercises.

2.1.4. Let $(X, \mathcal{S})$ be a measurable space and $f : X \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ be $\mathcal{S}$-measurable. We agree that $0^p = +\infty$ and $(+\infty)^p = 0$ if $p < 0$. Prove that, for all $p \in \mathbb{R}$, $p \neq 0$, the function $|f|^p$ is $\mathcal{S}$-measurable.

MAXIMUM AND MINIMUM.

Proposition 2.18. *Let $(X, \mathcal{S})$ be measurable space and $f, g : X \to \overline{\mathbb{R}}$ be $\mathcal{S}$-measurable. Then the functions* $\max\{f, g\}, \min\{f, g\} : X \to \overline{\mathbb{R}}$ *are $\mathcal{S}$-measurable.*

Proof. If $h = \max\{f, g\}$, then we have

$$\begin{aligned} h^{-1}((a, +\infty]) &= \{x \in A \,|\, a < h(x)\} = \{x \in X \,|\, a < f(x) \text{ or } a < g(x)\} \\ &= \{x \in X \,|\, a < f(x)\} \cup \{x \in X \,|\, a < g(x)\} \\ &= f^{-1}((a, +\infty]) \cup g^{-1}((a, +\infty]). \end{aligned}$$

Hence, $h^{-1}((a, +\infty]) \in \mathcal{S}$ for all $a \in \mathbb{R}$. Now, Proposition 2.11 implies that h is $\mathcal{S}$-measurable. And then we get that $\min\{f, g\} = -\max\{-f, -g\}$ is also $\mathcal{S}$-measurable. $\square$

The next result is about *comparison of measurable functions*.

Proposition 2.19. *Let $(X, \mathcal{S})$ be a measurable space and $f, g : X \to \overline{\mathbb{R}}$ be $\mathcal{S}$-measurable. Then $\{x \in X \,|\, f(x) = g(x)\} \in \mathcal{S}$ and $\{x \in X \,|\, f(x) < g(x)\} \in \mathcal{S}$.
If $f, g : X \to \overline{\mathbb{C}}$ is $\mathcal{S}$-measurable, then $\{x \in X \,|\, f(x) = g(x)\} \in \mathcal{S}$.*

Proof. Consider the set $A = \{x \in X \,|\, f(x) \in \mathbb{R},\, g(x) \in \mathbb{R}\} \in \mathcal{S}$. Then $f\rceil A, g\rceil A$ are $\mathcal{S}\rceil A$-measurable, and so $(f - g)\rceil A = (f\rceil A) - (g\rceil A)$ is $\mathcal{S}\rceil A$-measurable. Hence, the sets

$$\{x \in A \,|\, f(x) = g(x)\} = ((f - g)\rceil A)^{-1}(\{0\})$$

$$\{x \in A \,|\, f(x) < g(x)\} = ((f - g)\rceil A)^{-1}((-\infty, 0))$$

belong to $\mathcal{S}\rceil A$, and so they belong to $\mathcal{S}$. Therefore,

$$\begin{aligned} \{x \in X \,|\, f(x) = g(x)\} = \{x \in A \,|\, f(x) = g(x)\} &\cup \big(f^{-1}(\{-\infty\}) \cap g^{-1}(\{-\infty\})\big) \\ &\cup \big(f^{-1}(\{+\infty\}) \cap g^{-1}(\{+\infty\})\big) \in \mathcal{S}. \end{aligned}$$

In a similar manner,

$$\begin{aligned} \{x \in X \,|\, f(x) < g(x)\} = \{x \in A \,|\, f(x) < g(x)\} &\cup \big(f^{-1}(\{-\infty\}) \cap g^{-1}((-\infty, +\infty])\big) \\ &\cup \big(f^{-1}([-\infty, +\infty)) \cap g^{-1}(\{+\infty\})\big) \in \mathcal{S}. \end{aligned}$$

The case of $f, g : X \to \overline{\mathbb{C}}$ and $\{x \in X \,|\, f(x) = g(x)\}$ is even simpler. $\square$

TRUNCATION.

There are many possible truncations of a function.

Definition. *Let $f : X \to \overline{\mathbb{R}}$ and $\alpha, \beta \in \overline{\mathbb{R}}$ so that $\alpha \leq \beta$.*
We define $f_{(\alpha)}^{(\beta)} = \min\big\{ \max\{f, \alpha\}, \beta\big\}$.
We write $f^{(\beta)}$ instead of $f_{(-\infty)}^{(\beta)}$. I.e. $f^{(\beta)} = \min\{f, \beta\}$.
We write $f_{(\alpha)}$ instead of $f_{(\alpha)}^{(+\infty)}$. I.e. $f_{(\alpha)} = \max\{f, \alpha\}$.
The functions $f_{(\alpha)}^{(\beta)}, f^{(\beta)}, f_{(\alpha)}$ are called ***truncations*** *of f.*

In other words, we have: $f_{(\alpha)}^{(\beta)}(x) = f(x)$, if $\alpha \leq f(x) \leq \beta$, $f_{(\alpha)}^{(\beta)}(x) = \alpha$, if $f(x) < \alpha$, and $f_{(\alpha)}^{(\beta)}(x) = \beta$, if $\beta < f(x)$. Also: $f^{(\beta)}(x) = f(x)$, if $f(x) \leq \beta$, and $f^{(\beta)}(x) = \beta$, if $\beta < f(x)$. Finally: $f_{(\alpha)}(x) = f(x)$, if $\alpha \leq f(x)$, and $f_{(\alpha)}(x) = \alpha$, if $f(x) < \alpha$.

Proposition 2.20. *Let $(X, \mathcal{S})$ be a measurable space and $f : X \to \overline{\mathbb{R}}$ be $\mathcal{S}$-measurable. Then all truncations $f_{(\alpha)}^{(\beta)}$ are $\mathcal{S}$-measurable.*

Proof. The proof is obvious from $f_{(\alpha)}^{(\beta)} = \min\big\{ \max\{f, \alpha\}, \beta\big\}$ and Proposition 2.18. □

An important role is played by the following special truncations of $f : X \to \overline{\mathbb{R}}$. They are the functions $f^+ : X \to [0, +\infty]$ and $f^- : X \to [0, +\infty]$, which are defined by the formulas

$$f^+ = f_{(0)} = \max\{f, 0\}, \quad f^- = -f^{(0)} = -\min\{f, 0\} = \max\{-f, 0\},$$

and they are called, respectively, the **non-negative part** and the **non-positive part** of f.

If $\mathcal{S}$ is a σ-algebra of subsets of X and $f : X \to \overline{\mathbb{R}}$ is $\mathcal{S}$-measurable, then both f^+ and f^- are $\mathcal{S}$-measurable. It is also trivial to see that either $f^+(x) = 0$ or $f^-(x) = 0$ for every $x \in X$. I.e. $f^+f^- = 0$. Also

$$f^+ + f^- = |f|, \quad f^+ - f^- = f.$$

There is another type of truncations used mainly for extended-complex valued functions.

Definition. *Let $f : X \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ and $r \in [0, +\infty]$. We define ${}^{(r)}f(x) = f(x)$, if $|f(x)| \leq r$, and ${}^{(r)}f(x) = r\ \mathrm{sign}(f(x))$, if $r < |f(x)|$.*
The functions ${}^{(r)}f$ are also called ***truncations*** *of f.*

We observe that, if $f : X \to \overline{\mathbb{R}}$, then ${}^{(r)}f = f_{(-r)}^{(r)}$.

Proposition 2.21. *Let $(X, \mathcal{S})$ be a measurable space and $f : X \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ be $\mathcal{S}$-measurable. Then all truncations ${}^{(r)}f$ are $\mathcal{S}$-measurable.*

Proof. The case $f : X \to \overline{\mathbb{R}}$ is clear, since ${}^{(r)}f = f_{(-r)}^{(r)}$.
In the case $f : X \to \overline{\mathbb{C}}$ we consider the function $\phi_r : \overline{\mathbb{C}} \to \overline{\mathbb{C}}$ defined by: $\phi_r(x) = x$, if $|x| \leq r$, and $\phi_r(x) = r\ \mathrm{sign}(x)$, if $r < |x|$. We easily see that ϕ_r is Borel measurable. Now, ${}^{(r)}f = \phi_r \circ f$, and so ${}^{(r)}f$ is $\mathcal{S}$-measurable. □

Exercises.

2.1.5. Let $f : X \to \overline{\mathbb{R}}$. If $g, h : X \to \overline{\mathbb{R}}$ are such that $g, h \geq 0$ and $f = g - h$ on X, prove that $f^+ \leq g$ and $f^- \leq h$ on X.

2.1.6. Let $(X, \mathcal{S}, \mu)$ be a measure space, $f : X \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ be $\mathcal{S}$-measurable, and $0 \leq M < +\infty$. If $\mu(\{x \in X \,|\, |f(x)| = +\infty\}) = 0$ and $\mu(\{x \in X \,|\, |f(x)| > M\}) < +\infty$, prove that for every $\epsilon > 0$ there is a bounded $\mathcal{S}$-measurable $g : X \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ so that $\mu(\{x \in X \,|\, g(x) \neq f(x)\}) < \epsilon$.

LIMITS.

The next group of results is about various *limiting operations on measurable functions*. The rule is, roughly: *the supremum, the infimum and the limit of a sequence of measurable functions are measurable functions*.

Proposition 2.22. *Let $(X, \mathcal{S})$ be a measurable space and (f_j) be a sequence of $\mathcal{S}$-measurable functions $f_j : X \to \overline{\mathbb{R}}$. Then all functions $\sup_{j\in\mathbb{N}} f_j$, $\inf_{j\in\mathbb{N}} f_j$, $\overline{\lim}_{j\to+\infty} f_j$ and $\underline{\lim}_{j\to+\infty} f_j$ are $\mathcal{S}$-measurable.*

Proof. Let $h = \sup_{j\in\mathbb{N}} f_j : X \to \overline{\mathbb{R}}$. We have

$$\begin{aligned} h^{-1}((a, +\infty]) &= \{x \in A \,|\, a < h(x)\} = \{x \in X \,|\, a < f_j(x) \text{ for at least one } j\} \\ &= \textstyle\bigcup_{j=1}^{+\infty} \{x \in X \,|\, a < f_j(x)\} = \bigcup_{j=1}^{+\infty} f_j^{-1}((a, +\infty]), \end{aligned}$$

and so $h^{-1}((a, +\infty]) \in \mathcal{S}$ for every $a \in \mathbb{R}$. Now, Proposition 2.11 implies that h is $\mathcal{S}$-measurable. Therefore, $\inf_{j\in\mathbb{N}} f_j = -\sup_{j\in\mathbb{N}}(-f_j)$ is also $\mathcal{S}$-measurable.
And, finally, $\overline{\lim}_{j\to+\infty} f_j = \inf_{j\in\mathbb{N}}(\sup_{k\geq j} f_k)$ and $\underline{\lim}_{j\to+\infty} f_j = \sup_{j\in\mathbb{N}}(\inf_{k\geq j} f_k)$ are $\mathcal{S}$-measurable. ☐

Proposition 2.23. *Let $(X, \mathcal{S})$ be a measurable space and (f_j) be a sequence of $\mathcal{S}$-measurable functions $f_j : X \to \overline{\mathbb{R}}$. Then the set $A = \{x \in X \mid \lim_{j\to+\infty} f_j(x)$ exists in $\overline{\mathbb{R}}\}$ belongs to $\mathcal{S}$.*
(i) The function $\lim_{j\to+\infty} f_j : A \to \overline{\mathbb{R}}$ is $\mathcal{S}\rceil A$-measurable.
(ii) Let $h : A^c \to \overline{\mathbb{R}}$ be $\mathcal{S}\rceil A^c$-measurable. We define $(\lim_{j\to+\infty} f_j)(x) = \lim_{j\to+\infty} f_j(x)$, if $x \in A$, and $(\lim_{j\to+\infty} f_j)(x) = h(x)$, if $x \in A^c$. Then $\lim_{j\to+\infty} f_j : X \to \overline{\mathbb{R}}$ is $\mathcal{S}$-measurable.
Similar results hold if $f_j : X \to \overline{\mathbb{C}}$ for all j and $A = \{x \in X \mid \lim_{j\to+\infty} f_j(x)$ exists in $\overline{\mathbb{C}}\}$.

Proof. Suppose that $f_j : X \to \overline{\mathbb{R}}$ for all j.
Since $\lim_{j\to+\infty} f_j(x)$ exists if and only if $\overline{\lim}_{j\to+\infty} f_j(x) = \underline{\lim}_{j\to+\infty} f_j(x)$, we have that

$$A = \{x \in X \mid \overline{\lim}_{j\to+\infty} f_j(x) = \underline{\lim}_{j\to+\infty} f_j(x)\}.$$

Now, Proposition 2.22 implies that $\overline{\lim}_{j\to+\infty} f_j$ and $\underline{\lim}_{j\to+\infty} f_j$ are both $\mathcal{S}$-measurable, and then Proposition 2.19 implies $A \in \mathcal{S}$.
(i) It is clear that the function $\lim_{j\to+\infty} f_j : A \to \overline{\mathbb{R}}$ is just the restriction of $\overline{\lim}_{j\to+\infty} f_j$ (and of $\underline{\lim}_{j\to+\infty} f_j$) on A, and so it is $\mathcal{S}\rceil A$-measurable.
(ii) The proof of (ii) is a direct consequence of (i) and Proposition 2.7.
The case of complex valued (or extended complex valued) functions can be reduced to what we just proved and it is left as an exercise. ☐

SIMPLE FUNCTIONS.

Definition. *Let $E \subseteq X$. The function $\chi_E : X \to \mathbb{R}$ defined by $\chi_E(x) = 1$, if $x \in E$, and $\chi_E(x) = 0$, if $x \notin E$, is called the **characteristic function** of E.*

Of course, E determines its χ_E. But also, conversely, χ_E determines its corresponding E. Indeed, $E = \{x \in X \,|\, \chi_E(x) = 1\} = (\chi_E)^{-1}(\{1\})$.
The following are trivial:

$$\lambda\chi_E + \kappa\chi_F = \lambda\chi_{E\setminus F} + (\lambda + \kappa)\chi_{E\cap F} + \kappa\chi_{F\setminus E}, \quad \chi_E\chi_F = \chi_{E\cap F}, \quad \chi_{E^c} = 1 - \chi_E$$

for all $E, F \subseteq X$ and all $\lambda, \kappa \in \mathbb{C}$.

Proposition 2.24. *Let $(X, \mathcal{S})$ be a measurable space and $E \subseteq X$. Then χ_E is $\mathcal{S}$-measurable if and only if $E \in \mathcal{S}$.*

Proof. If χ_E is $\mathcal{S}$-measurable, then $E = (\chi_E)^{-1}(\{1\}) \in \mathcal{S}$.
Conversely, let $E \in \mathcal{S}$. Then for an arbitrary Borel set F in $\mathbb{R}$ or $\mathbb{C}$ we have: $(\chi_E)^{-1}(F) = \emptyset$ if $0 \notin F$, $1 \notin F$, and $(\chi_E)^{-1}(F) = E$ if $1 \in F$, $0 \notin F$, and $(\chi_E)^{-1}(F) = E^c$ if $1 \notin F$, $0 \in F$, and $(\chi_E)^{-1}(F) = X$ if $0 \in F$, $1 \in F$. In any case, $(\chi_E)^{-1}(F) \in \mathcal{S}$, and so χ_E is $\mathcal{S}$-measurable. $\square$

Definition. *A function defined on a non-empty set X is called a* ***simple function*** *on X if its range is a finite subset of $\mathbb{C}$. If, in particular, the range of the simple function is a subset of $\mathbb{R}$, then we may say that it is a real valued simple function. Also, if the range of the simple function is a subset of $[0, +\infty)$, then we may say that it is a non-negative simple function.*

We note that simple functions never take infinite values.
The following proposition describes completely the structure of simple functions.

Proposition 2.25. *(i) A function $\phi : X \to \mathbb{C}$ is a simple function on X if and only if it is a linear combination with complex coefficients of characteristic functions of subsets of X.*
(ii) For every simple function ϕ on X there are $m \in \mathbb{N}$, distinct $\kappa_1, \dots, \kappa_m \in \mathbb{C}$ and non-empty pairwise disjoint $E_1, \dots, E_m \subseteq X$ with $\bigcup_{j=1}^m E_j = X$ so that $\phi = \kappa_1\chi_{E_1} + \cdots + \kappa_m\chi_{E_m}$. This representation of ϕ is unique (apart from rearrangement).
(iii) If $\mathcal{S}$ is a σ-algebra of subsets of X, then the simple function ϕ on X is $\mathcal{S}$-measurable if and only if all E_k in the representation of ϕ described in (ii) belong to $\mathcal{S}$.

Proof. Let
$$\phi = \textstyle\sum_{j=1}^n \lambda_j\chi_{F_j},$$
where $\lambda_j \in \mathbb{C}$ and $F_j \subseteq X$ for all j. We consider any $x \in X$, and then either x belongs to no F_j, in which case $\phi(x) = 0$, or, by considering all the sets $F_{j_1}, \dots, F_{j_k}$ which contain x, we have that $\phi(x) = \lambda_{j_1} + \cdots + \lambda_{j_k}$. Hence, the range of ϕ contains at most all the possible sums $\lambda_{j_1} + \cdots + \lambda_{j_k}$ together with 0, and so it is a finite set. Thus, ϕ is simple on X.
Conversely, let ϕ be simple on X, and let the range of ϕ consist of the distinct $\kappa_1, \dots, \kappa_m \in \mathbb{C}$. We consider
$$E_j = \{x \in X \,|\, \phi(x) = \kappa_j\} = \phi^{-1}(\{\kappa_j\}).$$
Then every $x \in X$ belongs to exactly one of these sets, and so $E_1, \dots, E_m$ are pairwise disjoint and $X = E_1 \cup \cdots \cup E_m$. Now it is clear that
$$\phi = \textstyle\sum_{j=1}^m \kappa_j\chi_{E_j},$$
because both sides take the same value at every x.
If
$$\phi = \textstyle\sum_{i=1}^{m'} \kappa_i'\chi_{E_i'}$$
is another representation of ϕ with distinct κ_i' and non-empty pairwise disjoint E_i' covering X, then the range of ϕ is exactly the set $\{\kappa_1', \dots, \kappa_{m'}'\}$. Hence, $m' = m$ and, after rearrangement, $\kappa_1' = \kappa_1, \dots, \kappa_m' = \kappa_m$. Therefore,
$$E_j' = \phi^{-1}(\{\kappa_j'\}) = \phi^{-1}(\{\kappa_j\}) = E_j$$
for all j. We conclude that the representation is unique.
Now, if all E_j belong to $\mathcal{S}$, then, by Proposition 2.24, all χ_{E_j} are $\mathcal{S}$-measurable, and so ϕ is also $\mathcal{S}$-measurable. Conversely, if ϕ is $\mathcal{S}$-measurable, then all $E_j = \phi^{-1}(\{\kappa_j\})$ belong to $\mathcal{S}$. $\square$

Definition. *The unique representation $\phi = \kappa_1\chi_{E_1} + \cdots + \kappa_m\chi_{E_m}$ of the simple function ϕ, which is described in part (ii) of Proposition 2.25, is called the* ***standard representation*** *of ϕ.*

If one of the coefficients in the standard representation of a simple function is equal to 0, then we usually omit the corresponding term from the sum (but then the union of the pairwise disjoint sets which appear in the representation is not necessarily equal to the whole space).

Proposition 2.26. *Any linear combination with complex coefficients of simple functions is a simple function and any product of simple functions is a simple function. Also, the maximum and the minimum of real valued simple functions are real valued simple functions.*

Proof. Let ϕ, ψ be simple functions on X and $p, q \in \mathbb{C}$. Assume that $\lambda_1, \ldots, \lambda_n$ are the values of ϕ, and $\kappa_1, \ldots, \kappa_m$ are the values of ψ. It is obvious that the possible values of $p\phi + q\psi$ are among the nm numbers $p\lambda_i + q\kappa_j$, and that the possible values of $\phi\psi$ are among the nm numbers $\lambda_i\kappa_j$. Therefore, both functions $p\phi + q\psi, \phi\psi$ have a finite number of values. If ϕ, ψ are real valued, then the possible values of $\max\{\phi, \psi\}$ and $\min\{\phi, \psi\}$ are among the $n + m$ real numbers λ_i, κ_j. $\square$

Proposition 2.27. *(i) Given $f : X \to [0, +\infty]$, there exists an increasing sequence (ϕ_n) of non-negative simple functions on X which converges to f pointwise on X. Moreover, (ϕ_n) converges to f uniformly on every subset of X on which f is bounded.*
(ii) Given $f : X \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$, there is a sequence (ϕ_n) of real valued or complex valued, respectively, simple functions on X which converges to f pointwise on X and so that $(|\phi_n|)$ is increasing. Moreover, (ϕ_n) converges to f uniformly on every subset on which f is bounded.
If $\mathcal{S}$ is a σ-algebra of subsets of X and f is $\mathcal{S}$-measurable, then the ϕ_n in (i) and (ii) can be taken to be $\mathcal{S}$-measurable.

Proof. (i) For every $n, k \in \mathbb{N}$ with $0 \leq k \leq n^2 - 1$, we define the sets

$$E_{n,k} = f^{-1}\big(\big[\tfrac{k}{n}, \tfrac{k+1}{n}\big)\big), \quad F_n = f^{-1}([n, +\infty])$$

and the non-negative simple function

$$\psi_n = \textstyle\sum_{k=0}^{n^2-1} \frac{k}{n}\chi_{E_{n,k}} + n\chi_{F_n}.$$

For each n the sets

$$E_{n,0},\ E_{n,1},\ \ldots,\ E_{n,n^2-1},\ F_n$$

are pairwise disjoint and their union is X. Observe that if f is $\mathcal{S}$-measurable then all $E_{n,k}$ and F_n belong to $\mathcal{S}$, and so ψ_n is $\mathcal{S}$-measurable.
For every n we have

$$\psi_n = \tfrac{k}{n} \leq f < \tfrac{k+1}{n} = \psi_n + \tfrac{1}{n} \ \text{ on each } E_{n,k}, \quad \psi_n = n \leq f \ \text{ on } F_n.$$

Now, if $f(x) = +\infty$, then $x \in F_n$ and so $\phi_n(x) = n$ for every n. Hence $\lim_{n\to+\infty} \phi_n(x) = f(x)$. If $0 \leq f(x) < +\infty$, then for all large n we have $0 \leq f(x) < n$. So for each large n there is a unique k with $0 \leq k \leq n^2 - 1$ and $\frac{k}{n} \leq f(x) < \frac{k+1}{n}$. Then $x \in E_{n,k}$, and so $\psi_n(x) = \frac{k}{n}$. Hence

$$0 \leq f(x) - \psi_n(x) < \tfrac{1}{n}$$

for large n. This implies that $\lim_{n\to+\infty} \psi_n(x) = f(x)$.
Therefore,

$$\lim_{n\to+\infty} \psi_n = f \quad \text{pointwise on } X.$$

If $K \subseteq X$ and f is bounded on K, then there is an n_0 so that $f(x) < n_0$ for all $x \in K$. Hence, for all $n \geq n_0$ we have

$$0 \leq f(x) - \psi_n(x) < \tfrac{1}{n} \quad \text{for all } x \in K.$$

Thus,

$$\lim_{n\to+\infty} \psi_n = f \quad \text{uniformly on } K.$$

Now, for every n, we consider the simple function

$$\phi_n = \max\{\psi_1, \ldots, \psi_n\}$$

If f is $\mathcal{S}$-measurable, then every ψ_k is $\mathcal{S}$-measurable, and so every ϕ_n is $\mathcal{S}$-measurable. We have proved that $0 \leq \psi_k \leq f$ on X for every k, and so we have that $0 \leq \phi_n \leq f$ on X for every n. Moreover, $\phi_n = \max\{\psi_1, \ldots, \psi_n\} \geq \psi_n$, and so $\psi_n \leq \phi_n \leq f$ on X for every n.
Therefore,

$$\lim_{n\to+\infty} \phi_n = f \quad \text{pointwise on } X,$$

and, if $K \subseteq X$ and f is bounded on K, then

$$\lim_{n\to+\infty} \phi_n = f \quad \text{uniformly on } K.$$

Finally,

$$\phi_{n+1} = \max\{\psi_1, \ldots, \psi_n, \psi_{n+1}\} \geq \max\{\psi_1, \ldots, \psi_n\} = \phi_n$$

on X for every n, and so (ϕ_n) is increasing on X.
(ii) Let $f : X \to \overline{\mathbb{R}}$. We consider the functions $f^+, f^- : X \to [0, +\infty]$. If f is $\mathcal{S}$-measurable, then f^+, f^- are both $\mathcal{S}$-measurable.
By (i) there are increasing sequences (p_n) and (q_n) of non-negative simple functions on X converging to, respectively, f^+ and f^- pointwise on X and uniformly on every subset of X on which f is bounded (because on such a subset f^+, f^- are also bounded). Now it is obvious that, if we set

$$\phi_n = p_n - q_n,$$

then ϕ_n is a real valued simple function on X which is $\mathcal{S}$-measurable if f is $\mathcal{S}$-measurable. It is clear that (ϕ_n) converges to f pointwise on X and uniformly on every subset of X on which f is bounded. Since $0 \leq p_n \leq f^+$ and $0 \leq q_n \leq f^-$, we have that $p_n = \phi_n^+$ and $q_n = \phi_n^-$. Hence

$$|\phi_n| = p_n + q_n,$$

and so the sequence $(|\phi_n|)$ is increasing on X.
Now let $f : X \to \overline{\mathbb{C}}$. We consider $A = f^{-1}(\mathbb{C})$, the restriction $f \rceil A : A \to \mathbb{C}$, and the functions $\mathrm{Re}(f \rceil A), \mathrm{Im}(f \rceil A) : A \to \mathbb{R}$.
If f is $\mathcal{S}$-measurable, then these two functions are $\mathcal{S} \rceil A$-measurable.
By the previous case there are sequences (r_n) and (s_n) of real valued simple functions on A converging to, respectively, $\mathrm{Re}(f \rceil A)$ and $\mathrm{Im}(f \rceil A)$ pointwise on A and uniformly on every subset of A on which $f \rceil A$ is bounded. Now, if we set

$$\phi_n = r_n + i s_n,$$

then ϕ_n is a complex valued simple function on A which is $\mathcal{S} \rceil A$-measurable if f is $\mathcal{S}$-measurable. It is clear that (ϕ_n) converges to $f \rceil A$ pointwise on A and uniformly on every subset of A on which $f \rceil A$ is bounded. Also

$$|\phi_n| = \sqrt{r_n^2 + s_n^2},$$

and so the sequence $(|\phi_n|)$ is increasing on A.
If we also define $\phi_n = n$ on A^c, then the proof is complete. $\square$

Exercises.

2.1.7. (i) Prove that a Borel measurable $f : \mathbb{R} \to \mathbb{R}$ is also Lebesgue measurable.
(ii) Find a function $f : \mathbb{R} \to \mathbb{R}$ which is not Lebesgue measurable.
(iii) Using the Lebesgue but not Borel set constructed in exercise 1.4.18, find a function $f : \mathbb{R} \to \mathbb{R}$ which is Lebesgue measurable but not Borel measurable.

2.1.8. Give an example of a function $f : \mathbb{R} \to \mathbb{R}$ which is not Lebesgue measurable so that $|f|$ is Lebesgue measurable.

2.1.9. Starting with an appropriate function which is not Lebesgue measurable, give an example of an uncountable collection $\{f_i\}_{i\in I}$ of Lebesgue measurable functions $f_i : \mathbb{R} \to \mathbb{R}$ so that $\sup_{i\in I} f_i$ is not Lebesgue measurable.

2.1.10. (i) Prove that, if $G : \mathbb{R} \to \mathbb{R}$ is continuous and $H : \mathbb{R} \to \mathbb{R}$ is Borel measurable, then $H \circ G : \mathbb{R} \to \mathbb{R}$ is Borel measurable.
(ii) Using exercise 1.4.18, construct a continuous $G : \mathbb{R} \to \mathbb{R}$ and a Lebesgue measurable $H : \mathbb{R} \to \mathbb{R}$ so that $H \circ G : \mathbb{R} \to \mathbb{R}$ is not Lebesgue measurable.

2.1.11. We say that $\phi : X \to \mathbb{C}$ is an **elementary function** on X if it has countably many values. Is there a standard representation for an elementary function?
Prove that for any $f : X \to [0, +\infty)$, there is an increasing sequence (ϕ_n) of non-negative elementary functions on X so that $\lim_{n\to+\infty} \phi_n = f$ uniformly on X. If $\mathcal{S}$ is a σ-algebra of subsets of X and f is $\mathcal{S}$-measurable, prove that the ϕ_n can be taken to be $\mathcal{S}$-measurable.

2.2 The role of null sets.

Definition. *Let $(X, \mathcal{S}, \mu)$ be a measure space. We say that a property $P(x)$ holds **μ-almost everywhere** on X or for **μ-almost every** $x \in X$, if the set $\{x \in X \,|\, P(x)$ is not true$\}$ is included in a μ-null set.*

We may also say: $P(x)$ holds μ-a.e. on X and $P(x)$ holds for μ-a.e. $x \in X$. More simply: $P(x)$ holds a.e. on X and $P(x)$ holds for a.e. $x \in X$.
It is clear that, if $P(x)$ holds for a.e. $x \in X$ and μ is *complete*, then $\{x \in X \,|\, P(x)$ is not true$\}$ is contained in $\mathcal{S}$, and so its complement $\{x \in X \,|\, P(x)$ is true$\}$ is also contained in $\mathcal{S}$.

Proposition 2.28. *Let $(X, \mathcal{S}_X, \mu)$ be a measure space and $(X, \overline{\mathcal{S}_X}, \overline{\mu})$ be its completion, and assume $A \in \mathcal{S}_X$ has $\mu(A^c) = 0$. Let $(Y, \mathcal{S}_Y)$ be a measurable space and $f : A \to Y$ be $(\mathcal{S}_X\rceil A, \mathcal{S}_Y)$-measurable. If we extend f on X in an arbitrary manner as a function $F : X \to Y$, then the extended function F is $(\overline{\mathcal{S}_X}, \mathcal{S}_Y)$-measurable.*

Proof. We consider an arbitrary function $h : A^c \to Y$, and we define $F : X \to Y$ by $F(x) = f(x)$, if $x \in A$, and $F(x) = h(x)$, if $x \in A^c$.
We take an arbitrary $E \in \mathcal{S}_Y$, and we write

$$F^{-1}(E) = \{x \in A \,|\, f(x) \in E\} \cup \{x \in A^c \,|\, h(x) \in E\} = f^{-1}(E) \cup \{x \in A^c \,|\, h(x) \in E\}.$$

The first set belongs to $\mathcal{S}_X\rceil A$ and hence to $\mathcal{S}_X$, and the second set is a subset of A^c. Therefore, $F^{-1}(E) \in \overline{\mathcal{S}_X}$, and so F is $(\overline{\mathcal{S}_X}, \mathcal{S}_Y)$-measurable. $\square$

In other words, if $(X, \mathcal{S}_X, \mu)$ is a *complete* measure space, we get that, *if f is defined a.e. on X and it is measurable on its domain of definition, then any extension of f on X is measurable.*

Proposition 2.29. *Let $(X, \mathcal{S}_X, \mu)$ be a measure space and $(X, \overline{\mathcal{S}_X}, \overline{\mu})$ be its completion. Let $(Y, \mathcal{S}_Y)$ be a measurable space and $f : X \to Y$ be $(\mathcal{S}_X, \mathcal{S}_Y)$-measurable. If $g : X \to Y$ is equal to f a.e on X, then g is $(\overline{\mathcal{S}_X}, \mathcal{S}_Y)$-measurable.*

Proof. There exists $N \in \mathcal{S}_X$ so that $\{x \in X \,|\, f(x) \neq g(x)\} \subseteq N$ and $\mu(N) = 0$.
We consider $A = N^c \in \mathcal{S}_X$, and then $f\rceil A : A \to Y$ is $(\mathcal{S}_X\rceil A, \mathcal{S}_Y)$-measurable. Since $g = f$ on A, we see that $g : X \to Y$ is an extension of $f\rceil A : A \to Y$. Now, Proposition 2.28 implies that g is $(\overline{\mathcal{S}_X}, \mathcal{S}_Y)$-measurable. $\square$

In the particular case of a *complete* measure space $(X, \mathcal{S}_X, \mu)$, we get that, *if f is measurable and g is equal to f a.e., then g is also measurable.*

Proposition 2.30. *Let $(X, \mathcal{S}, \mu)$ be a measure space and $(X, \overline{\mathcal{S}}, \overline{\mu})$ be its completion. Let (f_j) be a sequence of $\mathcal{S}$-measurable functions $f_j : X \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$. If $f : X \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ is such that $f = \lim_{j\to+\infty} f_j$ a.e. on X, then f is $\overline{\mathcal{S}}$-measurable.*

Proof. There exists $N \in \mathcal{S}$ so that $\{x \in X \mid f(x) \neq \lim_{j\to+\infty} f_j(x)\} \subseteq N$ and $\mu(N) = 0$.
We consider $A = N^c \in \mathcal{S}$, and then $f\rceil A = \lim_{j\to+\infty} f_j\rceil A$ on A.
Now, every $f_j\rceil A$ is $\mathcal{S}\rceil A$-measurable, and so $f\rceil A$ is $\mathcal{S}\rceil A$-measurable. Since $f = f\rceil A$ on A, Proposition 2.28 implies that f is $\overline{\mathcal{S}}$-measurable. □

Again, in the particular case of a *complete* measure space $(X, \mathcal{S}, \mu)$ we get that, *if (f_j) is a sequence of measurable functions and its limit is equal to f a.e., then f is also measurable.*

Proposition 2.31. *Let $(X, \mathcal{S}, \mu)$ be a measure space and $(X, \overline{\mathcal{S}}, \overline{\mu})$ be its completion. If $g : X \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ is $\overline{\mathcal{S}}$-measurable, then there is a $\mathcal{S}$-measurable $f : X \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ so that $g = f$ a.e. on X.*

Proof. (a) Let $E \in \overline{\mathcal{S}}$. Then there are $A, M \in \mathcal{S}$ with $\mu(M) = 0$ so that $E = A \cup F$ for some $F \subseteq M$. Then $\chi_E \neq \chi_A$ only on $E \setminus A \subseteq M$, and so $\chi_E = \chi_A$ a.e. on X.
(b) Now, let $\phi : X \to \mathbb{R}$ or $\mathbb{C}$ be a $\overline{\mathcal{S}}$-measurable simple function with standard representation

$$\phi = \kappa_1\chi_{E_1} + \cdots + \kappa_m\chi_{E_m}.$$

Then $E_1, \dots, E_m \in \overline{\mathcal{S}}$, and by the result of (a), there are $A_1, \dots, A_m \in \mathcal{S}$ so that $\chi_{E_j} = \chi_{A_j}$ a.e. on X for every j. Then

$$\psi = \kappa_1\chi_{A_1} + \cdots + \kappa_m\chi_{A_m}$$

is a $\mathcal{S}$-measurable simple function. Since

$$\{x \in X \mid \phi(x) \neq \psi(x)\} \subseteq \textstyle\bigcup_{j=1}^m \{x \in X \mid \chi_{E_j}(x) \neq \chi_{A_j}(x)\},$$

we have that $\overline{\mu}(\{x \in X \mid \phi(x) \neq \psi(x)\}) = 0$, and so $\phi = \psi$ a.e. on X.
(c) Finally, let $g : X \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ be $\overline{\mathcal{S}}$-measurable. Proposition 2.27 implies that there are $\overline{\mathcal{S}}$-measurable simple functions $\phi_n : X \to \mathbb{R}$ or $\mathbb{C}$ so that $\lim_{n\to+\infty} \phi_n = g$ on X.
By (b), there are $\mathcal{S}$-measurable simple functions $\psi_n : X \to \mathbb{R}$ or $\mathbb{C}$ so that $\phi_n = \psi_n$ a.e. on X. We consider the set

$$B = \textstyle\bigcup_{n=1}^{+\infty} \{x \in X \mid \phi_n(x) \neq \psi_n(x)\}.$$

Then $B \in \overline{\mathcal{S}}$ and $\overline{\mu}(B) = 0$, and we have that $\phi_n = \psi_n$ for every n on B^c. Since $B^c \in \overline{\mathcal{S}}$, there are $A, M \in \mathcal{S}$ with $\mu(M) = 0$ so that $B^c = A \cup F$ for some $F \subseteq M$. Since $A \subseteq B^c$, we have that $\phi_n = \psi_n$ for every n on A, and so $\lim_{n\to+\infty} \psi_n = g$ on A.
Also, since $A \in \mathcal{S}$ and every ψ_n is $\mathcal{S}$-measurable, we have that every $\psi_n\rceil A$ is $\mathcal{S}\rceil A$-measurable, and so $g\rceil A = \lim_{n\to+\infty}(\psi_n\rceil A)$ is $\mathcal{S}\rceil A$-measurable.
Now, we consider $f : X \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ to be equal to $g\rceil A$ on A and equal to 0 on A^c. Then f is $\mathcal{S}$-measurable. Also, $A^c \subseteq B \cup M$ and

$$\overline{\mu}(B \cup M) \leq \overline{\mu}(B) + \overline{\mu}(M) = \overline{\mu}(B) + \mu(M) = 0.$$

Therefore, $g = f$ a.e. on X. □

Exercises.

2.2.1. Let $(X, \mathcal{S}, \mu)$ be a measure space.
(i) Let $f, g, h : X \to Y$. If $f = g$ a.e. on X and $g = h$ a.e. on X, prove that $f = h$ a.e. on X.
(ii) Let $f_1, f_2, g_1, g_2 : X \to \mathbb{R}$. If $f_1 = f_2$ a.e. on X and $g_1 = g_2$ a.e. on X, prove that $f_1 + g_1 = f_2 + g_2$ and $f_1g_1 = f_2g_2$ a.e. on X.
(iii) Let $f_j, g_j : X \to \overline{\mathbb{R}}$ so that $f_j = g_j$ a.e. on X for all $j \in \mathbb{N}$. Prove that $\sup_{j\in\mathbb{N}} f_j = \sup_{j\in\mathbb{N}} g_j$

a.e. on X. Similar results hold for inf, $\overline{\lim}$ and $\underline{\lim}$.
(iv) Let $f_j, g_j : X \to \overline{\mathbb{R}}$ so that $f_j = g_j$ a.e. on X for all $j \in \mathbb{N}$.
If $A = \{x \in X \mid \lim_{j\to+\infty} f_j(x) \text{ exists}\}$ and $B = \{x \in X \mid \lim_{j\to+\infty} g_j(x) \text{ exists}\}$, prove that $A \triangle B \subseteq N$ for some $N \in \mathcal{S}$ with $\mu(N) = 0$, and that $\lim_{j\to+\infty} f_j = \lim_{j\to+\infty} g_j$ a.e. on $A \cap B$.
If, moreover, we extend both $\lim_{j\to+\infty} f_j$ and $\lim_{j\to+\infty} g_j$ by a common function h on $(A \cap B)^c$, prove that $\lim_{j\to+\infty} f_j = \lim_{j\to+\infty} g_j$ a.e. on X.

2.3 Lusin's Theorem

A topological space X is called **locally compact** if for every $x \in X$ there is an open $V \subseteq X$ such that $x \in V$ and $\text{cl}(V)$ is compact.

Lemma 2.1. *Let the topological space X be locally compact and Hausdorff. For every $x \in X$ and every open $U \subseteq X$ with $x \in U$ there is an open $W \subseteq X$ such that $x \in W$,* $\text{cl}(W) \subseteq U$ *and* $\text{cl}(W)$ *is compact.*

Proof. There is an open $V \subseteq X$ such that $x \in V$ and $\text{cl}(V)$ is compact. Let $V_0 = V \cap U$. Then V_0 is open, and $x \in V_0 \subseteq U$. Since $\text{bd}(V_0) \subseteq \text{cl}(V_0) \subseteq \text{cl}(V)$, we have that $\text{bd}(V_0)$ is a closed subset of a compact set, and so $\text{bd}(V_0)$ is compact.
For every $y \in \text{bd}(V_0)$ we have $x \neq y$, and so there are open W_y, Y_y such that $x \in W_y$, $y \in Y_y$ and $W_y \cap Y_y = \emptyset$. Now, since $\text{bd}(V_0) \subseteq \bigcup_{y\in\text{bd}(V_0)} Y_y$, there are $y_1, \dots, y_n \in \text{bd}(V_0)$ such that

$$\text{bd}(V_0) \subseteq Y_{y_1} \cup \cdots \cup Y_{y_n}.$$

Now let $W = V_0 \cap W_{y_1} \cap \cdots \cap W_{y_n}$.
Then W is open, and $x \in W$. We also have that

$$W \cap (Y_{y_1} \cup \cdots \cup Y_{y_n}) = \emptyset.$$

Then, since $Y_{y_1} \cup \cdots \cup Y_{y_n}$ is open, we get that

$$\text{cl}(W) \cap (Y_{y_1} \cup \cdots \cup Y_{y_n}) = \emptyset,$$

and so $\text{cl}(W) \cap \text{bd}(V_0) = \emptyset$. Now, since $W \subseteq V_0$, we get $\text{cl}(W) \subseteq V_0$, and so $\text{cl}(W) \subseteq U$.
Finally, $\text{cl}(W) \subseteq V_0 \subseteq V \subseteq \text{cl}(V)$, and so $\text{cl}(W)$ is compact. □

Lemma 2.2. *Let the topological space X be locally compact and Hausdorff. If $K \subseteq X$ is compact and $U \subseteq X$ is open and $K \subseteq U$, then there is an open $W \subseteq X$ such that $K \subseteq W \subseteq$* $\text{cl}(W) \subseteq U$ *and* $\text{cl}(W)$ *is compact.*

Proof. By Lemma 2.1, for every $x \in K$ there is an open $W_x \subseteq X$ such that $x \in W_x$, $\text{cl}(W_x) \subseteq U$ and $\text{cl}(W_x)$ is compact. Since $K \subseteq \bigcup_{x\in K} W_x$, there are $x_1, \dots, x_n \in K$ such that

$$K \subseteq W_{x_1} \cup \cdots \cup W_{x_n}.$$

Let $W = W_{x_1} \cup \cdots \cup W_{x_n}$.
Then

$$\text{cl}(W) = \text{cl}(W_{x_1}) \cup \cdots \cup \text{cl}(W_{x_n}).$$

Therefore, W is open, $\text{cl}(W)$ is compact, and $K \subseteq W \subseteq \text{cl}(W) \subseteq U$. □

We know that, if X is a topological space and $f : X \to \mathbb{C}$ is continuous, then the set

$$\text{supp}(f) = \text{cl}(\{x \in X \mid f(x) \neq 0\})$$

is called the **support** of f. Clearly, $\text{supp}(f)$ is a closed subset of X. Also, clearly, $f(x) = 0$ for every $x \notin \text{supp}(f)$. Now, let us assume that F is a closed subset of X such that $f(x) = 0$ for every $x \notin F$. Then, $\{x \in X \mid f(x) \neq 0\} \subseteq F$, and, since F is closed, we get that $\text{supp}(f) \subseteq F$. Therefore, $\text{supp}(f)$ *is the smallest closed set outside of which* $f = 0$.

There is a more general version of the following result in Topology.

Urysohn's Lemma. *Let the topological space X be locally compact and Hausdorff. If $K \subseteq X$ is compact and $U \subseteq X$ is open and $K \subseteq U$, then there is a continuous $f : X \to [0,1]$ so that $f = 1$ on K and* $\mathrm{supp}(f)$ *is a compact subset of U.*

Proof. By Lemma 2.2, there is an open B_1 so that $\mathrm{cl}(B_1)$ is compact and

$$K \subseteq B_1 \subseteq \mathrm{cl}(B_1) \subseteq U.$$

Then there is some open $B_{1/2}$ so that $\mathrm{cl}(B_{1/2})$ is compact and

$$K \subseteq B_{1/2} \subseteq \mathrm{cl}(B_{1/2}) \subseteq B_1.$$

Similarly, there are some open $B_{1/4}$ and $B_{3/4}$ so that $\mathrm{cl}(B_{1/4})$ and $\mathrm{cl}(B_{3/4})$ are compact and

$$K \subseteq B_{1/4} \subseteq \mathrm{cl}(B_{1/4}) \subseteq B_{1/2} \subseteq \mathrm{cl}(B_{1/2}) \subseteq B_{3/4} \subseteq \mathrm{cl}(B_{3/4}) \subseteq B_1.$$

Continuing inductively, we see that to every rational of the form $r = \frac{k}{2^n}$ with $0 < k \leq 2^n$ corresponds some open set B_r so that $\mathrm{cl}(B_r)$ is compact and so that

$$K \subseteq B_r \subseteq \mathrm{cl}(B_r) \subseteq B_s \subseteq \mathrm{cl}(B_s) \subseteq U$$

for every two such rational r, s with $r < s$. Let $\mathbb{Q}_d$ be the set of all these rational numbers. It is easy to see that $\mathbb{Q}_d$ is dense in $[0,1]$.
Now, we define:

$$g(x) = \inf\{r \in \mathbb{Q}_d \,|\, x \in B_r\}, \text{ if } x \in B_1, \quad g(x) = 1, \text{ if } x \in B_1^c.$$

We see that $g = 0$ on K and that $g : X \to [0,1]$, and we shall prove that g is continuous on X.
Let $x \in X$ and $\epsilon > 0$.
If $0 < g(x) < 1$, there are $r, r', s \in \mathbb{Q}_d$ so that

$$g(x) - \epsilon < r < r' < g(x) < s < g(x) + \epsilon.$$

If $y \in B_s$, then $g(y) \leq s < g(x) + \epsilon$. If $y \notin \mathrm{cl}(B_r)$, then $y \notin B_r$, and so $g(y) \geq r > g(x) - \epsilon$. Also, $x \in B_s$ and $x \notin B_{r'}$, and so $x \in \mathrm{cl}(B_r)^c$. Hence, the open set $V = B_s \cap \mathrm{cl}(B_r)^c$ contains x, and we have that

$$g(x) - \epsilon < g(y) < g(x) + \epsilon \quad \text{for every } y \in V.$$

Therefore, g is continuous at x.
If $g(x) = 1$, we take, like before, $r, r' \in \mathbb{Q}_d$ so that

$$g(x) - \epsilon < r < r' < g(x).$$

Then we easily see that the open set $V = \mathrm{cl}(B_r)^c$ contains x, and that

$$g(x) - \epsilon < g(y) \leq 1 < g(x) + \epsilon \quad \text{for every } y \in V.$$

Hence, g is continuous at x.
Similarly, if $g(x) = 0$, we take $s \in \mathbb{Q}_d$ so that

$$0 < s < \epsilon.$$

Then we get that the open set $V = B_s$ contains x, and that

$$g(x) - \epsilon < 0 \leq g(y) < \epsilon = g(x) + \epsilon \quad \text{for every } y \in V.$$

Hence, g is continuous at x.
Finally, we take $f = 1 - g$. Then $f : X \to [0,1]$ is continuous on X, and $f = 1$ on K. Also $f = 0$ outside $\mathrm{cl}(B_1)$. Hence, $\mathrm{supp}(f)$ is contained in $\mathrm{cl}(B_1)$ which is a compact subset of U. $\square$

If X is a topological space and $U \subseteq X$ is open, then for a function f we write

$$f \prec U$$

whenever $f : X \to [0,1]$ is continuous on X and $\mathrm{supp}(f)$ is a compact subset of U.

Thus, Urysohn's Lemma says that, if X is locally compact and Hausdorff, $K \subseteq X$ is compact and $U \subseteq X$ is open and $K \subseteq U$, then there is a function f so that $f \prec U$ and $f = 1$ on K.

Lusin's Theorem. *Let the topological space X be locally compact and Hausdorff and μ be a regular Borel measure on X. If $f : X \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ is Borel measurable and f is finite a.e. on X and $f = 0$ outside a set of finite measure, then for every $\epsilon > 0$ there is a continuous $g : X \to \mathbb{R}$ or $\mathbb{C}$ so that $g = 0$ outside a compact set of finite measure and $\mu(\{x \in X \mid f(x) \neq g(x)\}) < \epsilon$. If f has certain bounds a.e. on X, then g can be chosen to have the same bounds on X.*

Proof. (a) Let E be any Borel set in X with $\mu(E) < +\infty$ and let $\epsilon > 0$. Then there is a compact K and an open U so that $K \subseteq E \subseteq U \subseteq X$ and $\mu(U \setminus K) < \epsilon$. By Urysohn's Lemma, there is a function g so that $g \prec U$ and $g = 1$ on K. Obviously, $g = \chi_E = 1$ on K, and $g = \chi_E = 0$ outside U. Therefore,

$$\mu(\{x \in X \mid \chi_E(x) \neq g(x)\}) \leq \mu(U \setminus K) < \epsilon.$$

We also observe that $\mathrm{supp}(g) \subseteq U$ is compact and $\mu(U) < +\infty$, and so g is 0 outside a compact set of finite measure.

(b) Now we consider a non-negative Borel measurable simple function $\phi : X \to [0,M]$ which is 0 outside some set of finite measure. We may write

$$\phi = \kappa_1 \chi_{E_1} + \cdots + \kappa_k \chi_{E_k},$$

where $E_1, \ldots, E_k$ are pairwise disjoint Borel sets of finite measure and $\kappa_1, \ldots, \kappa_k > 0$. Then from part (a) there are continuous $g_1, \ldots, g_k : X \to [0,1]$ so that $\mu(\{x \in X \mid \chi_{E_j}(x) \neq g_j(x)\}) < \frac{\epsilon}{k}$ for all j, and so that each g_j is 0 outside a compact set of finite measure. Now, we consider

$$h = \kappa_1 g_1 + \cdots + \kappa_k g_k$$

which is continuous and non-negative on X, and which is 0 outside a compact set of finite measure. Then

$$\{x \in X \mid \phi(x) \neq h(x)\} \subseteq \{x \in X \mid \chi_{E_1}(x) \neq g_1(x)\} \cup \cdots \cup \{x \in X \mid \chi_{E_k}(x) \neq g_k(x)\},$$

and so

$$\mu(\{x \in X \mid \phi(x) \neq h(x)\}) < \tfrac{\epsilon}{k} + \cdots + \tfrac{\epsilon}{k} = \epsilon.$$

Now, we take $g = h^{(M)} = \min\{h, M\}$.

Then $g : X \to [0,M]$ is continuous on X and 0 outside a compact set of finite measure. Since

$$\{x \in X \mid \phi(x) \neq g(x)\} \subseteq \{x \in X \mid \phi(x) \neq h(x)\},$$

we get $\mu(\{x \in X \mid \phi(x) \neq g(x)\}) < \epsilon$.

(c) Next let $f : X \to [0,M]$ be Borel measurable and 0 outside some set of finite measure. By Proposition 2.27, there is an *increasing* sequence (ϕ_k) of non-negative Borel measurable simple functions which converges uniformly to f on X. All ϕ_k are 0 outside the same set of finite measure. By taking an appropriate subsequence we may assume that

$$0 \leq f - \phi_k \leq \tfrac{1}{2^k} \quad \text{on } X$$

for every k. We consider the *non-negative* Borel measurable simple functions

$$\psi_1 = \phi_1, \quad \psi_k = \phi_k - \phi_{k-1} \text{ for } k \geq 2.$$

All ψ_k are 0 outside the same set of finite measure and it is clear that

$$\sum_{k=1}^{+\infty} \psi_k = f \quad \text{on } X.$$

Moreover,

$$\psi_k \le f - \phi_{k-1} \le \tfrac{1}{2^{k-1}} \quad \text{on } X \text{ for } k \ge 2.$$

Now, from part (b) there are continuous and non-negative g_k on X so that

$$\mu(\{x \in X \,|\, \psi_k(x) \ne g_k(x)\}) < \tfrac{\epsilon}{2^k}, \quad g_k \le \tfrac{1}{2^{k-1}} \text{ on } X \text{ for } k \ge 2.$$

We may also assume that all g_k are 0 outside the same compact set of finite measure. Then the series $\sum_{k=1}^{+\infty} g_k$ converges uniformly on X, and the function

$$\sum_{k=1}^{+\infty} g_k = h$$

is non-negative and continuous on X and is 0 outside a compact set of finite measure. We also have that

$$\{x \in X \,|\, f(x) \ne h(x)\} \subseteq \bigcup_{k=1}^{+\infty} \{x \in X \,|\, \psi_k(x) \ne g_k(x)\},$$

and so

$$\mu(\{x \in X \,|\, f(x) \ne h(x)\}) < \sum_{k=1}^{+\infty} \tfrac{\epsilon}{2^k} = \epsilon.$$

Finally, we consider $g = h^{(M)} = \min\{h, M\}$.
Then $g : X \to [0, M]$ is continuous on X and 0 outside a compact set of finite measure. Since

$$\{x \in X \,|\, f(x) \ne g(x)\} \subseteq \{x \in X \,|\, f(x) \ne h(x)\},$$

we get $\mu(\{x \in X \,|\, f(x) \ne g(x)\}) < \epsilon$.
(d) Let $f : X \to [0, +\infty]$ be Borel measurable and finite a.e. on X and 0 outside some set, say E, of finite measure.
We consider the sets

$$F_k = \{x \in X \,|\, k < f(x)\}$$

for $k \in \mathbb{N}$. Then (F_k) is decreasing and

$$\bigcap_{k=1}^{+\infty} F_k = \{x \in X \,|\, f(x) = +\infty\}.$$

Since $F_1 \subseteq E$, we have $\mu(F_1) < +\infty$, and so

$$\lim_{k\to+\infty} \mu(F_k) = \mu(\{x \in X \,|\, f(x) = +\infty\}) = 0.$$

Therefore, there is some M so that $\mu(F_M) < \frac{\epsilon}{2}$. Now we consider $f^{(M)} = \min\{f, M\}$, and then $f^{(M)} : X \to [0, M]$ is Borel measurable, $f^{(M)} = 0$ outside E, and

$$\mu(\{x \in X \,|\, f(x) \ne f^{(M)}(x)\}) = \mu(F_M) < \tfrac{\epsilon}{2}.$$

From part (c) there is a continuous $g : X \to [0, +\infty)$ which is 0 outside a compact set of finite measure so that

$$\mu(\{x \in X \,|\, f^{(M)}(x) \ne g(x)\}) < \tfrac{\epsilon}{2}.$$

Since

$$\{x \in X \,|\, f(x) \ne g(x)\} \subseteq \{x \in X \,|\, f^{(M)}(x) \ne g(x)\} \cup \{x \in X \,|\, f(x) \ne f^{(M)}(x)\},$$

we get $\mu(\{x \in X \,|\, f(x) \ne g(x)\}) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$.
We have finished the proof in the case of functions $f : X \to [0, +\infty]$. By considering the non-negative and non-positive parts of a function $f : X \to \overline{\mathbb{R}}$, and, after that, the real and imaginary parts of a function $f : X \to \overline{\mathbb{C}}$, we can easily finish the proof in the general case. We leave these last details as an exercise. $\square$

Loosely speaking, *every Borel measurable function which is finite a.e. on X and 0 outside a set of finite measure is equal to a continuous function with compact support except on a set of arbitrarily small measure*.

We recall that Theorem 1.2 gives conditions on a Hausdorff topological space X and a Borel measure μ on X so that μ is regular.

Exercises.

2.3.1. *Is it possible to nullify the set of non-equality in Lusin's Theorem?*
Take $\chi_{[0,+\infty)} : \mathbb{R} \to \mathbb{R}$ and prove that there is no continuous $g : \mathbb{R} \to \mathbb{R}$ so that $\chi_{[0,+\infty)} = g$ m_1-a.e. on $\mathbb{R}$

2.3.2. Let X, Y be topological spaces of which Y is Hausdorff, and μ be a Borel measure on X so that $\mu(U) > 0$ for every non-empty open $U \subseteq X$. Prove that, if $f, g : X \to Y$ are continuous and $f = g$ a.e. on X, then $f = g$ on X.

2.3.3. (a) Let μ be a Borel measure on the topological space X and $f : X \to \mathbb{C}$ be a Borel measurable function. In the spirit of exercise 1.5.4 about supports of Borel measures, a point $x \in X$ is called a **support point** for f if $\mu(\{y \in U_x \mid f(y) \neq 0\}) > 0$ for every open neighborhood U_x of x. The set

$$\text{supp}(f) = \{x \in X \mid x \text{ is a support point for } f\}$$

is called the **support** of f.

(i) Prove that $\text{supp}(f)$ is a closed set.

(ii) If X is Hausdorff, prove that $\mu(\{x \in K \mid f(x) \neq 0\}) = 0$ for all compact $K \subseteq (\text{supp}(f))^c$.

(iii) If X is Hausdorff and μ is regular, prove that $f = 0$ a.e on $(\text{supp}(f))^c$, and that $(\text{supp}(f))^c$ is the largest open set on which $f = 0$ a.e.

(b) Assume that the μ appearing in (a) has the additional property that $\mu(U) > 0$ for every open U. Use exercise 2.3.2 to prove that for any continuous $f : X \to \mathbb{C}$ the two definitions of $\text{supp}(f)$ (the usual one, which we mentioned in this section, and the one in (a)) coincide.

2.3.4. Let $f : \mathbb{R}^n \to \mathbb{R}$ be continuous at m_n-a.e. $x \in \mathbb{R}^n$. Prove that f is Lebesgue measurable.

2.3.5. Let X be a locally compact and Hausdorff topological space so that for every open set O there is an increasing sequence of compact subsets of O which cover O. If, moreover, μ is a Borel measure on X such that $\mu(K) < +\infty$ for every compact set K, prove that μ is regular.
Hint. Prove that there is an increasing sequence of open sets of finite μ-measure which cover X and then use Theorem 1.2.

2.3.6. Let X be a locally compact and Hausdorff topological space which is separable, i.e. there is a countable dense subset of X. Prove that for every open set O there is an increasing sequence of compact subsets of O whose interiors cover O. If, moreover, μ is a Borel measure on X such that $\mu(K) < +\infty$ for every compact set K, prove that μ is regular.

Chapter 3

Integrals.

3.1 Integrals.

In this whole section (except in the last subsection about point-mass distributions) $(X, \mathcal{S}, \mu)$ will be a general but fixed measure space. At some places we may also deal with a second measure space $(X, \mathcal{S}, \nu)$.

INTEGRALS OF NON-NEGATIVE SIMPLE FUNCTIONS.

Definition. *Let $\phi : X \to [0, +\infty)$ be a non-negative measurable simple function. If $\phi = \sum_{k=1}^{m} \kappa_k \chi_{E_k}$ is the standard representation of ϕ, we define*

$$\int_X \phi \, d\mu = \sum_{k=1}^{m} \kappa_k \mu(E_k).$$

*We say that $\int_X \phi \, d\mu$ is the **integral** of ϕ over X with respect to μ or, shortly, the **μ-integral** of ϕ. Sometimes we want to see the independent variable in the integral and we write $\int_X \phi(x) \, d\mu(x)$.*

If there is no danger of confusion, we shall simply say *integral* instead of μ-integral.

In the definition of $\int_X \phi \, d\mu$ we observe that if one of the values κ_k of ϕ is equal to 0, then, even if the corresponding set E_k has infinite measure, the product $\kappa_k \mu(E_k)$ is equal to 0. Therefore, *the set where $\phi = 0$ does not matter for the calculation of the integral of ϕ.*

Example. We consider the measure space $(X, \mathcal{P}(X), \delta_{x_0})$ for some $x_0 \in X$. Every simple function $\phi : X \to [0, +\infty)$ is measurable, and let $\phi = \sum_{k=1}^{m} \kappa_k \chi_{E_k}$ be the standard representation of ϕ. Then x_0 belongs to exactly one E_k, say E_{k_0}. Now, $\delta_{x_0}(E_{k_0}) = 1$, and $\delta_{x_0}(E_k) = 0$ for $k \neq k_0$. Also, $\chi_{E_{k_0}}(x_0) = 1$, and $\chi_{E_k}(x_0) = 0$ for $k \neq k_0$. Hence,

$$\int_X \phi \, d\delta_{x_0} = \sum_{k=1}^{m} \kappa_k \delta_{x_0}(E_k) = \kappa_{k_0} = \sum_{k=1}^{m} \kappa_k \chi_{E_k}(x_0) = \phi(x_0).$$

Proposition 3.1. *Let $\phi = \sum_{j=1}^{n} \lambda_j \chi_{F_j}$, where $0 \leq \lambda_j < +\infty$ for all j and the sets $F_j \in \mathcal{S}$ are pairwise disjoint. Then $\int_X \phi \, d\mu = \sum_{j=1}^{n} \lambda_j \mu(F_j)$.*

Proof. The representation $\phi = \sum_{j=1}^{n} \lambda_j \chi_{F_j}$ in the statement may not be the standard representation of the simple function ϕ. In fact, the numbers λ_j are not assumed different, and it is not assumed either that the sets F_j are non-empty or that they cover X.
(a) If all F_j are empty, then $\chi_{F_j} = 0$ on X for all j, and we get $\phi = 0 = 0 \, \chi_X$ as the standard representation of ϕ. Therefore

$$\int_X \phi \, d\mu = 0 \, \mu(X) = 0 = \sum_{j=1}^{n} \lambda_j \mu(F_j),$$

since $\mu(F_j) = 0$ for all j.
(b) If at least one F_j is non-empty, we rearrange so that $F_1 \neq \emptyset, \ldots, F_l \neq \emptyset, F_{l+1} = \emptyset, \ldots, F_n = \emptyset$. (We may have $l = n$.) Then

$$\phi = \sum_{j=1}^{l} \lambda_j \chi_{F_j}, \quad \sum_{j=1}^{n} \lambda_j \mu(F_j) = \sum_{j=1}^{l} \lambda_j \mu(F_j),$$

and the equality to be proved becomes $\int_X \phi\, d\mu = \sum_{j=1}^{l} \lambda_j \mu(F_j)$.
If the F_j do not cover X, we introduce the non-empty set $F_{l+1} = (F_1 \cup \cdots \cup F_l)^c$ and the value $\lambda_{l+1} = 0$. We can then write

$$\phi = \sum_{j=1}^{l+1} \lambda_j \chi_{F_j}, \quad \sum_{j=1}^{l} \lambda_j \mu(F_j) = \sum_{j=1}^{l+1} \lambda_j \mu(F_j),$$

and the equality to be proved becomes $\int_X \phi\, d\mu = \sum_{j=1}^{l+1} \lambda_j \mu(F_j)$.
In any case, using the symbol k for l or $l+1$, we have to prove that, if $\phi = \sum_{j=1}^{k} \lambda_j \chi_{F_j}$, where all $F_j \in \mathcal{S}$ are non-empty, pairwise disjoint and cover X, then $\int_X \phi\, d\mu = \sum_{j=1}^{k} \lambda_j \mu(F_j)$.
It is clear that $\lambda_1, \ldots, \lambda_k$ are *all* the values of ϕ on X, *perhaps with repetitions*. We rearrange in groups, so that

$$\begin{aligned} &\lambda_1 = \cdots = \lambda_{k_1} = \kappa_1, \\ &\lambda_{k_1+1} = \cdots = \lambda_{k_1+k_2} = \kappa_2, \\ &\cdots \\ &\lambda_{k_1+\cdots+k_{m-1}+1} = \cdots = \lambda_{k_1+\cdots+k_m} = \kappa_m \end{aligned}$$

are the *different* values of ϕ on X (and, of course, $k_1 + \cdots + k_m = k$). For every $i = 1, \ldots, m$ we define

$$E_i = \bigcup_{j=k_1+\cdots+k_{i-1}+1}^{k_1+\cdots+k_i} F_j = \{x \in X \mid \phi(x) = \kappa_i\},$$

and then $\phi = \sum_{i=1}^{m} \kappa_i \chi_{E_i}$ is the standard representation of ϕ.
By the definition of $\int_X \phi\, d\mu$, we get

$$\begin{aligned} \int_X \phi\, d\mu &= \sum_{i=1}^{m} \kappa_i \mu(E_i) = \sum_{i=1}^{m} \kappa_i \big(\sum_{j=k_1+\cdots+k_{i-1}+1}^{k_1+\cdots+k_i} \mu(F_j) \big) \\ &= \sum_{i=1}^{m} \big(\sum_{j=k_1+\cdots+k_{i-1}+1}^{k_1+\cdots+k_i} \lambda_j \mu(F_j) \big) = \sum_{j=1}^{k} \lambda_j \mu(F_j), \end{aligned}$$

and the proof is complete. □

Proposition 3.2. *(i) If $\phi, \psi : X \to [0, +\infty)$ are measurable simple functions and $0 \leq \lambda < +\infty$, then $\int_X (\phi + \psi)\, d\mu = \int_X \phi\, d\mu + \int_X \psi\, d\mu$ and $\int_X \lambda\phi\, d\mu = \lambda \int_X \phi\, d\mu$.*
(ii) If μ, ν are measures and $\phi : X \to [0, +\infty)$ is a measurable simple function and $0 \leq \lambda < +\infty$, then $\int_X \phi\, d(\mu + \nu) = \int_X \phi\, d\mu + \int_X \phi\, d\nu$ and $\int_X \phi\, d(\lambda\mu) = \lambda \int_X \phi\, d\mu$.

Proof. (i) If $\lambda = 0$, then $\lambda\phi = 0 = 0\,\chi_X$ is the standard representation of $\lambda\phi$, and so

$$\int_X \lambda\phi\, d\mu = 0\,\mu(X) = 0 = 0 \int_X \phi\, d\mu = \lambda \int_X \phi\, d\mu.$$

Now let $0 < \lambda < +\infty$. If $\phi = \sum_{j=1}^{m} \kappa_j \chi_{E_j}$ is the standard representation of ϕ, then $\lambda\phi = \sum_{j=1}^{m} \lambda\kappa_j \chi_{E_j}$ is the standard representation of $\lambda\phi$. Hence,

$$\int_X \lambda\phi\, d\mu = \sum_{j=1}^{m} \lambda\kappa_j \mu(E_j) = \lambda \sum_{j=1}^{m} \kappa_j \mu(E_j) = \lambda \int_X \phi\, d\mu.$$

Now, let $\phi = \sum_{j=1}^{m} \kappa_j \chi_{E_j}$ and $\psi = \sum_{i=1}^{n} \lambda_i \chi_{F_i}$ be the standard representations of ϕ and ψ. It is trivial to see that

$$X = \bigcup_{1 \leq j \leq m, 1 \leq i \leq n} (E_j \cap F_i)$$

and that the sets $E_j \cap F_i \in \mathcal{S}$ are pairwise disjoint. It is also clear that $\phi + \psi$ is constant $\kappa_j + \lambda_i$ on each $E_j \cap F_i$, and so

$$\phi + \psi = \sum_{1 \leq j \leq m, 1 \leq i \leq n} (\kappa_j + \lambda_i) \chi_{E_j \cap F_i}.$$

Proposition 3.1 implies

$$\begin{aligned}\int_X(\phi+\psi)\,d\mu &= \sum_{1\le j\le m,1\le i\le n}(\kappa_j+\lambda_i)\mu(E_j\cap F_i)\\ &= \sum_{1\le j\le m,1\le i\le n}\kappa_j\mu(E_j\cap F_i)+\sum_{1\le j\le m,1\le i\le n}\lambda_i\mu(E_j\cap F_i)\\ &= \sum_{j=1}^m\kappa_j\big(\sum_{i=1}^n\mu(E_j\cap F_i)\big)+\sum_{i=1}^n\lambda_i\big(\sum_{j=1}^m\mu(E_j\cap F_i)\big)\\ &= \sum_{j=1}^m\kappa_j\mu(E_j)+\sum_{i=1}^n\lambda_i\mu(F_i)=\int_X\phi\,d\mu+\int_X\psi\,d\mu,\end{aligned}$$

(ii) Let $\phi=\sum_{j=1}^m\kappa_j\chi_{E_j}$ be the standard representation of ϕ. Then

$$\begin{aligned}\int_X\phi\,d(\mu+\nu) &= \sum_{j=1}^m\kappa_j(\mu+\nu)(E_j)=\sum_{j=1}^m\kappa_j(\mu(E_j)+\nu(E_j))\\ &= \sum_{j=1}^m\kappa_j\mu(E_j)+\sum_{j=1}^m\kappa_j\nu(E_j)=\int_X\phi\,d\mu+\int_X\phi\,d\nu.\end{aligned}$$

Also,

$$\int_X\phi\,d(\lambda\mu)=\sum_{j=1}^m\kappa_j(\lambda\mu)(E_j)=\sum_{j=1}^m\kappa_j\lambda\,\mu(E_j)=\lambda\sum_{j=1}^m\kappa_j\mu(E_j)=\lambda\int_X\phi\,d\mu$$

and the proof is complete. □

Proposition 3.3. *(i) If $\phi,\psi:X\to[0,+\infty)$ are measurable simple functions and if $\phi\le\psi$ on X, then $\int_X\phi\,d\mu\le\int_X\psi\,d\mu$.*
(ii) If μ,ν are measures so that $\mu\le\nu$ and $\phi:X\to[0,+\infty)$ is a measurable simple function, then $\int_X\phi\,d\mu\le\int_X\phi\,d\nu$.

Proof. (i) Let $\phi=\sum_{j=1}^m\kappa_j\chi_{E_j}$ and $\psi=\sum_{i=1}^n\lambda_i\chi_{F_i}$ be the standard representations of ϕ and ψ. Whenever $E_j\cap F_i\ne\emptyset$, we take any $x\in E_j\cap F_i$, and we find

$$\kappa_j=\phi(x)\le\psi(x)=\lambda_i,$$

and so $\kappa_j\mu(E_j\cap F_i)\le\lambda_i\mu(E_j\cap F_i)$. The same is obviously true even when $E_j\cap F_i=\emptyset$. Therefore,

$$\begin{aligned}\int_X\phi\,d\mu &= \sum_{j=1}^m\kappa_j\mu(E_j)=\sum_{1\le j\le m,1\le i\le n}\kappa_j\mu(E_j\cap F_i)\\ &\le \sum_{1\le j\le m,1\le i\le n}\lambda_i\mu(E_j\cap F_i)=\sum_{i=1}^n\lambda_i\mu(F_i)=\int_X\psi\,d\mu.\end{aligned}$$

(ii) Let $\phi=\sum_{j=1}^m\kappa_j\chi_{E_j}$ be the standard representation of ϕ. Then

$$\int_X\phi\,d\mu=\sum_{j=1}^m\kappa_j\mu(E_j)\le\sum_{j=1}^m\kappa_j\nu(E_j)=\int_X\phi\,d\nu,$$

since $\mu(E_j)\le\nu(E_j)$ and $\kappa_j\ge0$ for all j. □

Proposition 3.4. *Let $\phi:X\to[0,+\infty)$ be a measurable simple function and (A_n) be an increasing sequence of measurable sets so that $\bigcup_{n=1}^{+\infty}A_n=X$. Then $\lim_{n\to+\infty}\int_X\phi\chi_{A_n}\,d\mu=\int_X\phi\,d\mu$.*

Proof. Let $\phi=\sum_{j=1}^m\kappa_j\chi_{E_j}$ be the standard representation of ϕ. Then we have

$$\phi\chi_{A_n}=\sum_{j=1}^m\kappa_j\chi_{E_j}\chi_{A_n}=\sum_{j=1}^m\kappa_j\chi_{E_j\cap A_n},$$

and Proposition 3.1 implies

$$\lim_{n\to+\infty}\int_X\phi\chi_{A_n}\,d\mu=\lim_{n\to+\infty}\sum_{j=1}^m\kappa_j\mu(E_j\cap A_n)=\sum_{j=1}^m\kappa_j\mu(E_j)=\int_X\phi\,d\mu,$$

since $(E_j\cap A_n)$ is increasing, $\bigcup_{n=1}^{+\infty}(E_j\cap A_n)=E_j$, and μ is continuous from below. □

Proposition 3.5. *Let $\phi,\phi_1,\phi_2,\dots:X\to[0,+\infty)$ be measurable simple functions so that the sequence (ϕ_n) is increasing on X.*
(i) If $\lim_{n\to+\infty}\phi_n\le\phi$ on X, then $\lim_{n\to+\infty}\int_X\phi_n\,d\mu\le\int_X\phi\,d\mu$.
(ii) If $\phi\le\lim_{n\to+\infty}\phi_n$ on X, then $\int_X\phi\,d\mu\le\lim_{n\to+\infty}\int_X\phi_n\,d\mu$.

Proof. By Proposition 3.3, the sequence $\big(\int_X \phi_n\,d\mu\big)$ is increasing, and so $\lim_{n\to+\infty}\int_X \phi_n\,d\mu$ exists in $[0,+\infty]$.
(i) Proposition 3.3 implies $\int_X \phi_n\,d\mu \le \int_X \phi\,d\mu$ for all n, and so $\lim_{n\to+\infty}\int_X \phi_n\,d\mu \le \int_X \phi\,d\mu$.
(ii) We consider an arbitrary $\alpha\in[0,1)$, and we define

$$A_n=\{x\in X\,|\,\alpha\phi(x)\le\phi_n(x)\}\in\mathcal{S}.$$

It is easy to see that (A_n) is increasing and $\bigcup_{n=1}^{+\infty}A_n=X$. Now, we have that $\alpha\phi\chi_{A_n}\le\phi_n$ on X, and Propositions 3.2, 3.3 and 3.4 imply

$$\alpha\int_X \phi\,d\mu=\int_X \alpha\phi\,d\mu=\lim_{n\to+\infty}\int_X \alpha\phi\chi_{A_n}\,d\mu\le\lim_{n\to+\infty}\int_X \phi_n\,d\mu.$$

From this we get $\int_X \phi\,d\mu\le\lim_{n\to+\infty}\int_X \phi_n\,d\mu$ by taking the limit as $\alpha\to 1-$, □

Proposition 3.6. *Let $\phi_1,\psi_1,\phi_2,\psi_2\ldots : X\to[0,+\infty)$ be measurable simple functions so that the sequences (ϕ_n) and (ψ_n) are increasing on X. If $\lim_{n\to+\infty}\phi_n=\lim_{n\to+\infty}\psi_n$ on X, then $\lim_{n\to+\infty}\int_X \phi_n\,d\mu=\lim_{n\to+\infty}\int_X \psi_n\,d\mu$.*

Proof. For each k we have

$$\psi_k\le\lim_{n\to+\infty}\psi_n=\lim_{n\to+\infty}\phi_n$$

on X. Now, Proposition 3.5 implies

$$\int_X \psi_k\,d\mu\le\lim_{n\to+\infty}\int_X \phi_n\,d\mu.$$

Taking the limit as $k\to+\infty$, we get $\lim_{n\to+\infty}\int_X \psi_n\,d\mu\le\lim_{n\to+\infty}\int_X \phi_n\,d\mu$.
The reverse inequality is proved symmetrically. □

Proposition 3.7. *Let $\phi : X\to[0,+\infty)$ be a measurable simple function. Then $\int_X \phi\,d\mu=0$ if and only if $\phi=0$ a.e. on X.*

Proof. If $\phi=\sum_{j=1}^m\kappa_j\chi_{E_j}$ is the standard representation of ϕ, then $\int_X \phi\,d\mu=\sum_{k=1}^m\kappa_k\mu(E_k)$. Hence, $\int_X \phi\,d\mu=0$ if and only if $\mu(E_k)=0$ for all k for which $\kappa_k>0$.
Now, since

$$\bigcup_{k:\,\kappa_k>0}E_k=\{x\in X\,|\,\phi(x)>0\},$$

we get

$$\sum_{k:\,\kappa_k>0}\mu(E_k)=\mu(\{x\in X\,|\,\phi(x)>0\}).$$

Thus, $\int_X \phi\,d\mu=0$ if and only if $\mu(\{x\in X\,|\,\phi(x)>0\})=0$ if and only if $\phi=0$ a.e. on X. □

INTEGRALS OF NON-NEGATIVE FUNCTIONS.

In this subsection we shall take for granted the notion of the integral $\int_X \phi\,d\mu$ for measurable simple functions $\phi : X\to[0,+\infty)$ and also all the relevant properties which we saw in the previous subsection.

Definition. *Let $f : X\to[0,+\infty]$ be a measurable function. We define the **integral** of f over X with respect to μ or, shortly, the **μ-integral** of f by*

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

where (ϕ_n) is any increasing sequence of non-negative measurable simple functions on X such that $\lim_{n\to+\infty}\phi_n=f$ on X.
*We say that f is **integrable** over X with respect to μ or **μ-integrable** over X if $\int_X f\,d\mu$ is finite, i.e. $\int_X f\,d\mu<+\infty$.*
We may use the symbol $\int_X f(x)\,d\mu(x)$ if we want to see the independent variable in the integral.

Proposition 3.6 guarantees that $\int_X f\,d\mu$ is well defined and Proposition 2.27 implies the existence of at least one sequence (ϕ_n) as in the definition.

If there is no danger of confusion, we shall simply say *integral* and *integrable* instead of μ-integral and μ-integrable.

Example. We consider the measure space $(X, \mathcal{P}(X), \delta_{x_0})$ for some $x_0 \in X$. Every function $f : X \to [0, +\infty]$ is measurable, and let (ϕ_n) be any increasing sequence of non-negative simple functions on X such that $\lim_{n\to+\infty} \phi_n = f$ on X. We have shown that $\int_X \phi_n\,d\delta_{x_0} = \phi_n(x_0)$ for every n, and we get

$$\int_X f\,d\delta_{x_0} = f(x_0)$$

by taking the limit as $n \to +\infty$.

Proposition 3.8. *(i) Let $f, g : X \to [0, +\infty]$ be measurable functions and let $\lambda \in [0, +\infty)$. Then $\int_X (f+g)\,d\mu = \int_X f\,d\mu + \int_X g\,d\mu$ and $\int_X \lambda f\,d\mu = \lambda \int_X f\,d\mu$.*
(ii) If μ, ν are measures and $f : X \to [0, +\infty]$ is a measurable function and $0 \le \lambda < +\infty$, then $\int_X f\,d(\mu+\nu) = \int_X f\,d\mu + \int_X f\,d\nu$ and $\int_X f\,d(\lambda\mu) = \lambda \int_X f\,d\mu$.

Proof. We consider increasing sequences (ϕ_n) and (ψ_n) of non-negative measurable simple functions on X so that $\lim_{n\to+\infty} \phi_n = f$ and $\lim_{n\to+\infty} \psi_n = g$ on X.
(i) Now, $(\phi_n + \psi_n)$ and $(\lambda\phi_n)$ are increasing sequences of non-negative measurable simple functions on X such that $\lim_{n\to+\infty}(\phi_n + \psi_n) = f + g$ and $\lim_{n\to+\infty} \lambda\phi_n = \lambda f$ on X.
We know that

$$\int_X (\phi_n + \psi_n)\,d\mu = \int_X \phi_n\,d\mu + \int_X \psi_n\,d\mu, \quad \int_X \lambda\phi_n\,d\mu = \lambda \int_X \phi_n\,d\mu$$

for all n. These imply $\int_X (f+g)\,d\mu = \int_X f\,d\mu + \int_X g\,d\mu$ and $\int_X \lambda f\,d\mu = \lambda \int_X f\,d\mu$, by taking the limit as $n \to +\infty$.
(ii) We have that

$$\int_X \phi_n\,d(\mu+\nu) = \int_X \phi_n\,d\mu + \int_X \phi_n\,d\nu, \quad \int_X \phi_n\,d(\lambda\mu) = \lambda \int_X \phi_n\,d\mu$$

for all n. These imply $\int_X f\,d(\mu+\nu) = \int_X f\,d\mu + \int_X f\,d\nu$ and $\int_X f\,d(\lambda\mu) = \lambda \int_X f\,d\mu$, by taking the limit as $n \to +\infty$. ☐

Proposition 3.9. *(i) Let $f, g : X \to [0, +\infty]$ be measurable functions such that $f \le g$ on X. Then $\int_X f\,d\mu \le \int_X g\,d\mu$.*
(ii) If μ, ν are measures so that $\mu \le \nu$ and $f : X \to [0, +\infty]$ is a measurable function, then $\int_X f\,d\mu \le \int_X f\,d\nu$.

Proof. We consider increasing sequences (ϕ_n) and (ψ_n) of non-negative measurable simple functions on X so that $\lim_{n\to+\infty} \phi_n = f$ and $\lim_{n\to+\infty} \psi_n = g$ on X.
(i) For every k we have that $\phi_k \le f \le g = \lim_{n\to+\infty} \psi_n$ on X, and Proposition 3.5 implies

$$\int_X \phi_k\,d\mu \le \lim_{n\to+\infty} \int_X \psi_n\,d\mu = \int_X g\,d\mu.$$

Taking the limit as $k \to +\infty$, we conclude that $\int_X f\,d\mu \le \int_X g\,d\mu$.
(ii) We have that

$$\int_X \phi_n\,d\mu \le \int_X \phi_n\,d\nu$$

for all n, and, taking the limit as $n \to +\infty$, we find $\int_X f\,d\mu \le \int_X f\,d\nu$. ☐

Proposition 3.10. *Let $f : X \to [0, +\infty]$ be measurable. Then $\int_X f\,d\mu = 0$ if and only if $f = 0$ a.e. on X.*

Proof. We take an increasing sequence (ϕ_n) of non-negative measurable simple functions on X so that $\lim_{n\to+\infty}\phi_n = f$ on X, and then $\lim_{n\to+\infty}\int_X \phi_n\,d\mu = \int_X f\,d\mu$.
Let $\int_X f\,d\mu = 0$. Since the sequence $\big(\int_X \phi_n\,d\mu\big)$ of non-negative numbers is increasing, we have that $\int_X \phi_n\,d\mu = 0$ for all n. Then $\phi_n = 0$ a.e. on X for all n, and so $f = 0$ a.e. on X.
Conversely, let $f = 0$ a.e. on X. For every n we have $0 \le \phi_n \le f$ on X, and so $\phi_n = 0$ a.e. on X. Then $\int_X \phi_n\,d\mu = 0$ for all n, and so $\int_X f\,d\mu = 0$. $\square$

Proposition 3.11. *Let $f : X \to [0,+\infty]$ be integrable. Then*
(i) $f(x) < +\infty$ for a.e. $x \in X$,
(ii) the set $\{x \in X \mid f(x) > 0\}$ is of σ-finite measure.

Proof. (i) We consider the set $B = \{x \in X \mid f(x) = +\infty\} \in \mathcal{S}$ and any $r \in (0,+\infty)$.
Now, we have that $r\chi_B \le f$ on X, and Proposition 3.9 implies

$$r\mu(B) = \int_X r\chi_B\,d\mu \le \int_X f\,d\mu < +\infty.$$

This implies $\mu(B) \le \frac{1}{r}\int_X f\,d\mu$, and, taking the limit as $r \to +\infty$, we find $\mu(B) = 0$.
(ii) We consider the sets $A = \{x \in X \mid f(x) \ne 0\}$ and $A_\epsilon = \{x \in X \mid |f(x)| \ge \epsilon\}$ for $\epsilon > 0$.
Then $\epsilon\chi_{A_\epsilon} \le f$ on X, and exactly as before, we get

$$\epsilon\mu(A_\epsilon) = \int_X \epsilon\chi_{A_\epsilon}\,d\mu \le \int_X f\,d\mu < +\infty.$$

Thus, $\mu(A_\epsilon) < +\infty$ for all $\epsilon > 0$. Since $A = \bigcup_{n=1}^{+\infty} A_{1/n}$, we get that A is of σ-finite measure. $\square$

Proposition 3.12. *Let $f, g : X \to [0,+\infty]$ be measurable and $f = g$ a.e. on X. Then*
(i) $\int_X g\,d\mu = \int_X f\,d\mu$,
(ii) if f is integrable, then g is integrable.

Proof. (i) We consider the set $A = \{x \in X \mid f(x) = g(x)\} \in \mathcal{S}$, and then $\mu(A^c) = 0$.
We have that $f\chi_{A^c} = 0$ a.e. on X, and Propositions 3.8 and 3.10 imply

$$\int_X f\,d\mu = \int_X (f\chi_A + f\chi_{A^c})\,d\mu = \int_X f\chi_A\,d\mu + \int_X f\chi_{A^c}\,d\mu = \int_X f\chi_A\,d\mu.$$

Similarly, we get $\int_X g\,d\mu = \int_X g\chi_A\,d\mu$.
Now, since $f\chi_A = g\chi_A$ on X, we find $\int_X f\,d\mu = \int_X g\,d\mu$.
(ii) If f is integrable, then $\int_X f\,d\mu < +\infty$. Now, (i) gives $\int_X g\,d\mu < +\infty$, and g is integrable. $\square$

Exercises.

3.1.1. Let $f : X \to [0,+\infty]$ be measurable. Let $\Delta = \{E_1,\dots,E_l\}$, where $l \in \mathbb{N}$ and the non-empty sets $E_1,\dots,E_l \in \mathcal{S}$ are pairwise disjoint and cover X. Such a Δ is called **$\mathcal{S}$-partition** of X. We define $\underline{S}(f,\Delta) = \sum_{j=1}^{l} m_j\mu(E_j)$, where $m_j = \inf\{f(x) \mid x \in E_j\}$.
Prove that $\int_X f\,d\mu = \sup\{\underline{S}(f,\Delta) \mid \Delta \text{ is a } \mathcal{S}\text{-partition of } X\}$.

INTEGRALS OF EXTENDED-REAL VALUED FUNCTIONS.

Now we shall take for granted the notion of the integral $\int_X f\,d\mu$ for measurable $f : X \to [0,+\infty]$ and also all the relevant properties which we saw in the two previous subsections.

Definition. *Let $f : X \to \overline{\mathbb{R}}$ be a measurable function and $f^+, f^- : X \to [0,+\infty]$ be the non-negative and non-positive parts of f. If at least one of $\int_X f^+\,d\mu$ and $\int_X f^-\,d\mu$ is finite, we define*

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

*Then $\int_X f\,d\mu$ is called the **integral** of f over X with respect to μ or, simply, the **μ-integral** of f. We say that f is **integrable** over X with respect to μ or **μ-integrable** over X if $\int_X f\,d\mu$ is finite. As in the case of non-negative functions, we may write $\int_X f(x)\,d\mu(x)$ if we want to see the independent variable in the integral.*

If there is no danger of confusion, we shall say *integral* and *integrable* instead of μ-integral and μ-integrable.

We note that, if $\int_X f^+ \, d\mu = \int_X f^- \, d\mu = +\infty$, then $\int_X f \, d\mu$ is *not* defined. On the other hand, if $\int_X f^+ \, d\mu = +\infty$ and $\int_X f^- \, d\mu < +\infty$, then $\int_X f \, d\mu = +\infty$. Also, if $\int_X f^+ \, d\mu < +\infty$ and $\int_X f^- \, d\mu = +\infty$, then $\int_X f \, d\mu = -\infty$. Finally, if $\int_X f^+ \, d\mu < +\infty$ and $\int_X f^- \, d\mu < +\infty$, then $\int_X f \, d\mu$ is a real number and so f is integrable.

Example. We consider the measure space $(X, \mathcal{P}(X), \delta_{x_0})$ for some $x_0 \in X$. Then every function $f : X \to \overline{\mathbb{R}}$ is measurable.
We know that $\int_X f^+ \, d\delta_{x_0} = f^+(x_0)$ and $\int_X f^- \, d\delta_{x_0} = f^-(x_0)$. Since at least one of $f^+(x_0)$ and $f^-(x_0)$ equals 0, we have that $\int_X f \, d\delta_{x_0}$ is defined. Subtracting the two equalities, we get

$$\int_X f \, d\delta_{x_0} = f(x_0).$$

Thus, *integration with respect to the Dirac measure at x_0 coincides with the so-called **point evaluation** at x_0.*

Proposition 3.13. *Let $f : X \to \overline{\mathbb{R}}$ be measurable. Then f is integrable if and only if f^+ and f^- are integrable if and only if $|f|$ is integrable.*

Proof. The first equivalence is clear from the definition. The second equivalence is due to Proposition 3.8 and the equality $f^+ + f^- = |f|$ on X. ☐

Proposition 3.14. *Let $f : X \to \overline{\mathbb{R}}$ be integrable. Then*
(i) $f(x) \in \mathbb{R}$ for a.e. $x \in X$,
(ii) the set $\{x \in X \mid f(x) \neq 0\}$ is of σ-finite measure.

Proof. Since the integrability of f implies the integrability of $|f|$, the result is immediate by applying Proposition 3.11 to $|f|$. ☐

Proposition 3.15. *Let $f, g : X \to \overline{\mathbb{R}}$ be measurable and $f = g$ a.e. on X. Then*
(i) if $\int_X f \, d\mu$ is defined, then $\int_X g \, d\mu$ is defined and $\int_X g \, d\mu = \int_X f \, d\mu$,
(ii) if f is integrable, then g is integrable.

Proof. From $f = g$ a.e. on X we get $f^+ = g^+$ a.e. on X and $f^- = g^-$ a.e. on X. Hence,

$$\int_X f^+ \, d\mu = \int_X g^+ \, d\mu, \quad \int_X f^- \, d\mu = \int_X g^- \, d\mu.$$

(i) Now, let $\int_X f \, d\mu$ be defined. Then either $\int_X f^+ \, d\mu$ is finite or $\int_X f^- \, d\mu$ is finite, and so either $\int_X g^+ \, d\mu$ is finite or $\int_X g^- \, d\mu$ is finite, and so $\int_X g \, d\mu$ is defined. Also,

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

(ii) If f is integrable, then $\int_X f \, d\mu$ is a real number. From (i) we have that $\int_X g \, d\mu$ is also a real number, and so g is integrable. ☐

Proposition 3.16. *Let $f, g : X \to \overline{\mathbb{R}}$ be measurable and let us consider any measurable definition of $f + g$. Then*
(i) if $\int_X f \, d\mu, \int_X g \, d\mu$ are both defined and they are not opposite infinities, then $\int_X (f + g) \, d\mu$ is defined and $\int_X (f + g) \, d\mu = \int_X f \, d\mu + \int_X g \, d\mu$,
(ii) if f, g are integrable, then $f + g$ is integrable.

Proof. (i) If $\int_X f \, d\mu, \int_X g \, d\mu$ are both defined and they are not opposite infinities, then *either* $\int_X f^- \, d\mu < +\infty, \int_X g^- \, d\mu < +\infty$ *or* $\int_X f^+ \, d\mu < +\infty, \int_X g^+ \, d\mu < +\infty$.
Let $\int_X f^+ \, d\mu < +\infty$ and $\int_X g^+ \, d\mu < +\infty$.

Proposition 3.11 implies that, if $A = \{x \in X \mid f(x) \neq +\infty,\ g(x) \neq +\infty\}$, then $\mu(A^c) = 0$.
We consider $F = f\chi_A$ and $G = g\chi_A$. Then $F, G : X \to [-\infty, +\infty)$ are measurable, and $F = f$ a.e. on X and $G = g$ a.e. on X. Also, $F^+ = f^+$ a.e. on X and $G^+ = g^+$ a.e. on X.
The advantage of F, G over f, g is that $F(x) + G(x)$ is defined for *every* $x \in X$.
We observe that for all measurable definitions of $f + g$ we have $F + G = f + g$ a.e. on X. Therefore, because of Proposition 3.15, it is enough to prove that $\int_X (F + G)\, d\mu$ is defined and that $\int_X (F + G)\, d\mu = \int_X F\, d\mu + \int_X G\, d\mu$.
From

$$F = F^+ - F^- \leq F^+, \quad G = G^+ - G^- \leq G^+$$

on X we get $F + G \leq F^+ + G^+$ on X, and so $(F + G)^+ \leq F^+ + G^+$ on X. Hence,

$$\int_X (F + G)^+\, d\mu \leq \int_X F^+\, d\mu + \int_X G^+\, d\mu = \int_X f^+\, d\mu + \int_X g^+\, d\mu < +\infty,$$

and so $\int_X (F + G)\, d\mu$ is defined.
We now have

$$(F + G)^+ - (F + G)^- = F + G = (F^+ + G^+) - (F^- + G^-)$$

or, equivalently,

$$(F + G)^+ + F^- + G^- = (F + G)^- + F^+ + G^+.$$

Hence,

$$\int_X (F + G)^+\, d\mu + \int_X F^-\, d\mu + \int_X G^-\, d\mu = \int_X (F + G)^-\, d\mu + \int_X F^+\, d\mu + \int_X G^+\, d\mu.$$

Because of the finiteness of the integrals $\int_X (F + G)^+\, d\mu$, $\int_X F^+\, d\mu$, $\int_X G^+\, d\mu$, we get

$$\begin{aligned}\int_X (F + G)\, d\mu &= \int_X (F + G)^+\, d\mu - \int_X (F + G)^-\, d\mu \\ &= \int_X F^+\, d\mu + \int_X G^+\, d\mu - \int_X F^-\, d\mu - \int_X G^-\, d\mu = \int_X F\, d\mu + \int_X G\, d\mu.\end{aligned}$$

If $\int_X f^-\, d\mu < +\infty$ and $\int_X g^-\, d\mu < +\infty$, then the proof is similar.
(ii) Let f, g be integrable. For every measurable definition of $f + g$ we have $|f + g| \leq |f| + |g|$ on X, and so

$$\int_X |f + g|\, d\mu \leq \int_X |f|\, d\mu + \int_X |g|\, d\mu < +\infty.$$

Hence, $f + g$ is integrable $\square$

Proposition 3.17. *Let $f : X \to \overline{\mathbb{R}}$ be measurable and μ, ν be two measures. Then*
(i) if $\int_X f\, d\mu$, $\int_X f\, d\nu$ are both defined and they are not opposite infinities, then $\int_X f\, d(\mu + \nu)$ is defined and $\int_X f\, d(\mu + \nu) = \int_X f\, d\mu + \int_X f\, d\nu$,
(ii) if f is μ-integrable and ν-integrable, then f is $(\mu + \nu)$-integrable.

Proof. (i) If $\int_X f\, d\mu$, $\int_X f\, d\nu$ are both defined and they are not opposite infinities, then *either* $\int_X f^-\, d\mu < +\infty$, $\int_X f^-\, d\nu < +\infty$ *or* $\int_X f^+\, d\mu < +\infty$, $\int_X f^+\, d\nu < +\infty$.
Let $\int_X f^+\, d\mu < +\infty$ and $\int_X f^+\, d\nu < +\infty$.
Then

$$\int_X f^+\, d(\mu + \nu) = \int_X f^+\, d\mu + \int_X f^+\, d\nu < +\infty,$$

and so $\int_X f\, d(\mu + \nu)$ is defined. We also have that

$$\int_X f^-\, d(\mu + \nu) = \int_X f^-\, d\mu + \int_X f^-\, d\nu,$$

and, subtracting these two equalities, we get $\int_X f\, d(\mu + \nu) = \int_X f\, d\mu + \int_X f\, d\nu$.
If $\int_X f^-\, d\mu < +\infty$ and $\int_X f^-\, d\nu < +\infty$, then the proof is similar.
(ii) Let f be μ-integrable and ν-integrable. Then

$$\int_X |f|\, d(\mu + \nu) = \int_X |f|\, d\mu + \int_X |f|\, d\nu < +\infty,$$

and so f is $(\mu + \nu)$-integrable. $\square$

Proposition 3.18. *Let $f : X \to \overline{\mathbb{R}}$ be measurable and $\lambda \in \mathbb{R}$. Then*
(i) if $\int_X f\,d\mu$ is defined, then $\int_X \lambda f\,d\mu$ is defined and $\int_X \lambda f\,d\mu = \lambda \int_X f\,d\mu$,
(ii) if f is integrable, then λf is integrable.

Proof. (i) Let $\int_X f\,d\mu$ be defined. Then at least one of $\int_X f^+\,d\mu$ and $\int_X f^-\,d\mu$ is finite.
If $\lambda > 0$, then $(\lambda f)^+ = \lambda f^+$ and $(\lambda f)^- = \lambda f^-$. Therefore, at least one of

$$\int_X (\lambda f)^+\,d\mu = \lambda \int_X f^+\,d\mu, \quad \int_X (\lambda f)^-\,d\mu = \lambda \int_X f^-\,d\mu$$

is finite. Hence, $\int_X \lambda f\,d\mu$ is defined, and

$$\int_X \lambda f\,d\mu = \int_X (\lambda f)^+\,d\mu - \int_X (\lambda f)^-\,d\mu = \lambda\big(\int_X f^+\,d\mu - \int_X f^-\,d\mu\big) = \lambda \int_X f\,d\mu.$$

If $\lambda < 0$, then $(\lambda f)^+ = -\lambda f^-$ and $(\lambda f)^- = -\lambda f^+$, and the previous argument can be repeated with no essential change.
If $\lambda = 0$, then the result is trivial.
(ii) If f is integrable, then $\int_X |\lambda f|\,d\mu = |\lambda| \int_X |f|\,d\mu < +\infty$, and so λf is integrable. $\square$

Proposition 3.19. *Let $f : X \to \overline{\mathbb{R}}$ be measurable and $\lambda \in [0,+\infty)$.*
(i) If $\int_X f\,d\mu$ is defined, then $\int_X f\,d(\lambda\mu)$ is defined and $\int_X f\,d(\lambda\mu) = \lambda \int_X f\,d\mu$.
(ii) If f is μ-integrable, then f is $\lambda\mu$-integrable.

Proof. (i) *Either $\int_X f^-\,d\mu < +\infty$ or $\int_X f^+\,d\mu < +\infty$.*
Let $\int_X f^+\,d\mu < +\infty$.
Then

$$\int_X f^+\,d(\lambda\mu) = \lambda \int_X f^+\,d\mu < +\infty,$$

and so $\int_X f\,d(\lambda\mu)$ is defined. We also have that

$$\int_X f^-\,d(\lambda\mu) = \lambda \int_X f^-\,d\mu,$$

and subtracting these two equalities we get the equality in (i).
If $\int_X f^-\,d\mu < +\infty$, then the proof is similar.
(ii) If f is μ-integrable, then $\int_X |f|\,d(\lambda\mu) = \lambda \int_X |f|\,d\mu < +\infty$, and so f is $\lambda\mu$-integrable. $\square$

Proposition 3.20. *Let $f, g : X \to \overline{\mathbb{R}}$ be measurable. If $\int_X f\,d\mu$ and $\int_X g\,d\mu$ are defined and if $f \le g$ on X, then $\int_X f\,d\mu \le \int_X g\,d\mu$.*

Proof. If $\int_X f\,d\mu = -\infty$ or $\int_X g\,d\mu = +\infty$, then the inequality $\int_X f\,d\mu \le \int_X g\,d\mu$ is obviously true. So we assume that $\int_X f^-\,d\mu < +\infty$ and $\int_X g^+\,d\mu < +\infty$.
From $f \le g = g^+ - g^- \le g^+$ we get $f^+ \le g^+$ on X. Similarly, we get $g^- \le f^-$ on X. Therefore,

$$\int_X f^+\,d\mu \le \int_X g^+\,d\mu < +\infty, \quad \int_X g^-\,d\mu \le \int_X f^-\,d\mu < +\infty.$$

So we can subtract the inequalities $\int_X f^+\,d\mu \le \int_X g^+\,d\mu$ and $\int_X g^-\,d\mu \le \int_X f^-\,d\mu$, and then we get $\int_X f\,d\mu \le \int_X g\,d\mu$. $\square$

Proposition 3.21. *Let $f : X \to \overline{\mathbb{R}}$ be measurable. If $\int_X f\,d\mu$ is defined, then $\big|\int_X f\,d\mu\big| \le \int_X |f|\,d\mu$.*

Proof. We have that

$$\big|\int_X f\,d\mu\big| = \big|\int_X f^+\,d\mu - \int_X f^-\,d\mu\big| \le \int_X f^+\,d\mu + \int_X f^-\,d\mu = \int_X (f^+ + f^-)\,d\mu = \int_X |f|\,d\mu$$

since $\int_X f^+\,d\mu \ge 0$ and $\int_X f^-\,d\mu \ge 0$. $\square$

Exercises.

3.1.2. If $f, g, h : X \to \overline{\mathbb{R}}$ are measurable, g, h are integrable and $g \le f \le h$ a.e. on X, prove that f is also integrable.
Hint. Prove that $f^- \le g^-$ a.e. on X and $f^+ \le h^+$ a.e. on X.

INTEGRALS OF EXTENDED-COMPLEX VALUED FUNCTIONS.

Now we shall take for granted the notion of the integral $\int_X f\,d\mu$ for measurable $f : X \to \overline{\mathbb{R}}$ and also all the relevant properties which we saw in the three previous subsections.

Definition. *Let $f : X \to \overline{\mathbb{C}}$ be measurable. Then $|f| : X \to [0, +\infty]$ is measurable, and we say that f is* ***integrable*** *over X with respect to μ or, simply,* ***μ-integrable*** *over X, if $\int_X |f|\,d\mu < +\infty$.*

If there is no danger of confusion, we shall say *integrable* instead of μ-integrable.

Proposition 3.22. *Let $f : X \to \overline{\mathbb{C}}$ be integrable. Then*
(i) $f(x) \in \mathbb{C}$ for a.e. $x \in X$,
(ii) the set $\{x \in X \mid f(x) \neq 0\}$ is of σ-finite measure.

Proof. Immediate application of Proposition 3.11 to $|f|$. □

Let $f : X \to \overline{\mathbb{C}}$ be *integrable*. By Proposition 3.22, the set D_f, defined by

$$D_f = \{x \in X \mid f(x) \in \mathbb{C}\} = f^{-1}(\mathbb{C}) \in \mathcal{S}$$

has a null complement. Thus, the function $f\chi_{D_f}$ is measurable, and

$$f\chi_{D_f} = f \quad \text{a.e. on } X.$$

The advantage of $f\chi_{D_f}$ over f is that $f\chi_{D_f}$ is *complex valued*, i.e. $f\chi_{D_f} : X \to \mathbb{C}$. Therefore, the real and imaginary parts of $f\chi_{D_f}$, namely $\mathrm{Re}(f\chi_{D_f}) : X \to \mathbb{R}$ and $\mathrm{Im}(f\chi_{D_f}) : X \to \mathbb{R}$, are defined on X. We also have that

$$|\mathrm{Re}(f\chi_{D_f})| \leq |f\chi_{D_f}| \leq |f|, \quad |\mathrm{Im}(f\chi_{D_f})| \leq |f\chi_{D_f}| \leq |f|$$

on X. Hence,

$$\int_X |\mathrm{Re}(f\chi_{D_f})|\,d\mu \leq \int_X |f|\,d\mu < +\infty, \quad \int_X |\mathrm{Im}(f\chi_{D_f})|\,d\mu \leq \int_X |f|\,d\mu < +\infty.$$

Thus, $\mathrm{Re}(f\chi_{D_f})$ and $\mathrm{Im}(f\chi_{D_f})$ are integrable real valued functions, and so $\int_X \mathrm{Re}(f\chi_{D_f})\,d\mu$ and $\int_X \mathrm{Im}(f\chi_{D_f})\,d\mu$ are defined and they are real numbers.

Definition. *Let $f : X \to \overline{\mathbb{C}}$ be integrable and $D_f = \{x \in X \mid f(x) \in \mathbb{C}\}$. We define*

$$\int_X f\,d\mu = \int_X \mathrm{Re}(f\chi_{D_f})\,d\mu + i\int_X \mathrm{Im}(f\chi_{D_f})\,d\mu,$$

and we call it the ***integral*** *of f over X with respect to μ or the* ***μ-integral*** *of f over X. If we want to see the independent variable in the integral we may write $\int_X f(x)\,d\mu(x)$.*

If there is no danger of confusion, we shall say *integral* instead of μ-integral.

We shall make a few comments regarding this definition.

(i) The integral of an extended-complex valued function is defined *only if* the function is integrable, and then the value of the integral is a complex number. On the contrary, the integral of an extended-real valued function is defined *either* when the function is integrable (and then the value of the integral is a real number) *or* in certain other cases (and then the value of the integral is either $+\infty$ or $-\infty$).

(ii) We used the function $f\chi_{D_f}$, which is equal to f on D_f and equal to 0 on D_f^c, simply because we need complex values in order to be able to consider their real and imaginary parts. We may allow more freedom and use a function F which is equal to f on D_f and equal to h on D_f^c, where h is an arbitrary $\mathcal{S}\rceil D_f^c$-measurable *complex valued* function on D_f^c. Then we have that $F = f\chi_{D_f}$ a.e. on X, and so $\mathrm{Re}(F) = \mathrm{Re}(f\chi_{D_f})$ and $\mathrm{Im}(F) = \mathrm{Im}(f\chi_{D_f})$ a.e. on X. Now, Proposition 3.15 implies

$$\int_X \mathrm{Re}(F)\,d\mu = \int_X \mathrm{Re}(f\chi_{D_f})\,d\mu, \quad \int_X \mathrm{Im}(F)\,d\mu = \int_X \mathrm{Im}(f\chi_{D_f})\,d\mu.$$

Therefore, there is no difference between the possible definition

$$\int_X f\,d\mu = \int_X \mathrm{Re}(F)\,d\mu + i\int_X \mathrm{Im}(F)\,d\mu$$

and the one we have given. Of course, the function 0 on D_f^c is the simplest of all possible choices for h.
(iii) If $f : X \to \mathbb{C}$ is complex valued on X, then $D_f = X$, and so the definition of $\int_X f\,d\mu$ takes the simpler form:

$$\int_X f\,d\mu = \int_X \mathrm{Re}(f)\,d\mu + i\int_X \mathrm{Im}(f)\,d\mu.$$

In the same case we also have that

$$\mathrm{Re}\,\big(\int_X f\,d\mu\big) = \int_X \mathrm{Re}(f)\,d\mu, \quad \mathrm{Im}\,\big(\int_X f\,d\mu\big) = \int_X \mathrm{Im}(f)\,d\mu.$$

Example. Again, we consider the measure space $(X, \mathcal{P}(X), \delta_{x_0})$ for some $x_0 \in X$. Then every function $f : X \to \overline{\mathbb{C}}$ is measurable.
We know that $\int_X |f|\,d\delta_{x_0} = |f|(x_0) = |f(x_0)|$, and so f is integrable if and only if $f(x_0) \in \mathbb{C}$. In this case, we have that $x_0 \in D_f$, and so

$$\int_X \mathrm{Re}(f\chi_{D_f})\,d\delta_{x_0} = \mathrm{Re}(f\chi_{D_f})(x_0) = \mathrm{Re}(f(x_0)),$$

$$\int_X \mathrm{Im}(f\chi_{D_f})\,d\delta_{x_0} = \mathrm{Im}(f\chi_{D_f})(x_0) = \mathrm{Im}(f(x_0)).$$

Combining the two equalities, we get

$$\int_X f\,d\delta_{x_0} = f(x_0).$$

We find again that *integration with respect to the Dirac measure at x_0 coincides with point evaluation at x_0*.

The next result is obviously helpful and we shall make use of it very often.

Lemma 3.1. *If $f : X \to \overline{\mathbb{C}}$ is integrable, there is an integrable $F : X \to \mathbb{C}$ so that $F = f$ a.e. on X and $\int_X F\,d\mu = \int_X f\,d\mu$.*

Proof. We just consider $F = f\chi_{D_f}$, where $D_f = f^{-1}(\mathbb{C})$. □

Proposition 3.23. *Let $f, g : X \to \overline{\mathbb{C}}$ be measurable and $f = g$ a.e. on X. If f is integrable, then g is integrable and $\int_X g\,d\mu = \int_X f\,d\mu$.*

Proof. Let $f = g$ a.e. on X and f be integrable. Then $|f| = |g|$ a.e. on X, and so g is integrable. Now, Lemma 3.1 says that there are $F, G : X \to \mathbb{C}$ so that $F = f$ a.e. on X and $G = g$ a.e. on X and also

$$\int_X F\,d\mu = \int_X f\,d\mu, \quad \int_X G\,d\mu = \int_X g\,d\mu.$$

From $f = g$ a.e. on X we get that $F = G$ a.e. on X. This implies that $\mathrm{Re}(F) = \mathrm{Re}(G)$ a.e. on X and $\mathrm{Im}(F) = \mathrm{Im}(G)$ a.e. on X. Hence,

$$\int_X F\,d\mu = \int_X \mathrm{Re}(F)\,d\mu + i\int_X \mathrm{Im}(F)\,d\mu = \int_X \mathrm{Re}(G)\,d\mu + i\int_X \mathrm{Im}(G)\,d\mu = \int_X G\,d\mu,$$

and so $\int_X f\,d\mu = \int_X g\,d\mu$. □

Proposition 3.24. *Let $f, g : X \to \overline{\mathbb{C}}$ be integrable and let us consider any measurable definition of $f + g$. Then $f + g$ is integrable and $\int_X (f + g)\,d\mu = \int_X f\,d\mu + \int_X g\,d\mu$.*

Proof. Let f, g be integrable. For every measurable definition of $f+g$ we have $|f+g| \leq |f|+|g|$ on X, and so $\int_X |f+g|\, d\mu \leq \int_X |f|\, d\mu + \int_X |g|\, d\mu < +\infty$. Hence, $f+g$ is integrable, and so there are integrable $F, G : X \to \mathbb{C}$ so that $F = f$ a.e. on X and $G = g$ a.e. on X. This implies that for all measurable definitions of $f+g$ we have $F+G = f+g$ a.e. on X, and so, by Proposition 3.23, we have

$$\int_X f\, d\mu = \int_X F\, d\mu, \quad \int_X g\, d\mu = \int_X G\, d\mu, \quad \int_X (f+g)\, d\mu = \int_X (F+G)\, d\mu.$$

Therefore, it is enough to prove that $\int_X (F+G)\, d\mu = \int_X F\, d\mu + \int_X G\, d\mu$.
Now,

$$\int_X \mathrm{Re}(F+G)\, d\mu = \int_X \mathrm{Re}(F)\, d\mu + \int_X \mathrm{Re}(G)\, d\mu,$$
$$\int_X \mathrm{Im}(F+G)\, d\mu = \int_X \mathrm{Im}(F)\, d\mu + \int_X \mathrm{Im}(G)\, d\mu.$$

Combining, we get $\int_X (F+G)\, d\mu = \int_X F\, d\mu + \int_X G\, d\mu$. □

Proposition 3.25. *Let $f : X \to \overline{\mathbb{C}}$ be μ-integrable and ν-integrable. Then f is $(\mu+\nu)$-integrable and $\int_X f\, d(\mu+\nu) = \int_X f\, d\mu + \int_X f\, d\nu$.*

Proof. If f is μ-integrable and ν-integrable, then $\int_X |f|\, d(\mu+\nu) = \int_X |f|\, d\mu + \int_X |f|\, d\nu < +\infty$, and so f is $(\mu+\nu)$-integrable. Then there is a $(\mu+\nu)$-integrable $F : X \to \mathbb{C}$ so that $F = f$ a.e. on X. From Proposition 3.23 we get

$$\int_X f\, d\mu = \int_X F\, d\mu, \quad \int_X f\, d\nu = \int_X F\, d\nu, \quad \int_X f\, d(\mu+\nu) = \int_X F\, d(\mu+\nu).$$

Now,

$$\int_X \mathrm{Re}(F)\, d(\mu+\nu) = \int_X \mathrm{Re}(F)\, d\mu + \int_X \mathrm{Re}(F)\, d\nu,$$
$$\int_X \mathrm{Im}(F)\, d(\mu+\nu) = \int_X \mathrm{Im}(F)\, d\mu + \int_X \mathrm{Im}(F)\, d\nu.$$

Then $\int_X F\, d(\mu+\nu) = \int_X F\, d\mu + \int_X F\, d\nu$, and so $\int_X f\, d(\mu+\nu) = \int_X f\, d\mu + \int_X f\, d\nu$. □

Proposition 3.26. *Let $f : X \to \overline{\mathbb{C}}$ be integrable and $\lambda \in \mathbb{C}$. Then λf is integrable and $\int_X \lambda f\, d\mu = \lambda \int_X f\, d\mu$.*

Proof. Let f be integrable. Then $\int_X |\lambda f|\, d\mu = |\lambda| \int_X |f|\, d\mu < +\infty$, and so λf is also integrable. Then there is an integrable $F : X \to \mathbb{C}$ so that $F = f$ a.e. on X. Then, $\lambda F = \lambda f$ a.e. on X, and Proposition 3.23 implies

$$\int_X \lambda f\, d\mu = \int_X \lambda F\, d\mu, \quad \int_X f\, d\mu = \int_X F\, d\mu.$$

From $\mathrm{Re}(\lambda F) = \mathrm{Re}(\lambda)\,\mathrm{Re}(F) - \mathrm{Im}(\lambda)\,\mathrm{Im}(F)$ and $\mathrm{Im}(\lambda F) = \mathrm{Re}(\lambda)\,\mathrm{Im}(F) + \mathrm{Im}(\lambda)\,\mathrm{Re}(F)$ we get

$$\int_X \mathrm{Re}(\lambda F)\, d\mu = \mathrm{Re}(\lambda) \int_X \mathrm{Re}(F)\, d\mu - \mathrm{Im}(\lambda) \int_X \mathrm{Im}(F)\, d\mu,$$
$$\int_X \mathrm{Im}(\lambda F)\, d\mu = \mathrm{Re}(\lambda) \int_X \mathrm{Im}(F)\, d\mu + \mathrm{Im}(\lambda) \int_X \mathrm{Re}(F)\, d\mu.$$

From these two equalities we easily get

$$\int_X \lambda F\, d\mu = \lambda \int_X \mathrm{Re}(F)\, d\mu + i\lambda \int_X \mathrm{Im}(F)\, d\mu = \lambda \int_X F\, d\mu.$$

Hence, $\int_X \lambda f\, d\mu = \lambda \int_X f\, d\mu$. □

Proposition 3.27. *Let $f : X \to \overline{\mathbb{C}}$ be μ-integrable and $\lambda \in [0, +\infty)$. Then f is $\lambda\mu$-integrable and $\int_X f\, d(\lambda\mu) = \lambda \int_X f\, d\mu$.*

Proof. Let f be μ-integrable. Then $\int_X |f|\, d(\lambda\mu) = \lambda \int_X |f|\, d\mu < +\infty$, and so f is also $\lambda\mu$-integrable. Then there is a μ-integrable $F : X \to \mathbb{C}$ so that $F = f$ μ-a.e. on X. Of course, this implies that $F = f$ $\lambda\mu$-a.e. on X, and Proposition 3.23 implies

$$\int_X f\, d\mu = \int_X F\, d\mu, \quad \int_X f\, d(\lambda\mu) = \int_X F\, d(\lambda\mu).$$

Now,

$$\int_X \mathrm{Re}(F)\, d(\lambda\mu) = \lambda \int_X \mathrm{Re}(F)\, d\mu, \quad \int_X \mathrm{Im}(F)\, d(\lambda\mu) = \lambda \int_X \mathrm{Im}(F)\, d\mu.$$

Hence, $\int_X F\, d(\lambda\mu) = \lambda \int_X F\, d\mu$, and so $\int_X f\, d(\lambda\mu) = \lambda \int_X f\, d\mu$. □

Proposition 3.28. *Let $f : X \to \overline{\mathbb{C}}$ be integrable. Then $\big| \int_X f\, d\mu \big| \leq \int_X |f|\, d\mu$.*

Proof. There is an integrable $F : X \to \mathbb{C}$ so that $F = f$ a.e. on X. By Proposition 3.23, it is enough to prove $\big| \int_X F\, d\mu \big| \leq \int_X |F|\, d\mu$.
We consider the complex number

$$\lambda = \overline{\mathrm{sign}(\int_X F\, d\mu)},$$

and we get

$$\begin{aligned} \Big| \int_X F\, d\mu \Big| &= \lambda \int_X F\, d\mu = \int_X \lambda F\, d\mu = \mathrm{Re}\left(\int_X \lambda F\, d\mu \right) = \int_X \mathrm{Re}(\lambda F)\, d\mu \leq \int_X |\mathrm{Re}(\lambda F)|\, d\mu \\ &\leq \int_X |\lambda F|\, d\mu \leq \int_X |F|\, d\mu, \end{aligned}$$

since $|\lambda| \leq 1$. □

THE LIMIT THEOREMS.

The next five theorems are probably the most important results of integration theory.

Monotone Convergence Theorem (Lebesgue, Levi). *Let $f, f_1, f_2, \ldots : X \to [0, +\infty]$ be measurable so that $f_n \leq f_{n+1}$ a.e. on X for all n and $\lim_{n\to+\infty} f_n = f$ a.e. on X. Then*

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

Proof. (a) Assume that $f_n \leq f_{n+1}$ *everywhere* on X for all n and $\lim_{n\to+\infty} f_n = f$ *everywhere* on X.
The sequence $\left(\int_X f_n\, d\mu \right)$ is increasing and it is bounded above by $\int_X f\, d\mu$. Hence, the limit $\lim_{n\to+\infty} \int_X f_n\, d\mu$ exists and $\lim_{n\to+\infty} \int_X f_n\, d\mu \leq \int_X f\, d\mu$.
Now we consider an increasing sequence (ϕ_n) of non-negative measurable simple functions on X so that $\lim_{n\to+\infty} \phi_n = f$ on X. Then for each k we have $\phi_k \leq f = \lim_{n\to+\infty} f_n$.
We consider an arbitrary $\alpha \in [0, 1)$, and, for any fixed k, we define

$$A_n = \{x \in X \mid \alpha\phi_k(x) \leq f_n(x)\} \in \mathcal{S}.$$

Then (A_n) is increasing, $\bigcup_{n=1}^{+\infty} A_n = X$, and $\alpha\phi_k\chi_{A_n} \leq f_n$ on X. Hence,

$$\alpha \int_X \phi_k\, d\mu = \int_X \alpha\phi_k\, d\mu = \lim_{n\to+\infty} \int_X \alpha\phi_k\chi_{A_n}\, d\mu \leq \lim_{n\to+\infty} \int_X f_n\, d\mu,$$

where we used Proposition 3.4 for the second equality. Taking the limit as $\alpha \to 1-$ and then taking the limit as $k \to +\infty$, we conclude that $\int_X f\, d\mu \leq \lim_{n\to+\infty} \int_X f_n\, d\mu$, and the proof is complete.
Here is an alternative proof of the last inequality.
For each k we consider an increasing sequence $(\psi_{k,n})$ of non-negative measurable simple functions on X so that $\lim_{n\to+\infty} \psi_{k,n} = f_k$ on X. We define the non-negative measurable simple functions

$$\phi_n = \max\{\psi_{1,n}, \ldots, \psi_{n,n}\}.$$

Then we have

$$\phi_n = \max\{\psi_{1,n}, \ldots, \psi_{n,n}\} \leq \max\{f_1, \ldots, f_n\} = f_n \leq f$$

on X. Also,

$$\phi_n = \max\{\psi_{1,n}, \dots, \psi_{n,n}\} \le \max\{\psi_{1,n+1}, \dots, \psi_{n,n+1}\}$$
$$\le \max\{\psi_{1,n+1}, \dots, \psi_{n,n+1}, \psi_{n+1,n+1}\} = \phi_{n+1}$$

on X. Therefore, $\lim_{n\to+\infty} \phi_n$ exists, and $\lim_{n\to+\infty} \phi_n \le f$ on X. Moreover, if $k \le n$, we have

$$\phi_n = \max\{\psi_{1,n}, \dots, \psi_{n,n}\} \ge \max\{\psi_{1,n}, \dots, \psi_{k,n}\},$$

and, taking the limit as $n \to +\infty$, we get

$$\lim_{n\to+\infty} \phi_n \ge \max\{f_1, \dots, f_k\} = f_k$$

for every k. Now, taking the limit as $k \to +\infty$, we get $\lim_{n\to+\infty} \phi_n \ge f$ on X.
We conclude that (ϕ_n) is increasing and $\lim_{n\to+\infty} \phi_n = f$ on X, and so

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

(b) In the general case there is some $A \in \mathcal{S}$ with $\mu(A^c) = 0$ so that $f_n \le f_{n+1}$ on A for all n and $\lim_{n\to+\infty} f_n = f$ on A. Then $f_n\chi_A \le f_{n+1}\chi_A$ on X for all n and $\lim_{n\to+\infty} f_n\chi_A = f\chi_A$ on X, and so from part (a) we get

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

Since $f = f\chi_A$ a.e. on X and $f_n = f_n\chi_A$ a.e. on X for every n, Proposition 3.12 finally implies $\lim_{n\to+\infty} \int_X f_n\,d\mu = \int_X f\,d\mu$. $\square$

Non-negative Series Theorem. *Let $s, f_1, f_2, \dots : X \to [0, +\infty]$ be measurable and $\sum_{n=1}^{+\infty} f_n = s$ a.e. on X. Then*

$$\sum_{n=1}^{+\infty} \int_X f_n\,d\mu = \int_X s\,d\mu.$$

Proof. We consider the partial sums $s_k = f_1 + \cdots + f_k$. Then $s_k \le s_{k+1}$ on X for all k and $\lim_{k\to+\infty} s_k = s$ a.e. on X, and so

$$\sum_{n=1}^{+\infty} \int_X f_n\,d\mu = \lim_{k\to+\infty} \sum_{n=1}^{k} \int_X f_n\,d\mu = \lim_{k\to+\infty} \int_X s_k\,d\mu = \int_X s\,d\mu$$

by the Monotone Convergence Theorem. $\square$

Fatou's Lemma. *Let $f, f_1, f_2, \dots : X \to [0, +\infty]$ be measurable and $f = \underline{\lim}_{n\to+\infty} f_n$ a.e. on X. Then*

$$\int_X f\,d\mu \le \underline{\lim}_{n\to+\infty} \int_X f_n\,d\mu.$$

Proof. We define $g_n = \inf_{k\ge n} f_k$ for each n. Then every $g_n : X \to [0, +\infty]$ is measurable, and we have that $g_n \le f_n$ on X and $g_n \le g_{n+1}$ on X for all n and $\lim_{n\to+\infty} g_n = f$ a.e. on X. Then

$$\int_X f\,d\mu = \lim_{n\to+\infty} \int_X g_n\,d\mu \le \underline{\lim}_{n\to+\infty} \int_X f_n\,d\mu$$

by the Monotone Convergence Theorem. $\square$

Dominated Convergence Theorem (Lebesgue). *Let $f, f_1, f_2, \dots : X \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ and $g : X \to [0, +\infty]$ be measurable. Let also $\lim_{n\to+\infty} f_n = f$ a.e. on X, $|f_n| \le g$ a.e. on X for all n and $\int_X g\,d\mu < +\infty$. Then all f, f_n are integrable and*

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

Proof. From $|f_n| \leq g$ a.e. on X, we get $\int_X |f_n| \, d\mu \leq \int_X g \, d\mu < +\infty$, and so all f_n are integrable. Also, from $|f_n| \leq g$ a.e. on X and $\lim_{n\to+\infty} f_n = f$ a.e. on X, we get $|f| \leq g$ a.e. on X, and so f is also integrable.
Now, there are integrable $F, F_1, F_2, \ldots : X \to \mathbb{R}$ or $\mathbb{C}$ so that $F = f$ a.e. on X and $F_n = f_n$ a.e. on X for all n. Then

$$\textstyle\int_X f_n \, d\mu = \int_X F_n \, d\mu, \quad \int_X f \, d\mu = \int_X F \, d\mu,$$

and so is enough to prove $\lim_{n\to+\infty} \int_X F_n \, d\mu = \int_X F \, d\mu$.
We have that $|F_n| \leq g$ a.e. on X for all n, and $\lim_{n\to+\infty} F_n = F$ a.e. on X.
(a) Let $F, F_n : X \to \mathbb{R}$.
Since $0 \leq g + F_n$ a.e. on X and $0 \leq g - F_n$ a.e. on X for all n, Fatou's Lemma implies

$$\textstyle\int_X (g \pm F) \, d\mu \leq \underline{\lim}_{n\to+\infty} \int_X (g \pm F_n) \, d\mu,$$

and so

$$\textstyle\int_X g \, d\mu \pm \int_X F \, d\mu \leq \int_X g \, d\mu + \underline{\lim}_{n\to+\infty} (\pm \int_X F_n \, d\mu).$$

Since $\int_X g \, d\mu$ is finite, we get

$$\textstyle\pm \int_X F \, d\mu \leq \underline{\lim}_{n\to+\infty} (\pm \int_X F_n \, d\mu).$$

Therefore,

$$\textstyle\overline{\lim}_{n\to+\infty} \int_X F_n \, d\mu \leq \int_X F \, d\mu \leq \underline{\lim}_{n\to+\infty} \int_X F_n \, d\mu,$$

and this implies $\lim_{n\to+\infty} \int_X F_n \, d\mu = \int_X F \, d\mu$.
(b) Let $F, F_n : X \to \mathbb{C}$.
From $|\operatorname{Re}(F_n)| \leq |F_n| \leq g$ a.e. on X for all n and $\lim_{n\to+\infty} \operatorname{Re}(F_n) = \operatorname{Re}(F)$ a.e. on X, and from part (a) we have that

$$\textstyle\lim_{n\to+\infty} \int_X \operatorname{Re}(F_n) \, d\mu = \int_X \operatorname{Re}(F) \, d\mu.$$

Similarly,

$$\textstyle\lim_{n\to+\infty} \int_X \operatorname{Im}(F_n) \, d\mu = \int_X \operatorname{Im}(F) \, d\mu.$$

Therefore, $\lim_{n\to+\infty} \int_X F_n \, d\mu = \int_X F \, d\mu$. ☐

Series Theorem. *Let $f, f_1, f_2, \ldots : X \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ be measurable. If $\sum_{n=1}^{+\infty} \int_X |f_n| \, d\mu < +\infty$, then*
(i) $\sum_{n=1}^{+\infty} f_n(x)$ converges for a.e. $x \in X$,
(ii) if $\sum_{n=1}^{+\infty} f_n = s$ a.e. on X, then

$$\textstyle\sum_{n=1}^{+\infty} \int_X f_n \, d\mu = \int_X s \, d\mu.$$

Proof. (i) We define $S = \sum_{n=1}^{+\infty} |f_n|$. From the Non-negative Series Theorem we have

$$\textstyle\int_X S \, d\mu = \sum_{n=1}^{+\infty} \int_X |f_n| \, d\mu < +\infty.$$

This implies $S(x) < +\infty$ for a.e. $x \in X$. Therefore, the series $\sum_{n=1}^{+\infty} f_n(x)$ converges absolutely, and hence converges, for a.e. $x \in X$.
(ii) We consider the partial sums $s_k = f_1 + \cdots + f_k$. Then $|s_k| \leq |f_1| + \cdots + |f_k| \leq S$ a.e. on X for all k, and $\lim_{k\to+\infty} s_k = s$ a.e. on X. Hence,

$$\textstyle\sum_{n=1}^{+\infty} \int_X f_n \, d\mu = \lim_{k\to+\infty} \sum_{n=1}^{k} \int_X f_n \, d\mu = \lim_{k\to+\infty} \int_X s_k \, d\mu = \int_X s \, d\mu$$

by the Dominated Convergence Theorem. ☐

Exercises.

3.1.3. Let $f, f_n : X \to [0,+\infty]$ be measurable with $f_n \le f$ a.e. on X for all n and $\lim_{n\to+\infty} f_n = f$ a.e. on X. Prove that $\lim_{n\to+\infty} \int_X f_n\,d\mu = \int_X f\,d\mu$.
Hint. Use Fatou's Lemma.

3.1.4. Let $f, f_n : X \to [0,+\infty]$ be measurable and $\lim_{n\to+\infty} f_n = f$ a.e. on X. If there is $M < +\infty$ so that $\int_X f_n\,d\mu \le M$ for infinitely many n, prove that $\int_X f\,d\mu < +\infty$.

3.1.5. Let $f, f_n : X \to [0,+\infty]$ be measurable so that $f_{n+1} \le f_n$ a.e. on X for all n and $\lim_{n\to+\infty} f_n = f$ a.e. on X and $\int_X f_1\,d\mu < +\infty$. Prove that $\lim_{n\to+\infty} \int_X f_n\,d\mu = \int_X f\,d\mu$.
Hint. Use the Monotone Convergence Theorem.

3.1.6. Use either Fatou's Lemma or the Series Theorem to prove the Monotone Convergence Theorem.

3.1.7. Let μ be σ-finite. Prove that there is $f : X \to [0,+\infty]$ such that $f(x) > 0$ for every $x \in X$ and $\int_X f\,d\mu < +\infty$.
Hint. Consider pairwise disjoint $X_1, X_2, \ldots \in \mathcal{S}$ which cover X and so that $0 < \mu(X_j) < +\infty$ for all j. Then let $f = a_j$ on X_j, where the $a_j > 0$ are chosen appropriately.

3.1.8. Assume that $f : X \to [0,+\infty]$ is measurable, $0 < \int_X f\,d\mu < +\infty$, and $0 < \alpha < +\infty$. Prove that the limit $I = \lim_{n\to+\infty} n \int_X \log\big(1+\big(\frac{f}{n}\big)^\alpha\big)\,d\mu$ exists, and that: $I = \int_X f\,d\mu$, if $\alpha = 1$, and $I = +\infty$, if $0 < \alpha < 1$, and $I = 0$, if $1 < \alpha < +\infty$.
Hint. Consider the case $\alpha = 1$ first, using the Monotone Convergence Theorem.

3.1.9. *Uniform Convergence Theorem.*
Let $f_n : X \to \mathbb{R}$ or $\mathbb{C}$ be integrable and let $\lim_{n\to+\infty} f_n = f$ uniformly on X. If $\mu(X) < +\infty$, prove that f is integrable and that $\lim_{n\to+\infty} \int_X f_n\,d\mu = \int_X f\,d\mu$.

3.1.10. *Bounded Convergence Theorem.*
Let $f, f_n : X \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ be measurable. If $\mu(X) < +\infty$ and there is $M < +\infty$ so that $|f_n| \le M$ a.e. on X for all n and $\lim_{n\to+\infty} f_n = f$ a.e. on X, prove that $\lim_{n\to+\infty} \int_X f_n\,d\mu = \int_X f\,d\mu$.

3.1.11. Let $f, f_n : X \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ be measurable and $g : X \to [0,+\infty]$ be integrable. If $|f_n| \le g$ a.e. on X for all n and $\lim_{n\to+\infty} f_n = f$ a.e. on X, prove that $\lim_{n\to+\infty} \int_X |f_n - f|\,d\mu = 0$.
Hint. Prove that $|f_n - f| \le 2g$ a.e. on X.

3.1.12. Let $f, g, f_n : X \to \overline{\mathbb{R}}$ be measurable and $\int_X g^-\,d\mu < +\infty$. If $g \le f_n$ a.e. on X for all n and $f = \underline{\lim}_{n\to+\infty} f_n$ a.e. on X, prove that $\int_X f\,d\mu \le \underline{\lim}_{n\to+\infty} \int_X f_n\,d\mu$.
Hint. Prove that $f_n + g^- \ge 0$ a.e. on X.

3.1.13. Let $f, f_n : X \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ and $g, g_n : X \to [0,+\infty]$ be all measurable. If $|f_n| \le g_n$ a.e. on X for all n, if $\lim_{n\to+\infty} \int_X g_n\,d\mu = \int_X g\,d\mu < +\infty$ and if $\lim_{n\to+\infty} f_n = f$ a.e. on X and $\lim_{n\to+\infty} g_n = g$ a.e. on X, prove that $\lim_{n\to+\infty} \int_X f_n\,d\mu = \int_X f\,d\mu$.
Hint. Aplly Fatou's Lemma to $(g_n + f_n)$ and to $(g_n - f_n)$.

3.1.14. Let $f, f_n : X \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ be integrable and $\lim_{n\to+\infty} f_n = f$ a.e. on X. Prove that $\lim_{n\to+\infty} \int_X |f_n - f|\,d\mu = 0$ if and only if $\lim_{n\to+\infty} \int_X |f_n|\,d\mu = \int_X |f|\,d\mu$.
Hint. One direction is trivial. For the other direction, use $|f_n - f| \le |f_n| + |f|$ and the result of exercise 3.1.13.

3.1.15. *Continuity of an integral as a function of a parameter.*
Let $f : X \times (a,b) \to \mathbb{R}$ and $g : X \to [0,+\infty]$ be such that
(i) g is integrable and, for every $t \in (a,b)$, $f(\cdot,t)$ is measurable,
(ii) for a.e. $x \in X$, $f(x,t)$ is continuous as a function of t on (a,b),

(iii) for every $t \in (a,b)$, $|f(x,t)| \le g(x)$ for a.e. $x \in X$.
Define $F(t) = \int_X f(x,t)\,d\mu(x)$ for all $t \in (a,b)$ and prove that F is continuous on (a,b).
Hint. Assume $\lim_{n\to+\infty} t_n = t$ and use the Dominated Convergence Theorem to prove that $\lim_{n\to+\infty} F(t_n) = F(t)$.

3.1.16. *Differentiability of an integral as a function of a parameter*.
Let $f : X \times (a,b) \to \mathbb{R}$ and $g : X \to [0,+\infty]$ be such that
(i) g is integrable and, for every $t \in (a,b)$, $f(\cdot,t)$ is measurable,
(ii) for at least one $t_0 \in (a,b)$, $f(\cdot,t_0)$ is integrable,
(iii) for a.e. $x \in X$, $f(x,t)$ is differentiable as a function of t on (a,b) and $\left|\frac{\partial f}{\partial t}(x,t)\right| \le g(x)$ for every $t \in (a,b)$. Thus, $\frac{\partial f}{\partial t} : A \times (a,b) \to \mathbb{R}$ for some $A \in \mathcal{S}$ with $\mu(X \setminus A) = 0$.
Define $F(t) = \int_X f(x,t)\,d\mu(x)$ for all $t \in (a,b)$ and prove that F is differentiable on (a,b) and that $\frac{dF}{dt}(t) = \int_X \frac{\partial f}{\partial t}(x,t)\,d\mu(x)$ for all $t \in (a,b)$.
Hint. Assume $\lim_{n\to+\infty} t_n = t$ and use the Dominated Convergence Theorem to prove that $\lim_{n\to+\infty} \frac{F(t_n)-F(t)}{t_n-t} = \int_X \frac{\partial f}{\partial t}(x,t)\,d\mu(x)$.

APPROXIMATION BY SIMPLE FUNCTIONS.

Proposition 3.29. *Let $f : X \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ be integrable. Then for every $\epsilon > 0$ there is an integrable simple function $\phi : X \to \mathbb{R}$ or $\mathbb{C}$ so that $\int_X |f-\phi|\,d\mu < \epsilon$.*

Proof. (a) If $f : X \to [0,+\infty]$ is integrable, there is an increasing sequence (ϕ_n) of non-negative measurable simple functions so that $\lim_{n\to+\infty} \phi_n = f$ on X and $\lim_{n\to+\infty} \int_X \phi_n\,d\mu = \int_X f\,d\mu$. Then for some n we have

$$\int_X f\,d\mu - \epsilon < \int_X \phi_n\,d\mu \le \int_X f\,d\mu,$$

and so ϕ_n is integrable and

$$\int_X |f-\phi_n|\,d\mu = \int_X (f-\phi_n)\,d\mu < \epsilon.$$

(b) Now, if $f : X \to \overline{\mathbb{R}}$ is integrable, then $\int_X f^+\,d\mu < +\infty$ and $\int_X f^-\,d\mu < +\infty$. By (a) we have that there are non-negative integrable simple functions χ, ψ so that

$$\int_X |f^+-\chi|\,d\mu < \tfrac{\epsilon}{2}, \quad \int_X |f^--\psi|\,d\mu < \tfrac{\epsilon}{2}.$$

We consider the integrable simple function $\phi = \chi - \psi : X \to \mathbb{R}$, and we get

$$\int_X |f-\phi|\,d\mu \le \int_X |f^+-\chi|\,d\mu + \int_X |f^--\psi|\,d\mu < \epsilon.$$

(c) Finally, let $f : X \to \overline{\mathbb{C}}$ be integrable. Then there is an integrable $F : X \to \mathbb{C}$ so that $F = f$ a.e. on X. The functions $\mathrm{Re}(F), \mathrm{Im}(F) : X \to \mathbb{R}$ are both integrable. By (b) we know that there are real valued integrable simple functions χ, ψ so that

$$\int_X |\mathrm{Re}(F)-\chi|\,d\mu < \tfrac{\epsilon}{2}, \quad \int_X |\mathrm{Im}(F)-\psi|\,d\mu < \tfrac{\epsilon}{2}.$$

We consider the integrable simple function $\phi = \chi + i\psi : X \to \mathbb{C}$, and we get

$$\int_X |f-\phi|\,d\mu = \int_X |F-\phi|\,d\mu \le \int_X |\mathrm{Re}(F)-\chi|\,d\mu + \int_X |\mathrm{Im}(F)-\psi|\,d\mu < \epsilon.$$

So the proof is complete in all cases. $\square$

INTEGRALS OVER SUBSETS.

Let $A \in \mathcal{S}$ and $f : X \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ be measurable. In order to define *an integral of f over A* we have two *natural* choices. One choice is to take $f\chi_A$, which is equal to f on A and equal to 0 on A^c, and consider $\int_X f\chi_A \, d\mu$. Another choice is to take the restriction $f\rceil A$ of f on A, and consider $\int_A (f\rceil A) \, d(\mu\rceil A)$ with respect to the restricted measure $\mu\rceil A$ on the restricted σ-algebra $(A, \mathcal{S}\rceil A)$. The following lemma says that the two procedures are equivalent and that they give the same results.

Lemma 3.2. *Let $A \in \mathcal{S}$ and $f : X \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ be measurable.*
(i) If $f : X \to \overline{\mathbb{R}}$ and either $\int_X f\chi_A \, d\mu$ or $\int_A (f\rceil A) \, d(\mu\rceil A)$ is defined, then the other is also defined and $\int_X f\chi_A \, d\mu = \int_A (f\rceil A) \, d(\mu\rceil A)$.
(ii) If $f : X \to \overline{\mathbb{C}}$ and either $\int_X |f\chi_A| \, d\mu$ or $\int_A |f\rceil A| \, d(\mu\rceil A)$ is finite, then the other is also finite and $\int_X f\chi_A \, d\mu = \int_A (f\rceil A) \, d(\mu\rceil A)$.

Proof. (a) We take a non-negative measurable simple function ϕ on X with its standard representation $\phi = \sum_{j=1}^m \kappa_j \chi_{E_j}$.
Then we have $\phi\chi_A = \sum_{j=1}^m \kappa_j \chi_{E_j \cap A}$, with

$$\int_X \phi\chi_A \, d\mu = \sum_{j=1}^m \kappa_j \mu(E_j \cap A).$$

On the other hand, $\phi\rceil A = \sum_{j=1}^m \kappa_j \chi_{E_j \cap A}$ has

$$\int_A (\phi\rceil A) \, d(\mu\rceil A) = \sum_{j=1}^m \kappa_j (\mu\rceil A)(E_j \cap A) = \sum_{j=1}^m \kappa_j \mu(E_j \cap A).$$

(b) Now let $f : X \to [0, +\infty]$ be measurable. We consider an increasing sequence (ϕ_n) of non-negative measurable simple functions on X so that $\lim_{n\to+\infty} \phi_n = f$ on X.
Then $(\phi_n\chi_A)$ is increasing and $\lim_{n\to+\infty} \phi_n\chi_A = f\chi_A$ on X. Also, $(\phi_n\rceil A)$ is increasing and $\lim_{n\to+\infty} \phi_n\rceil A = f\rceil A$ on A. Now, by part (a) we get

$$\int_X f\chi_A \, d\mu = \lim_{n\to+\infty} \int_X \phi_n\chi_A \, d\mu = \lim_{n\to+\infty} \int_A (\phi_n\rceil A) \, d(\mu\rceil A) = \int_A (f\rceil A) \, d(\mu\rceil A).$$

(c) If $f : X \to \overline{\mathbb{R}}$ is measurable, then $f^+\chi_A = (f\chi_A)^+$ and $f^-\chi_A = (f\chi_A)^-$ on X, and also $(f\rceil A)^+ = f^+\rceil A$ and $(f\rceil A)^- = f^-\rceil A$ on A. Hence, by part (b) we get

$$\int_X (f\chi_A)^+ \, d\mu = \int_X f^+\chi_A \, d\mu = \int_A (f^+\rceil A) \, d(\mu\rceil A) = \int_A (f\rceil A)^+ \, d(\mu\rceil A)$$

and, similarly,

$$\int_X (f\chi_A)^- \, d\mu = \int_A (f\rceil A)^- \, d(\mu\rceil A).$$

These prove (i).
(d) Finally, let $f : X \to \overline{\mathbb{C}}$ be measurable. Then $|f\chi_A| = |f|\chi_A$ on X and $|f\rceil A| = |f|\rceil A$ on A. By part (b) we have

$$\int_X |f\chi_A| \, d\mu = \int_X |f|\chi_A \, d\mu = \int_A (|f|\rceil A) \, d(\mu\rceil A) = \int_A |f\rceil A| \, d(\mu\rceil A),$$

and so $f\chi_A$ and $f\rceil A$ are simultaneously integrable or non-integrable.
Assuming integrability, there is an integrable $F : X \to \mathbb{C}$ so that $F = f\chi_A$ a.e. on X. It is clear that $F\chi_A = f\chi_A$ a.e. on X and $F\rceil A = f\rceil A$ a.e. on A. Therefore, it is enough to prove that $\int_X F\chi_A \, d\mu = \int_A (F\rceil A) \, d(\mu\rceil A)$. Now, part (c) implies

$$\int_X \mathrm{Re}(F\chi_A) \, d\mu = \int_X \mathrm{Re}(F)\chi_A \, d\mu = \int_A (\mathrm{Re}(F)\rceil A) \, d(\mu\rceil A) = \int_A \mathrm{Re}(F\rceil A) \, d(\mu\rceil A).$$

Similarly,

$$\int_X \mathrm{Im}(F\chi_A) \, d\mu = \int_A \mathrm{Im}(F\rceil A) \, d(\mu\rceil A).$$

Thus, $\int_X F\chi_A \, d\mu = \int_A (F\rceil A) \, d(\mu\rceil A)$. $\square$

Definition. *Let $A \in \mathcal{S}$ and $f : X \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ be measurable.*
(i) If $f : X \to \overline{\mathbb{R}}$ and $\int_X f\chi_A \, d\mu$ or, equivalently, $\int_A (f\rceil A)\, d(\mu\rceil A)$ is defined, then we say that $\int_A f \, d\mu$ is defined as

$$\int_A f \, d\mu = \int_X f\chi_A \, d\mu = \int_A (f\rceil A)\, d(\mu\rceil A).$$

(ii) If $f : X \to \overline{\mathbb{C}}$ and $f\chi_A$ is integrable over X or, equivalently, $f\rceil A$ is integrable over A, then we say that f is ***integrable*** *over A and we define $\int_A f \, d\mu$ exactly as in (i).*

Proposition 3.30. *Let $f : X \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ be measurable.*
(i) If $f : X \to \overline{\mathbb{R}}$ and $\int_X f \, d\mu$ is defined, then $\int_A f \, d\mu$ is defined for every $A \in \mathcal{S}$.
(ii) If $f : X \to \overline{\mathbb{C}}$ is integrable, then f is integrable over every $A \in \mathcal{S}$.

Proof. (i) Let $\int_X f \, d\mu$ be defined. We have $(f\chi_A)^+ = f^+\chi_A \leq f^+$ and $(f\chi_A)^- = f^-\chi_A \leq f^-$ on X. Thus, either $\int_X (f\chi_A)^+ \, d\mu \leq \int_X f^+ \, d\mu < +\infty$ or $\int_X (f\chi_A)^- \, d\mu \leq \int_X f^- \, d\mu < +\infty$. Therefore, $\int_X f\chi_A \, d\mu$ is defined, and so $\int_A f \, d\mu$ is also defined.
(ii) Let f be integrable. Then $\int_X |f\chi_A| \, d\mu \leq \int_X |f| \, d\mu < +\infty$, and so $f\chi_A$ is integrable. $\square$

Proposition 3.31. *Let $f : X \to \overline{\mathbb{R}}$ be measurable and $\int_X f \, d\mu$ be defined. Then either $\int_A f \, d\mu > -\infty$ for all $A \in \mathcal{S}$ or $\int_A f \, d\mu < +\infty$ for all $A \in \mathcal{S}$.*

Proof. Let $\int_X f^- \, d\mu < +\infty$. We have $(f\chi_A)^- = f^-\chi_A \leq f^-$ on X. Then

$$\int_X (f\chi_A)^- \, d\mu \leq \int_X f^- \, d\mu < +\infty,$$

and so $\int_A f \, d\mu = \int_X f\chi_A \, d\mu > -\infty$ for all $A \in \mathcal{S}$.
Similarly, if $\int_X f^+ \, d\mu < +\infty$, then $\int_A f \, d\mu < +\infty$ for all $A \in \mathcal{S}$. $\square$

Theorem 3.1. *Let $f : X \to \overline{\mathbb{R}}$ be measurable and $\int_X f \, d\mu$ be defined, or let $f : X \to \overline{\mathbb{C}}$ be integrable.*
(i) $\int_A f \, d\mu = 0$ for all $A \in \mathcal{S}$ with $\mu(A) = 0$,
(ii) $\sum_{n=1}^{+\infty} \int_{A_n} f \, d\mu = \int_A f \, d\mu$ for all pairwise disjoint $A_1, A_2, \ldots \in \mathcal{S}$ with $A = \bigcup_{n=1}^{+\infty} A_n$,
(iii) $\lim_{n\to+\infty} \int_{A_n} f \, d\mu = \int_A f \, d\mu$ for all $A_1, A_2, \ldots \in \mathcal{S}$ such that (A_n) is increasing and $\bigcup_{n=1}^{+\infty} A_n = A$,
(iv) $\lim_{n\to+\infty} \int_{A_n} f \, d\mu = \int_A f \, d\mu$ for all $A_1, A_2, \ldots \in \mathcal{S}$ such that (A_n) is decreasing and $\bigcap_{n=1}^{+\infty} A_n = A$ and $\left| \int_{A_N} f \, d\mu \right| < +\infty$ for some N.

Proof. (i) This is easy because, if $\mu(A) = 0$, then $f\chi_A = 0$ a.e. on X.
(ii) Let $A_1, A_2, \ldots \in \mathcal{S}$ be pairwise disjoint and $A = \bigcup_{n=1}^{+\infty} A_n$.
(a) If $f : X \to [0, +\infty]$ is measurable, then, since $\sum_{n=1}^{+\infty} f\chi_{A_n} = f\chi_A$ on X, the Non-negative Series Theorem gives

$$\sum_{n=1}^{+\infty} \int_{A_n} f \, d\mu = \sum_{n=1}^{+\infty} \int_X f\chi_{A_n} \, d\mu = \int_X f\chi_A \, d\mu = \int_A f \, d\mu.$$

(b) If $f : X \to \overline{\mathbb{R}}$ and $\int_X f^- \, d\mu < +\infty$, we apply (a) and we get

$$\sum_{n=1}^{+\infty} \int_{A_n} f^+ \, d\mu = \int_A f^+ \, d\mu, \quad \sum_{n=1}^{+\infty} \int_{A_n} f^- \, d\mu = \int_A f^- \, d\mu < +\infty.$$

Subtracting, we find $\sum_{n=1}^{+\infty} \int_{A_n} f \, d\mu = \int_A f \, d\mu$.
If $f : X \to \overline{\mathbb{R}}$ and $\int_X f^+ \, d\mu < +\infty$, then the proof is similar.
(c) If $f : X \to \overline{\mathbb{C}}$ and f is integrable, we have by (a) that

$$\sum_{n=1}^{+\infty} \int_X |f\chi_{A_n}| \, d\mu = \sum_{n=1}^{+\infty} \int_{A_n} |f| \, d\mu = \int_A |f| \, d\mu < +\infty.$$

Now, since $\sum_{n=1}^{+\infty} f\chi_{A_n} = f\chi_A$ on X, we get

$$\sum_{n=1}^{+\infty} \int_{A_n} f\,d\mu = \sum_{n=1}^{+\infty} \int_X f\chi_{A_n}\,d\mu = \int_X f\chi_A\,d\mu = \int_A f\,d\mu$$

by the Series Theorem.
(iii) We have that $A = A_1 \cup \big(\bigcup_{k=2}^{+\infty}(A_k \setminus A_{k-1})\big)$, where the sets in the union are pairwise disjoint. We apply (ii) and we get

$$\begin{aligned}\int_A f\,d\mu &= \int_{A_1} f\,d\mu + \sum_{k=2}^{+\infty} \int_{A_k\setminus A_{k-1}} f\,d\mu = \int_{A_1} f\,d\mu + \lim_{n\to+\infty} \sum_{k=2}^{n} \int_{A_k\setminus A_{k-1}} f\,d\mu \\ &= \lim_{n\to+\infty} \int_{A_n} f\,d\mu.\end{aligned}$$

(iv) We have that $(A_N \setminus A_n)$ is increasing and $\bigcup_{n=1}^{+\infty}(A_N \setminus A_n) = A_N \setminus A$. So (iii) implies

$$\lim_{n\to+\infty} \int_{A_N\setminus A_n} f\,d\mu = \int_{A_N\setminus A} f\,d\mu. \tag{3.1}$$

Now, from the equality

$$\int_{A_N\setminus A} f\,d\mu + \int_A f\,d\mu = \int_{A_N} f\,d\mu$$

and from $\big|\int_{A_N} f\,d\mu\big| < +\infty$ we get $\big|\int_A f\,d\mu\big| < +\infty$. From the same equality we then get

$$\int_{A_N\setminus A} f\,d\mu = \int_{A_N} f\,d\mu - \int_A f\,d\mu.$$

Similarly,

$$\int_{A_N\setminus A_n} f\,d\mu = \int_{A_N} f\,d\mu - \int_{A_n} f\,d\mu$$

for all $n \geq N$, and now (3.1) implies

$$\int_{A_N} f\,d\mu - \lim_{n\to+\infty} \int_{A_n} f\,d\mu = \int_{A_N} f\,d\mu - \int_A f\,d\mu.$$

Because of $\big|\int_{A_N} f\,d\mu\big| < +\infty$ again, we get $\lim_{n\to+\infty} \int_{A_n} f\,d\mu = \int_A f\,d\mu$. $\square$

We must say that all results we have proved about integrals $\int_X$ over X hold without change for integrals $\int_A$ over an arbitrary $A \in \mathcal{S}$. To see this we either repeat all proofs, making the necessary *minor* changes, or we just apply those results to the functions multiplied by χ_A or to their restrictions on A. As an example let us look at the following version of the Dominated Convergence Theorem.

Let $f, f_1, f_2, \ldots : X \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ and $g : X \to [0,+\infty]$ be measurable. Let also $\lim_{n\to+\infty} f_n = f$ a.e. on A, $|f_n| \leq g$ a.e. on A for all n and $\int_A g\,d\mu < +\infty$. Then $\lim_{n\to+\infty} \int_A f_n\,d\mu = \int_A f\,d\mu$.

Indeed, the assumptions imply that $\lim_{n\to+\infty} f_n\chi_A = f\chi_A$ a.e. on X, $|f_n\chi_A| \leq g\chi_A$ a.e. on X for all n and $\int_X g\chi_A\,d\mu < +\infty$. Then the Dominated Convergence Theorem (for X) implies

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

Alternatively, the same assumptions imply $\lim_{n\to+\infty} f_n\rceil A = f\rceil A$ a.e on A, $|f_n\rceil A| \leq g\rceil A$ a.e. on A for all n and $\int_A (g\rceil A)\,d(\mu\rceil A) < +\infty$. Again, the Dominated Convergence Theorem (for A) implies

$$\lim_{n\to+\infty} \int_A f_n\,d\mu = \lim_{n\to+\infty} \int_A (f_n\rceil A)\,d(\mu\rceil A) = \int_A (f\rceil A)\,d(\mu\rceil A) = \int_A f\,d\mu.$$

Exercises.

3.1.17. Consider the measure space $(X, \mathcal{P}(X), \delta_{x_0})$ for some $x_0 \in X$ and any $f : X \to \overline{\mathbb{R}}$. Prove that $\int_A f\,d\delta_{x_0} = f(x_0)$, if $x_0 \in A$, and $\int_A f\,d\delta_{x_0} = 0$, if $x_0 \notin A$.

3.1.18. Let $f, f_n : X \to [0, +\infty]$ be measurable. Assume that $\lim_{n\to+\infty} f_n = f$ a.e. on X and $\lim_{n\to+\infty} \int_X f_n \, d\mu = \int_X f \, d\mu < +\infty$, and prove that $\lim_{n\to+\infty} \int_A f_n \, d\mu = \int_A f \, d\mu$ for every $A \in \mathcal{S}$.
Hint. Aplly Fatou's Lemma over both A and A^c.

3.1.19. (i) Let $f : X \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ be integrable. Prove that for every $\epsilon > 0$ there is $E \in \mathcal{S}$ with $\mu(E) < +\infty$ and $\int_{E^c} |f| \, d\mu < \epsilon$.
Hint. Consider $E = \{x \in X \,\big|\, |f(x)| \geq \frac{1}{n}\}$ for large $n \in \mathbb{N}$.
(ii) Let f be Lebesgue integrable over $\mathbb{R}^n$. Prove that for every $\epsilon > 0$ there is a compact $K \subseteq \mathbb{R}^n$ so that $\int_{K^c} |f| \, dm_n < \epsilon$.
Hint. Consider K to be a large closed ball in $\mathbb{R}^n$ with center 0.

3.1.20. Let $f : X \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ be integrable. Prove that for every $\epsilon > 0$ there is $\delta > 0$ so that: $\big| \int_E f \, d\mu \big| < \epsilon$ for all $E \in \mathcal{S}$ with $\mu(E) < \delta$.
Hint. One may prove it first for simple functions and then use Proposition 3.20.

3.1.21. *Mean values.*
Let $f : X \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ be integrable and F be a closed subset of $\mathbb{R}$ or $\mathbb{C}$. If $\frac{1}{\mu(E)} \int_E f \, d\mu \in F$ for every $E \in \mathcal{S}$ with $\mu(E) > 0$, prove that $f(x) \in F$ for a.e. $x \in X$.
Hint. If $E \in \mathcal{S}, 0 < \mu(E) < +\infty, |f(x) - y_0| \leq r_0$ for all $x \in E$, then $\big| \frac{1}{\mu(E)} \int_E f \, d\mu - y_0 \big| \leq r_0$. Now, consider the open set $U = \mathbb{R} \setminus F$ or $\mathbb{C} \setminus F$, and prove that $\mu(\{x \in X \mid f(x) \in U\}) = 0$, using a covering of U by countably many closed intevals or closed discs which are contained in U.

POINT-MASS DISTRIBUTIONS.

Consider the point-mass distribution μ induced by a function $a : X \to [0, +\infty]$ through the formula

$$\mu(E) = \textstyle\sum_{x\in E} a_x$$

for all $E \subseteq X$.

We observe that *all* functions $f : X \to Y$, no matter what the measure space $(Y, \mathcal{S}_Y)$ is, are $(\mathcal{P}(X), \mathcal{S}_Y)$-measurable.

Proposition 3.32. *If* $f : X \to [0, +\infty]$ *then* $\int_X f \, d\mu = \sum_{x\in X} f(x) \, a_x$.

Proof. If ϕ is a non-negative simple function on X with standard representation $\phi = \sum_{j=1}^n \kappa_j \chi_{E_j}$, then

$$\begin{aligned} \textstyle\int_X \phi \, d\mu &= \textstyle\sum_{j=1}^n \kappa_j \mu(E_j) = \sum_{j=1}^n \kappa_j \big(\sum_{x\in E_j} a_x \big) = \sum_{j=1}^n \big(\sum_{x\in E_j} \kappa_j a_x \big) \\ &= \textstyle\sum_{j=1}^n \big(\sum_{x\in E_j} \phi(x) a_x \big) = \sum_{x\in X} \phi(x) \, a_x, \end{aligned}$$

where the last equality is implied by Proposition 1.23.
Now, we take an increasing sequence (ϕ_n) of non-negative simple functions so that $\lim_{n\to+\infty} \phi_n = f$ on X, and then $\lim_{n\to+\infty} \int_X \phi_n \, d\mu = \int_X f \, d\mu$. Since

$$\textstyle\int_X \phi_n \, d\mu = \sum_{x\in X} \phi_n(x) \, a_x \leq \sum_{x\in X} f(x) \, a_x,$$

we find $\int_X f \, d\mu \leq \sum_{x\in X} f(x) \, a_x$ by taking the limit as $n \to +\infty$.
If F is any finite subset of X, then

$$\textstyle\sum_{x\in F} \phi_n(x) \, a_x \leq \sum_{x\in X} \phi_n(x) \, a_x = \int_X \phi_n \, d\mu.$$

Taking the limit as $n \to +\infty$, we get $\sum_{x\in F} f(x) \, a_x \leq \int_X f \, d\mu$. Now, taking the supremum over the finite subsets of X, we find $\sum_{x\in X} f(x) \, a_x \leq \int_X f \, d\mu$. $\square$

We would like to extend the validity of Proposition 3.32 to (extended-) real valued or complex valued functions, but we do not have a definition for sums of (extended-) real valued or complex valued terms! We can give such a definition in a straightforward manner, but we prefer to use the theory of the integral developed so far.

The amusing thing is that any series $\sum_{i\in I} b_i$ of *non-negative* terms over the general index set I can be written as an integral

$$\sum_{i\in I} b_i = \int_I b\, d\sharp,$$

where $\sharp$ is the counting measure on I (and we freely write $b_i = b(i)$). This is a simple application of Proposition 3.32: we just take $X = I$, $f = b$, and $a_i = 1$ for all $i \in I$.

Using properties of integrals, we may prove corresponding properties of sums. For example, it is true that

$$\sum_{i\in I}(b_i + c_i) = \sum_{i\in I} b_i + \sum_{i\in I} c_i, \quad \sum_{i\in I} \lambda b_i = \lambda \sum_{i\in I} b_i$$

for every non-negative b_i, c_i and λ. The proof consists in rewriting

$$\int_I (b + c)\, d\sharp = \int_I b\, d\sharp + \int_I c\, d\sharp, \quad \int_I \lambda b\, d\sharp = \lambda \int_I b\, d\sharp$$

in terms of sums.

For every $b \in \overline{\mathbb{R}}$ we write $b^+ = \max\{b, 0\}$ and $b^- = -\min\{b, 0\}$, and then we have that $b = b^+ - b^-$ and $|b| = b^+ + b^-$.

Definition. *If I is any index set and $b : I \to \overline{\mathbb{R}}$, we define the **sum** of $(b_i)_{i\in I}$ over I by*

$$\sum_{i\in I} b_i = \sum_{i\in I} b_i^+ - \sum_{i\in I} b_i^-$$

*only when either $\sum_{i\in I} b_i^+ < +\infty$ or $\sum_{i\in I} b_i^- < +\infty$. We say that $(b_i)_{i\in I}$ is **summable** (over I) if $\sum_{i\in I} b_i$ is finite or, equivalently, if both $\sum_{i\in I} b_i^+$ and $\sum_{i\in I} b_i^-$ are finite.*

Since we can write

$$\sum_{i\in I} b_i = \sum_{i\in I} b_i^+ - \sum_{i\in I} b_i^- = \int_I b^+\, d\sharp - \int_I b^-\, d\sharp = \int_I b\, d\sharp$$

and also

$$\sum_{i\in I} |b_i| = \sum_{i\in I} b_i^+ + \sum_{i\in I} b_i^- = \int_I b^+\, d\sharp + \int_I b^-\, d\sharp = \int_I |b|\, d\sharp,$$

we may say that $(b_i)_{i\in I}$ *is summable over I if and only if b is integrable over I with respect to the counting measure $\sharp$ or, equivalently, if and only if $\sum_{i\in I} |b_i| = \int_I |b|\, d\sharp < +\infty$.* Also, $\sum_{i\in I} b_i$ *is defined if and only if $\int_I b\, d\sharp$ is defined and, in this case, they are equal.*

Further exploiting the analogy between sums and integrals, we have:

Definition. *If I is any index set and $b : I \to \overline{\mathbb{C}}$, we say that $(b_i)_{i\in I}$ is **summable** over I if $\sum_{i\in I} |b_i| < +\infty$.*

This is the same condition as in the case of $b : I \to \overline{\mathbb{R}}$.

Proposition 3.33. *Let $b : I \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$. Then $(b_i)_{i\in I}$ is summable if and only if $\{i \in I \mid b_i \neq 0\}$ is countable and $\sum_{k=1}^{+\infty} |b_{i_k}| < +\infty$, where $\{i_1, i_2, \dots\}$ is an arbitrary enumeration of I.*

Proof. An application of Propositions 1.20 and 1.21. □

In particular, if $(b_i)_{i\in I}$ is summable then b_i is *finite* for all i. This allows us to give the

Definition. *Let $b : I \to \overline{\mathbb{C}}$ be summable over I. We define the **sum** of $(b_i)_{i\in I}$ over I as*

$$\sum_{i\in I} b_i = \sum_{i\in I} \mathrm{Re}(b_i) + i \sum_{i\in I} \mathrm{Im}(b_i).$$

Therefore, the sum of complex valued terms is defined only when the sum is summable and, in this case, this sum has a finite value. Again, we can say that if $b : I \to \overline{\mathbb{C}}$ is summable over I (which is equivalent to b being integrable over I with respect to the counting measure) then

$$\sum_{i \in I} b_i = \int_I b \, d\,\sharp.$$

We shall see now the form that some of the important results on general integrals take when we specialize them to sums. They are simple and straightforward formulations of known results but, since they are very important when one is working with sums, we shall state them explicitly. Their content is *the interchange of limits and sums*. It should be stressed that it is very helpful to be able to recognize the underlying *integral* theorem behind a property of *sums*.

Monotone Convergence Theorem. *Let $b, b_1, b_2, \ldots : I \to [0, +\infty]$. If $(b_{n,i})$ is increasing for all $i \in I$ and $\lim_{n\to+\infty} b_{n,i} = b_i$ for all $i \in I$, then $\lim_{n\to+\infty} \sum_{i\in I} b_{n,i} = \sum_{i\in I} b_i$.*

Non-negative Series Theorem. *Let $b_1, b_2, \ldots : I \to [0, +\infty]$. Then $\sum_{i\in I} \big(\sum_{n=1}^{+\infty} b_{n,i} \big) = \sum_{n=1}^{+\infty} \big(\sum_{i\in I} b_{n,i} \big)$.*

Fatou's Lemma. *Let $b, b_1, b_2, \ldots : I \to [0, +\infty]$. If $b_i = \underline{\lim}_{n\to+\infty} b_{n,i}$ for all $i \in I$, then $\sum_{i\in I} b_i \leq \underline{\lim}_{n\to+\infty} \sum_{i\in I} b_{n,i}$.*

Dominated Convergence Theorem. *Let $b, b_1, b_2, \ldots : I \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ and $c : I \to [0, +\infty]$. If $\lim_{n\to+\infty} b_{n,i} = b_i$ for all $i \in I$, and $|b_{n,i}| \leq c_i$ for all $i \in I$ and $n \in \mathbb{N}$, and $\sum_{i\in I} c_i < +\infty$, then $\lim_{n\to+\infty} \sum_{i\in I} b_{n,i} = \sum_{i\in I} b_i$.*

Series Theorem. *Let $b_1, b_2, \ldots : I \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$. If $\sum_{n=1}^{+\infty} \big(\sum_{i\in I} |b_{n,i}| \big) < +\infty$, then $\sum_{n=1}^{+\infty} b_{n,i}$ converges for every $i \in I$ and $\sum_{i\in I} \big(\sum_{n=1}^{+\infty} b_{n,i} \big) = \sum_{n=1}^{+\infty} \big(\sum_{i\in I} b_{n,i} \big)$.*

Observe that $\emptyset$ is the only $\sharp$-null set. Therefore, saying that a property holds $\sharp$-a.e. on I is equivalent to saying that it holds at every point of I.

Now we go back to the general case, where μ is the point-mass distribution induced by the function $a : X \to [0, +\infty]$, and $f : X \to \overline{\mathbb{R}}$. Using Proposition 3.32, we get

$$\int_X f^+ \, d\mu = \sum_{x\in X} f^+(x) a_x, \quad \int_X f^- \, d\mu = \sum_{x\in X} f^-(x) a_x.$$

Then $\int_X f \, d\mu$ is defined if and only if either $\int_X f^+ \, d\mu < +\infty$ or $\int_X f^- \, d\mu < +\infty$, and in this case we have

$$\int_X f \, d\mu = \int_X f^+ \, d\mu - \int_X f^- \, d\mu = \sum_{x\in X} f^+(x) a_x - \sum_{x\in X} f^-(x) a_x = \sum_{x\in X} f(x) a_x.$$

Moreover, f is integrable if and only if

$$\int_X |f| \, d\mu = \sum_{x\in X} |f(x)| a_x < +\infty.$$

This is also true when $f : X \to \overline{\mathbb{C}}$, and in this case we have

$$\int_X f \, d\mu = \sum_{x\in X} \mathrm{Re}(f(x)\chi_{D_f}(x)) a_x + i \sum_{x\in X} \mathrm{Im}(f(x)\chi_{D_f}(x)) a_x,$$

where

$$D_f = \{x \in X \mid f(x) \neq \infty\}.$$

Since $\sum_{x\in X} |f(x)| a_x < +\infty$, it is clear that $f(x) = \infty$ can happen only if $a_x = 0$, and $a_x = +\infty$ can happen only if $f(x) = 0$. But, then $f(x) a_x \in \mathbb{C}$ for *all* $x \in X$ and, moreover, $f(x)\chi_{D_f}(x) a_x = f(x) a_x$ for *all* $x \in X$. Therefore, we get

$$\int_X f \, d\mu = \sum_{x\in X} \mathrm{Re}(f)(x) a_x + i \sum_{x\in X} \mathrm{Im}(f)(x) a_x = \sum_{x\in X} f(x) a_x.$$

Now we have arrived at the complete interpretation of sums as integrals.

Proposition 3.34. *Let μ be the point-mass distribution induced by $a : X \to [0, +\infty]$.*

(i) If $f : X \to \overline{\mathbb{R}}$, then $\int_X f \, d\mu$ is defined if and only if $\sum_{x\in X} f(x) a_x$ is defined and, in this case, we have $\int_X f \, d\mu = \sum_{x\in X} f(x) a_x$.

(ii) If $f : X \to \overline{\mathbb{C}}$, then f is μ-integrable if and only if $\sum_{x\in X} |f(x)| a_x < +\infty$ and, in this case, the equality in (i) is true.

3.2 Lebesgue integral.

A function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ is Lebesgue integrable if it is Lebesgue measurable and also integrable with respect to m_n. For example, it is easy to see that every continuous $f : \mathbb{R}^n \to \mathbb{R}$ or $\mathbb{C}$ which is 0 outside some bounded set is Lebesgue integrable. Indeed, f is then Borel measurable, and, if Q is any closed interval in $\mathbb{R}^n$ outside of which f is 0, then $|f| \leq M\chi_Q$, where $M = \max\{|f(x)| \,|\, x \in Q\} < +\infty$. Therefore,

$$\int_{\mathbb{R}^n} |f| \, dm_n \leq M \int_{\mathbb{R}^n} \chi_Q \, dm_n = M m_n(Q) < +\infty.$$

LEBESGUE INTEGRAL VS RIEMANN INTEGRAL.

We shall now investigate the relation between the Lebesgue integral and the Riemann integral. We recall the definition of the latter.

We consider a bounded closed interval $Q = [a_1, b_1] \times \cdots \times [a_n, b_n]$ in $\mathbb{R}^n$, and a *bounded* function $f : Q \to \mathbb{R}$. If $l \in \mathbb{N}$ is arbitrary and $Q_1, \ldots, Q_l$ are arbitrary closed intervals which have *pairwise disjoint interiors* and so that $Q = Q_1 \cup \cdots \cup Q_l$, then we say that $\Delta = \{Q_1, \ldots Q_l\}$ is a **partition** of Q. If $P, P_1, \ldots, P_l$ are the open-closed intervals with the same sides as, respectively, $Q, Q_1, \ldots, Q_l$, then $\{Q_1, \ldots, Q_l\}$ is a partition of Q if and only if the $P_1, \ldots, P_l$ are pairwise disjoint and $P = P_1 \cup \cdots \cup P_l$. Now, since f is bounded, for each Q_j we may consider the real numbers $m_j = \inf\{f(x) \,|\, x \in Q_j\}$ and $M_j = \sup\{f(x) \,|\, x \in Q_j\}$. We then define the **lower Darboux sum** and the **upper Darboux sum** of f with respect to the partition Δ as, respectively,

$$\underline{\Sigma}(f;\Delta) = \sum_{j=1}^{l} m_j \ \mathrm{vol}_n(Q_j), \quad \overline{\Sigma}(f;\Delta) = \sum_{j=1}^{l} M_j \ \mathrm{vol}_n(Q_j).$$

If $m = \inf\{f(x) \,|\, x \in Q\}$, $M = \sup\{f(x) \,|\, x \in Q\}$, we have that

$$m \leq m_j \leq M_j \leq M$$

for every j. Using Proposition 1.30 (and working with the corresponding open-closed intervals), we see that

$$m \ \mathrm{vol}_n(Q) \leq \underline{\Sigma}(f;\Delta) \leq \overline{\Sigma}(f;\Delta) \leq M \ \mathrm{vol}_n(Q).$$

If $\Delta_1 = \{Q_1^{(1)}, \ldots, Q_{l_1}^{(1)}\}$ and $\Delta_2 = \{Q_1^{(2)}, \ldots, Q_{l_2}^{(2)}\}$ are two partitions of Q, we say that Δ_2 is **finer** than Δ_1 if every $Q_i^{(2)}$ is included in some $Q_j^{(1)}$. Then it is obvious that, for every $Q_j^{(1)}$ of Δ_1, the $Q_i^{(2)}$ of Δ_2 which are included in $Q_j^{(1)}$ form a partition of $Q_j^{(1)}$. Therefore, from Proposition 1.30 again,

$$\begin{aligned} m_j^{(1)} \ \mathrm{vol}_n(Q_j^{(1)}) &\leq \sum_{i:\, Q_i^{(2)} \subseteq Q_j^{(1)}} m_i^{(2)} \ \mathrm{vol}_n(Q_i^{(2)}) \\ &\leq \sum_{i:\, Q_i^{(2)} \subseteq Q_j^{(1)}} M_i^{(2)} \ \mathrm{vol}_n(Q_i^{(2)}) \leq M_j^{(1)} \ \mathrm{vol}_n(Q_j^{(1)}). \end{aligned}$$

Summing over all $j = 1, \ldots, l_1$ we find

$$\underline{\Sigma}(f;\Delta_1) \leq \underline{\Sigma}(f;\Delta_2) \leq \overline{\Sigma}(f;\Delta_2) \leq \overline{\Sigma}(f;\Delta_1).$$

Now, if $\Delta_1 = \{Q_1^{(1)}, \ldots, Q_{l_1}^{(1)}\}$ and $\Delta_2 = \{Q_1^{(2)}, \ldots, Q_{l_2}^{(2)}\}$ are any two partitions of Q, we form their common refinement $\Delta = \{Q_j^{(1)} \cap Q_i^{(2)} \,|\, 1 \leq j \leq l_1, 1 \leq i \leq l_2\}$, and we get

$$\underline{\Sigma}(f;\Delta_1) \leq \underline{\Sigma}(f;\Delta) \leq \overline{\Sigma}(f;\Delta) \leq \overline{\Sigma}(f;\Delta_2).$$

We conclude that

$$m \ \mathrm{vol}_n(Q) \leq \underline{\Sigma}(f;\Delta_1) \leq \overline{\Sigma}(f;\Delta_2) \leq M \ \mathrm{vol}_n(Q)$$

for all partitions Δ_1, Δ_2 of Q. Now, we define

$$(\mathcal{R}_n)\underline{\int}_Q f = \sup\{\underline{\Sigma}(f;\Delta) \mid \Delta \text{ partition of } Q\}, \quad (\mathcal{R}_n)\overline{\int}_Q f = \inf\{\overline{\Sigma}(f;\Delta) \mid \Delta \text{ partition of } Q\}$$

and we call them, respectively, the **lower Riemann integral** and the **upper Riemann integral** of f over Q. It is then clear that

$$m \ \mathrm{vol}_n(Q) \le (\mathcal{R}_n)\underline{\int}_Q f \le (\mathcal{R}_n)\overline{\int}_Q f \le M \ \mathrm{vol}_n(Q).$$

We say that f is **Riemann integrable** over Q if $(\mathcal{R}_n)\underline{\int}_Q f = (\mathcal{R}_n)\overline{\int}_Q f$. In this case we define

$$(\mathcal{R}_n) \int_Q f = (\mathcal{R}_n)\underline{\int}_Q f = (\mathcal{R}_n)\overline{\int}_Q f$$

and we call it the **Riemann integral** of f over Q.

Lemma 3.3. *The bounded $f : Q \to \mathbb{R}$ is Riemann integrable over the bounded closed interval Q if and only if for every $\epsilon > 0$ there is a partition Δ of Q so that $\overline{\Sigma}(f;\Delta) - \underline{\Sigma}(f;\Delta) < \epsilon$.*

Proof. For the sufficiency, we take an arbitrary $\epsilon > 0$. Then for the corresponding Δ we have

$$0 \le (\mathcal{R}_n)\overline{\int}_Q f - (\mathcal{R}_n)\underline{\int}_Q f \le \overline{\Sigma}(f;\Delta) - \underline{\Sigma}(f;\Delta) < \epsilon.$$

This implies the equality of the upper and lower Riemann integrals of f over Q.
For the necessity, we assume $(\mathcal{R}_n)\underline{\int}_Q f = (\mathcal{R}_n)\overline{\int}_Q f$, and then for each $\epsilon > 0$ we take partitions Δ_1, Δ_2 of Q so that

$$(\mathcal{R}_n) \int_Q f - \tfrac{\epsilon}{2} < \underline{\Sigma}(f;\Delta_1), \quad \overline{\Sigma}(f;\Delta_2) < (\mathcal{R}_n) \int_Q f + \tfrac{\epsilon}{2}.$$

Then

$$\overline{\Sigma}(f;\Delta) - \underline{\Sigma}(f;\Delta) \le \overline{\Sigma}(f;\Delta_2) - \underline{\Sigma}(f;\Delta_1) < \epsilon$$

for the common refinement Δ of Δ_1 and Δ_2. $\square$

Proposition 3.35. *If $f : Q \to \mathbb{R}$ is continuous on the bounded closed interval Q, then f is Riemann integrable over Q.*

Proof. By uniform continuity of f on Q, for any $\epsilon > 0$ there is a $\delta > 0$ so that $|f(x) - f(y)| < \frac{\epsilon}{\mathrm{vol}_n(Q)}$ for all $x, y \in Q$ whose distance is $< \delta$. We take any partition $\Delta = \{Q_1, \dots, Q_l\}$ of Q, so that every Q_j has diameter $< \delta$. Then $|f(x) - f(y)| < \frac{\epsilon}{\mathrm{vol}_n(Q)}$ for all x, y on the same Q_j. This implies that for every Q_j we have $M_j - m_j < \frac{\epsilon}{\mathrm{vol}_n(Q)}$. Hence

$$\overline{\Sigma}(f;\Delta) - \underline{\Sigma}(f;\Delta) = \textstyle\sum_{j=1}^l (M_j - m_j) \ \mathrm{vol}_n(Q_j) < \frac{\epsilon}{\mathrm{vol}_n(Q)} \sum_{j=1}^l \mathrm{vol}_n(Q_j) = \epsilon,$$

and Lemma 3.3 implies that f is Riemann integrable over Q. $\square$

Theorem 3.2. *If $f : Q \to \mathbb{R}$ is Riemann integrable over the bounded closed interval Q and we extend f as 0 outside Q, then f is Lebesgue integrable and $\int_{\mathbb{R}^n} f \, dm_n = \int_Q f \, dm_n = (\mathcal{R}_n) \int_Q f$.*

Proof. Lemma 3.3 implies that, for every $k \in \mathbb{N}$, there is a partition $\Delta_k = \{Q_1^{(k)}, \dots, Q_{l_k}^{(k)}\}$ of Q so that

$$\overline{\Sigma}(f;\Delta_k) - \underline{\Sigma}(f;\Delta_k) < \tfrac{1}{k}.$$

We consider the simple functions

$$\psi_k = \textstyle\sum_{j=1}^{l_k} m_j^{(k)} \chi_{P_j^{(k)}}, \quad \phi_k = \sum_{j=1}^{l_k} M_j^{(k)} \chi_{P_j^{(k)}},$$

where $m_j^{(k)} = \inf\{f(x)\,|\,x \in Q_j^{(k)}\}$ and $M_j^{(k)} = \sup\{f(x)\,|\,x \in Q_j^{(k)}\}$ and $P_j^{(k)}$ is the open-closed interval with the same sides as $Q_j^{(k)}$. Clearly, all ψ_k, ϕ_k are Borel measurable. Now, we have

$$\lim_{k\to+\infty} \overline{\Sigma}(f;\Delta_k) = (\mathcal{R}_n)\int_Q f, \quad \lim_{k\to+\infty} \underline{\Sigma}(f;\Delta_k) = (\mathcal{R}_n)\int_Q f.$$

It is clear that $\psi_k \le f\chi_P \le \phi_k$ on $\mathbb{R}^n$ for all k, where P is the open-closed interval with the same sides as Q. It is also clear that

$$\int_{\mathbb{R}^n} \psi_k\, dm_n = \sum_{j=1}^{l_k} m_j^{(k)}\ \mathrm{vol}_n(P_j^{(k)}) = \sum_{j=1}^{l_k} m_j^{(k)}\ \mathrm{vol}_n(Q_j^{(k)}) = \underline{\Sigma}(f;\Delta_k)$$

$$\int_{\mathbb{R}^n} \phi_k\, dm_n = \sum_{j=1}^{l_k} M_j^{(k)}\ \mathrm{vol}_n(P_j^{(k)}) = \sum_{j=1}^{l_k} M_j^{(k)}\ \mathrm{vol}_n(Q_j^{(k)}) = \overline{\Sigma}(f;\Delta_k).$$

Hence,

$$\lim_{k\to+\infty} \int_{\mathbb{R}^n} \psi_k\, dm_n = \lim_{k\to+\infty} \int_{\mathbb{R}^n} \phi_k\, dm_n = (\mathcal{R}_n)\int_Q f.$$

We define

$$g = \overline{\lim_{k\to+\infty}}\ \psi_k, \quad h = \underline{\lim_{k\to+\infty}}\ \phi_k,$$

and then g, h are Borel measurable, and $g \le f\chi_P \le h$ on $\mathbb{R}^n$.
Fatou's Lemma implies

$$0 \le \int_{\mathbb{R}^n}(h-g)\, dm_n \le \underline{\lim}_{k\to+\infty} \int_{\mathbb{R}^n}(\phi_k - \psi_k)\, dm_n = 0.$$

Hence, $g = h$ a.e. on $\mathbb{R}^n$, and so $f\chi_P = g = h$ a.e. on $\mathbb{R}^n$. Since g, h are Borel measurable, we have that $f\chi_P$ is Lebesgue measurable. Since $f = 0$ outside Q, we have that $f \ne f\chi_P$ only on a subset of $Q \setminus P$, and so $f = f\chi_P$ a.e. on $\mathbb{R}^n$. Hence, f is Lebesgue measurable.
Now, f is bounded and $f = 0$ outside Q, and so $|f| \le K\chi_Q$, where $K = \sup\{|f(x)|\,|\,x \in Q\}$. Thus, $\int_{\mathbb{R}^n} |f|\, dm_n \le K m_n(Q) < +\infty$, and so f is Lebesgue integrable.
Another application of Fatou's Lemma gives

$$\int_{\mathbb{R}^n}(h - f\chi_P)\, dm_n \le \underline{\lim}_{k\to+\infty} \int_{\mathbb{R}^n}(\phi_k - f\chi_P)\, dm_n = (\mathcal{R}_n)\int_Q f - \int_{\mathbb{R}^n} f\chi_P\, dm_n,$$

$$\int_{\mathbb{R}^n}(f\chi_P - g)\, dm_n \le \underline{\lim}_{k\to+\infty} \int_{\mathbb{R}^n}(f\chi_P - \psi_k)\, dm_n = \int_{\mathbb{R}^n} f\chi_P\, dm_n - (\mathcal{R}_n)\int_Q f.$$

Hence,

$$\int_{\mathbb{R}^n} h\, dm_n \le (\mathcal{R}_n)\int_Q f \le \int_{\mathbb{R}^n} g\, dm_n.$$

Since $f = g = h$ a.e. on $\mathbb{R}^n$, we conclude that $(\mathcal{R}_n)\int_Q f = \int_{\mathbb{R}^n} f\, dm_n$. □

The converse of Theorem 3.2 does not hold. There are bounded $f : Q \to \mathbb{R}$ which are Lebesgue integrable but not Riemann integrable over Q.

Example. We define $f(x) = 1$, if $x \in Q$ has only rational coordinates, and $f(x) = 0$, if $x \in Q$ has at least one irrational coordinate. If $\Delta = \{Q_1, \dots, Q_k\}$ is any partition of Q, then all Q_j with non-empty interior (the rest do not matter because they have zero volume) contain at least one x with $f(x) = 0$ and at least one x with $f(x) = 1$. Hence, for all such Q_j we have $m_j = 0$ and $M_j = 1$, and so $\underline{\Sigma}(f;\Delta) = 0$ and $\overline{\Sigma}(f;\Delta) = \mathrm{vol}_n(Q)$ for every Δ. Thus, $(\mathcal{R}_1)\underline{\int}_Q f = 0$ and $(\mathcal{R}_n)\overline{\int}_Q f = \mathrm{vol}_n(Q)$, and so f is not Riemann integrable over Q.
On the other hand, if we extend f as 0 outside Q, then $f = 0$ a.e on $\mathbb{R}^n$, and so f is Lebesgue integrable over $\mathbb{R}^n$ with $\int_{\mathbb{R}^n} f\, dm_n = \int_Q f\, dm_n = 0$.

Theorem 3.2 incorporates the notion of Riemann integral in the notion of Lebesgue integral. It says that the collection of Riemann integrable functions is included in the collection of Lebesgue integrable functions and that the Riemann integral is the restriction of the Lebesgue integral on the collection of Riemann integrable functions. This provides us with greater flexibility over the symbol we may use for the Lebesgue integral, at least in the case of bounded intervals $[a, b]$ in the one-dimensional space $\mathbb{R}$. The standard symbol used in Infinitesimal Calculus for the Riemann integral $(\mathcal{R}_1)\int_{[a,b]} f$ is the familiar

$$\int_a^b f(x)\,dx.$$

We may now use the same symbol for the Lebesgue integral

$$\int_{[a,b]} f\,dm_1 = \int_{[a,b]} f(x)\,dm_1(x),$$

without the danger of confusion between the Riemann and the Lebesgue integrals when the function is integrable both in the Riemann and in the Lebesgue sense. Now, since the one-point sets $\{a\}$, $\{b\}$ have zero Lebesgue measure, the Lebesgue integrals $\int_{[a,b]} f\,dm_1$, $\int_{(a,b]} f\,dm_1$, $\int_{[a,b)} f\,dm_1$ and $\int_{(a,b)} f\,dm_1$ are all the same. Therefore, we may use the symbol $\int_a^b f(x)\,dx$ for all these Lebesgue integrals. This is extended to cases where the Riemann integral does not apply. For example, we may use the symbol

$$\int_{-\infty}^{+\infty} f(x)\,dx$$

for the Lebesgue integral $\int_{\mathbb{R}} f\,dm_1$ and, likewise, the symbol $\int_a^{+\infty} f(x)\,dx$ for the Lebesgue integral $\int_{[a,+\infty)} f\,dm_1$ and the symbol $\int_{-\infty}^b f(x)\,dx$ for the Lebesgue integral $\int_{(-\infty,b]} f\,dm_1$.

Theorem 3.2 provides us with a powerful tool to *calculate* Lebesgue integrals, at least in the case of $\mathbb{R}$. If a function f is Riemann integrable over a closed interval $[a,b]\subseteq\mathbb{R}$, we have many techniques (integration by parts, change of variable, primitives etc) to calculate its $\int_a^b f(x)\,dx$ which is the same as $\int_{[a,b]} f(x)\,dm_1(x)$. Moreover, if the given f is Riemann integrable over intervals $[a_k,b_k]$ with $\lim_{k\to+\infty} a_k = -\infty$ and $\lim_{k\to+\infty} b_k = +\infty$ and if we can calculate the integrals $\int_{a_k}^{b_k} f(x)\,dx = \int_{[a_k,b_k]} f(x)\,dm_1(x)$, then it is a matter of being able to justify the limit $\lim_{k\to+\infty}\int_{[a_k,b_k]} f(x)\,dm_1(x) = \int_{\mathbb{R}} f(x)\,dm_1(x)$ in order to calculate the Lebesgue integral over $\mathbb{R}$. To do this we may try to use the Monotone Convergence Theorem or the Dominated Convergence Theorem.

Exercises.

3.2.1. *The graph and the volume under the graph of a function.*
Let $f:\mathbb{R}^n\to[0,+\infty]$. If $A_f=\{(x_1,\dots,x_n,x_{n+1})\,|\,0\le x_{n+1}<f(x_1,\dots,x_n)\}\subseteq\mathbb{R}^{n+1}$ and $G_f=\{(x_1,\dots,x_n,x_{n+1})\,|\,x_{n+1}=f(x_1,\dots,x_n)\}\subseteq\mathbb{R}^{n+1}$ and if f is Lebesgue measurable, prove that $A_f,G_f\in\mathcal{L}_{n+1}$ and $m_{n+1}(A_f)=\int_{\mathbb{R}^n} f\,dm_n$ and $m_{n+1}(G_f)=0$.
Hint. Prove $m_{n+1}(A_\phi)=\int_{\mathbb{R}^n}\phi\,dm_n$ when $\phi:\mathbb{R}^n\to[0,+\infty)$ is a Lebesgue measurable simple function.

3.2.2. *Improper Integrals.*
Let $f:[a,b)\to\mathbb{R}$, where $-\infty<a<b\le+\infty$. If f is Riemann integrable over $[a,c]$ for every $c\in(a,b)$ and the limit $\lim_{c\to b-}\int_a^c f(x)\,dx$ exists in $\overline{\mathbb{R}}$, we say that the **improper integral** of f over $[a,b)$ exists and we define it as $\int_a^{\to b} f(x)\,dx=\lim_{c\to b-}\int_a^c f(x)\,dx$.
We have similar terminology and definition for $\int_{a\leftarrow}^b f(x)\,dx$, the improper integral of f over $(a,b]$.
(i) Let $f:[a,b)\to[0,+\infty)$ be Riemann integrable over $[a,c]$ for every $c\in(a,b)$. Prove that the Lebesgue integral $\int_a^b f(x)\,dx$ and the improper integral $\int_a^{\to b} f(x)\,dx$ both exist and they are equal.
(ii) Let $f:[a,b)\to\mathbb{R}$ be Riemann integrable over $[a,c]$ for every $c\in(a,b)$. Prove that, if the Lebesgue integral $\int_a^b f(x)\,dx$ exists, then $\int_a^{\to b} f(x)\,dx$ also exists and the two integrals are equal.
(iii) Prove that the converse of (ii) is not true in general: look at the fourth function in exercise 3.2.4 or at (ii) and (iii) of exercise 3.2.18.
(iv) If $\int_a^{\to b}|f(x)|\,dx<+\infty$ (then we say that the improper integral is **absolutely convergent**), prove that $\int_a^{\to b} f(x)\,dx$ exists and is a real number (then we say that the improper integral is **convergent**.)

3.2.3. Using improper integrals (see exercise 3.2.2), find the Lebesgue integral $\int_{-\infty}^{+\infty} f(x)\,dx$ (if it exists), where $f(x)$ is any of the functions:

$$\tfrac{1}{1+x^2},\quad e^{-|x|},\quad \tfrac{1}{x^2}\chi_{[0,+\infty)}(x),\quad \tfrac{1}{x},\quad \tfrac{1}{|x|},\quad \tfrac{1}{|x|^{1/2}}\chi_{[-1,1]}(x),\quad \textstyle\sum_{n=1}^{+\infty}\tfrac{1}{2^n}\chi_{[n,n+1)}(x),$$

$$\textstyle\sum_{n=1}^{+\infty}\tfrac{(-1)^{n+1}}{2^n}\chi_{[n,n+1)}(x),\quad \sum_{n=1}^{+\infty}\tfrac{1}{n}\chi_{[n,n+1)}(x),\quad \sum_{n=1}^{+\infty}\tfrac{(-1)^{n+1}}{n}\chi_{[n,n+1)}(x).$$

3.2.4. Apply the Fatou's Lemma for Lebesgue measure on $\mathbb{R}$ and the sequences (f_n), where $f_n(x)$ is any of the functions:

$$\chi_{(n,n+1)}(x),\quad \chi_{(n,+\infty)}(x),\quad n\chi_{(0,\frac{1}{n})}(x),\quad 1+\text{sign}\big(\sin\tfrac{2^n x}{2\pi}\big).$$

3.2.5. If f is Lebesgue integrable over $[-1,1]$, prove that $\lim_{n\to+\infty}\int_{-1}^{1} x^n f(x)\,dx = 0$.

3.2.6. Prove that the limit $I=\lim_{t\to+\infty}\frac{1}{\pi}\int_a^{+\infty}\frac{t}{1+t^2x^2}\,dx$ exists, and that: $I=0$, if $0<a$, and $I=\frac{1}{2}$, if $a=0$, and $I=1$, if $a<0$.

3.2.7. Prove that the limit $I=\lim_{n\to+\infty}\int_0^n\big(1+\frac{x}{n}\big)^n e^{-\alpha x}\,dx$ exists, and that: $I=\frac{1}{\alpha-1}$, if $1<\alpha$, and $I=+\infty$, if $\alpha\le 1$.

3.2.8. Let $f:\mathbb{R}\to\overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ be Lebesgue integrable. Prove $F(x)=\int_{-\infty}^{x} f(t)\,dt$ is a continuous function of x on $\mathbb{R}$.

3.2.9. *Continuity of translations.*
If $f:\mathbb{R}^n\to\overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ is Lebesgue integrable, prove $\lim_{h\to 0}\int_{\mathbb{R}^n}|f(x-h)-f(x)|\,dm_n(x)=0$.
Hint. Prove it first for continuous functions which are 0 outside a bounded set, and then use Theorem 3.12.

3.2.10. Let $Q\subseteq\mathbb{R}^n$ be a bounded closed interval and $x_0\in Q$. If $f:Q\to\mathbb{R}$ is Riemann integrable over Q and $g:Q\to\mathbb{R}$ coincides with f on $Q\setminus\{x_0\}$, prove that g is also Riemann integrable over Q and that $(\mathcal{R}_n)\int_Q g=(\mathcal{R}_n)\int_Q f$.

3.2.11. Let $Q\subseteq\mathbb{R}^n$ be a bounded closed interval, $\lambda\in\mathbb{R}$ and $f,g:Q\to\mathbb{R}$ be Riemann integrable over Q. Prove that $f+g,\lambda f$ and fg are all Riemann integrable over Q and $(\mathcal{R}_n)\int_Q(f+g)=(\mathcal{R}_n)\int_Q f+(\mathcal{R}_n)\int_Q g$ and $(\mathcal{R}_n)\int_Q\lambda f=\lambda(\mathcal{R}_n)\int_Q f$.

3.2.12. Let $Q\subseteq\mathbb{R}^n$ be a bounded closed interval.
(i) If the bounded functions $f,f_k:Q\to\mathbb{R}$ are all Riemann integrable over Q and $f_k\uparrow f$ on Q, prove that $(\mathcal{R}_n)\int_Q f_k\uparrow(\mathcal{R}_n)\int_Q f$.
(ii) Find bounded functions $f,f_k:Q\to\mathbb{R}$ so that $f_k\uparrow f$ on Q and so that all f_k are Riemann integrable over Q, but f is not Riemann integrable over Q.

3.2.13. Consider the functions $f(x)=\frac{1}{2}\big(\int_0^x e^{-\frac{1}{2}t^2}\,dt\big)^2$ and $h(x)=\int_0^1\frac{e^{-\frac{1}{2}x^2(t^2+1)}}{t^2+1}\,dt$.
(i) Using exercise 3.1.16, prove that $f'(x)+h'(x)=0$ and, hence, $f(x)+h(x)=\frac{\pi}{4}$ for every x.
(ii) Prove that

$$\int_{-\infty}^{+\infty}e^{-\frac{1}{2}t^2}\,dt=\sqrt{2\pi}.$$

3.2.14. (i) Using exercise 3.1.16, prove that the function $F(t)=\int_0^{+\infty}e^{-tx}\frac{\sin x}{x}\,dx$ is differentiable on $(0,+\infty)$, and that $\frac{dF}{dt}(t)=-\frac{1}{1+t^2}$ for every $t>0$. Find the $\lim_{t\to+\infty}F(t)$, and conclude that $F(t)=\arctan\frac{1}{t}$ for every $t>0$.
(ii) Prove that the function $\frac{\sin x}{x}$ is not Lebesgue integrable over $(0,+\infty)$.
(iii) Prove that the improper integral (exercise 3.2.2) $\int_0^{\to+\infty}\frac{\sin x}{x}\,dx$ exists.
(iv) Justify the equality $\lim_{t\to 0+}F(t)=\int_0^{\to+\infty}\frac{\sin x}{x}\,dx$.

(v) Conclude that

$$\int_0^{\to+\infty} \frac{\sin x}{x}\, dx = \frac{\pi}{2}.$$

(vi) Prove that the limit $I = \lim_{t\to+\infty} \frac{1}{\pi} \int_a^{\to+\infty} \frac{\sin(tx)}{x}\, dx$ exists, and that: $I = 0$, if $0 < a$, and $I = \frac{1}{2}$, if $a = 0$, and $I = 1$, if $a < 0$.

3.2.15. Let $\mathbb{H}_+ = \{z = x + iy \in \mathbb{C} \,|\, x > 0\}$, and consider $\Gamma : \mathbb{H}_+ \to \mathbb{C}$ defined by

$$\Gamma(z) = \int_0^{+\infty} t^{z-1} e^{-t}\, dt.$$

This is called the **gamma-function**.

(i) Prove that this Lebesgue integral exists and is finite for every $z \in \mathbb{H}_+$.

(ii) Using exercises 3.1.15 and 3.1.16, prove that $\frac{\partial\Gamma}{\partial x}$, $\frac{\partial\Gamma}{\partial y}$ are continuous on $\mathbb{H}_+$, and that $\frac{\partial\Gamma}{\partial x}(z) = -i\frac{\partial\Gamma}{\partial y}(z)$ for every $z \in \mathbb{H}_+$. This means that Γ is holomorphic on $\mathbb{H}_+$.

(iii) Prove that $\Gamma(z+1) = z\Gamma(z)$ for every $z \in \mathbb{H}_+$, and that $\Gamma(1) = 1$.
Prove that $\Gamma(n) = (n-1)!$ for every $n \in \mathbb{N}$.

(iv) Prove that $\Gamma(\frac{1}{2}) = \sqrt{\pi}$.

3.2.16. Let $E \subseteq R^n$ be bounded.
We define the **inner Jordan content** $c_n^{(i)}(E)$ of E to be the supremum of $\sum_{j=1}^m \mathrm{vol}_n(R_j)$ for all $m \in \mathbb{N}$ and all pairwise disjoint open intervals $R_1, \dots, R_m$ with $\bigcup_{j=1}^m R_j \subseteq E$. We also define the **outer Jordan content** $c_n^{(o)}(E)$ of E to be the infimum of $\sum_{j=1}^m \mathrm{vol}_n(R_j)$ for all $m \in \mathbb{N}$ and all open intervals $R_1, \dots, R_m$ with $E \subseteq \bigcup_{j=1}^m R_j$.

(i) Prove that the values of $c_n^{(i)}(E)$ and $c_n^{(o)}(E)$ remain the same if in the above definitions we use closed intervals instead of open intervals.

(ii) Prove that $c_n^{(i)}(E) \leq c_n^{(o)}(E)$ for every bounded $E \subseteq \mathbb{R}^n$.

Now, E is called a **Jordan set** if $c_n^{(i)}(E) = c_n^{(o)}(E)$, and the value $c_n(E) = c_n^{(i)}(E) = c_n^{(o)}(E)$ is called the **Jordan content** of E.

(iii) If $c_n^{(o)}(E) = 0$, prove that E is a Jordan set.

(iv) Prove that all bounded intervals $S \subseteq \mathbb{R}^n$ are Jordan sets and $c_n(S) = \mathrm{vol}_n(S)$.

(v) Prove that E is a Jordan set if and only if for every $\epsilon > 0$ there exist pairwise disjoint open intervals $R_1, \dots, R_m$ and open intervals $R_1', \dots, R_k'$ so that $\bigcup_{j=1}^m R_j \subseteq E \subseteq \bigcup_{i=1}^k R_i'$ and $\sum_{i=1}^k \mathrm{vol}_n(R_i') - \sum_{j=1}^m \mathrm{vol}_n(R_j) < \epsilon$.

(vi) Prove that E is a Jordan set if and only if $c_n^{(o)}(\mathrm{bd}(E)) = 0$, where $\mathrm{bd}(E)$ is the boundary of E.

(vii) Prove that the collection of bounded Jordan sets is closed under finite unions and set-theoretic differences. If $E_1, \dots, E_l$ are pairwise disjoint Jordan sets, prove that $c_n(E) = \sum_{j=1}^l c_n(E_j)$.

(viii) If E is closed, prove that $m_n(E) = 0$ implies $c_n(E) = 0$. If E is not closed, then this result may not be true. For example, if $E = \mathbb{Q} \cap [0,1] \subseteq \mathbb{R}$, then $m_1(E) = 0$, but $c_1^{(i)}(E) = 0 < 1 = c_1^{(o)}(E)$, and so E is not a Jordan set. (See exercise 1.4.6.)

(ix) If E is a Jordan set, prove that E is a Lebesgue set and $m_n(E) = c_n(E)$.

(x) Let $E \subseteq Q$, where Q is any bounded closed interval. Prove that E is a Jordan set if and only if χ_E is Riemann integrable over Q, and that, in this case, $c_n(E) = (\mathcal{R}_n) \int_Q \chi_E$.

(xi) Let Q be a bounded closed interval, $f, g : Q \to \mathbb{R}$ be bounded and $E \subseteq Q$ be a Jordan set with $c_n(E) = 0$. If f is Riemann integrable over Q and $f = g$ on $Q \setminus E$, prove that g is also Riemann integrable over Q, and that $(\mathcal{R}_n) \int_Q f = (\mathcal{R}_n) \int_Q g$.

3.2.17. *Lebesgue's characterisation of Riemann integrable functions.*

Let $Q \subseteq \mathbb{R}^n$ be a bounded closed interval and $f : Q \to \mathbb{R}$ be bounded. For any $x \in Q$ we define

$$\omega_f(x) = \lim_{\delta\to 0+} \sup\{|f(x') - f(x'')| \,|\, x', x'' \in Q, |x' - x| < \delta, |x'' - x| < \delta\},$$

and we call it the **oscillation** of f at x.

(i) Prove that f is continuous at x if and only if $\omega_f(x)=0$.

(ii) Prove that for every $\epsilon>0$ the set $\{x\in Q\,|\,\omega_f(x)\geq\epsilon\}$ is closed.

(iii) Assume that $\{x\in Q\,|\,f$ is discontinuous at $x\}$ is a m_n-null set.
Take any $\epsilon>0$ and prove that there are closed subintervals $Q_1',\dots,Q_l'$ of Q with pairwise disjoint interiors so that $\{x\in Q\,|\,\omega_f(x)\geq\epsilon\}\subseteq Q_1'\cup\cdots\cup Q_l'$ and $\mathrm{vol}_n(Q_1')+\cdots+\mathrm{vol}_n(Q_l')<\epsilon$. Then prove that there are closed subintervals $Q_1'',\dots,Q_m''$ of Q so that $Q_1',\dots,Q_l',Q_1'',\dots,Q_m''$ form a partition Δ of Q, and then prove that $\overline{\Sigma}(f;\Delta)-\underline{\Sigma}(f;\Delta)<(M-m+\mathrm{vol}_n(Q))\epsilon$, where $m=\inf\{f(x)\,|\,x\in Q\}$, $M=\sup\{f(x)\,|\,x\in Q\}$.
Conclude that f is Riemann integrable over Q.

(iv) Assume that f is Riemann integrable over Q.
Take any $\epsilon>0$ and consider a partition $\Delta=\{Q_1,\dots,Q_k\}$ of Q so that $\overline{\Sigma}(f;\Delta)-\underline{\Sigma}(f;\Delta)<\epsilon^2$. Consider those subintervals among the $Q_1,\dots,Q_k$ which intersect the set $\{x\in Q\,|\,\omega_f(x)\geq\epsilon\}$, and prove that the sum of their volumes is $<\epsilon$. Thus, $m_n(\{x\in Q\,|\,\omega_f(x)\geq\epsilon\})<\epsilon$.
Conclude that $\{x\in Q\,|\,f$ is discontinuous at $x\}$ is a m_n-null set.

LEBESGUE INTEGRAL AND SIMPLE TRANSFORMATIONS.

Another topic is the effect on Lebesgue integrals of translations and linear transformations of the space.

Proposition 3.36. *Let $A\in\mathcal{L}_n$ and $f:A\to\overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ be $\mathcal{L}_n\rceil A$-measurable.*
(i) If $\int_A f\,dm_n$ is defined, then $\int_{\tau_z(A)}\tau_z(f)\,dm_n$ is defined and

$$\int_{\tau_z(A)}\tau_z(f)\,dm_n=\int_A f\,dm_n.$$

(ii) If f is Lebesgue integrable over A, then $\tau_z(f)$ is Lebesgue integrable over $\tau_z(A)$ and the equality in (i) is again true.

Proof. Let $\phi:A\to[0,+\infty)$ be a Lebesgue measurable simple function and let $\phi=\sum_{j=1}^m\kappa_j\chi_{E_j}$ be its standard representation. Then

$$\int_A\phi\,dm_n=\sum_{j=1}^m\kappa_j m_n(E_j).$$

It is clear that

$$\tau_z(\phi)(x)=\phi(x-z)=\sum_{j=1}^m\kappa_j\chi_{E_j}(x-z)=\sum_{j=1}^m\kappa_j\chi_{E_j+z}(x)=\sum_{j=1}^m\kappa_j\chi_{\tau_z(E_j)}(x),$$

from which we get

$$\int_{\tau_z(A)}\tau_z(\phi)\,dm_n=\sum_{j=1}^m\kappa_j m_n(\tau_z(E_j))=\sum_{j=1}^m\kappa_j m_n(E_j)=\int_A\phi\,dm_n.$$

Now we pass to the case of $f:\mathbb{R}^n\to[0,+\infty]$ by considering an increasing sequence of simple functions, and then we pass to the case of $f:\mathbb{R}^n\to\overline{\mathbb{R}}$ by considering the non-positive part and the non-negative part of f, and then to the case of $f:\mathbb{R}^n\to\overline{\mathbb{C}}$ by considering the real part and the imaginary part of f. All this is already standard and it is left as an exercise. ☐

The equality $\int_{\tau_z(A)}\tau_z(f)\,dm_n=\int_A f\,dm_n$ can be written

$$\int_{A+z}f(x-z)\,dm_n(x)=\int_A f(y)\,dm_n(y).$$

We can view this as *change of variable formula*. We write $y=(\tau_z)^{-1}(x)=x-z$ or, equivalently, $x=\tau_z(y)=y+z$, and we employ the *informal rule* for the change of differentials:

$$dm_n(x)=dm_n(y).$$

Proposition 3.37. *Let $T:\mathbb{R}^n\to\mathbb{R}^n$ be a linear transformation with* $\det(T)\neq 0$ *and $A\in\mathcal{L}_n$ and $f:A\to\overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ be $\mathcal{L}_n\rceil A$-measurable.*
(i) If $\int_A f\,dm_n$ is defined, then $\int_{T(A)} T(f)\,dm_n$ is defined and

$$\int_{T(A)} T(f)\,dm_n = |\det(T)|\int_A f\,dm_n.$$

(ii) If f is Lebesgue integrable over A, then $T(f)$ is Lebesgue integrable over $T(A)$ and the equality in (i) is again true.

Proof. Let $\phi:A\to[0,+\infty)$ be a Lebesgue measurable simple function and let $\phi=\sum_{j=1}^m \kappa_j\chi_{E_j}$ be its standard representation. Then

$$\int_A \phi\,dm_n = \sum_{j=1}^m \kappa_j m_n(E_j).$$

It is clear that

$$T(\phi)(x)=\phi(T^{-1}(x))=\sum_{j=1}^m \kappa_j\chi_{E_j}(T^{-1}(x))=\sum_{j=1}^m \kappa_j\chi_{T(E_j)}(x),$$

from which we get

$$\int_{T(A)} T(\phi)\,dm_n=\sum_{j=1}^m \kappa_j m_n(T(E_j))=|\det(T)|\sum_{j=1}^m \kappa_j m_n(E_j)=|\det(T)|\int_A\phi\,dm_n.$$

As in the proof of Proposition 3.36, we pass to the case of $f:\mathbb{R}^n\to[0,+\infty]$ by considering an increasing sequence of simple functions, and then we pass to the case of $f:\mathbb{R}^n\to\overline{\mathbb{R}}$ by considering the non-positive part and the non-negative part of f, and then we pass to the case of $f:\mathbb{R}^n\to\overline{\mathbb{C}}$ by considering the real part and the imaginary part of f. $\square$

The equality $\int_{T(A)} T(f)\,dm_n=|\det(T)|\int_A f\,dm_n$ can be written

$$\int_{T(A)} f(T^{-1}(x))\,dm_n(x)=|\det(T)|\int_A f(y)\,dm_n(y).$$

Again, this expresses a *change of variable formula*. We write $y=T^{-1}(x)$ or, equivalently, $x=T(y)$, and we employ the *informal rule* for the change of differentials:

$$dm_n(x)=|\det(T)|\,dm_n(y).$$

As special cases of linear transformations we consider the dilations and the reflection, and we get the equalities

$$\frac{1}{\lambda^n}\int_{\lambda A} f\big(\tfrac{x}{\lambda}\big)\,dm_n(x)=\int_A f(y)\,dm_n(y),\quad \int_{-A} f(-x)\,dm_n(x)=\int_A f(y)\,dm_n(y)$$

for all $\lambda>0$.

Exercises.

3.2.18. Let $\mathbb{Q}\cap[0,1]=\{r_1,r_2,\dots\}$ and $\sum_{n=1}^{+\infty}|a_n|<+\infty$. Prove that the series $\sum_{n=1}^{+\infty}\frac{a_n}{|x-r_n|^{1/2}}$ converges absolutely for m_1-a.e. $x\in[0,1]$.

3.2.19. Let $\mathbb{Q}=\{r_1,r_2,\dots\}$. Prove that the series $\sum_{n=1}^{+\infty} e^{-n^2|x-r_n|}$ converges for m_1-a.e. $x\in\mathbb{R}$.

3.2.20. *The Fourier transforms of Lebesgue integrable functions.*
Let $f:\mathbb{R}^n\to\overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ be Lebesgue integrable over $\mathbb{R}^n$. We define the function $\widehat{f}:\mathbb{R}^n\to\mathbb{C}$ by

$$\widehat{f}(\xi)=\int_{\mathbb{R}^n} e^{-2\pi i x\cdot\xi} f(x)\,dm_n(x),$$

where $x\cdot\xi=x_1\xi_1+\cdots x_n\xi_n$ is the Euclidean inner product. The function $\widehat{f}$ is called the **Fourier transform** of f.

(i) Prove that $\widehat{f_1 + f_2} = \widehat{f_1} + \widehat{f_2}$ and $\widehat{\lambda f} = \lambda \widehat{f}$.
(ii) If $g(x) = f(x - a)$ for m_n-a.e. $x \in \mathbb{R}^n$, prove that $\widehat{g}(\xi) = e^{-2\pi i a\cdot\xi}\widehat{f}(\xi)$ for all $\xi \in \mathbb{R}^n$.
(iii) If $g(x) = e^{-2\pi i a\cdot x} f(x)$ for m_n-a.e. $x \in \mathbb{R}^n$, prove that $\widehat{g}(\xi) = \widehat{f}(\xi + a)$ for all $\xi \in \mathbb{R}^n$.
(iv) If $g(x) = \overline{f(x)}$ for m_n-a.e. $x \in \mathbb{R}^n$, prove that $\widehat{g}(\xi) = \overline{\widehat{f}(-\xi)}$ for all $\xi \in \mathbb{R}^n$.
(v) If $T : \mathbb{R}^n \to \mathbb{R}^n$ is a linear transformation with $\det(T) \neq 0$ and $g(x) = f(Tx)$ for m_n-a.e. $x \in \mathbb{R}^n$, prove that $\widehat{g}(\xi) = \frac{1}{|\det(T)|} \widehat{f}\big((T^*)^{-1}(\xi)\big)$ for all $\xi \in \mathbb{R}^n$, where T^* is the adjoint of T.
(vi) Prove that $\widehat{f}$ is continuous on $\mathbb{R}^n$.
(vii) Prove that $|\widehat{f}(\xi)| \leq \int_{\mathbb{R}^n} |f(x)|\, dm_n(x)$ for every $\xi \in \mathbb{R}^n$.
(viii) (The **Riemann-Lebesgue Lemma**) Prove that $\widehat{f}(\xi) \to 0$ as $|\xi| \to +\infty$.
Hint. Prove that $\int_{\mathbb{R}^n} e^{-2\pi i x\cdot\xi} f(x)\, dm_n(x) = \frac{1}{2}\int_{\mathbb{R}^n} e^{-2\pi i x\cdot\xi}\big(f\big(x - \frac{\xi}{2|\xi|^2}\big) - f(x)\big)\, dm_n(x)$ and then use the result of exercise 3.2.12.

3.2.21. Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be an isometric linear transformation (see exercise 1.4.5). Prove that $\int_{\mathbb{R}^n} f \circ T^{-1}\, dm_n = \int_{\mathbb{R}^n} f\, dm_n$ for every Lebesgue measurable $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$, provided that at least one of the two integrals exists.

3.2.22. Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$. We say that f is 1-periodic if $f = f \circ \tau_k^{-1}$ for every $k \in \mathbb{Z}^n$. In other words, f is 1-periodic if $f(x_1, \dots, x_n) = f(x_1 - k_1, \dots, x_n - k_n)$ for every $(x_1, \dots, x_n) \in \mathbb{R}^n$ and every $(k_1, \dots, k_n) \in \mathbb{Z}^n$.
(i) Let f be 1-periodic, $A \in \mathcal{L}_n$ and $k \in \mathbb{Z}$. Prove that $\int_A f\, dm_n$ exists if and only if $\int_{\tau_k^{-1}(A)} f\, dm_n$ exists and, in this case, we have $\int_A f\, dm_n = \int_{\tau_k^{-1}(A)} f\, dm_n$.
(ii) Let f be 1-periodic, and $y \in \mathbb{R}^n$. Prove that $\int_{[0,1)^n} f\, dm_n$ exists if and only if $\int_{\tau_y^{-1}([0,1)^n)} f\, dm_n$ exists and, in this case, we have $\int_{[0,1)^n} f\, dm_n = \int_{\tau_y^{-1}([0,1)^n)} f\, dm_n$.

3.3 Lebesgue-Stieltjes integrals.

Let $-\infty \leq a_0 < b_0 \leq +\infty$. We know that every continuous $f : (a_0, b_0) \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ is Borel measurable. On the other hand, also every monotone $f : (a_0, b_0) \to \overline{\mathbb{R}}$ is Borel measurable. This is seen by observing that $f^{-1}(I)$ is an interval, and hence a Borel set, for every interval I in $\overline{\mathbb{R}}$. Now, if $F : (a_0, b_0) \to \mathbb{R}$ is an increasing function and μ_F is the induced Borel measure, then f, in both cases, satisfies the necessary measurability condition, and the integral $\int_{(a_0,b_0)} f\, d\mu_F$ exists provided, as usual, that either $\int_{(a_0,b_0)} f^+\, d\mu_F < +\infty$ or $\int_{(a_0,b_0)} f^-\, d\mu_F < +\infty$ in the case of $f : (a_0, b_0) \to \overline{\mathbb{R}}$, and that $\int_{(a_0,b_0)} |f|\, d\mu_F < +\infty$ in the case of $f : (a_0, b_0) \to \overline{\mathbb{C}}$.

In particular, *if f, besides being continuous or monotone, is also bounded on an interval $S \subseteq (a_0, b_0)$ with $\mu_F(S) < +\infty$, then it is integrable over S with respect to μ_F.*

We shall prove three classical results about Lebesgue-Stieltjes integrals.

Observe that the cases $[a, b]$, $[a, b)$, $(a, b]$ and (a, b) for the interval S may give different corresponding integrals $\int_S f\, d\mu_F$. This is because the one-point integral

$$\int_{\{x\}} f\, d\mu_F = f(x)\mu_F(\{x\}) = f(x)(F(x+) - F(x-))$$

may not be zero.

Proposition 3.38. (Integration by parts) *Let $F, G : (a_0, b_0) \to \mathbb{R}$ be two increasing functions and μ_F, μ_G be the induced Lebesgue-Stieltjes measures. Then*

$$\int_{(a,b]} G(x+)\, d\mu_F(x) + \int_{(a,b]} F(x-)\, d\mu_G(x) = G(b+)F(b+) - G(a+)F(a+)$$

for all $a, b \in (a_0, b_0)$ with $a \leq b$. In this equality we may interchange F with G.
Similar equalities hold for the other types of intervals, provided we use the appropriate limits of F, G at a, b at the right side of the above equality.

Proof. We consider a sequence of partitions $\Delta_k = \{c_0^{(k)}, \dots, c_{l_k}^{(k)}\}$ of $[a,b]$ so that

$$\lim_{k\to+\infty} \max\{c_j^{(k)} - c_{j-1}^{(k)} \,|\, 1 \leq j \leq l_k\} = 0.$$

We also introduce the simple functions

$$g_k = \sum_{j=1}^{l_k} G(c_j^{(k)}+)\chi_{(c_{j-1}^{(k)}, c_j^{(k)}]}, \quad f_k = \sum_{j=1}^{l_k} F(c_{j-1}^{(k)}+)\chi_{(c_{j-1}^{(k)}, c_j^{(k)}]}.$$

It is clear that

$$G(a+) \leq g_k \leq G(b+), \quad F(a+) \leq f_k \leq F(b-)$$

for all k.
For any $x \in (a,b]$ we consider the interval $(c_{j-1}^{(k)}, c_j^{(k)}]$ containing x (where j depends upon both k and x). Then $g_k(x) = G(c_j^{(k)}+)$ and $f_k(x) = F(c_{j-1}^{(k)}+)$. Since $\lim_{k\to+\infty}(c_j^{(k)} - c_{j-1}^{(k)}) = 0$, we have that $\lim_{k\to+\infty} c_{j-1}^{(k)} = \lim_{k\to+\infty} c_j^{(k)} = x$, and so

$$\lim_{k\to+\infty} g_k(x) = G(x+), \quad \lim_{k\to+\infty} f_k(x) = F(x-).$$

Now, we have that

$$\sum_{j=1}^{l_k} G(c_j^{(k)}+)(F(c_j^{(k)}+) - F(c_{j-1}^{(k)}+)) = \int_{(a,b]} g_k(x)\, d\mu_F(x),$$

$$\sum_{j=1}^{l_k} F(c_{j-1}^{(k)}+)(G(c_j^{(k)}+) - G(c_{j-1}^{(k)}+)) = \int_{(a,b]} f_k(x)\, d\mu_G(x).$$

We apply the Dominated Convergence Theorem and we get

$$\lim_{k\to+\infty} \sum_{j=1}^{l_k} G(c_j^{(k)}+)(F(c_j^{(k)}+) - F(c_{j-1}^{(k)}+)) = \int_{(a,b]} G(x+)\, d\mu_F(x),$$

$$\lim_{k\to+\infty} \sum_{j=1}^{l_k} F(c_{j-1}^{(k)}+)(G(c_j^{(k)}+) - G(c_{j-1}^{(k)}+)) = \int_{(a,b]} F(x-)\, d\mu_G(x).$$

Adding, we find

$$G(b+)F(b+) - G(a+)F(a+) = \int_{(a,b]} G(x+)\, d\mu_F(x) + \int_{(a,b]} F(x-)\, d\mu_G(x).$$

We work in the same way for all other types of intervals. $\square$

The next two results concern the reduction of Lebesgue-Stieltjes integrals to Lebesgue integrals. This makes the calculation of the former more accessible in many situations.

Proposition 3.39. *Let $F : (a_0, b_0) \to \mathbb{R}$ be increasing and have a continuous derivative on (a_0, b_0). Then*

$$\mu_F(E) = \int_E F'(x)\, dm_1(x)$$

for every Borel set $E \subseteq (a_0, b_0)$. Also

$$\int_{(a_0,b_0)} f(x)\, d\mu_F(x) = \int_{(a_0,b_0)} f(x)F'(x)\, dm_1(x)$$

for every Borel measurable $f : (a_0, b_0) \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ for which either of the two integrals exists.

Proof. The assumptions on F imply that it is continuous and that $F' \geq 0$ on (a_0, b_0). The Fundamental Theorem of Calculus for Riemann integrals implies that for every $[a,b] \subseteq (a_0, b_0)$ we have

$$\int_{[a,b]} F'(x)\, dm_1(x) = F(b) - F(a) = \mu_F([a,b]).$$

By the continuity of F, this equality holds for all intervals $(a,b], [a,b), (a,b)$ in (a_0, b_0).
Now we define the Borel measure μ on (a_0, b_0) by

$$\mu(E) = \int_E F'(x)\, dm_1(x)$$

for every Borel set $E \subseteq (a_0, b_0)$. It is easy to see that μ is a measure indeed. Clearly, $\mu(\emptyset) = 0$, and $\mu(E) \geq 0$ for all Borel $E \subseteq (a_0, b_0)$. Also, the σ-additivity of μ is an immediate consequence of Theorem 3.1.
Now we have $\mu(S) = \mu_F(S)$ for every bounded interval $S \subseteq (a_0, b_0)$. Then Proposition 1.45 implies that $\mu = \mu_F$, and so

$$\mu_F(E) = \int_E F'(x)\, dm_1(x)$$

for every Borel set $E \subseteq (a_0, b_0)$.
Considering arbitrary linear combinations of characteristic functions, we get

$$\int_{(a_0,b_0)} \phi(x)\, d\mu_F(x) = \int_{(a_0,b_0)} \phi(x) F'(x)\, dm_1(x)$$

for all Borel measurable simple functions $\phi : (a_0, b_0) \to [0, +\infty)$.
The rest is a standard exercise. □

Proposition 3.40. *Let $F : (a_0, b_0) \to \mathbb{R}$ be increasing and $G : (a, b) \to \mathbb{R}$ be bounded and have a continuous derivative which is Lebesgue integrable over (a, b), where $a_0 < a < b < b_0$. Then,*

$$\begin{aligned}\int_{(a,b)} G(x)\, d\mu_F(x) &= G(b-)F(b-) - G(a+)F(a+) - \int_{(a,b)} F(x-)G'(x)\, dm_1(x)\\ &= G(b-)F(b-) - G(a+)F(a+) - \int_{(a,b)} F(x+)G'(x)\, dm_1(x).\end{aligned}$$

Proof. (a) Let us assume that G is also *increasing* on (a, b). Then its extension as $G(a+)$ on $(a_0, a]$ and as $G(b-)$ on $[b, b_0)$ is increasing on (a_0, b_0). We apply Proposition 3.38 and we get

$$\int_{(a,b)} G(x)\, d\mu_F(x) = G(b-)F(b-) - G(a+)F(a+) - \int_{(a,b)} F(x-)\, d\mu_G(x).$$

Now, the integral $\int_{(a,b)} F(x-)G'(x)\, dm_1(x)$ exists, since $F(x-)$ is bounded on (a, b) and G' is Lebesgue integrable over (a, b), and Proposition 3.39 implies

$$\int_{(a,b)} G(x)\, d\mu_F(x) = G(b-)F(b-) - G(a+)F(a+) - \int_{(a,b)} F(x-)G'(x)\, dm_1(x).$$

(b) In the general case, we take an arbitrary $x_0 \in (a, b)$, and we have that

$$G(x) = G(x_0) + \int_{(x_0,x)} G'(t)\, dm_1(t)$$

for every $x \in (a, b)$. Now, $(G')^+$ and $(G')^-$ are non-negative, continuous and Lebesgue integrable over (a, b), and we have $G = G_1 - G_2$ on (a, b), where

$$G_1(x) = G(x_0) + \int_{(x_0,x)} (G')^+(t)\, dm_1(t), \quad G_2(x) = \int_{(x_0,x)} (G')^-(t)\, dm_1(t)$$

for all $x \in (a, b)$. By the continuity of $(G')^+$ and $(G')^-$ and the Fundamental Theorem of Calculus, we have that $G_1' = (G')^+ \geq 0$ and $G_2' = (G')^- \geq 0$ on (a, b). Hence, G_1 and G_2 are both increasing with a continuous derivative which is Lebesgue integrable over (a, b), and so from (a) we have

$$\int_{(a,b)} G_i(x)\, d\mu_F(x) = G_i(b-)F(b-) - G_i(a+)F(a+) - \int_{(a,b)} F(x-)G_i'(x)\, dm_1(x)$$

for $i = 1, 2$. We subtract these two equalities and we get the desired equality. □

From the proof of Proposition 3.40 it is worth keeping in mind the fact that an arbitrary G with a continuous derivative integrable over an interval (a, b) can be decomposed as a difference, $G = G_1 - G_2$, of two *increasing* functions with continuous derivatives integrable over (a, b).

Exercises.

3.3.1. Consider the function $g : \mathbb{R} \to \mathbb{R}$ defined by $g(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}t^2}\, dt$.
(i) Prove that g is continuous, strictly increasing, with $g(-\infty) = 0$ and $g(+\infty) = 1$ (see exercise 3.2.13), and with continuous derivative $g'(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$.
(ii) The Lebesgue-Stieltjes measure μ_g induced by g is called the **distribution** or **measure of Gauss**. Prove that $\mu_g(\mathbb{R}) = 1$, that $\mu_g(E) = \frac{1}{\sqrt{2\pi}} \int_E e^{-\frac{1}{2}x^2}\, dx$ for every Borel set in $\mathbb{R}$, and that $\int_{\mathbb{R}} f(x)\, d\mu_g(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(x) e^{-\frac{1}{2}x^2}\, dx$ for every Borel measurable $f : \mathbb{R} \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ for which either of the two integrals exists.

3.3.2. (i) Consider the Cantor set C and the $I_0 = [0,1], I_1, I_2, \ldots$ which were used for its construction. Prove that the 2^{k-1} subintervals of $I_{k-1} \setminus I_k$ are the $\big(\frac{a_1}{3} + \cdots + \frac{a_{k-1}}{3^{k-1}} + \frac{1}{3^k}, \frac{a_1}{3} + \cdots + \frac{a_{k-1}}{3^{k-1}} + \frac{2}{3^k}\big)$, where each of $a_1, \ldots, a_{k-1}$ takes the values 0 and 2.
(ii) Prove that the Cantor function f is constant $f(x) = \frac{a_1}{2^2} + \cdots + \frac{a_{k-1}}{2^k} + \frac{1}{2^k}$ on the above subinterval $\big(\frac{a_1}{3} + \cdots + \frac{a_{k-1}}{3^{k-1}} + \frac{1}{3^k}, \frac{a_1}{3} + \cdots + \frac{a_{k-1}}{3^{k-1}} + \frac{2}{3^k}\big)$.
(iii) If $G : (0,1) \to \mathbb{R}$ has continuous derivative which is Lebesgue integrable over $(0,1)$, prove:

$$\begin{aligned} \sum_{k=1}^{+\infty} \sum_{a_1,\ldots,a_{k-1} \in \{0,2\}} \Big(\tfrac{a_1}{2^2} + \cdots + \tfrac{a_{k-1}}{2^k} + \tfrac{1}{2^k}\Big) \Big(G\Big(\tfrac{a_1}{3} + \cdots + \tfrac{a_{k-1}}{3^{k-1}} + \tfrac{2}{3^k}\Big) \\ - G\Big(\tfrac{a_1}{3} + \cdots + \tfrac{a_{k-1}}{3^{k-1}} + \tfrac{1}{3^k}\Big)\Big) \\ = G(1-) - \int_{(0,1)} G(x)\, d\mu_f(x). \end{aligned}$$

(iv) In particular, prove that $\int_{(0,1)} x\, d\mu_f(x) = \frac{1}{2}$.
(v) Prove that $\int_{(0,1)} e^{-2\pi i \xi x}\, d\mu_f(x) = e^{-\pi i \xi} \lim_{k \to +\infty} \prod_{j=1}^{k} \cos\big(\frac{2\pi\xi}{3^j}\big)$ for every $\xi \in \mathbb{R}$.

3.3.3. Let $F, G : \mathbb{R} \to \mathbb{R}$ be increasing and assume that FG is also increasing.
(i) Prove that $\mu_{FG}(E) = \int_E G(x+)\, d\mu_F(x) + \int_E F(x-)\, d\mu_G(x)$ for every Borel set $E \subseteq \mathbb{R}$.
(ii) Prove that $\int_{\mathbb{R}} f(x)\, d\mu_{FG}(x) = \int_{\mathbb{R}} f(x) G(x+)\, d\mu_F(x) + \int_{\mathbb{R}} f(x) F(x-)\, d\mu_G(x)$ for every Borel measurable $f : \mathbb{R} \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ for which at least two of the three integrals exist.

3.3.4. If $F : \mathbb{R} \to \mathbb{R}$ is increasing and continuous and $f : \mathbb{R} \to [0, +\infty]$ is Borel measurable, prove that $\int_{\mathbb{R}} f(F(x))\, d\mu_F(x) = \int_{F(-\infty)}^{F(+\infty)} f(t)\, dt$.
Show, by example, that this may not be true if F is not continuous.

3.3.5. *Riemann's criterion for convergence of a series.*
Assume $F : \mathbb{R} \to [0, +\infty)$ is increasing and $g : (0, +\infty) \to [0, +\infty)$ is decreasing. Let $a_n \geq 0$ for all n and $\sharp\{n \,|\, a_n \geq g(x)\} \leq F(x)$ for all $x \in (0, +\infty)$ and $\int_{(0,+\infty)} g(x)\, d\mu_F(x) < +\infty$. Prove that $\sum_{n=1}^{+\infty} a_n < +\infty$.

REDUCTION TO INTEGRALS OVER $\mathbb{R}$.

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

Definition. *Let $f : X \to [0, +\infty]$ be measurable. Then the function $\lambda_f : [0, +\infty) \to [0, +\infty]$, defined by*

$$\lambda_f(t) = \mu(\{x \in X \,|\, t < f(x)\}),$$

*is called the **distribution function** of f.*

Some properties of λ_f are easy to prove. It is obvious that λ_f is *non-negative* and *decreasing* on $[0, +\infty)$. Also, continuity of μ from below implies that λ_f is *continuous from the right* on $[0, +\infty)$. Hence, there exists some $t_0 \in [0, +\infty]$ with the property that λ_f is $+\infty$ on the interval $[0, t_0)$ (which may be empty) and λ_f is finite on the interval $(t_0, +\infty)$ (which may be empty).

Proposition 3.41. (Chebychev) *If $f : X \to [0, +\infty]$ is measurable, then*

$$\lambda_f(t) \leq \tfrac{1}{t} \int_X f \, d\mu$$

for every $t \in (0, +\infty)$.

Proof. We consider the set $A = \{x \in X \mid t < f(x)\} \in \mathcal{S}$. Then

$$t\lambda_f(t) = t\mu(A) = t \int_X \chi_A \, d\mu \leq \int_X f \, d\mu,$$

since $t\chi_A \leq f$ on X. □

Proposition 3.42. *Let $f : X \to [0, +\infty]$ be measurable and $G : \mathbb{R} \to \mathbb{R}$ be increasing with $G(0-) = 0$. Then*

$$\int_X G(f(x)-) \, d\mu(x) = \int_{[0,+\infty)} \lambda_f(t) \, d\mu_G(t).$$

Moreover, if G has continuous derivative on $(0, +\infty)$, then

$$\int_X G(f(x)) \, d\mu(x) = \int_{(0,+\infty)} \lambda_f(t) G'(t) \, dm_1(t) + \lambda_f(0) G(0+).$$

In particular,

$$\int_X f(x) \, d\mu(x) = \int_{(0,+\infty)} \lambda_f(t) \, dm_1(t).$$

Proof. (a) Let ϕ be a non-negative measurable simple function on X with standard representation $\phi = \sum_{j=1}^m \kappa_j \chi_{E_j}$, where we omit the value 0. We rearrange so that $0 < \kappa_1 < \cdots < \kappa_m$, and then

$$\lambda_\phi(t) = \begin{cases} \mu(E_1) + \mu(E_2) + \cdots + \mu(E_m), & \text{if } 0 \leq t < \kappa_1 \\ \mu(E_2) + \cdots + \mu(E_m), & \text{if } \kappa_1 \leq t < \kappa_2 \\ \cdots & \\ \mu(E_m), & \text{if } \kappa_{m-1} \leq t < \kappa_m \\ 0, & \text{if } \kappa_m \leq t \end{cases}$$

Then

$$\begin{aligned} \int_{[0,+\infty)} \lambda_\phi(t) \, d\mu_G(t) &= \big(\mu(E_1) + \mu(E_2) + \cdots + \mu(E_m)\big)\big(G(\kappa_1-) - G(0-)\big) \\ &\quad + \big(\mu(E_2) + \cdots + \mu(E_m)\big)\big(G(\kappa_2-) - G(\kappa_1-)\big) \\ &\quad \cdots \\ &\quad + \mu(E_m)\big(G(\kappa_m-) - G(\kappa_{m-1}-)\big) \\ &= G(\kappa_1-)\mu(E_1) + G(\kappa_2-)\mu(E_2) + \cdots + G(\kappa_m-)\mu(E_m) \\ &= \int_X G(\phi(x)-) \, d\mu(x), \end{aligned}$$

since $G(\phi(x)-)$ is a simple function taking value $G(\kappa_j-)$ on each E_j and value $G(0-) = 0$ on $(E_1 \cup \cdots \cup E_m)^c$.
(b) Now we consider any measurable $f : X \to [0, +\infty]$ and any increasing sequence (ϕ_n) of non-negative measurable simple functions on X so that $\lim_{n\to+\infty} \phi_n = f$ on X. Then $(G(\phi_n(x)-))$ is an increasing sequence of functions so that

$$\lim_{n\to+\infty} G(\phi_n(x)-) = G(f(x)-)$$

for every $x \in X$, and so

$$\lim_{n\to+\infty} \int_X G(\phi_n(x)-) \, d\mu(x) = \int_X G(f(x)-) \, d\mu(x)$$

by the Monotone Convergence Theorem.
Also (λ_{ϕ_n}) is an increasing sequence of functions so that $\lim_{n\to+\infty}\lambda_{\phi_n}=\lambda_f$ on $[0,+\infty)$. Hence,

$$\lim_{n\to+\infty}\int_{[0,+\infty)}\lambda_{\phi_n}(t)\,d\mu_G(t)=\int_{[0,+\infty)}\lambda_f(t)\,d\mu_G(t)$$

by the Monotone Convergence Theorem.
Applying the result of (a) to each ϕ_n, we get $\int_X G(f(x)-)\,d\mu(x)=\int_{[0,+\infty)}\lambda_f(t)\,d\mu_G(t)$.
Now, Proposition 3.39 implies the second equality of the statement, and the special case $G(t)=t$ implies the last equality. $\square$

Exercises.

3.3.6. Let $(X,\mathcal{S},\mu)$ be a measure space and $f: X\to[0,+\infty]$ be μ-integrable. Prove that $\lim_{t\to+\infty}t\lambda_f(t)=0$.

3.3.7. Let $(X,\mathcal{S},\mu)$ be a measure space and $f: X\to[0,+\infty]$ be measurable. Prove that

$$\frac{1}{2}\sum_{n\in\mathbb{Z}}2^n\lambda_f(2^n)\leq\int_X f(x)\,d\mu(x)\leq\sum_{n\in\mathbb{Z}}2^n\lambda_f(2^n).$$

Conclude that f is integrable if and only if $\sum_{n\in\mathbb{Z}}2^n\lambda_f(2^n)$ is finite.

3.3.8. Let $(X,\mathcal{S},\mu)$ be a measure space, $f: X\to[0,+\infty]$ be measurable and $0<p<+\infty$. Prove that $\int_X f(x)^p\,d\mu(x)=p\int_0^{+\infty}t^{p-1}\lambda_f(t)\,dt$.

3.3.9. Let $(X,\mathcal{S},\mu)$ be a measure space and $f,g: X\to[0,+\infty]$ be measurable. The f,g are called **equidistributed** if $\lambda_f(t)=\lambda_g(t)$ for every $t\in[0,+\infty)$.
If f,g are equidistributed, prove that $\int_X f(x)^p\,d\mu(x)=\int_X g(x)^p\,d\mu(x)$ for every $p>0$.
Hint. See exercise 3.3.8.

3.3.10. Let $(X,\mathcal{S},\mu)$ be a measure space and $\phi,\psi: X\to[0,+\infty)$ be two measurable simple functions, and let $\phi=\sum_{j=1}^m\kappa_j\chi_{E_j}$ and $\psi=\sum_{i=1}^n\lambda_i\chi_{F_i}$ be their standard representations so that $0<\kappa_1<\cdots<\kappa_m$ and $0<\lambda_1<\cdots<\lambda_n$ (where we omit the possible value 0).
If ϕ and ψ are integrable, prove that they are equidistributed (see exercise 3.3.9) if and only if $m=n$, $\kappa_1=\lambda_1,\ldots,\kappa_m=\lambda_m$ and $\mu(E_1)=\mu(F_1),\ldots,\mu(E_m)=\mu(F_m)$.

3.4 Integrals on Borel measure spaces.

Let X be a Hausdorff topological space and μ be a Borel measure on X. It is easy to see that every continuous $f: X\to\mathbb{R}$ or $\mathbb{C}$, which is 0 outside some compact set of finite measure, is integrable with respect to μ. Indeed, since f is continuous, it is Borel measurable. Also, let K be a compact set with $\mu(K)<+\infty$ outside of which f is 0. Then $|f|\leq M\chi_K$, where $M=\max\{|f(x)|\,|\,x\in K\}<+\infty$. Therefore,

$$\int_X|f|\,d\mu\leq M\int_X\chi_K\,d\mu=M\mu(K)<+\infty,$$

and so f is integrable.

APPROXIMATION BY CONTINUOUS FUNCTIONS.

Theorem 3.3. *Let the topological space X be locally compact and Hausdorff and μ be a regular Borel measure on X and let the Borel measurable $f: X\to\overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ be integrable. Then for every $\epsilon>0$ there is a continuous $g: X\to\mathbb{R}$ or $\mathbb{C}$ which is 0 outside some compact set of finite measure so that $\int_X|g-f|\,d\mu<\epsilon$.*

Proof. By Proposition 3.29, there is an integrable Borel measurable simple function $\phi : X \to \mathbb{R}$ or $\mathbb{C}$ so that

$$\int_X |\phi - f| \, d\mu < \tfrac{\epsilon}{2}.$$

Let $\phi = \sum_{k=1}^m \kappa_k \chi_{E_k}$, where $E_1, \dots, E_m$ are pairwise disjoint Borel sets and all κ_k are $\neq 0$. Since ϕ is integrable, we have that $\mu(E_k) < +\infty$ for all k.
From the regularity of μ, we have that there are compact K_k and open U_k so that $K_k \subseteq E_k \subseteq U_k$ and $\mu(U_k \setminus K_k) < \eta$ for all k, where $\eta > 0$ will be chosen appropriately in a moment.
Urysohn's Lemma implies that there are continuous functions $g_k : X \to [0, 1]$ so that $g_k = 1$ on K_k and supp(g_k) is a compact subset of U_k.
Now we consider $g = \sum_{k=1}^m \kappa_k g_k$.
Then $g : X \to \mathbb{R}$ or $\mathbb{C}$ is continuous and equal to 0 outside some compact set of finite measure. Indeed, $g = 0$ outside the compact $K = \bigcup_{k=1}^m \operatorname{supp}(g_k)$ with

$$\mu(K) \leq \sum_{k=1}^m \mu(U_k) \leq \sum_{k=1}^m (\mu(E_k) + \eta) < +\infty.$$

Moreover, we have $g_k = 1 = \chi_{E_k}$ on K_k, and $g_k = 0 = \chi_{E_k}$ on U_k^c, and $|g_k - \chi_{E_k}| \leq 1$ on $U_k \setminus K_k$. Hence,

$$\int_X |g_k - \chi_{E_k}| \, d\mu \leq \mu(U_k \setminus K_k) < \eta.$$

Therefore,

$$\int_X |g - \phi| \, d\mu \leq \sum_{k=1}^m |\kappa_k| \int_X |g_k - \chi_{E_k}| \, d\mu < \eta \sum_{k=1}^m |\kappa_k|.$$

Now we choose $\eta = \frac{\epsilon}{2 \sum_{k=1}^m |\kappa_k|}$, and we get

$$\int_X |g - \phi| \, d\mu < \tfrac{\epsilon}{2}.$$

Hence,

$$\int_X |g - f| \, d\mu \leq \int_X |g - \phi| \, d\mu + \int_X |\phi - f| \, d\mu < \epsilon$$

and the proof is complete. $\square$

We recall that Theorem 1.2 gives conditions on a Hausdorff topological space X and a Borel measure μ on X so that μ is regular.

Chapter 4

Product measures.

4.1 Product σ-algebra.

If I is a general set of indices, then the elements of the cartesian product $\prod_{i\in I} X_i$ are the functions $x : I \to \bigcup_{i\in I} X_i$ with the property: $x(i) \in X_i$ for every $i \in I$. It is customary to use the notation x_i, instead of $x(i)$, for the value of x at $i \in I$ and, accordingly, to use the notation $(x_i)_{i\in I}$ for the element $x \in \prod_{i\in I} X_i$.

If I is a finite set, say $I = \{1, \dots, n\}$, we use the traditional notation $(x_1, \dots, x_n)$ for the element $(x_i)_{i\in\{1,\dots,n\}}$ and we use the notation $\prod_{i=1}^n X_i$ or $X_1 \times \cdots \times X_n$ for $\prod_{i\in\{1,\dots,n\}} X_i$. And if I is countable, say $I = \mathbb{N} = \{1, 2, \dots\}$, we write $(x_1, x_2, \dots)$ for the element $(x_i)_{i\in\mathbb{N}}$ and we write $\prod_{i=1}^{+\infty} X_i$ or $X_1 \times X_2 \times \cdots$ for $\prod_{i\in\mathbb{N}} X_i$.

Definition. *Let $(X_i, \mathcal{S}_i)$ be a measurable space for every $i \in I$. We consider the σ-algebra of subsets of the cartesian product $\prod_{i\in I} X_i$ which is generated by the collection*

$$\mathcal{C} = \left\{ \prod_{i\in I} A_i \,\middle|\, A_i \neq X_i \textit{ for at most finitely many } i \in I, \textit{ and } A_i \in \mathcal{S}_i \textit{ if } A_i \neq X_i \right\}.$$

This σ-algebra $\mathcal{S}(\mathcal{C})$ is called the ***product σ-algebra*** *of $\mathcal{S}_i$ and it is denoted by*

$$\bigotimes_{i\in I} \mathcal{S}_i.$$

In particular, $\bigotimes_{i=1}^n \mathcal{S}_i$ is generated by the collection of all sets of the form $A_1 \times \cdots \times A_n$, where $A_i \in \mathcal{S}_i$ for all $i = 1, \dots, n$. Similarly, $\bigotimes_{i=1}^{+\infty} \mathcal{S}_i$ is generated by the collection of all sets of the form $A_1 \times \cdots \times A_n \times X_{n+1} \times X_{n+2} \times \cdots$, where $n \in \mathbb{N}$ and $A_i \in \mathcal{S}_i$ for all $i = 1, \dots, n$.

Proposition 4.1. *Let $(X_i, \mathcal{S}_i)$ be a measurable space for every $i \in I$ and $\mathcal{C}_i$ be a collection of subsets of X_i so that $\mathcal{S}_i = \mathcal{S}(\mathcal{C}_i)$ for every $i \in I$. Then $\bigotimes_{i\in I} \mathcal{S}_i = \mathcal{S}(\tilde{\mathcal{C}})$, where*

$$\tilde{\mathcal{C}} = \left\{ \prod_{i\in I} A_i \,\middle|\, A_i \neq X_i \textit{ for at most finitely many } i \in I, \textit{ and } A_i \in \mathcal{C}_i \textit{ if } A_i \neq X_i \right\}.$$

Proof. If $\mathcal{C}$ is the collection in the definition of $\bigotimes_{i\in I} \mathcal{S}_i$, then $\tilde{\mathcal{C}} \subseteq \mathcal{C}$, and so $\mathcal{S}(\tilde{\mathcal{C}}) \subseteq \mathcal{S}(\mathcal{C})$.
Now we fix some $j \in I$, and for every $A_j \subseteq X_j$ we define $A_j^* = \prod_{i\in I} Y_i$, where $Y_i = X_i$ for $i \neq j$ and $Y_j = A_j$. We then consider the collection

$$\mathcal{S}_j^* = \{A_j \,|\, A_j \subseteq X_j \text{ and } A_j^* \in \mathcal{S}(\tilde{\mathcal{C}})\}.$$

We can easily show that $\mathcal{S}_j^*$ is a σ-algebra of subsets of X_j and that $\mathcal{C}_j \subseteq \mathcal{S}_j^*$. Therefore, $\mathcal{S}_j = \mathcal{S}(\mathcal{C}_j) \subseteq \mathcal{S}_j^*$. This means that for every $A_j \in \mathcal{S}_j$ we have $A_j^* \in \mathcal{S}(\tilde{\mathcal{C}})$.
Now, every element of $\mathcal{C}$ is a finite intersection (i.e. for a finite collection of indices $j \in I$) of sets of the form A_j^*, and so $\mathcal{C} \subseteq \mathcal{S}(\tilde{\mathcal{C}})$. Hence, $\mathcal{S}(\mathcal{C}) \subseteq \mathcal{S}(\tilde{\mathcal{C}})$. $\square$

In particular, $\bigotimes_{i=1}^n \mathcal{S}_i$ is generated by the collection of all sets of the form $A_1 \times \cdots \times A_n$, where $A_i \in \mathcal{C}_i$ for all $i = 1, \dots, n$. Also, $\bigotimes_{i=1}^{+\infty} \mathcal{S}_i$ is generated by the collection of all sets of the form $A_1 \times \cdots \times A_n \times X_{n+1} \times X_{n+2} \times \cdots$, where $n \in \mathbb{N}$ and $A_i \in \mathcal{C}_i$ for all $i = 1, \dots, n$.

SECTIONS OF SETS AND FUNCTIONS.

Let $x \in \prod_{i\in I} X_i$. Then x is a function with domain of definition I and values $x_i \in X_i$ for all $i \in I$. Now, if $J \subseteq I$, then we may consider the restriction x_J of x on J. Then x_J is a function with domain of definition J and values $(x_J)_i = x_i \in X_i$ for all $i \in J$. In other words, $x_J \in \prod_{i\in J} X_i$.

If $I = \{1, \dots, n\}$, then we use the notation $x = (x_1, \dots, x_n) \in X_1 \times \cdots \times X_n$. Now, if $J = \{i_1, \dots, i_m\}$ with $1 \leq i_1 < \cdots < i_m \leq n$ is a subset of I, then, accordingly, we use the notation $x_J = (x_{i_1}, \dots, x_{i_m}) \in X_{i_1} \times \cdots \times X_{i_m}$. For example, if $x = (x_1, x_2, x_3, x_4, x_5)$, then $x_{\{1,3,5\}} = (x_1, x_3, x_5)$.

We may also consider the complement $J^c = I \setminus J$ of $J \subseteq I$. Then, besides the restriction x_J of x on J, we may also consider the restriction x_{J^c} of x on J^c. We have $x_J \in \prod_{i\in J} X_i$ and $x_{J^c} \in \prod_{i\in J^c} X_i$. Also, $(x_J)_i = x_i \in X_i$ for all $i \in J$ and $(x_{J^c})_i = x_i \in X_i$ for all $i \in J^c$. Clearly, x uniquely determines x_J and x_{J^c}. Conversely, x is uniquely determined by its restrictions x_J and x_{J^c}. Indeed, if $y \in \prod_{i\in J} X_i$ and $z \in \prod_{i\in J^c} X_i$ are given, then there is a unique $x \in \prod_{i\in I} X_i$ so that $x_J = y$ and $x_{J^c} = z$: we define $x_i = y_i$, if $i \in J$, and $x_i = z_i$, if $i \in J^c$.

For example, if $x = (x_1, x_2, x_3, x_4, x_5)$, then $x_{\{1,3,5\}} = (x_1, x_3, x_5)$ and $x_{\{2,4\}} = (x_2, x_4)$. It is obvious that $x = (x_1, x_2, x_3, x_4, x_5)$ uniquely determines the restrictions $y = (x_1, x_3, x_5)$ and $z = (x_2, x_4)$ and is uniquely determined by them.

Thus, we have an identification between $\prod_{i\in I} X_i$ and $\big(\prod_{i\in J} X_i\big) \times \big(\prod_{i\in J^c} X_i\big)$. We identify the element x of the first space with the pair (y, z) of the second space, whenever $y = x_J$ and $z = x_{J^c}$. For example, we identify $x = (x_1, x_2, x_3, x_4, x_5)$ with $(y, z) = \big((x_1, x_3, x_5), (x_2, x_4)\big)$. It must be stressed that these are formal identifications (logically supported by underlying bijections) and *not* actual equalities.

Definition. *Let $A \subseteq \prod_{i\in I} X_i$ and $J \subseteq I$ and $z \in \prod_{i\in J^c} X_i$. We define*

$$A_z = \big\{y \in \textstyle\prod_{i\in J} X_i \,\big|\, (y, z) \in A\big\}.$$

*We call A_z the **z-section** of A.*

It is clear that every z-section of A is a subset of $\prod_{i\in J} X_i$.

For example, if $A \subseteq X_1 \times X_2 \times X_3 \times X_4 \times X_5$ and $(x_2, x_4) \in X_2 \times X_4$, then we have $A_{(x_2,x_4)} = \{(x_1, x_3, x_5) \,|\, (x_1, x_2, x_3, x_4, x_5) \in A\} \subseteq X_1 \times X_3 \times X_5$.

Definition. *Let $f : \prod_{i\in I} X_i \to Y$ and $J \subseteq I$ and $z \in \prod_{i\in J^c} X_i$. We define $f_z : \prod_{i\in J} X_i \to Y$ by*

$$f_z(y) = f(y, z) \quad \text{for all } y \in \textstyle\prod_{i\in J} X_i.$$

*We call f_z the **z-section** of f.*

For example, if $f : X_1 \times X_2 \times X_3 \times X_4 \times X_5 \to Y$ and $(x_2, x_4) \in X_2 \times X_4$, then $f_{(x_2,x_4)} : X_1 \times X_3 \times X_5 \to Y$ is defined by $f_{(x_2,x_4)}(x_1, x_3, x_5) = f(x_1, x_2, x_3, x_4, x_5)$ for all $(x_1, x_3, x_5) \in X_1 \times X_3 \times X_5$.

Whenever $J^c = \{j\}$ is a one-point set, then, for simplicity, we prefer to write A_{x_j} and f_{x_j}, instead of $A_{(x_j)}$ and $f_{(x_j)}$.

Proposition 4.2. *Let $(X_i, \mathcal{S}_i)$ be a measurable space for every $i \in I$ and let $J \subseteq I$ and $z \in \prod_{i\in J^c} X_i$. If $A \subseteq \prod_{i\in I} X_i$ belongs to $\bigotimes_{i\in I} \mathcal{S}_i$, then $A_z \subseteq \prod_{i\in J} X_i$ belongs to $\bigotimes_{i\in J} \mathcal{S}_i$.*

Proof. We fix $z \in \prod_{i\in J^c} X_i$ and we consider the collection $\mathcal{S}$ of all $A \subseteq \prod_{i\in I} X_i$ with the property that $A_z \in \bigotimes_{i\in J} \mathcal{S}_i$.
We shall prove that $\mathcal{S}$ is a σ-algebra of subsets of $\prod_{i\in I} X_i$.
For the $\emptyset \subseteq \prod_{i\in I} X_i$ we have $\emptyset_z = \emptyset \in \bigotimes_{i\in J} \mathcal{S}_i$, and so $\emptyset \in \mathcal{S}$.
Let $A \in \mathcal{S}$. Then $A_z \in \bigotimes_{i\in J} \mathcal{S}_i$. Hence,

$$(A^c)_z = (A_z)^c \in \textstyle\bigotimes_{i\in J} \mathcal{S}_i,$$

and so $A^c \in \mathcal{S}$.
Let $A_n \in \mathcal{S}$ for all $n \in \mathbb{N}$. Then $(A_n)_z \in \bigotimes_{i\in J} \mathcal{S}_i$ for all $n \in \mathbb{N}$. Hence,

$$\left(\bigcup_{n=1}^{+\infty} A_n\right)_z = \bigcup_{n=1}^{+\infty} (A_n)_z \in \bigotimes_{i\in J} \mathcal{S}_i,$$

and so $\bigcup_{n=1}^{+\infty} A_n \in \mathcal{S}$.
Now we fix a $k \in I$ and an $A_k \in \mathcal{S}_k$ and we consider the set $A_k^* = \prod_{i\in I} Y_i$, where $Y_i = X_i$, if $i \neq k$, and $Y_k = A_k$. We observe that, if $k \in J$, then $(A_k^*)_z = \prod_{i\in J} Y_i$, where $Y_i = X_i$ if $i \in J$, $i \neq k$, and $Y_k = A_k$. Also, if $k \in J^c$, then $(A_k^*)_z = \prod_{i\in J} Y_i$, where $Y_i = X_i$ for all $i \in J$. In both cases we have that $(A_k^*)_z \in \bigotimes_{i\in J} \mathcal{S}_i$, and so $A_k^* \in \mathcal{S}$.
Now we observe that every element in the original collection $\mathcal{C}$ which generates $\bigotimes_{i\in I} \mathcal{S}_i$ is a finite intersection of sets A_k^* (for a finite collection of indices $k \in I$), and so $\mathcal{C} \subseteq \mathcal{S}$. Therefore, $\bigotimes_{i\in I} \mathcal{S}_i \subseteq \mathcal{S}$. Thus, if $A \in \bigotimes_{i\in I} \mathcal{S}_i$, then $A \in \mathcal{S}$, and so $A_z \in \bigotimes_{i\in J} \mathcal{S}_i$. $\square$

Proposition 4.3. *Let $(X_i, \mathcal{S}_i), (Y, \mathcal{S})$ be measurable spaces for every $i \in I$ and let $J \subseteq I$ and $z \in \prod_{i\in J^c} X_i$. If $f : \prod_{i\in I} X_i \to Y$ is $(\bigotimes_{i\in I} \mathcal{S}_i, \mathcal{S})$-measurable, then $f_z : \prod_{i\in J} X_i \to Y$ is $(\bigotimes_{i\in J} \mathcal{S}_i, \mathcal{S})$-measurable.*

Proof. Let $B \in \mathcal{S}$. Then $f^{-1}(B) \in \bigotimes_{i\in I} \mathcal{S}_i$. Since

$$(f_z)^{-1}(B) = (f^{-1}(B))_z,$$

Theorem 4.1 implies that $(f_z)^{-1}(B) \in \bigotimes_{i\in J} \mathcal{S}_i$. $\square$

The last two theorems say, in informal language, that *sets or functions which are measurable on a product space have all their sections measurable on the appropriate product subspaces*.

PRODUCTS OF BOREL σ-ALGEBRAS.

Example. We consider $\mathbb{R}^n = \prod_{i=1}^n \mathbb{R}$, and, for each copy of $\mathbb{R}$, we consider the collection of all bounded 1-dimensional intervals as a generator of $\mathcal{B}_1$. Proposition 4.1 implies that the collection of all bounded n-dimensional intervals is a generator of $\bigotimes_{i=1}^n \mathcal{B}_1$. But we already know that the same collection is a generator of $\mathcal{B}_n$. Therefore,

$$\mathcal{B}_n = \bigotimes_{i=1}^n \mathcal{B}_1.$$

This can be generalised.
If $n_1 + \cdots + n_k = n$, we formally identify the typical element $(x_1, \ldots, x_n) \in \mathbb{R}^n$ with

$$\big((x_1, \ldots, x_{n_1}), (x_{n_1+1}, \ldots, x_{n_1+n_2}), \ldots, (x_{n_1+\cdots+n_{k-1}+1}, \ldots, x_{n_1+\cdots+n_k})\big),$$

i.e. with the typical element of $\prod_{j=1}^k \mathbb{R}^{n_j}$. In other words, we consider the identification:

$$\mathbb{R}^n = \prod_{j=1}^k \mathbb{R}^{n_j}.$$

Now, Proposition 4.1 implies that $\bigotimes_{j=1}^k \mathcal{B}_{n_j}$ is generated by the collection of all $\prod_{j=1}^k A_j$, where A_j is an n_j-dimensional bounded interval. By the above identification, $\prod_{j=1}^k A_j$ is the typical n-dimensional bounded interval, and so $\bigotimes_{j=1}^k \mathcal{B}_{n_j}$ is generated by the collection of all bounded intervals in $\mathbb{R}^n$. But the same collection generates $\mathcal{B}_n$, and we conclude that

$$\mathcal{B}_n = \bigotimes_{j=1}^k \mathcal{B}_{n_j}.$$

Let X be any non-empty set. We recall that a *topology* $\mathcal{T}$ of X is any collection of subsets of X which contains $\emptyset$ and X and which is closed under arbitrary unions and under finite intersections. The elements of a topology $\mathcal{T}$ of X are called *open* subsets of X, and the complements of the open subsets of X are called *closed* subsets of X. A set X with a topology of X is called *topological space*.

It is well known (and trivial to show) that any intersection of topologies of X is a topology of X. Now, let $\mathcal{C}$ be an arbitrary collection of subsets of X. We consider all the topologies of X which include $\mathcal{C}$, and we take their intersection. This is, clearly, the smallest topology of X which includes $\mathcal{C}$, it is called the topology of X **generated** by $\mathcal{C}$ and it is denoted $\mathcal{T}(\mathcal{C})$.

Definition. *Let X_i be a topological space with topology $\mathcal{T}_i$ for every $i \in I$. We consider the collection*

$$\mathcal{C} = \left\{ \textstyle\prod_{i\in I} U_i \,\middle|\, U_i \neq X_i \text{ for at most finitely many } i \in I, \text{ and } U_i \in \mathcal{T}_i \text{ if } U_i \neq X_i \right\}.$$

Then $\mathcal{T}(\mathcal{C})$ is called the ***product topology*** *of $\prod_{i\in I} X_i$.*

We say that a topological space X is **second countable** if there is a *countable* collection of open subsets of X such that every open subset of X can be written as a (necessarily, countable) union of open sets contained in this collection.

Proposition 4.4. *Let X_i be a topological space for every $i \in I$ and let $X = \prod_{i\in I} X_i$ have the product topology. Then $\bigotimes_{i\in I} \mathcal{B}_{X_i} \subseteq \mathcal{B}_X$. If, moreover, I is countable and every X_i is second countable, then $\bigotimes_{i\in I} \mathcal{B}_{X_i} = \mathcal{B}_X$.*

Proof. Let $\mathcal{T}_i$ be the topology of X_i for every $i \in I$. Then $\mathcal{B}_{X_i} = \mathcal{S}(\mathcal{T}_i)$. Proposition 4.1 implies that $\bigotimes_{i\in I} \mathcal{B}_{X_i} = \mathcal{S}(\mathcal{C})$, where

$$\mathcal{C} = \left\{ \textstyle\prod_{i\in I} U_i \,\middle|\, U_i \neq X_i \text{ for at most finitely many } i \in I, \text{ and } U_i \in \mathcal{T}_i \text{ if } U_i \neq X_i \right\}.$$

But, by the definition of the product topology $\mathcal{T}$ of X, we have that $\mathcal{C} \subseteq \mathcal{T} \subseteq \mathcal{S}(\mathcal{T}) = \mathcal{B}_X$. Therefore, $\bigotimes_{i\in I} \mathcal{B}_{X_i} \subseteq \mathcal{B}_X$.
Now, let I be countable and every X_i be second countable. Since X_i is second countable, there is a countable collection $\mathcal{C}_i$ of open subsets of X_i so that every open subset of X_i can be written as a countable union of sets contained in $\mathcal{C}_i$. We consider

$$\tilde{\mathcal{C}} = \left\{ \textstyle\prod_{i\in I} U_i \,\middle|\, U_i \neq X_i \text{ for at most finitely many } i \in I, \text{ and } U_i \in \mathcal{C}_i \text{ if } U_i \neq X_i \right\}.$$

Then $\tilde{\mathcal{C}}$ is countable. Moreover, since $\tilde{\mathcal{C}} \subseteq \mathcal{C}$, we get $\mathcal{S}(\tilde{\mathcal{C}}) \subseteq \mathcal{S}(\mathcal{C}) = \bigotimes_{i\in I} \mathcal{B}_{X_i}$.
Now, we consider the collection $\mathcal{T}^*$ which contains $\emptyset$ and all unions (necessarily, countable) of elements of $\tilde{\mathcal{C}}$. It is clear that $\mathcal{T}^* \subseteq \mathcal{S}(\tilde{\mathcal{C}}) \subseteq \bigotimes_{i\in I} \mathcal{B}_{X_i}$. Every finite intersection of elements of $\mathcal{T}^*$ is either $\emptyset$ or a countable union of finite intersections of elements of $\tilde{\mathcal{C}}$. But it is easy to see that every finite intersection of elements of $\tilde{\mathcal{C}}$ belongs to $\mathcal{C}$ and so it is a countable union of elements of $\tilde{\mathcal{C}}$. Therefore, $\mathcal{T}^*$ is closed under finite intersections, and, since it is obviously closed under arbitrary unions, it is a topology of X.
We have already mentioned that every element of $\mathcal{C}$ is a union of elements of $\tilde{\mathcal{C}}$, and so $\mathcal{C} \subseteq \mathcal{T}^*$. Thus, $\mathcal{T} = \mathcal{T}(\mathcal{C}) \subseteq \mathcal{T}^* \subseteq \bigotimes_{i\in I} \mathcal{B}_{X_i}$. Therefore, $\mathcal{B}_X = \mathcal{S}(\mathcal{T}) \subseteq \bigotimes_{i\in I} \mathcal{B}_{X_i}$. □

Example. We consider $\mathbb{R}^n = \prod_{i=1}^n \mathbb{R}$, where $I = \{1, \ldots, n\}$. We also consider $J = \{i_1, \ldots, i_m\}$ with $1 \leq i_1 < \cdots < i_m \leq n$. We take $k = n - m$ and we write $J^c = \{i'_1, \ldots, i'_k\}$ with $1 \leq i'_1 < \cdots < i'_k \leq n$.
We naturally identify $\prod_{i\in J} \mathbb{R}$ with $\mathbb{R}^m$, writing each $y = (x_{i_1}, \ldots, x_{i_m})$ as $y = (y_1, \ldots, y_m)$. We also identify $\prod_{i\in J^c} \mathbb{R}$ with $\mathbb{R}^k$, writing each $z = (x_{i'_1}, \ldots, x_{i'_k})$ as $z = (z_1, \ldots, z_k)$.
Therefore, $\bigotimes_{i\in J} \mathcal{B}_1 = \mathcal{B}_m$ and $\bigotimes_{i\in J^c} \mathcal{B}_1 = \mathcal{B}_k$. Also, $\mathcal{B}_n = \mathcal{B}_m \otimes \mathcal{B}_k$.
Now, if A is a Borel set in $\mathbb{R}^n$, then, for arbitrary $z \in \prod_{i\in J^c} \mathbb{R} = \mathbb{R}^k$, the z-section A_z of A is a Borel set in $\mathbb{R}^m$.

Example. We consider any $E \subseteq \mathbb{R}$, which is *not* a Borel set, and $A = \{(x,x) \in \mathbb{R}^2 \,|\, x \in E\}$. Since all 1-dimensional sections of A are either empty or one-point sets, they are Borel sets in $\mathbb{R}$. We shall see that A is *not* a Borel set in $\mathbb{R}^2$.
Indeed, we assume that A is a Borel set in $\mathbb{R}^2$, and we consider the invertible linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ defined by $T(x_1, x_2) = \big(\frac{x_1+x_2}{2}, \frac{x_1-x_2}{2}\big)$.
Then $T(A) = \{(x,0) \,|\, x \in E\}$ is a Borel set in $\mathbb{R}^2$, and so all 1-dimensional sections of $T(A)$ must be Borel sets in $\mathbb{R}$. In particular, the (horizontal) section $T(A)_0 = \{x \,|\, x \in E\} = E$ must be a Borel set in $\mathbb{R}$, and so we arrive at a contradiction.

Exercises.

4.1.1. (i) The function $\pi_j : \prod_{i \in I} X_i \to X_j$ defined by $\pi_j(x) = x_j$ for all $x = (x_i)_{i\in I} \in \prod_{i\in I} X_i$, is called the **j-th projection** of $\prod_{i\in I} X_i$ or the projection of $\prod_{i\in I} X_i$ onto its j-th component X_j. If $A_j \subseteq X_j$, prove that $\pi_j^{-1}(A_j) = \prod_{i\in I} Y_i$, where $Y_i = X_i$, if $i \in I$, $i \neq j$, and $Y_j = A_j$.
(ii) Let $(X_i, \mathcal{S}_i)$ be a measurable space for every $i \in I$. Prove that the product σ-algebra $\bigotimes_{i\in I} \mathcal{S}_i$ is the smallest σ-algebra $\mathcal{S}$ of subsets of $\prod_{i\in I} X_i$ such that $\pi_j : \prod_{i\in I} X_i \to X_j$ is $(\mathcal{S}, \mathcal{S}_j)$-measurable for every $j \in I$.
(iii) Let $(X_i, \mathcal{S}_i)$ be a measurable space for every $i \in I$, and $(Y, \mathcal{S})$ be a measurable space, and $i_0 \in I$, and $g : X_{i_0} \to Y$ be $(\mathcal{S}_{i_0}, \mathcal{S})$-measurable. If we define $f : \prod_{i\in I} X_i \to Y$ by $f\big((x_i)_{i\in I}\big) = g(x_{i_0})$, prove that f is $(\bigotimes_{i\in I} \mathcal{S}_i, \mathcal{S})$-measurable.
(iv) Let $(X_i, \mathcal{T}_i)$ be a topological space for every $i \in I$. Prove that the product topology is the smallest topology on $\prod_{i\in I} X_i$ such that $\pi_j : \prod_{i\in I} X_i \to X_j$ is continuous for every $j \in I$.

4.1.2. Let $(X_i, \mathcal{S}_i)$ be a measurable space for every $i \in I$ and let $\mathcal{C}_i$ be a collection of subsets of X_i so that $\mathcal{S}_i = \mathcal{S}(\mathcal{C}_i)$ for every $i \in I$. Prove that $\bigotimes_{i\in I} \mathcal{S}_i = \mathcal{S}(\tilde{\mathcal{C}})$, where

$$\tilde{\mathcal{C}} = \{\textstyle\prod_{i\in I} E_i \,|\, E_i \neq X_i \text{ for at most countably many } i \in I, \text{ and } E_i \in \mathcal{C}_i \text{ if } E_i \neq X_i\}.$$

4.1.3. Let $\mathbb{R}^n_* = \mathbb{R}^n \setminus \{0\}$.
(i) If U is open in $\mathbb{R}^n_*$, prove that $\mathbb{R}^+ U = \{rx \,|\, r > 0, x \in U\}$ is open in $\mathbb{R}^n_*$.
(ii) If A is a Borel set in $\mathbb{R}^n_*$, prove that $\mathbb{R}^+ A$ is a Borel set in $\mathbb{R}^n_*$.

4.2 Product measure.

In this section we shall limit ourselves to cartesian products of finitely many spaces and, for simplicity, we shall work with two measure spaces.

We fix the measure spaces $(X_1, \mathcal{S}_1, \mu_1)$, $(X_2, \mathcal{S}_2, \mu_2)$ and $(X_1 \times X_2, \mathcal{S}_1 \otimes \mathcal{S}_2)$. We know that $\mathcal{S}_1 \otimes \mathcal{S}_2$, by its definition, is generated by the collection

$$\mathcal{C} = \{A_1 \times A_2 \,|\, A_1 \in \mathcal{S}_1,\ A_2 \in \mathcal{S}_2\}.$$

We observe that $X_1 \times X_2$ and $\emptyset \times \emptyset = \emptyset$ belong to $\mathcal{C}$.

The elements of $\mathcal{C}$ play the same role that n-dimensional intervals play for the introduction of Lebesgue measure on $\mathbb{R}^n$. We agree to call these sets $(\mathcal{S}_1 \otimes \mathcal{S}_2)$**-measurable intervals** or, for simplicity, just **measurable intervals** in $X_1 \times X_2$, a term which will be justified by Theorem 4.3, and denote them by

$$R = A_1 \times A_2.$$

Proposition 4.5. *The* $\mathcal{A} = \{R_1 \cup \cdots \cup R_m \,|\, m \in \mathbb{N}, R_1, \ldots, R_m$ *pairwise disjoint elements of* $\mathcal{C}\}$ *is an algebra of subsets of* $X_1 \times X_2$.

Proof. Similar to the proof of Proposition 1.11. We first prove that $R' \cap R'' \in \mathcal{C}$ for all $R', R'' \in \mathcal{C}$. This implies that $\mathcal{A}$ is closed under finite intersections. Then we prove that $R^c \in \mathcal{A}$ for every $R \in \mathcal{C}$. This implies that $\mathcal{A}$ is closed under complements. Finally, we prove that $\mathcal{A}$ is closed under finite unions. The details are left to the reader as an exercise. $\square$

For each $R = A_1 \times A_2 \in \mathcal{C}$, we define the quantity

$$\tau(R) = \mu_1(A_1)\mu_2(A_2),$$

which plays the role of *volume* of the measurable interval R.

Definition. *For every $E \subseteq X_1 \times X_2$ we define*

$$\mu^*(E) = \inf\left\{ \textstyle\sum_{i=1}^{+\infty} \tau(R_i) \,\middle|\, R_i \in \mathcal{C} \text{ for all } i, \text{ and } E \subseteq \bigcup_{i=1}^{+\infty} R_i \right\}$$

Theorem 1.8 implies that $\mu^* : \mathcal{P}(X_1 \times X_2) \to [0, +\infty]$ is an outer measure on $X_1 \times X_2$.

Proposition 4.6. *Let R, R_i be measurable intervals for every $i \in \mathbb{N}$.*
(i) If $R \subseteq \bigcup_{i=1}^{+\infty} R_i$, then $\tau(R) \le \sum_{i=1}^{+\infty} \tau(R_i)$.
(ii) If $R = \bigcup_{i=1}^{+\infty} R_i$ and the R_i are pairwise disjoint, then $\tau(R) = \sum_{i=1}^{+\infty} \tau(R_i)$.

Proof. (i) Let $R = A_1 \times A_2$ and $R_i = A_{i,1} \times A_{i,2}$.
From $A_1 \times A_2 \subseteq \bigcup_{i=1}^{+\infty}(A_{i,1} \times A_{i,2})$, we get that

$$\chi_{A_1}(x_1)\chi_{A_2}(x_2) = \chi_{A_1\times A_2}(x_1, x_2) \le \textstyle\sum_{i=1}^{+\infty} \chi_{A_{i,1}\times A_{i,2}}(x_1, x_2) = \sum_{i=1}^{+\infty} \chi_{A_{i,1}}(x_1)\chi_{A_{i,2}}(x_2)$$

for every $x_1 \in X_1, x_2 \in X_2$. Integrating over X_1 with respect to μ_1, we find

$$\mu_1(A_1)\chi_{A_2}(x_2) \le \textstyle\sum_{i=1}^{+\infty} \mu_1(A_{i,1})\chi_{A_{i,2}}(x_2)$$

for every $x_2 \in X_2$. Integrating the last relation over X_2 with respect to μ_2, we get

$$\mu_1(A_1)\mu_2(A_2) \le \textstyle\sum_{i=1}^{+\infty} \mu_1(A_{i,1})\mu_2(A_{i,2}).$$

(ii) We use equalities everywhere in the above calculations. $\square$

The next result justifies the term *measurable interval* for each $R \in \mathcal{C}$.

Theorem 4.1. *Every $R \in \mathcal{C}$ is μ^*-measurable, and $\mu^*(R) = \tau(R)$. Moreover, $\mathcal{S}_1 \otimes \mathcal{S}_2$ is included in the σ-algebra $\mathcal{S}_{\mu^*}$ of μ^*-measurable subsets of $X_1 \times X_2$.*

Proof. (a) Let $R \in \mathcal{C}$. From $R \subseteq R$, we get $\mu^*(R) \le \tau(R)$. Proposition 4.4 implies $\tau(R) \le \sum_{i=1}^{+\infty} \tau(R_i)$ for every covering $R \subseteq \bigcup_{i=1}^{+\infty} R_i$ with $R_i \in \mathcal{C}$. Hence, $\tau(R) \le \mu^*(R)$, and we conclude that $\mu^*(R) = \tau(R)$.
(b) Let $R, R' \in \mathcal{C}$. Proposition 4.3 implies that there are pairwise disjoint $R_1, \dots, R_m \in \mathcal{C}$ so that $R' \setminus R = R_1 \cup \dots \cup R_m$. By the subadditivity of μ^*, the result of (a) and Proposition 4.4, we get

$$\begin{aligned}\mu^*(R' \cap R) + \mu^*(R' \setminus R) &\le \mu^*(R' \cap R) + \mu^*(R_1) + \dots + \mu^*(R_n)\\ &= \tau(R' \cap R) + \tau(R_1) + \dots + \tau(R_n) = \tau(R').\end{aligned}$$

(c) Let $R \in \mathcal{C}$ and $E \subseteq X_1 \times X_2$ with $\mu^*(E) < +\infty$. For any $\epsilon > 0$ there is a covering $E \subseteq \bigcup_{i=1}^{+\infty} R_i$ with $R_i \in \mathcal{C}$ such that $\sum_{i=1}^{+\infty} \tau(R_i) < \mu^*(E) + \epsilon$. By the result of (b) and the subadditivity of μ^*, we get

$$\mu^*(E \cap R) + \mu^*(E \setminus R) \le \textstyle\sum_{i=1}^{+\infty} \big(\mu^*(R_i \cap R) + \mu^*(R_i \setminus R)\big) \le \sum_{i=1}^{+\infty} \tau(R_i) < \mu^*(E) + \epsilon.$$

Since ϵ is arbitrary, we get $\mu^*(E \cap R) + \mu^*(E \setminus R) \le \mu^*(E)$, and we conclude that R is μ^*-measurable.
Therefore, $\mathcal{C} \subseteq \mathcal{S}_{\mu^*}$. Since $\mathcal{S}_1 \otimes \mathcal{S}_2$ is generated by $\mathcal{C}$, we have that $\mathcal{S}_1 \otimes \mathcal{S}_2 \subseteq \mathcal{S}_{\mu^*}$. $\square$

Definition. *The measure induced from μ^* is called the* **product measure** *of μ_1, μ_2 and it is denoted*

$$\mu_1 \otimes \mu_2.$$

We denote by

$$\mathcal{S}_{\mu_1 \otimes \mu_2}$$

the σ-algebra $\mathcal{S}_{\mu^*}$ of μ^*-measurable subsets of $X_1 \times X_2$.
Thus, $(X_1 \times X_2, \mathcal{S}_{\mu_1\otimes\mu_2}, \mu_1 \otimes \mu_2)$ is a complete measure space.

Theorem 4.3 implies that

$$\mathcal{S}_1 \otimes \mathcal{S}_2 \subseteq \mathcal{S}_{\mu_1\otimes\mu_2}$$

and

$$(\mu_1 \otimes \mu_2)(A_1 \times A_2) = \mu_1(A_1)\mu_2(A_2) \quad \text{for all } A_1 \in \mathcal{S}_1, A_2 \in \mathcal{S}_2.$$

It is very common to consider the restriction, also denoted $\mu_1 \otimes \mu_2$, of $\mu_1 \otimes \mu_2$ on $\mathcal{S}_1 \otimes \mathcal{S}_2$.

Theorem 4.2. *If μ_1 and μ_2 are σ-finite measures, then*
(i) $\mu_1 \otimes \mu_2$ is the unique measure μ on $(X_1 \times X_2, \mathcal{S}_1 \otimes \mathcal{S}_2)$ such that $\mu(A_1 \times A_2) = \mu_1(A_1)\mu_2(A_2)$ for every $A_1 \in \mathcal{S}_1, A_2 \in \mathcal{S}_2$,
(ii) $(X_1 \times X_2, \mathcal{S}_{\mu_1\otimes\mu_2}, \mu_1 \otimes \mu_2)$ is the completion of $(X_1 \times X_2, \mathcal{S}_1 \otimes \mathcal{S}_2, \mu_1 \otimes \mu_2)$.

Proof. (i) We consider the algebra $\mathcal{A}$ of subsets of $X_1 \times X_2$ which is described in Proposition 4.3. Let μ be any measure on $(X_1 \times X_2, \mathcal{S}_1 \otimes \mathcal{S}_2)$ such that $\mu(R) = (\mu_1 \otimes \mu_2)(R)$ for every $R \in \mathcal{C}$. Then, by additivity of measures, we have that $\mu(R_1 \cup \cdots \cup R_m) = (\mu_1 \otimes \mu_2)(R_1 \cup \cdots \cup R_m)$ for all pairwise disjoint $R_1, \ldots, R_m \in \mathcal{C}$. Therefore, the measures μ and $\mu_1 \otimes \mu_2$ are equal on $\mathcal{A}$.
Since μ_1, μ_2 are σ-finite, there exist $A_{i,1} \in \mathcal{S}_1$, $A_{i,2} \in \mathcal{S}_2$ with $\mu_1(A_{i,1}) < +\infty$, $\mu_2(A_{i,2}) < +\infty$ and $A_{i,1} \uparrow X_1$, $A_{i,2} \uparrow X_2$. This implies that the $R_i = A_{i,1} \times A_{i,2} \in \mathcal{C}$ have the property that $R_i \uparrow X_1 \times X_2$ and that $\mu(R_i) = (\mu_1 \otimes \mu_2)(R_i) = \mu_1(A_{i,1})\mu_2(A_{i,2}) < +\infty$ for every i.
Since $\mathcal{S}_1 \otimes \mathcal{S}_2 = \mathcal{S}(\mathcal{C}) = \mathcal{S}(\mathcal{A})$, Theorem 1.7 implies that μ and $\mu_1 \otimes \mu_2$ are equal on $\mathcal{S}_1 \otimes \mathcal{S}_2$.
(ii) $(X_1 \times X_2, \mathcal{S}_{\mu_1\otimes\mu_2}, \mu_1 \otimes \mu_2)$ is a complete extension of $(X_1 \times X_2, \mathcal{S}_1 \otimes \mathcal{S}_2, \mu_1 \otimes \mu_2)$, and so it is an extension of the completion $(X, \overline{\mathcal{S}_1 \otimes \mathcal{S}_2}, \overline{\mu_1 \otimes \mu_2})$. Hence, it is enough to prove that $\mathcal{S}_{\mu_1\otimes\mu_2} \subseteq \overline{\mathcal{S}_1 \otimes \mathcal{S}_2}$.
Let $A \in \mathcal{S}_{\mu_1\otimes\mu_2}$ with $(\mu_1 \otimes \mu_2)(A) < +\infty$.
For each $k \in \mathbb{N}$ there is a covering $A \subseteq \bigcup_{i=1}^{+\infty} R_{k,i}$ by measurable intervals so that

$$\sum_{i=1}^{+\infty} \tau(R_{k,i}) < (\mu_1 \otimes \mu_2)(A) + \tfrac{1}{k}.$$

We define $B_k = \bigcup_{i=1}^{+\infty} R_{k,i} \in \mathcal{S}_1 \otimes \mathcal{S}_2$, and we have that $A \subseteq B_k$ and

$$(\mu_1 \otimes \mu_2)(A) \leq (\mu_1 \otimes \mu_2)(B_k) < (\mu_1 \otimes \mu_2)(A) + \tfrac{1}{k}.$$

Now, we define $B = \bigcap_{k=1}^{+\infty} B_k \in \mathcal{S}_1 \otimes \mathcal{S}_2$. Then $A \subseteq B$ and $(\mu_1 \otimes \mu_2)(A) = (\mu_1 \otimes \mu_2)(B)$. Therefore, $(\mu_1 \otimes \mu_2)(B \setminus A) = 0$.
Now, let $A \in \mathcal{S}_{\mu_1\otimes\mu_2}$ with $(\mu_1 \otimes \mu_2)(A) = +\infty$.
We consider the specific measurable intervals R_i which we used in the proof of part (i), and the $A_i = A \cap R_i$. These sets have $(\mu_1 \otimes \mu_2)(A_i) < +\infty$, and, by the previous paragraph, there are $B_i \in \mathcal{S}_1 \otimes \mathcal{S}_2$ so that $A_i \subseteq B_i$ and $(\mu_1 \otimes \mu_2)(B_i \setminus A_i) = 0$. We define $B = \bigcup_{i=1}^{+\infty} B_i \in \mathcal{S}_1 \otimes \mathcal{S}_2$. Then $A \subseteq B$, and, since $B \setminus A \subseteq \bigcup_{i=1}^{+\infty}(B_i \setminus A_i)$, we conclude that $(\mu_1 \otimes \mu_2)(B \setminus A) = 0$.
We proved that for each $A \in \mathcal{S}_{\mu_1\otimes\mu_2}$ there is $B \in \mathcal{S}_1 \otimes \mathcal{S}_2$ so that $A \subseteq B$ and $(\mu_1 \otimes \mu_2)(B \setminus A) = 0$. Considering $B \setminus A$, there is $C \in \mathcal{S}_1 \otimes \mathcal{S}_2$ so that $B \setminus A \subseteq C$ and $(\mu_1 \otimes \mu_2)\big(C \setminus (B \setminus A)\big) = 0$. Of course, $(\mu_1 \otimes \mu_2)(C) = 0$.
Now we observe that $A = (B \setminus C) \cup (A \cap C)$, where $B \setminus C \in \mathcal{S}_1 \otimes \mathcal{S}_2$ and $A \cap C \subseteq C \in \mathcal{S}_1 \otimes \mathcal{S}_2$ with $(\mu_1 \otimes \mu_2)(C) = 0$. This says that $A \in \overline{\mathcal{S}_1 \otimes \mathcal{S}_2}$. $\square$

Now we shall examine the influence to the product measure space of replacing the measure spaces $(X_1, \mathcal{S}_1, \mu_1), (X_2, \mathcal{S}_2, \mu_2)$ with their completions $(X_1, \overline{\mathcal{S}_1}, \overline{\mu_1}), (X_2, \overline{\mathcal{S}_2}, \overline{\mu_2})$.

Theorem 4.3. *(i) The measure spaces $(X_1, \mathcal{S}_1, \mu_1), (X_2, \mathcal{S}_2, \mu_2)$ induce the same product measure space as their completions $(X_1, \overline{\mathcal{S}_1}, \overline{\mu_1}), (X_2, \overline{\mathcal{S}_2}, \overline{\mu_2})$. Namely,*

$$(X_1 \times X_2, \mathcal{S}_{\mu_1 \otimes \mu_2}, \mu_1 \otimes \mu_2) = (X_1 \times X_2, \mathcal{S}_{\overline{\mu_1} \otimes \overline{\mu_2}}, \overline{\mu_1} \otimes \overline{\mu_2}).$$

Moreover, the above product measure space is an extension of both $(X_1 \times X_2, \mathcal{S}_1 \otimes \mathcal{S}_2, \mu_1 \otimes \mu_2)$ and $(X_1 \times X_2, \overline{\mathcal{S}_1} \otimes \overline{\mathcal{S}_2}, \overline{\mu_1} \otimes \overline{\mu_2})$, of which the second is an extension of the first.
(ii) If μ_1 and μ_2 are σ-finite measures, then $(X_1 \times X_2, \mathcal{S}_{\mu_1 \otimes \mu_2}, \mu_1 \otimes \mu_2)$ is the completion of both $(X_1 \times X_2, \mathcal{S}_1 \otimes \mathcal{S}_2, \mu_1 \otimes \mu_2)$ and $(X_1 \times X_2, \overline{\mathcal{S}_1} \otimes \overline{\mathcal{S}_2}, \overline{\mu_1} \otimes \overline{\mu_2})$.

Proof. (i) We recall that to construct the product measure space $(X_1 \times X_2, \mathcal{S}_{\mu_1 \otimes \mu_2}, \mu_1 \otimes \mu_2)$, we consider all $(\mathcal{S}_1 \otimes \mathcal{S}_2)$-measurable intervals of the form $R = A_1 \times A_2$ for any $A_1 \in \mathcal{S}_1, A_2 \in \mathcal{S}_2$, and we define the outer measure

$$\mu^*(E) = \inf\big\{ \textstyle\sum_{i=1}^{+\infty} \tau(R_i) \,\big|\, R_i \text{ are } (\mathcal{S}_1 \otimes \mathcal{S}_2)\text{-measurable intervals, and } E \subseteq \bigcup_{i=1}^{+\infty} R_i \big\},$$

where $\tau(R) = \mu_1(A_1)\mu_2(A_2)$ for all $R = A_1 \times A_2$. Similarly, to construct the product measure space $(X_1 \times X_2, \mathcal{S}_{\overline{\mu_1} \otimes \overline{\mu_2}}, \overline{\mu_1} \otimes \overline{\mu_2})$, we consider all $(\overline{\mathcal{S}_1} \otimes \overline{\mathcal{S}_2})$-measurable intervals of the form $R = A_1 \times A_2$ for any $A_1 \in \overline{\mathcal{S}_1}, A_2 \in \overline{\mathcal{S}_2}$, and we define the outer measure

$$\nu^*(E) = \inf\big\{ \textstyle\sum_{i=1}^{+\infty} \sigma(R_i) \,\big|\, R_i \text{ are } (\overline{\mathcal{S}_2} \otimes \overline{\mathcal{S}_2})\text{-measurable intervals, and } E \subseteq \bigcup_{i=1}^{+\infty} R_i \big\},$$

where $\sigma(R) = \overline{\mu_1}(A_1)\overline{\mu_2}(A_2)$ for all $R = A_1 \times A_2$.
Our first task will be to prove that the two outer measures μ^* and ν^* are identical.
We observe that all $(\mathcal{S}_1 \otimes \mathcal{S}_2)$-measurable intervals R are at the same time $(\overline{\mathcal{S}_1} \otimes \overline{\mathcal{S}_2})$-measurable and that $\sigma(R) = \tau(R)$ for them. Hence, $\nu^*(E) \le \mu^*(E)$ for every $E \subseteq X_1 \times X_2$.
Now let $E \subseteq X_1 \times X_2$ with $\nu^*(E) < +\infty$, and let $\epsilon > 0$. Then there is a covering $E \subseteq \bigcup_{i=1}^{+\infty} R_i$ with $(\overline{\mathcal{S}_1} \otimes \overline{\mathcal{S}_2})$-measurable intervals R_i so that

$$\textstyle\sum_{i=1}^{+\infty} \sigma(R_i) < \nu^*(E) + \epsilon.$$

For all i we have $R_i = A_{i,1} \times A_{i,2}$ with $A_{i,1} \in \overline{\mathcal{S}_1}, A_{i,2} \in \overline{\mathcal{S}_2}$. Then there are $B_{i,1} \in \mathcal{S}_1, B_{i,2} \in \mathcal{S}_2$ so that $A_{i,1} \subseteq B_{i,1}, A_{i,2} \subseteq B_{i,2}$ and $\overline{\mu_1}(A_{i,1}) = \mu_1(B_{i,1}), \overline{\mu_2}(A_{i,2}) = \mu_2(B_{i,2})$. We form the $(\mathcal{S}_1 \otimes \mathcal{S}_2)$-measurable intervals $R_i' = B_{i,1} \times B_{i,2}$ and we have $R_i \subseteq R_i'$ and $\sigma(R_i) = \tau(R_i')$ for all i. We now have a covering $E \subseteq \bigcup_{i=1}^{+\infty} R_i'$ with $(\mathcal{S}_1 \otimes \mathcal{S}_2)$-measurable intervals R_i', and this implies

$$\mu^*(E) \le \textstyle\sum_{i=1}^{+\infty} \tau(R_i') = \sum_{i=1}^{+\infty} \sigma(R_i) < \nu^*(E) + \epsilon.$$

Hence, $\mu^*(E) \le \nu^*(E)$. If $\nu^*(E) = +\infty$, then $\mu^*(E) \le \nu^*(E)$ is obviously true.
We conclude that $\mu^*(E) = \nu^*(E)$ for every $E \subseteq X_1 \times X_2$.
The next step in forming the product measures is to apply Caratheodory's Theorem to the common outer measure $\mu^* = \nu^*$, and we find the common complete product measure space

$$(X_1 \times X_2, \mathcal{S}_{\mu_1 \otimes \mu_2}, \mu_1 \otimes \mu_2) = (X_1 \times X_2, \mathcal{S}_{\overline{\mu_1} \otimes \overline{\mu_2}}, \overline{\mu_1} \otimes \overline{\mu_2})$$

where $\mathcal{S}_{\mu_1 \otimes \mu_2} = \mathcal{S}_{\overline{\mu_1} \otimes \overline{\mu_2}}$ is the symbol we use for $\mathcal{S}_{\mu^*} = \mathcal{S}_{\nu^*}$, and $\mu_1 \otimes \mu_2 = \overline{\mu_1} \otimes \overline{\mu_2}$ is the restriction of $\mu^* = \nu^*$ on $\mathcal{S}_{\mu^*} = \mathcal{S}_{\nu^*}$.
Finally, Theorem 4.3 says that $\mathcal{S}_1 \otimes \mathcal{S}_2$ and $\overline{\mathcal{S}_1} \otimes \overline{\mathcal{S}_2}$ are included in $\mathcal{S}_{\mu_1 \otimes \mu_2}$ and, since every $(\mathcal{S}_1 \otimes \mathcal{S}_2)$-measurable interval is also a $(\overline{\mathcal{S}_1} \otimes \overline{\mathcal{S}_2})$-measurable interval, we have that $\mathcal{S}_1 \otimes \mathcal{S}_2$ is included in $\overline{\mathcal{S}_1} \otimes \overline{\mathcal{S}_2}$. Thus, $\mathcal{S}_1 \otimes \mathcal{S}_2 \subseteq \overline{\mathcal{S}_2} \otimes \overline{\mathcal{S}_2} \subseteq \mathcal{S}_{\mu_1 \otimes \mu_2}$.
(ii) The proof is immediate from Theorem 4.4. $\square$

The most basic application of Theorem 4.5 (in its formulation with more than two components) is related to the n-dimensional Lebesgue measure. The next result is no surprise, since the n-dimensional Lebesgue measure of any interval in $\mathbb{R}^n$ is equal to the product of the 1-dimensional Lebesgue measures of its edges: $m_n\big(\prod_{j=1}^n [a_j, b_j]\big) = \prod_{j=1}^n m_1([a_j, b_j])$.

Theorem 4.4. *Let $n = n_1 + \cdots + n_k$.*
(i) The Lebesgue measure space $(\mathbb{R}^n, \mathcal{L}_n, m_n)$ is the product measure space of the Borel measure spaces $(\mathbb{R}^{n_j}, \mathcal{B}_{n_j}, m_{n_j})$ and, at the same time, the product measure space of the Lebesgue measure spaces $(\mathbb{R}^{n_j}, \mathcal{L}_{n_j}, m_{n_j})$.
(ii) The Lebesgue measure space $(\mathbb{R}^n, \mathcal{L}_n, m_n)$ is the completion of both $(\mathbb{R}^n, \bigotimes_{j=1}^k \mathcal{B}_{n_j}, m_n) = (\mathbb{R}^n, \mathcal{B}_n, m_n)$ and $(\mathbb{R}^n, \bigotimes_{j=1}^k \mathcal{L}_{n_j}, m_n)$, of which the second is an extension of the first.

Proof. We know that $\bigotimes_{j=1}^k \mathcal{B}_{n_j} = \mathcal{B}_n$, that $(\mathbb{R}^{n_j}, \mathcal{L}_{n_j}, m_{n_j})$ is the completion of $(\mathbb{R}^{n_j}, \mathcal{B}_{n_j}, m_{n_j})$, and that m_{n_j} is a σ-finite measure. Hence, Theorem 4.5 implies that the Borel measure spaces $(\mathbb{R}^{n_j}, \mathcal{B}_{n_j}, m_{n_j})$ and the Lebesgue measure spaces $(\mathbb{R}^{n_j}, \mathcal{L}_{n_j}, m_{n_j})$ induce the same product measure space $(\mathbb{R}^n, \mathcal{S}_{\bigotimes_{j=1}^k m_{n_j}}, \bigotimes_{j=1}^k m_{n_j})$, and that this is the completion of both measure spaces $(\mathbb{R}^n, \mathcal{B}_n, \bigotimes_{j=1}^k m_{n_j})$ and $(\mathbb{R}^n, \bigotimes_{j=1}^k \mathcal{L}_{n_j}, \bigotimes_{j=1}^k m_{n_j})$, of which the second is an extension of the first.
Theorem 4.3 implies that $\big(\bigotimes_{j=1}^k m_{n_j}\big)(R) = \prod_{j=1}^k m_{n_j}(A_j)$ for every Borel measurable interval $R = \prod_{j=1}^k A_j$. In particular, $\big(\bigotimes_{j=1}^k m_{n_j}\big)(S) = \mathrm{vol}_n(S)$ for every bounded interval S in $\mathbb{R}^n$. Now Theorem 1.13 implies that $\bigotimes_{j=1}^k m_{n_j} = m_n$ on $\mathcal{B}_n$. Hence, $(\mathbb{R}^n, \mathcal{B}_n, \bigotimes_{j=1}^k m_{n_j}) = (\mathbb{R}^n, \mathcal{B}_n, m_n)$.
The proof finishes, since, by Theorem 1.13, $(\mathbb{R}^n, \mathcal{L}_n, m_n)$ is the completion of $(\mathbb{R}^n, \mathcal{B}_n, m_n)$. $\square$

It is, perhaps, surprising that, although the Lebesgue measure spaces $(\mathbb{R}^{n_j}, \mathcal{L}_{n_j}, m_{n_j})$ are complete, the product $(\mathbb{R}^n, \bigotimes_{j=1}^k \mathcal{L}_{n_j}, m_n)$ is *not* complete (when $k \geq 2$, of course).

Example. We consider any non-Lebesgue set $A \subseteq \mathbb{R}$, and the $E = A \times \{0\} \times \cdots \times \{0\} \subseteq \mathbb{R}^n$. We also consider the Lebesgue measurable interval $R = \mathbb{R} \times \{0\} \times \cdots \times \{0\} \subseteq \mathbb{R}^n$. We have that $E \subseteq R$ and $m_n(R) = m_1(\mathbb{R}) m_1(\{0\}) \cdots m_1(\{0\}) = 0$. If we assume that $(\mathbb{R}^n, \bigotimes_{j=1}^n \mathcal{L}_1, m_n)$ is complete, then we conclude that $E \in \bigotimes_{j=1}^n \mathcal{L}_1$. We now take $z = (0, \dots, 0) \in \mathbb{R}^{n-1}$ and, then, the section $E_z = A$ must belong to $\mathcal{L}_1$. This is not true, and so we arrive at a contradiction.

Exercises.

4.3 Multiple integrals.

The purpose of this section is to describe the mechanism which reduces the calculation of product measures of subsets of cartesian products and of integrals of functions defined on cartesian products to the calculation of the measures or, respectively, the integrals of their sections. The gain is obvious: the reduced calculations are performed over sets of lower dimension.

As in the previous section, for the sake of simplicity, we restrict to the case of two measure spaces.

We recall that, if $E \subseteq X_1 \times X_2$ and $x_1 \in X_1$, then the x_1-section of E is defined by

$$E_{x_1} = \{x_2 \in X_2 \,|\, (x_1, x_2) \in E\} \subseteq X_2.$$

Similarly, if $x_2 \in X_2$, then the x_2-section of E is defined by

$$E^{x_2} = \{x_1 \in X_1 \,|\, (x_1, x_2) \in E\} \subseteq X_1.$$

We do not write E_{x_2} in order to avoid confusion with E_{x_1}.

Theorem 4.5. *Let $(X_1, \mathcal{S}_1, \mu_1)$ and $(X_2, \mathcal{S}_2, \mu_2)$ be σ-finite and let $(X_1 \times X_2, \mathcal{S}_1 \otimes \mathcal{S}_2, \mu_1 \otimes \mu_2)$ be their restricted product measure space.*
If $E \in \mathcal{S}_1 \otimes \mathcal{S}_2$, then $E_{x_1} \in \mathcal{S}_2$ for every $x_1 \in X_1$ and $E^{x_2} \in \mathcal{S}_1$ for every $x_2 \in X_2$. Moreover, the function $x_1 \mapsto \mu_2(E_{x_1})$ is $\mathcal{S}_1$-measurable, the function $x_2 \mapsto \mu_1(E^{x_2})$ is $\mathcal{S}_2$-measurable, and

$$(\mu_1 \otimes \mu_2)(E) = \int_{X_1} \mu_2(E_{x_1})\, d\mu_1(x_1) = \int_{X_2} \mu_1(E^{x_2})\, d\mu_2(x_2).$$

Proof. We observe that the first statement, namely that $E_{x_1} \in \mathcal{S}_2$ for every $x_1 \in X_1$ and $E^{x_2} \in \mathcal{S}_1$ for every $x_2 \in X_2$, is a direct consequence of Theorem 4.1 and it holds without the assumption about the σ-finiteness of μ_1, μ_2.
We denote $\mathcal{N}$ the collection of all $E \in \mathcal{S}_1 \otimes \mathcal{S}_2$ which have all the properties in the conclusion of the theorem.
(a) Every measurable interval $R = A_1 \times A_2$ belongs to $\mathcal{N}$.
Indeed, $R_{x_1} = \emptyset$, if $x_1 \notin A_1$, and $R_{x_1} = A_2$, if $x_1 \in A_1$. Hence, $\mu_2(R_{x_1}) = \mu_2(A_2)\chi_{A_1}(x_1)$ for every $x_1 \in X_1$, and so the function $x_1 \mapsto \mu_2(R_{x_1})$ is $\mathcal{S}_1$-measurable. Moreover, we have

$$\int_{X_1} \mu_2(R_{x_1})\, d\mu_1(x_1) = \mu_2(A_2) \int_{X_1} \chi_{A_1}\, d\mu_1 = \mu_2(A_2)\mu_1(A_1) = (\mu_1 \otimes \mu_2)(R).$$

The same arguments hold for x_2-sections.
(b) Let $E_1, \dots, E_m \in \mathcal{N}$ be pairwise disjoint. Then $E = E_1 \cup \cdots \cup E_m \in \mathcal{N}$.
Indeed, from $E_{x_1} = (E_1)_{x_1} \cup \cdots \cup (E_m)_{x_1}$ for every $x_1 \in X_1$, we have that $E_{x_1} \in \mathcal{S}_2$ for every $x_1 \in X_1$ and $\mu_2(E_{x_1}) = \mu_2((E_1)_{x_1}) + \cdots + \mu_2((E_m)_{x_1})$ for every $x_1 \in X_1$. Then the function $x_1 \mapsto \mu_2(E_{x_1})$ is $\mathcal{S}_1$-measurable, and

$$\int_{X_1} \mu_2(E_{x_1})\, d\mu_1(x_1) = \sum_{j=1}^m \int_{X_1} \mu_2((E_j)_{x_1})\, d\mu_1(x_1) = \sum_{j=1}^m (\mu_1 \otimes \mu_2)(E_j) = (\mu_1 \otimes \mu_2)(E).$$

The same arguments hold for x_2-sections.
(c) Let $E_n \in \mathcal{N}$ for every $n \in \mathbb{N}$ and $E_n \uparrow E$. Then $E \in \mathcal{N}$.
Indeed, from $(E_n)_{x_1} \uparrow E_{x_1}$ for every $x_1 \in X_1$, we have that $E_{x_1} \in \mathcal{S}_2$ for every $x_1 \in X_1$. Continuity of μ_2 from below implies that $\mu_2((E_n)_{x_1}) \uparrow \mu_2(E_{x_1})$ for every $x_1 \in X_1$, and so the function $x_1 \mapsto \mu_2(E_{x_1})$ is $\mathcal{S}_1$-measurable. By continuity of $\mu_1 \otimes \mu_2$ from below and by the Monotone Convergence Theorem, we get $(\mu_1 \otimes \mu_2)(E) = \int_{X_1} \mu_2(E_{x_1})\, d\mu_1(x_1)$.
The same can be proved, symmetrically, for x_2-sections.
(d) Let $R = A_1 \times A_2$ be any measurable interval with $\mu_1(A_1) < +\infty$ and $\mu_2(A_2) < +\infty$, and let $\mathcal{N}_R$ be the collection of all sets $E \in \mathcal{S}_1 \otimes \mathcal{S}_2$ for which $E \cap R \in \mathcal{N}$.
If $E_n \in \mathcal{N}_R$ for all n and $E_n \downarrow E$, then $E \in \mathcal{N}_R$.
Indeed, $E_n \cap R \downarrow E \cap R$, and so $(E_n \cap R)_{x_1} \downarrow (E \cap R)_{x_1}$ for every $x_1 \in X_1$. Hence, $(E \cap R)_{x_1} \in \mathcal{S}_2$ for every $x_1 \in X_1$. Now, for every $x_1 \in X_1$ we have $(E_1 \cap R)_{x_1} \subseteq R_{x_1}$. Since either $R_{x_1} = A_2$ or $R_{x_1} = \emptyset$, and since $\mu_2(A_2) < +\infty$, by the continuity of μ_2 from above, we find $\mu_2\big((E_n \cap R)_{x_1}\big) \downarrow \mu_2\big((E \cap R)_{x_1}\big)$ for every $x_1 \in X_1$. Hence, the function $x_1 \mapsto \mu_2\big((E \cap R)_{x_1}\big)$ is $\mathcal{S}_1$-measurable. Now, from our calculations in (a) we have that $\mu_2\big((E_n \cap R)_{x_1}\big) \le \mu_2(R_{x_1}) = \mu_2(A_2)\chi_{A_1}(x_1)$. Since $\int_{X_1} \mu_2(A_2)\chi_{A_1}(x_1)\, d\mu_1(x_1) = \mu_1(A_1)\mu_2(A_2) < +\infty$, the Dominated Convergence Theorem implies that

$$\int_{X_1} \mu_2\big((E_n \cap R)_{x_1}\big)\, d\mu_1(x_1) \downarrow \int_{X_1} \mu_2\big((E \cap R)_{x_1}\big)\, d\mu_1(x_1).$$

We also have that $(\mu_1 \otimes \mu_2)(E_1 \cap R) \le (\mu_1 \otimes \mu_2)(R) = \mu_1(A_1)\mu_2(A_2) < +\infty$. Hence, continuity of $\mu_1 \otimes \mu_2$ from above implies that $(\mu_1 \otimes \mu_2)(E_n \cap R) \downarrow (\mu_1 \otimes \mu_2)(E \cap R)$. Therefore,

$$(\mu_1 \otimes \mu_2)(E \cap R) = \int_{X_1} \mu_2\big((E \cap R)_{x_1}\big)\, d\mu_1(x_1).$$

Since all arguments hold for x_2-sections as well, we get that $E \cap R \in \mathcal{N}$, and so $E \in \mathcal{N}_R$.
If $E_n \in \mathcal{N}_R$ for all n and $E_n \uparrow E$, then $E_n \cap R \uparrow E \cap R$, and the result of (c) implies that $E \in \mathcal{N}_R$.

Therefore, the collection $\mathcal{N}_R$ is a monotone class of subsets of $X_1 \times X_2$.
Now, let $E_1, \dots, E_m \in \mathcal{N}_R$ be pairwise disjoint and $E = E_1 \cup \cdots \cup E_m$. Then the result of (b) implies that $E \cap R = (E_1 \cap R) \cup \cdots \cup (E_m \cap R) \in \mathcal{N}$, and so $E \in \mathcal{N}_R$. From (a), we have that $\mathcal{N}_R$ contains all measurable intervals, and so $\mathcal{N}_R$ contains all elements of the algebra $\mathcal{A}$ of Proposition 4.3. Therefore, $\mathcal{N}_R$ includes the monotone class generated by $\mathcal{A}$, which, by Theorem 1.1, is the same as the σ-algebra generated by $\mathcal{A}$, namely $\mathcal{S}_1 \otimes \mathcal{S}_2$.
Thus, $E \cap R \in \mathcal{N}$ for all $E \in \mathcal{S}_1 \otimes \mathcal{S}_2$ and all measurable intervals $R = A_1 \times A_2$ with $\mu_1(A_1) < +\infty$ and $\mu_2(A_2) < +\infty$.
(e) Since μ_1 is σ-finite, there are $A_{1,n} \in \mathcal{S}_1$ so that $A_{1,n} \uparrow X_1$ and $\mu_1(A_{1,n}) < +\infty$ for every n. Similarly, there are $A_{2,n} \in \mathcal{S}_2$ so that $A_{2,n} \uparrow X_2$ and $\mu_2(A_{2,n}) < +\infty$ for every n. Now, we form the measurable intervals $R_n = A_{1,n} \times A_{2,n}$.
We consider any $E \in \mathcal{S}_1 \otimes \mathcal{S}_2$. From the result of (d), we have that $E_n = E \cap R_n \in \mathcal{N}$ for every n. Since $E_n \uparrow E$, the result of (c) implies $E \in \mathcal{N}$. □

Theorem 4.6. *Let $(X_1, \mathcal{S}_1, \mu_1)$ and $(X_2, \mathcal{S}_2, \mu_2)$ be σ-finite and let $(X_1 \times X_2, \mathcal{S}_{\mu_1 \otimes \mu_2}, \mu_1 \otimes \mu_2)$ be their product measure space.*
If $E \in \mathcal{S}_{\mu_1 \otimes \mu_2}$, then $E_{x_1} \in \overline{\mathcal{S}_2}$ for μ_1-a.e. $x_1 \in X_1$ and $E^{x_2} \in \overline{\mathcal{S}_1}$ for μ_2-a.e. $x_2 \in X_2$. Moreover, the μ_1-almost everywhere defined function $x_1 \mapsto \overline{\mu_2}(E_{x_1})$ is $\overline{\mathcal{S}_1}$-measurable, the μ_2-almost everywhere defined function $x_2 \mapsto \overline{\mu_1}(E^{x_2})$ is $\overline{\mathcal{S}_2}$-measurable, and

$$(\mu_1 \otimes \mu_2)(E) = \int_{X_1} \overline{\mu_2}(E_{x_1})\, d\overline{\mu_1}(x_1) = \int_{X_2} \overline{\mu_1}(E^{x_2})\, d\overline{\mu_2}(x_2).$$

Proof. Let $E \in \mathcal{S}_{\mu_1 \otimes \mu_2}$. Since, by Theorems 4.4 and 4.5, $(X_1 \times X_2, \mathcal{S}_{\mu_1 \otimes \mu_2}, \mu_1 \otimes \mu_2)$ is the completion of $(X_1 \times X_2, \mathcal{S}_1 \otimes \mathcal{S}_2, \mu_1 \otimes \mu_2)$, there are $A, M \in \mathcal{S}_1 \otimes \mathcal{S}_2$ so that $(\mu_1 \otimes \mu_2)(M) = 0$ and $E = A \cup F$ for some $F \subseteq M$.
Now, we apply Theorem 4.7 to A and M.
We have that $A_{x_1}, M_{x_1} \in \mathcal{S}_2$ for every $x_1 \in X_1$. Moreover, the function $x_1 \mapsto \mu_2(M_{x_1})$ is $\mathcal{S}_1$-measurable, and

$$\int_{X_1} \mu_2(M_{x_1})\, d\mu_1(x_1) = (\mu_1 \otimes \mu_2)(M) = 0.$$

Hence, $\mu_2(M_{x_1}) = 0$ for μ_1-a.e. $x_1 \in X_1$. Since $E_{x_1} = A_{x_1} \cup F_{x_1}$ and $F_{x_1} \subseteq M_{x_1}$ for every $x_1 \in X_1$, we have that $E_{x_1} \in \overline{\mathcal{S}_2}$ and $\overline{\mu_2}(E_{x_1}) = \mu_2(A_{x_1})$ for μ_1-a.e. $x_1 \in X_1$.
Also, since the function $x_1 \mapsto \mu_2(A_{x_1})$ is $\mathcal{S}_1$-measurable, we have that the function $x_1 \mapsto \overline{\mu_2}(E_{x_1})$ is $\overline{\mathcal{S}_1}$-measurable.
Finally, by Theorem 4.7 again, we have

$$(\mu_1 \otimes \mu_2)(E) = (\mu_1 \otimes \mu_2)(A) = \int_{X_1} \mu_2(A_{x_1})\, d\mu_1(x_1) = \int_{X_1} \overline{\mu_2}(E_{x_1})\, d\overline{\mu_1}(x_1).$$

All these arguments hold for x_2-sections as well. □

Example. Let us think about the difference between Theorems 4.7 and 4.8.
We have $\mathbb{R}^n = \mathbb{R}^k \times \mathbb{R}^l$, where $n = k + l$.
We know that $\mathcal{B}_n = \mathcal{B}_k \otimes \mathcal{B}_l$, and let us consider any Borel set E in $\mathbb{R}^n$. Theorem 4.7 implies that *for every* $z \in \mathbb{R}^l$ the section E_z is a Borel set in $\mathbb{R}^k$, that the *everywhere* defined function $z \mapsto m_k(E_z)$ is Borel measurable, and that

$$m_n(E) = \int_{\mathbb{R}^l} m_k(E_z)\, dm_l(z).$$

On the other hand, we cannot have the same result for Lebesgue sets, because $\mathcal{L}_n \supsetneq \mathcal{L}_k \otimes \mathcal{L}_l$. The relevant fact is that $\mathcal{L}_n = \mathcal{S}_{m_k \otimes m_l}$ (i.e. $\mathcal{L}_n$ is the completion of $\mathcal{L}_k \otimes \mathcal{L}_l$). So let us consider any Lebesgue set E in $\mathbb{R}^n$. Theorem 4.8 implies that *for m_l-a.e. every* $z \in \mathbb{R}^l$ the section E_z is a Lebesgue set in $\mathbb{R}^k$, that the m_l-almost everywhere defined function $z \mapsto m_k(E_z)$ is Lebesgue measurable, and that

$$m_n(E) = \int_{\mathbb{R}^l} m_k(E_z)\, dm_l(z).$$

We recall that, if $f : X_1 \times X_2 \to Y$ and $x_1 \in X_1$, $x_2 \in X_2$, then the corresponding sections $f_{x_1} : X_2 \to Y$, $f^{x_2} : X_1 \to Y$ of f are defined by

$$f_{x_1}(x_2) = f(x_1, x_2) \text{ for } x_2 \in X_2, \quad f^{x_2}(x_1) = f(x_1, x_2) \text{ for } x_1 \in X_1.$$

Theorem 4.7. *Let $(X_1, \mathcal{S}_1, \mu_1)$ and $(X_2, \mathcal{S}_2, \mu_2)$ be σ-finite and let $(X_1 \times X_2, \mathcal{S}_1 \otimes \mathcal{S}_2, \mu_1 \otimes \mu_2)$ be their restricted product measure space.*
(i) (**Tonelli's Theorem**) *If $f : X_1 \times X_2 \to [0, +\infty]$ is $\mathcal{S}_1 \otimes \mathcal{S}_2$-measurable, then f_{x_1} is $\mathcal{S}_2$-measurable for every $x_1 \in X_1$ and f^{x_2} is $\mathcal{S}_1$-measurable for every $x_2 \in X_2$. Moreover, the function $x_1 \mapsto \int_{X_2} f_{x_1}(x_2)\, d\mu_2(x_2)$ is $\mathcal{S}_1$-measurable, the function $x_2 \mapsto \int_{X_1} f^{x_2}(x_1)\, d\mu_1(x_1)$ is $\mathcal{S}_2$-measurable, and*

$$\begin{aligned}\int_X f(x)\, d(\mu_1 \otimes \mu_2)(x) &= \int_{X_1} \big(\int_{X_2} f(x_1, x_2)\, d\mu_2(x_2)\big)\, d\mu_1(x_1) \\ &= \int_{X_2} \big(\int_{X_1} f(x_1, x_2)\, d\mu_1(x_1)\big)\, d\mu_2(x_2).\end{aligned}$$

(ii) (**Fubini's Theorem**) *If $f : X_1 \times X_2 \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ is $\mathcal{S}_1 \otimes \mathcal{S}_2$-measurable and $\mu_1 \otimes \mu_2$-integrable, then f_{x_1} is $\mathcal{S}_2$-measurable for every $x_1 \in X_1$ and μ_2-integrable for μ_1-a.e. $x_1 \in X_1$, and f^{x_2} is $\mathcal{S}_1$-measurable for every $x_2 \in X_2$ and μ_1-integrable for μ_2-a.e. $x_2 \in X_2$. Moreover, the μ_1-almost everywhere defined function $x_1 \mapsto \int_{X_2} f_{x_1}(x_2)\, d\mu_2(x_2)$ is $\mathcal{S}_1$-measurable and μ_1-integrable, the μ_2-almost everywhere defined function $x_2 \mapsto \int_{X_1} f^{x_2}(x_1)\, d\mu_1(x_1)$ is $\mathcal{S}_2$-measurable and μ_2-integrable, and the equalities in (i) are true.*

Proof. We observe that in both (i) and (ii) the measurability of the sections f_{x_1} and f^{x_2} is an immediate application of Theorem 4.2 and does not need the assumption about σ-finiteness.
(i) We consider the characteristic function χ_E of an $E \in \mathcal{S}_1 \otimes \mathcal{S}_2$.
Theorem 4.7 implies that $(\chi_E)_{x_1} = \chi_{E_{x_1}}$ is $\mathcal{S}_2$-measurable for every $x_1 \in X_1$ and the function

$$x_1 \mapsto \int_{X_2} (\chi_E)_{x_1}(x_2)\, d\mu_2(x_2) = \int_{X_2} \chi_{E_{x_1}}(x_2)\, d\mu_2(x_2) = \mu_2(E_{x_1})$$

is $\mathcal{S}_1$-measurable. Finally, we have

$$\begin{aligned}\int_X \chi_E(x)\, d(\mu_1 \otimes \mu_2)(x) = (\mu_1 \otimes \mu_2)(E) &= \int_{X_1} \mu_2(E_{x_1})\, d\mu_1(x_1) \\ &= \int_{X_1} \big(\int_{X_2} (\chi_E)_{x_1}(x_2)\, d\mu_2(x_2)\big)\, d\mu_1(x_1) \\ &= \int_{X_1} \big(\int_{X_2} \chi_E(x_1, x_2)\, d\mu_2(x_2)\big)\, d\mu_1(x_1).\end{aligned}$$

The argument for x_2-sections is the same.
Now, we consider a $\mathcal{S}_1 \otimes \mathcal{S}_2$-measurable simple $\phi : X_1 \times X_2 \to [0, +\infty)$ with standard representation $\phi = \sum_{j=1}^m \kappa_j \chi_{E_j}$. Then $\phi_{x_1} = \sum_{j=1}^m \kappa_j (\chi_{E_j})_{x_1}$ for every $x_1 \in X_1$. By the results in the case of a single characteristic function, we get that ϕ_{x_1} is $\mathcal{S}_2$-measurable for every $x_1 \in X_1$ and that the function

$$x_1 \mapsto \int_{X_2} \phi_{x_1}(x_2)\, d\mu_2(x_2) = \sum_{j=1}^m \kappa_j \int_{X_2} (\chi_{E_j})_{x_1}(x_2)\, d\mu_2(x_2)$$

is $\mathcal{S}_1$-measurable. Also,

$$\begin{aligned}\int_X \phi(x)\, d(\mu_1 \otimes \mu_2)(x) &= \sum_{j=1}^m \kappa_j \int_X \chi_{E_j}(x)\, d(\mu_1 \otimes \mu_2)(x) \\ &= \sum_{j=1}^m \kappa_j \int_{X_1} \big(\int_{X_2} \chi_{E_j}(x_1, x_2)\, d\mu_2(x_2)\big)\, d\mu_1(x_1) \\ &= \int_{X_1} \big(\int_{X_2} \phi(x_1, x_2)\, d\mu_2(x_2)\big)\, d\mu_1(x_1).\end{aligned}$$

The argument for x_2-sections is the same.
Finally, we consider a $\mathcal{S}_1 \otimes \mathcal{S}_2$-measurable $f : X_1 \times X_2 \to [0, +\infty]$. Then there are $\mathcal{S}_1 \otimes \mathcal{S}_2$-measurable simple $\phi_n : X_1 \times X_2 \to [0, +\infty]$ so that $\phi_n \uparrow f$ on $X_1 \times X_2$. From our results so far, every ϕ_n satisfies the conclusions of (i), and, since $(\phi_n)_{x_1} \uparrow f_{x_1}$ for every $x_1 \in X_1$ and

$(\phi_n)^{x_2} \uparrow f^{x_2}$ for every $x_2 \in X_2$, an application of the Monotone Convergence Theorem implies that f also satisfies the conclusions of (i).

(ii) If $f : X_1 \times X_2 \to [0, +\infty]$ is $\mathcal{S}_1 \otimes \mathcal{S}_2$-measurable and $\mu_1 \otimes \mu_2$-integrable, then (i) implies

$$\begin{aligned}\int_{X_1} \big(\int_{X_2} f(x_1, x_2)\, d\mu_2(x_2)\big)\, d\mu_1(x_1) &= \int_{X_2} \big(\int_{X_1} f(x_1, x_2)\, d\mu_1(x_1)\big)\, d\mu_2(x_2) \\ &= \int_X f(x)\, d(\mu_1 \otimes \mu_2)(x) < +\infty.\end{aligned}$$

Hence, $\int_{X_2} f_{x_1}(x_2)\, d\mu_2(x_2) < +\infty$ for μ_1-a.e. $x_1 \in X_1$ and $\int_{X_1} f^{x_2}(x_1)\, d\mu_1(x_1) < +\infty$ for μ_2-a.e. $x_2 \in X_2$. Therefore, the conclusion of the theorem is true for non-negative functions.

If $f : X_1 \times X_2 \to \overline{\mathbb{R}}$ is $\mathcal{S}_1 \otimes \mathcal{S}_2$-measurable and $\mu_1 \otimes \mu_2$-integrable, the same is true for f^+ and f^- and, by the result for non-negative functions, the conclusion is true for these two functions. Since $f_{x_1} = (f^+)_{x_1} - (f^-)_{x_1}$ for every $x_1 \in X_1$ and $f^{x_2} = (f^+)^{x_2} - (f^-)^{x_2}$ for every $x_2 \in X_2$, the conclusion is, by linearity, true also for extended real valued functions.

If $f : X_1 \times X_2 \to \mathbb{C}$ is $\mathcal{S}_1 \otimes \mathcal{S}_2$-measurable and $\mu_1 \otimes \mu_2$-integrable, the same is true for $\mathrm{Re}(f)$ and $\mathrm{Im}(f)$. By the result for real valued functions, the conclusion is true for $\mathrm{Re}(f)$ and $\mathrm{Im}(f)$. Since $f_{x_1} = \mathrm{Re}(f)_{x_1} + i\,\mathrm{Im}(f)_{x_1}$ for every $x_1 \in X_1$ and $f^{x_2} = \mathrm{Re}(f)^{x_2} + i\,\mathrm{Im}(f)^{x_2}$ for every $x_2 \in X_2$, the conclusion is, by linearity, true also for complex valued functions.

Finally, let $f : X_1 \times X_2 \to \overline{\mathbb{C}}$ be $\mathcal{S}_1 \otimes \mathcal{S}_2$-measurable and $\mu_1 \otimes \mu_2$-integrable. Then the set $E = f^{-1}(\{\infty\}) \in \mathcal{S}_1 \otimes \mathcal{S}_2$ has $(\mu_1 \otimes \mu_2)(E) = 0$. Theorem 4.7 implies that $\mu_2(E_{x_1}) = 0$ for μ_1-a.e. $x_1 \in X_1$ and $\mu_1(E^{x_2}) = 0$ for μ_2-a.e. $x_2 \in X_2$.

If we define $F = f\chi_{E^c}$, then $F : X_1 \times X_2 \to \mathbb{C}$ is $\mathcal{S}_1 \otimes \mathcal{S}_2$-measurable and $\mu_1 \otimes \mu_2$-integrable, and so, by the result for complex valued functions, the conclusion of the theorem holds for F. Since $F = f$ holds $(\mu_1 \otimes \mu_2)$-a.e. on $X_1 \times X_2$, we have

$$\int_{X_1 \times X_2} F(x)\, d(\mu_1 \otimes \mu_2)(x) = \int_{X_1 \times X_2} f(x)\, d(\mu_1 \otimes \mu_2)(x).$$

We, also, have that $F_{x_1} = f_{x_1}$ on $X_2 \setminus E_{x_1}$, and so $F_{x_1} = f_{x_1}$ holds μ_2-a.e. on X_2 for μ_1-a.e. $x_1 \in X_1$. Therefore, f_{x_1} is μ_2-integrable and $\int_{X_2} f_{x_1}(x_2)\, d\mu_2(x_2) = \int_{X_2} F_{x_1}(x_2)\, d\mu_2(x_2)$ for μ_1-a.e. $x_1 \in X_1$. This implies

$$\int_{X_1} \big(\int_{X_2} f(x_1, x_2)\, d\mu_2(x_2)\big)\, d\mu_1(x_1) = \int_{X_1} \big(\int_{X_2} F(x_1, x_2)\, d\mu_2(x_2)\big)\, d\mu_1(x_1)$$

and, equating the corresponding integrals of F, we get

$$\int_X f(x)\, d(\mu_1 \otimes \mu_2)(x) = \int_{X_1} \big(\int_{X_2} f(x_1, x_2)\, d\mu_2(x_2)\big)\, d\mu_1(x_1).$$

The argument is the same for x_2-sections. □

Theorem 4.8. *Let $(X_1, \mathcal{S}_1, \mu_1)$ and $(X_2, \mathcal{S}_2, \mu_2)$ be σ-finite and let $(X_1 \times X_2, \mathcal{S}_{\mu_1 \otimes \mu_2}, \mu_1 \otimes \mu_2)$ be their product measure space.*

(i) (**Tonelli's Theorem**) *If $f : X_1 \times X_2 \to [0, +\infty]$ is $\mathcal{S}_{\mu_1 \otimes \mu_2}$-measurable, then f_{x_1} is $\overline{\mathcal{S}_2}$-measurable for μ_1-a.e. $x_1 \in X_1$ and f^{x_2} is $\overline{\mathcal{S}_1}$-measurable for μ_2-a.e. $x_2 \in X_2$. Moreover, the μ_1-almost everywhere defined function $x_1 \mapsto \int_{X_2} f_{x_1}(x_2)\, d\overline{\mu_2}(x_2)$ is $\overline{\mathcal{S}_1}$-measurable, the μ_2-almost everywhere defined function $x_2 \mapsto \int_{X_1} f^{x_2}(x_1)\, d\overline{\mu_1}(x_1)$ is $\overline{\mathcal{S}_2}$-measurable, and*

$$\begin{aligned}\int_{X_1 \times X_2} f(x)\, d(\mu_1 \otimes \mu_2)(x) &= \int_{X_1} \big(\int_{X_2} f(x_1, x_2)\, d\overline{\mu_2}(x_2)\big)\, d\overline{\mu_1}(x_1) \\ &= \int_{X_2} \big(\int_{X_1} f(x_1, x_2)\, d\overline{\mu_1}(x_1)\big)\, d\overline{\mu_2}(x_2).\end{aligned}$$

(ii) (**Fubini's Theorem**) *If $f : X_1 \times X_2 \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ is $\mathcal{S}_{\mu_1 \otimes \mu_2}$-measurable and $\mu_1 \otimes \mu_2$-integrable, then f_{x_1} is $\overline{\mathcal{S}_2}$-measurable and $\overline{\mu_2}$-integrable for μ_1-a.e. $x_1 \in X_1$, and f^{x_2} is $\overline{\mathcal{S}_1}$-measurable and $\overline{\mu_1}$-integrable for μ_2-a.e. $x_2 \in X_2$. Moreover, the μ_1-almost everywhere defined function $x_1 \mapsto \int_{X_2} f_{x_1}(x_2)\, d\overline{\mu_2}(x_2)$ is $\overline{\mathcal{S}_1}$-measurable and $\overline{\mu_1}$-integrable, the μ_2-almost everywhere defined function $x_2 \mapsto \int_{X_1} f^{x_2}(x_1)\, d\overline{\mu_1}(x_1)$ is $\overline{\mathcal{S}_2}$-measurable and $\overline{\mu_2}$-integrable, and the equalities in (i) are true.*

Proof. The proof follows the same line as the proof of Theorem 4.9, starting with the characteristic function χ_E of a set $E \in \mathcal{S}_{\mu_1 \otimes \mu_2}$ and using, for this case, Theorem 4.8 instead of Theorem 4.7. The rest of the proof is the same, with only minor modifications, which are left to the reader as an exercise. $\square$

Example. Let us think about the difference between Theorems 4.9 and 4.10, based on the equality $\mathbb{R}^n = \mathbb{R}^k \times \mathbb{R}^l$, where $n = k + l$, as we did for the difference between Theorems 4.7 and 4.8.
We know that $\mathcal{B}_n = \mathcal{B}_k \otimes \mathcal{B}_l$, and let us consider any Borel measurable $f : \mathbb{R}^n \to \overline{\mathbb{C}}$. Theorem 4.9 implies that the section $f_z : \mathbb{R}^k \to \overline{\mathbb{C}}$ is Borel measurable for every $z \in \mathbb{R}^l$ and Lebesgue integrable for m_l-a.e. $z \in \mathbb{R}^l$, that the m_l-almost everywhere defined function $z \mapsto \int_{\mathbb{R}^k} f_z(y)\, dm_k(y)$ is Borel measurable and Lebesgue integrable, and that

$$\int_{\mathbb{R}^n} f(x)\, dm_n(x) = \int_{\mathbb{R}^l} \big(\int_{\mathbb{R}^k} f(y,z)\, dm_k(y) \big)\, dm_l(z).$$

On the other hand, we cannot have the same result for Lebesgue measurable $f : \mathbb{R}^n \to \overline{\mathbb{C}}$. Since $\mathcal{L}_n = \mathcal{S}_{m_k \otimes m_l}$ (i.e. $\mathcal{L}_n$ is the completion of $\mathcal{L}_k \otimes \mathcal{L}_l$), Theorem 4.10 implies that the section $f_z : \mathbb{R}^k \to \overline{\mathbb{C}}$ is Lebesgue measurable and Lebesgue integrable for m_l-a.e. $z \in \mathbb{R}^l$, that the m_l-almost everywhere defined function $z \mapsto \int_{\mathbb{R}^k} f_z(y)\, dm_k(y)$ is Lebesgue measurable and Lebesgue integrable, and that

$$\int_{\mathbb{R}^n} f(x)\, dm_n(x) = \int_{\mathbb{R}^l} \big(\int_{\mathbb{R}^k} f(y,z)\, dm_k(y) \big)\, dm_l(z).$$

The power of the Theorems of Tonelli and of Fubini lies in the resulting *successive integration formula* for the calculation of integrals over product spaces and in the *interchange of successive integrations*. The function f to which we may want to apply Fubini's Theorem must be $\mathcal{S}_{\mu_1 \otimes \mu_2}$-measurable and $\mu_1 \otimes \mu_2$-integrable. The Theorem of Tonelli is applied to non-negative functions f which must be $\mathcal{S}_{\mu_1 \otimes \mu_2}$-measurable. Thus, the assumptions of the Theorem of Tonelli are, except for the sign, weaker than the assumptions of the Theorem of Fubini.

The strategy, in order to calculate the integral of f over the product space by means of successive integrations or in order to interchange successive integrations, consists of three steps. The first is to prove that f is $\mathcal{S}_{\mu_1 \otimes \mu_2}$-measurable. The second step is to apply Tonelli's Theorem to $|f|$ to get

$$\begin{aligned} \int_{X_1 \times X_2} |f(x)|\, d(\mu_1 \otimes \mu_2)(x) &= \int_{X_1} \big(\int_{X_2} |f(x_1,x_2)|\, d\overline{\mu_2}(x_2) \big)\, d\overline{\mu_1}(x_1) \\ &= \int_{X_2} \big(\int_{X_1} |f(x_1,x_2)|\, d\overline{\mu_1}(x_1) \big)\, d\overline{\mu_2}(x_2). \end{aligned}$$

We need to estimate one of the two successive integrals, to see if $\int_{X_1 \times X_2} |f(x)|\, d(\mu_1 \otimes \mu_2)(x)$ is finite. If this is true, i.e. if f is $\mu_1 \otimes \mu_2$-integrable, then we take the third step: we apply Fubini's Theorem to find the desired

$$\begin{aligned} \int_{X_1 \times X_2} f(x)\, d(\mu_1 \otimes \mu_2)(x) &= \int_{X_1} \big(\int_{X_2} f(x_1,x_2)\, d\overline{\mu_2}(x_2) \big)\, d\overline{\mu_1}(x_1) \\ &= \int_{X_2} \big(\int_{X_1} f(x_1,x_2)\, d\overline{\mu_1}(x_1) \big)\, d\overline{\mu_2}(x_2) \end{aligned}$$

by calculating one of the two successive integrals.

Of the three steps the first, namely proving the $\mathcal{S}_{\mu_1 \otimes \mu_2}$-measurability of f, is more subtle and sometimes difficult to do.

Exercises.

4.3.1. Consider the measure spaces $(\mathbb{R}, \mathcal{B}_1, m_1)$ and $(\mathbb{R}, \mathcal{P}(\mathbb{R}), \sharp)$, where $\sharp$ is the counting measure. If $E = \{(x_1,x_2) \,|\, 0 \leq x_1 = x_2 \leq 1\}$, prove that all numbers $(m_1 \otimes \sharp)(E)$, $\int_{\mathbb{R}} \sharp(E_{x_1})\, dm_1(x_1)$ and $\int_{\mathbb{R}} m_1(E_{x_2})\, d\sharp(x_2)$ are different.

4.3.2. Consider $a_{m,n} = 1$ if $m = n$, $a_{m,n} = -1$ if $m = n+1$ and $a_{m,n} = 0$ in any other case. Then $\sum_{n=1}^{+\infty} \big(\sum_{m=1}^{+\infty} a_{m,n} \big) \neq \sum_{m=1}^{+\infty} \big(\sum_{n=1}^{+\infty} a_{m,n} \big)$. Explain, through the Theorem of Fubini.

4.3.3. Suppose that $(X, \mathcal{S}, \mu)$ is a measure space and $f : X \to [0, +\infty]$ is $\mathcal{S}$-measurable. Consider $A_f = \{(x, y) \in X \times \mathbb{R} \,|\, 0 \leq y < f(x)\}$ and $G_f = \{(x, y) \in X \times \mathbb{R} \,|\, y = f(x)\}$ and prove that both A_f and G_f are $\mathcal{S} \otimes \mathcal{B}_1$-measurable. If, moreover, μ is σ-finite, prove that $(\mu \otimes m_1)(A_f) = \int_X f \, d\mu$ and $(\mu \otimes m_1)(G_f) = 0$. A special case appears in exercise 3.2.1.

4.3.4. Let $(X, \mathcal{S}, \mu)$ be a σ-finite measure space and $f : X \to [0, +\infty]$ be $\mathcal{S}$-measurable. Calculating the measure $\mu \otimes \mu_G$ of the set $A_f = \{(x, y) \in X \times \mathbb{R} \,|\, 0 \leq y < f(x)\}$, prove Proposition 3.14.

4.3.5. Consider measure spaces $(X_1, \mathcal{S}_1, \mu_1)$ and $(X_2, \mathcal{S}_2, \mu_2)$, a $\mathcal{S}_1$-measurable $f_1 : X_1 \to \mathbb{C}$ and a $\mathcal{S}_2$-measurable $f_2 : X_2 \to \mathbb{C}$. Consider the function $f : X_1 \times X_2 \to \mathbb{C}$ defined by $f(x_1, x_2) = f_1(x_1) f_2(x_2)$.
Prove that f is $\mathcal{S}_1 \otimes \mathcal{S}_2$-measurable.
If f_1 is integrable with respect to μ_1 and f_2 is integrable with respect to μ_2, prove that f is integrable with respect to $\mu_1 \otimes \mu_2$ and that $\int_{X_1 \times X_2} f \, d(\mu_1 \otimes \mu_2) = \int_{X_1} f_1 \, d\mu_1 \int_{X_2} f_2 \, d\mu_2$.

4.3.6. From $\int_0^n \frac{\sin x}{x} \, dx = \int_0^n \big(\int_0^{+\infty} e^{-xt} \, dt \big) \sin x \, dx$, prove that $\int_0^{\to +\infty} \frac{\sin x}{x} \, dx = \frac{\pi}{2}$. (See also exercises 3.2.2 and 3.2.18.)

4.3.7. Let $f, g : \mathbb{R}^n \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ be $\mathcal{L}_n$-measurable.
(i) Prove that $H : \mathbb{R}^n \times \mathbb{R}^n \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ defined by $H(x, y) = f(x - y) g(y)$ is $\mathcal{L}_{2n}$-measurable.
Now, let f and g be integrable with respect to m_n.
(ii) Prove that H is integrable with respect to m_{2n} and $\int_{\mathbb{R}^{2n}} |H| \, dm_{2n} \leq \int_{\mathbb{R}^n} |f| \, dm_n \int_{\mathbb{R}^n} |g| \, dm_n$.
(iii) Prove that for m_n-a.e. $x \in \mathbb{R}^n$ the function $f(x - \cdot) g(\cdot)$ is integrable with respect to m_n.
The a.e. defined function $f * g : \mathbb{R}^n \to \mathbb{R}$ or $\mathbb{C}$ by $(f * g)(x) = \int_{\mathbb{R}^n} f(x - y) g(y) \, dm_n(y)$ is called the **convolution** of f and g.
(iv) Prove that $f * g$ is integrable with respect to m_n, that $\int_{\mathbb{R}^n} (f * g) \, dm_n = \int_{\mathbb{R}^n} f \, dm_n \int_{\mathbb{R}^n} g \, dm_n$ and $\int_{\mathbb{R}^n} |f * g| \, dm_n \leq \int_{\mathbb{R}^n} |f| \, dm_n \int_{\mathbb{R}^n} |g| \, dm_n$.
(v) Prove that, for every f, g, h, f_1, f_2 which are Lebesgue integrable, we have m_n-a.e. on $\mathbb{R}^n$ that $f * g = g * f$, $(f * g) * h = f * (g * h)$, $(\lambda f) * g = \lambda (f * g)$ and $(f_1 + f_2) * g = f_1 * g + f_2 * g$.
(vi) Prove that $\widehat{f * g} = \widehat{f} \, \widehat{g}$, where $\widehat{f}$ is the Fourier transform of f (exercise 3.2.13).

4.3.8. Let K be a set in $[0, 1]$ of the type considered in exercise 1.4.14 with $m_1(K) > 0$. Prove that $\{(x, y) \in [0, 1] \times [0, 1] \,|\, x - y \in K\}$ is a compact subset of $\mathbb{R}^2$ with positive m_2-measure, which does not contain any measurable interval of positive m_2-measure.

4.3.9. Let μ and ν be two locally finite Borel measures on $\mathbb{R}^n$, which are translation invariant. Namely: $\mu(A + x) = \mu(A)$ and $\nu(A + x) = \nu(A)$ for every $x \in \mathbb{R}^n$ and every $A \in \mathcal{B}_n$.
Working with $\int_{\mathbb{R}^n \times \mathbb{R}^n} \chi_A(x) \chi_B(x + y) \, d(\mu \otimes \nu)(x, y)$, prove that either $\mu = \lambda \nu$ or $\nu = \lambda \mu$ for some $\lambda \in [0, +\infty)$.
Conclude that the only translation invariant locally finite Borel measure on $\mathbb{R}^n$ which has value 1 at the unit cube $[0, 1]^n$ is the Lebesgue measure m_n.

4.3.10. Let $E \subseteq [0, 1] \times [0, 1]$ have the property that every horizontal section E_y is countable and every vertical section E_x has countable complementary set $[0, 1] \setminus E_x$. Prove that E is not Lebesgue measurable.

4.3.11. Let $(X, \mathcal{S}_X, \mu)$ be a measure space and $(Y, \mathcal{S}_Y)$ be a measurable space. Suppose that for every $x \in X$ there is a measure ν_x on $(Y, \mathcal{S}_Y)$ so that for every $B \in \mathcal{S}_Y$ the function $x \mapsto \nu_x(B)$ is $\mathcal{S}_X$-measurable.
We define $\nu(B) = \int_X \nu_x(B) \, d\mu(x)$ for every $B \in \mathcal{S}_Y$.
(i) Prove that ν is a measure on $(Y, \mathcal{S}_Y)$.
(ii) If $g : Y \to [0, +\infty]$ is $\mathcal{S}_Y$-measurable and if $f(x) = \int_Y g \, d\nu_x$ for every $x \in X$, prove that f is $\mathcal{S}_X$-measurable and $\int_X f \, d\mu = \int_Y g \, d\nu$.

4.3.12. If I_1, I_2 are two sets of indices with their counting measures, prove that the product measure on $I_1 \times I_2$ is its counting measure. Applying the Theorems of Tonelli and Fubini, derive results about the validity of $\sum_{i_1\in I_1, i_2\in I_2} c_{i_1,i_2} = \sum_{i_1\in I_1}(\sum_{i_2\in I_2} c_{i_1,i_2}) = \sum_{i_2\in I_2}(\sum_{i_1\in I_1} c_{i_1,i_2})$.

4.3.13. Consider the interval $R = (a,b] \times (a,b]$, and partition it into $\Delta_1 = \{(t,s) \in R \,|\, t \leq s\}$ and $\Delta_2 = \{(t,s) \in R \,|\, s < t\}$. Writing $(\mu_G \otimes \mu_F)(R) = (\mu_G \otimes \mu_F)(\Delta_1) + (\mu_G \otimes \mu_F)(\Delta_2)$, prove Proposition 3.11.

4.3.14. (i) Prove Theorem 4.8 replacing the σ-finiteness of $(X_1, \mathcal{S}_1, \mu_1)$ and $(X_2, \mathcal{S}_2, \mu_2)$ with the σ-finiteness of $E \in \mathcal{S}_{\mu_1\otimes\mu_2}$.
(ii) In Theorem 4.10, prove Tonelli's Theorem replacing the σ-finiteness of $(X_1, \mathcal{S}_1, \mu_1)$ and of $(X_2, \mathcal{S}_2, \mu_2)$ with the assumption that the set $f^{-1}((0,+\infty])$ has σ-finite $(\mu_1 \otimes \mu_2)$-measure.
Also, prove Fubini's Theorem omitting the σ-finiteness of $(X_1, \mathcal{S}_1, \mu_1)$ and of $(X_2, \mathcal{S}_2, \mu_2)$.

4.4 Surface measure on $\mathbb{S}^{n-1}$.

For every $x = (x_1, \ldots, x_n) \in \mathbb{R}^n_* = \mathbb{R}^n \setminus \{0\}$ we write

$$r = \|x\| = (x_1^2 + \cdots + x_n^2)^{1/2} \in \mathbb{R}^+ = (0,+\infty), \quad y = \tfrac{x}{\|x\|} \in \mathbb{S}^{n-1},$$

where $\mathbb{S}^{n-1} = \{y \in \mathbb{R}^n \,|\, \|y\| = 1\}$ is the **unit sphere** of $\mathbb{R}^n$.

The mapping $\Phi : \mathbb{R}^n_* \to \mathbb{R}^+ \times \mathbb{S}^{n-1}$, defined by

$$\Phi(x) = (r,y) = \big(\|x\|, \tfrac{x}{\|x\|}\big),$$

is one-to-one and onto, and its inverse $\Phi^{-1} : \mathbb{R}^+ \times \mathbb{S}^{n-1} \to \mathbb{R}^n_*$ is given by

$$\Phi^{-1}(r,y) = x = ry.$$

The elements $r = \|x\|$ and $y = \frac{x}{\|x\|}$ are called the **polar coordinates** of x and the mappings Φ and Φ^{-1} determine an *identification* of $\mathbb{R}^n_*$ with the cartesian product $\mathbb{R}^+ \times \mathbb{S}^{n-1}$, where every point $x \neq 0$ is identified with the pair (r,y) of its polar coordinates.

As usual, we consider $\mathbb{S}^{n-1}$ as a metric subspace of $\mathbb{R}^n$. This means that the distance between points $y', y'' \in \mathbb{S}^{n-1}$ is their Euclidean distance $\|y' - y''\|$ considered as points of the larger space $\mathbb{R}^n$. The open ball in $\mathbb{S}^{n-1}$ with center $y \in \mathbb{S}^{n-1}$ and radius $\delta > 0$ is the *spherical cap* $S(y;\delta) = \{y' \in \mathbb{S}^{n-1} \,|\, \|y' - y\| < \delta\}$, which is the intersection with $\mathbb{S}^{n-1}$ of the Euclidean ball $B(y;\delta) = \{x \in \mathbb{R}^n \,|\, \|x - y\| < \delta\}$. In fact, the intersection with $\mathbb{S}^{n-1}$ of an arbitrary Euclidean open ball in $\mathbb{R}^n$ is, if non-empty, a spherical cap of $\mathbb{S}^{n-1}$.

It is easy to see that there is a countable collection of spherical caps with the property that every open set in $\mathbb{S}^{n-1}$ is a union (countable, necessarily) of spherical caps from this collection. Indeed, such is the collection of the (non-empty) intersections with $\mathbb{S}^{n-1}$ of all open balls in $\mathbb{R}^n$ with rational centers (i.e. points in $\mathbb{R}^n$ with all their coordinates being rational) and rational radii.

If we equip $\mathbb{R}^+ \times \mathbb{S}^{n-1}$ with the *product metric* $d((r,y),(r',y')) = \max\{|r - r'|, \|y - y'\|\}$, then $\Phi : \mathbb{R}^n_* \to \mathbb{R}^+ \times \mathbb{S}^{n-1}$ and $\Phi^{-1} : \mathbb{R}^+ \times \mathbb{S}^{n-1} \to \mathbb{R}^n_*$ are both continuous.

Proposition 4.7 contains information about the Borel structures of $\mathbb{R}^n_*$ and of $\mathbb{R}^+$, $\mathbb{S}^{n-1}$ and their product $\mathbb{R}^+ \times \mathbb{S}^{n-1}$.

Proposition 4.7. *(i)* $\mathcal{B}_{\mathbb{R}^+\times\mathbb{S}^{n-1}} = \mathcal{B}_{\mathbb{R}^+} \otimes \mathcal{B}_{\mathbb{S}^{n-1}}$.
(ii) $\Phi(E)$ is a Borel set in $\mathbb{R}^+ \times \mathbb{S}^{n-1}$ for every Borel set E in $\mathbb{R}^n_*$, and $\Phi^{-1}(F)$ is a Borel set in $\mathbb{R}^n_*$ for every Borel set F in $\mathbb{R}^+ \times \mathbb{S}^{n-1}$.
(iii) $MA = \{ry \,|\, r \in M, y \in A\}$ is a Borel set in $\mathbb{R}^n_*$ for every Borel set M in $\mathbb{R}^+$ and every Borel set A in $\mathbb{S}^{n-1}$.

Proof. (i) A corollary of Proposition 4.2.
(ii) Since Φ is continuous, it is $(\mathcal{B}_{\mathbb{R}^n_*}, \mathcal{B}_{\mathbb{R}^+\times\mathbb{S}^{n-1}})$-measurable. Hence, $\Phi^{-1}(F)$ is a Borel set in $\mathbb{R}^n_*$ for every Borel set F in $\mathbb{R}^+ \times \mathbb{S}^{n-1}$. Similarly, the second statement is a consequence of the continuity of Φ^{-1}.
(iii) Let M be a Borel set in $\mathbb{R}^+$ and A be a Borel set in $\mathbb{S}^{n-1}$. Then $M\times A$ is a Borel set (measurable interval) in $\mathbb{R}^+ \times \mathbb{S}^{n-1}$. From (ii) we have that $MA = \Phi^{-1}(M \times A)$ is a Borel set in $\mathbb{R}^n_*$. □

A set $\Gamma \subseteq \mathbb{R}^n_*$ is called a **positive cone** if $rx \in \Gamma$ for every $r \in \mathbb{R}^+$ and every $x \in \Gamma$ or, equivalently, if Γ is closed under multiplication by positive numbers or, equivalently, if Γ is invariant under dilations. If $B \subseteq \mathbb{R}^n_*$, then the set $\mathbb{R}^+B = \{rb \,|\, r \in \mathbb{R}^+, b \in B\}$ is, obviously, a positive cone and it is called the positive cone **determined by** B. It is easy to see that, if Γ is a positive cone and $A = \Gamma \cap \mathbb{S}^{n-1}$, then Γ is the positive cone determined by A, and, conversely, if $A \subseteq \mathbb{S}^{n-1}$ and Γ is the positive cone determined by A, then $\Gamma \cap \mathbb{S}^{n-1} = A$. In other words there is a one-to-one correspondence between subsets of $\mathbb{S}^{n-1}$ and positive cones of $\mathbb{R}^n$.

The next result expresses a simple characterization of open and of Borel subsets of $\mathbb{S}^{n-1}$ in terms of the corresponding positive cones.

Proposition 4.8. *Let $A \subseteq \mathbb{S}^{n-1}$.*
(i) A is open in $\mathbb{S}^{n-1}$ if and only if the cone $\mathbb{R}^+A$ is open in $\mathbb{R}^n_*$.
(ii) A is a Borel set in $\mathbb{S}^{n-1}$ if and only if $\mathbb{R}^+A$ is a Borel set in $\mathbb{R}^n_*$ if and only if $(0,1]A$ is a Borel set in $\mathbb{R}^n_*$.

Proof. (i) Let A be open in $\mathbb{S}^{n-1}$. Then $\mathbb{R}^+ \times A$ is open in $\mathbb{R}^+ \times \mathbb{S}^{n-1}$. Now, the continuity of Φ implies that $\mathbb{R}^+A = \Phi^{-1}(\mathbb{R}^+ \times A)$ is an open set in $\mathbb{R}^n_*$.
Conversely, if $\mathbb{R}^+A$ is open in $\mathbb{R}^n_*$, then $A = (\mathbb{R}^+A) \cap \mathbb{S}^{n-1}$ is open in $\mathbb{S}^{n-1}$.
(ii) If A is a Borel set in $\mathbb{S}^{n-1}$, Proposition 4.5 implies that $\mathbb{R}^+A$ and $(0,1]A$ are Borel sets in $\mathbb{R}^n_*$. Conversely, if either $\mathbb{R}^+A$ is a Borel set in $\mathbb{R}^n_*$ or $(0,1]A$ is a Borel set in $\mathbb{R}^n_*$, then $A = (\mathbb{R}^+A) \cap \mathbb{S}^{n-1} = ((0,1]A) \cap \mathbb{S}^{n-1}$ is a Borel set in $\mathbb{S}^{n-1}$. □

Proposition 4.9. *If we define*
$$\sigma_{n-1}(A) = n\, m_n\big((0,1]A\big)$$
for every $A \in \mathcal{B}_{\mathbb{S}^{n-1}}$, then σ_{n-1} is a measure on $(\mathbb{S}^{n-1}, \mathcal{B}_{\mathbb{S}^{n-1}})$.

Proof. By Proposition 4.5 (or 4.6), for every Borel set A in $\mathbb{S}^{n-1}$ the set $(0,1]A$ is a Borel set in $\mathbb{R}^n_*$, and so $\sigma_{n-1}(A)$ is well defined.
We have $\sigma_{n-1}(\emptyset) = n\, m_n\big((0,1]\emptyset\big) = n\, m_n(\emptyset) = 0$.
Moreover, if $A_1, A_2, \ldots \in \mathcal{B}_{\mathbb{S}^{n-1}}$ are pairwise disjoint, then the sets $(0,1]A_1, (0,1]A_2, \ldots$ are also pairwise disjoint. Hence,
$$\begin{aligned}\sigma_{n-1}\big(\textstyle\bigcup_{j=1}^{+\infty} A_j\big) &= n\, m_n\big((0,1]\textstyle\bigcup_{j=1}^{+\infty} A_j\big) = n\, m_n\big(\textstyle\bigcup_{j=1}^{+\infty}((0,1]A_j)\big)\\ &= \textstyle\sum_{j=1}^{+\infty} n\, m_n\big((0,1]A_j\big) = \textstyle\sum_{j=1}^{+\infty} \sigma_{n-1}(A_j),\end{aligned}$$
and so σ_{n-1} is a measure on $(\mathbb{S}^{n-1}, \mathcal{B}_{\mathbb{S}^{n-1}})$. □

Definition. *The measure σ_{n-1} on $(\mathbb{S}^{n-1}, \mathcal{B}_{\mathbb{S}^{n-1}})$, which is defined in Proposition 4.7, is called the $(n-1)$-dimensional* ***surface measure*** *on $\mathbb{S}^{n-1}$.*

Lemma 4.1. *If we define*
$$\rho(N) = \textstyle\int_N r^{n-1}\, dr$$
for every $N \in \mathcal{B}_{\mathbb{R}^+}$, then ρ is a measure on $(\mathbb{R}^+, \mathcal{B}_{\mathbb{R}^+})$.

Proof. A simple consequence of Theorem 3.9. □

Lemma 4.2. *If we define $\widetilde{m_n}(F) = m_n(\Phi^{-1}(F))$ for every Borel set F in $\mathbb{R}^+ \times \mathbb{S}^{n-1}$, then $\widetilde{m_n}$ is a measure on the measurable space $(\mathbb{R}^+ \times \mathbb{S}^{n-1}, \mathcal{B}_{\mathbb{R}^+ \times \mathbb{S}^{n-1}})$.*

Proof. $\Phi^{-1}(F)$ is a Borel set in $\mathbb{R}^n_*$ for every Borel set F in $\mathbb{R}^+ \times \mathbb{S}^{n-1}$, and so $\widetilde{m_n}(F)$ is well defined.
Clearly, $\widetilde{m_n}(\emptyset) = m_n(\Phi^{-1}(\emptyset)) = m_n(\emptyset) = 0$.
If $F_1, F_2, \dots$ are pairwise disjoint, then $\Phi^{-1}(F_1), \Phi^{-1}(F_2), \dots$ are also pairwise disjoint, and

$$\begin{aligned}\widetilde{m_n}(\textstyle\bigcup_{j=1}^{+\infty} F_j) &= m_n(\Phi^{-1}(\textstyle\bigcup_{j=1}^{+\infty} F_j)) = m_n(\textstyle\bigcup_{j=1}^{+\infty} \Phi^{-1}(F_j)) = \textstyle\sum_{j=1}^{+\infty} m_n(\Phi^{-1}(F_j)) \\ &= \textstyle\sum_{j=1}^{+\infty} \widetilde{m_n}(F_j),\end{aligned}$$

and so $\widetilde{m_n}$ is a measure on $(\mathbb{R}^+ \times \mathbb{S}^{n-1}, \mathcal{B}_{\mathbb{R}^+ \times \mathbb{S}^{n-1}})$. □

Lemma 4.3. *On the measurable space $(\mathbb{R}^+ \times \mathbb{S}^{n-1}, \mathcal{B}_{\mathbb{R}^+ \times \mathbb{S}^{n-1}}) = (\mathbb{R}^+ \times \mathbb{S}^{n-1}, \mathcal{B}_{\mathbb{R}^+} \otimes \mathcal{B}_{\mathbb{S}^{n-1}})$ the measures $\widetilde{m_n}$ and $\rho \otimes \sigma_{n-1}$ are identical.*

Proof. The equality $\mathcal{B}_{\mathbb{R}^+ \times \mathbb{S}^{n-1}} = \mathcal{B}_{\mathbb{R}^+} \otimes \mathcal{B}_{\mathbb{S}^{n-1}}$ is in Proposition 4.5.
If A is a Borel set in $\mathbb{S}^{n-1}$, then the sets $(0,b]A$ and $(0,1]A$ are Borel sets in $\mathbb{R}^n_*$ and the first is a dilate of the second by the factor $b > 0$. Hence, $m_n((0,b]A) = b^n m_n((0,1]A)$ for every $b > 0$, and so

$$m_n((a,b]A) = m_n\big(((0,b]A) \setminus ((0,a]A)\big) = m_n((0,b]A) - m_n((0,a]A) = (b^n - a^n) m_n((0,1]A)$$

for every a, b with $0 \le a < b < +\infty$.
Therefore, if A is a Borel set in $\mathbb{S}^{n-1}$, then

$$\begin{aligned}\widetilde{m_n}((a,b] \times A) &= m_n(\Phi^{-1}((a,b] \times A)) = m_n((a,b]A) = (b^n - a^n) m_n((0,1]A) \\ &= \tfrac{b^n - a^n}{n}\, \sigma_{n-1}(A) = \textstyle\int_{(a,b]} r^{n-1}\, dr\ \sigma_{n-1}(A) = \rho((a,b])\ \sigma_{n-1}(A) \\ &= (\rho \otimes \sigma_{n-1})\ ((a,b] \times A).\end{aligned}$$

Now, we define

$$\mu(N) = \widetilde{m_n}(N \times A), \quad \nu(N) = (\rho \otimes \sigma_{n-1})(N \times A)$$

for every Borel set N in $\mathbb{R}^+$. It is easy to see that μ, ν are Borel measures on $\mathbb{R}^+$, and, by what we just proved, they satisfy $\mu((a,b]) = \nu((a,b])$ for every interval in $\mathbb{R}^+$. This, obviously, extends to all finite unions of pairwise disjoint open-closed subintervals of $\mathbb{R}^+$. Now, Theorem 1.7 implies that the two measures are equal on the σ-algebra generated by the collection of all these sets, which is equal to $\mathcal{B}_{\mathbb{R}^+}$. Therefore, $\widetilde{m_n}(N \times A) = (\rho \otimes \sigma_{n-1})(N \times A)$ for every Borel set N in $\mathbb{R}^+$ and every Borel set A in $\mathbb{S}^{n-1}$.
Finally, Theorem 4.4 implies the equality of the two measures, since ρ and σ_{n-1} are σ-finite. □

If $E \subseteq \mathbb{R}^n_*$, we consider the set $\Phi(E) \subseteq \mathbb{R}^+ \times \mathbb{S}^{n-1}$. We also consider the r-sections

$$\Phi(E)_r = \{y \in \mathbb{S}^{n-1} \,|\, (r,y) \in \Phi(E)\} = \{y \in \mathbb{S}^{n-1} \,|\, ry \in E\}$$

and the y-sections

$$\Phi(E)^y = \{r \in \mathbb{R}^+ \,|\, (r,y) \in \Phi(E)\} = \{r \in \mathbb{R}^+ \,|\, ry \in E\}$$

of $\Phi(E)$. We extend the notation as follows.

Definition. *If $E \subseteq \mathbb{R}^n$, we define, for every $r \in \mathbb{R}^+$ and every $y \in \mathbb{S}^{n-1}$,*

$$E_r = \{y \in \mathbb{S}^{n-1} \,|\, ry \in E\}, \quad E^y = \{r \in \mathbb{R}^+ \,|\, ry \in E\}$$

*and call them the **r-sections** and the **y-sections** of E, respectively.*

Observe that E may contain 0, but this plays no role: the sections of E are the same as the corresponding sections of $E \setminus \{0\}$. Thus, the sections of E are, by definition, exactly the same as the sections of $\Phi(E \setminus \{0\})$. This is justified by the informal identification of $E \setminus \{0\}$ with $\Phi(E \setminus \{0\})$.

Theorem 4.9. *Let E be any Borel set in $\mathbb{R}^n$. Then E_r is a Borel set in $\mathbb{S}^{n-1}$ for every $r \in \mathbb{R}^+$ and E^y is a Borel set in $\mathbb{R}^+$ for every $y \in \mathbb{S}^{n-1}$. Moreover, the function $r \mapsto \sigma_{n-1}(E_r)$ is $\mathcal{B}_{\mathbb{R}^+}$-measurable and the function $y \mapsto \int_{E^y} r^{n-1}\, dr$ is $\mathcal{B}_{\mathbb{S}^{n-1}}$-measurable. Also,*

$$m_n(E) = \int_0^{+\infty} \sigma_{n-1}(E_r) r^{n-1}\, dr = \int_{\mathbb{S}^{n-1}} \big(\int_{E^y} r^{n-1}\, dr \big)\, d\sigma_{n-1}(y).$$

Proof. The set $E \setminus \{0\}$ is a Borel set in $\mathbb{R}^n_*$. Proposition 4.5 implies that $\Phi(E \setminus \{0\})$ is a Borel set in $\mathbb{R}^+ \times \mathbb{S}^{n-1}$, and Lemmas 4.2 and 4.3 imply

$$m_n(E) = m_n(E \setminus \{0\}) = \widetilde{m_n}(\Phi(E \setminus \{0\})) = (\rho \otimes \sigma_{n-1})(\Phi(E \setminus \{0\})).$$

Also $E_r = \Phi(E \setminus \{0\})_r$ and $E^y = \Phi(E \setminus \{0\})^y$. The rest is a consequence of Theorem 4.7. $\square$

We shall see a simple description of the completion of the measure space $(\mathbb{S}^{n-1}, \mathcal{B}_{\mathbb{S}^{n-1}}, \sigma_{n-1})$ in terms of positive cones.

Definition. *Let $(\mathbb{S}^{n-1}, \mathcal{S}_{n-1}, \sigma_{n-1})$ be the completion of the measure space $(\mathbb{S}^{n-1}, \mathcal{B}_{\mathbb{S}^{n-1}}, \sigma_{n-1})$.*

Proposition 4.10. *If $A \subseteq \mathbb{S}^{n-1}$, then*
(i) $A \in \mathcal{S}_{n-1}$ if and only if $\mathbb{R}^+ A \in \mathcal{L}_n$ if and only if $(0,1]A \in \mathcal{L}_n$,
(ii) $\sigma_{n-1}(A) = n\, m_n\big((0,1]A\big)$ for every $A \in \mathcal{S}_{n-1}$.

Proof. (i) If $A \in \mathcal{S}_{n-1}$, there exist $A_1, A_2 \in \mathcal{B}_{\mathbb{S}^{n-1}}$ with $\sigma_{n-1}(A_2) = 0$ so that $A_1 \subseteq A$ and $A \setminus A_1 \subseteq A_2$. Proposition 4.5 implies that the positive cones $\mathbb{R}^+ A_1$ and $\mathbb{R}^+ A_2$ are Borel sets in $\mathbb{R}^n$ with $\mathbb{R}^+ A_1 \subseteq \mathbb{R}^+ A$ and $(\mathbb{R}^+ A) \setminus (\mathbb{R}^+ A_1) \subseteq \mathbb{R}^+ A_2$. Lemmas 4.2 and 4.3 (or Theorem 4.11) imply

$$m_n(\mathbb{R}^+ A_2) = \widetilde{m_n}(\Phi(\mathbb{R}^+ A_2)) = \widetilde{m_n}(\mathbb{R}^+ \times A_2) = (\rho \otimes \sigma_{n-1})(\mathbb{R}^+ \times A_2) = \rho(\mathbb{R}^+)\sigma_{n-1}(A_2) = 0.$$

Hence, $\mathbb{R}^+ A \in \mathcal{L}_n$.
Conversely, let $\mathbb{R}^+ A \in \mathcal{L}_n$. Then, there are Borel sets $B_1, B_2 \subseteq \mathbb{R}^n$ with $m_n(B_2) = 0$, so that $B_1 \subseteq \mathbb{R}^+ A$ and $(\mathbb{R}^+ A) \setminus B_1 \subseteq B_2$. For every $r \in \mathbb{R}^+$ we have that $(B_1)_r \subseteq A$ and $A \setminus (B_1)_r \subseteq (B_2)_r$. Theorem 4.11 implies that

$$\int_0^{+\infty} \sigma_{n-1}((B_2)_r) r^{n-1}\, dr = m_n(B_2) = 0,$$

and so $\sigma_{n-1}((B_2)_r) = 0$ for m_1-a.e. $r \in (0, +\infty)$. If we consider such an r, since $(B_1)_r$ and $(B_2)_r$ are Borel sets in $\mathbb{S}^{n-1}$, we conclude that $A \in \mathcal{S}_{n-1}$.
If $\mathbb{R}^+ A \in \mathcal{L}_n$, then $(0,1]A = (\mathbb{R}^+ A) \cap \mathbb{B}_n \in \mathcal{L}_n$, where $\mathbb{B}_n$ is the closed unit ball of $\mathbb{R}^n$ centered at 0. Conversely, if $(0,1]A \in \mathcal{L}_n$, then $\mathbb{R}^+ A = \bigcup_{k=1}^{+\infty} k\big((0,1]A\big) \in \mathcal{L}_n$.
(ii) Let $A \in \mathcal{S}_{n-1}$. Then there are $A_1, A_2 \in \mathcal{B}_{\mathbb{S}^{n-1}}$ with $\sigma_{n-1}(A_2) = 0$ so that $A_1 \subseteq A$ and $A \setminus A_1 \subseteq A_2$. Then the sets $(0,1]A_1$ and $(0,1]A_2$ are Borel sets in $\mathbb{R}^n$ with $(0,1]A_1 \subseteq (0,1]A$ and $(0,1]A \setminus (0,1]A_1 \subseteq (0,1]A_2$. We conclude that

$$\sigma_{n-1}(A) = \sigma_{n-1}(A_1) = n\, m_n((0,1]A_1) = n\, m_n((0,1]A),$$

since $m_n((0,1]A_2) = \frac{1}{n}\sigma_{n-1}(A_2) = 0$. $\square$

The next result is an extension of Theorem 4.11 to Lebesgue sets.

Theorem 4.10. *Let $E \in \mathcal{L}_n$. Then $E_r \in \mathcal{S}_{n-1}$ for m_1-a.e. $r \in \mathbb{R}^+$ and $E^y \in \mathcal{L}_1$ for σ_{n-1}-a.e. $y \in \mathbb{S}^{n-1}$. Moreover, the m_1-almost everywhere defined function $r \mapsto \sigma_{n-1}(E_r)$ is $\mathcal{L}_1$-measurable and the σ_{n-1}-almost everywhere defined function $y \mapsto \int_{E^y} r^{n-1}\, dr$ is $\mathcal{S}_{n-1}$-measurable. Also,*

$$m_n(E) = \int_0^{+\infty} \sigma_{n-1}(E_r) r^{n-1}\, dr = \int_{\mathbb{S}^{n-1}} \big(\int_{E^y} r^{n-1}\, dr \big)\, d\sigma_{n-1}(y).$$

Proof. Since $E \in \mathcal{L}_n$, there are Borel sets B_1, B_2 in $\mathbb{R}^n$ with $m_n(B_2) = 0$ so that $B_1 \subseteq E$ and $E \setminus B_1 \subseteq B_2$. Theorem 4.11 implies that, for every $r \in \mathbb{R}^+$, $(B_1)_r$ and $(B_2)_r$ are Borel sets in $\mathbb{S}^{n-1}$ with $(B_1)_r \subseteq E_r$ and $E_r \setminus (B_1)_r \subseteq (B_2)_r$. From Theorem 4.11 again,

$$\int_0^{+\infty} \sigma_{n-1}((B_2)_r) r^{n-1}\, dr = m_n(B_2) = 0,$$

and so $\sigma_{n-1}((B_2)_r) = 0$ for m_1-a.e. $r \in \mathbb{R}^+$. Hence, $E_r \in \mathcal{S}_{n-1}$ and $\sigma_{n-1}(E_r) = \sigma_{n-1}((B_1)_r)$ for m_1-a.e. $r \in \mathbb{R}^+$.
Similarly, for every $y \in \mathbb{S}^{n-1}$, $(B_1)^y$ and $(B_2)^y$ are Borel sets in $\mathbb{R}^+$ with $(B_1)^y \subseteq E^y$ and $E^y \setminus (B_1)^y \subseteq (B_2)^y$. Also

$$\int_{\mathbb{S}^{n-1}} \big(\int_{(B_2)^y} r^{n-1}\, dr \big)\, d\sigma_{n-1}(y) = m_n(B_2) = 0,$$

and so $\int_{(B_2)^y} r^{n-1}\, dr = 0$ for σ_{n-1}-a.e. $y \in \mathbb{S}^{n-1}$. This implies that $m_1((B_2)^y) = 0$ for σ_{n-1}-a.e. $y \in \mathbb{S}^{n-1}$, and so $E^y \in \mathcal{L}_1$ and $\int_{E^y} r^{n-1}\, dr = \int_{(B_1)^y} r^{n-1}\, dr$ for σ_{n-1}-a.e. $y \in \mathbb{S}^{n-1}$.
Finally,

$$m_n(E) = m_n(B_1) = \int_0^{+\infty} \sigma_{n-1}((B_1)_r) r^{n-1}\, dr = \int_0^{+\infty} \sigma_{n-1}(E_r) r^{n-1}\, dr,$$

$$m_n(E) = m_n(B_1) = \int_{\mathbb{S}^{n-1}} \big(\int_{(B_1)^y} r^{n-1}\, dr \big)\, d\sigma_{n-1}(y) = \int_{\mathbb{S}^{n-1}} \big(\int_{E^y} r^{n-1}\, dr \big)\, d\sigma_{n-1}(y)$$

from Theorem 4.11. □

The rest of this section consists of a series of theorems which describe the so-called method of **integration by polar coordinates**.

Definition. *Let $f : \mathbb{R}^n \to Z$. For every $r \in \mathbb{R}^+$ and every $y \in \mathbb{S}^{n-1}$ we define the functions $f_r : \mathbb{S}^{n-1} \to Z$ and $f^y : \mathbb{R}^+ \to Z$ by*

$$f_r(y) = f^y(r) = f(ry).$$

*f_r is called the r-**section** of f and f^y is called the y-**section** of f.*

Theorem 4.13 treats integration by polar coordinates for Borel measurable functions.

Theorem 4.11. *(i) If $f : \mathbb{R}^n \to [0, +\infty]$ is $\mathcal{B}_{\mathbb{R}^n}$-measurable, then f_r is $\mathcal{B}_{\mathbb{S}^{n-1}}$-measurable for every $r \in \mathbb{R}^+$ and f^y is $\mathcal{B}_{\mathbb{R}^+}$-measurable for every $y \in \mathbb{S}^{n-1}$. Moreover, the function $r \mapsto \int_{\mathbb{S}^{n-1}} f_r(y)\, d\sigma_{n-1}(y)$ is $\mathcal{B}_{\mathbb{R}^+}$-measurable, and the function $y \mapsto \int_0^{+\infty} f^y(r) r^{n-1}\, dr$ is $\mathcal{B}_{\mathbb{S}^{n-1}}$-measurable. Also*

$$\begin{aligned} \int_{\mathbb{R}^n} f(x)\, dm_n(x) &= \int_0^{+\infty} \big(\int_{\mathbb{S}^{n-1}} f(ry)\, d\sigma_{n-1}(y) \big) r^{n-1}\, dr \\ &= \int_{\mathbb{S}^{n-1}} \big(\int_0^{+\infty} f(ry) r^{n-1}\, dr \big)\, d\sigma_{n-1}(y). \end{aligned}$$

(ii) If $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ is $\mathcal{B}_{\mathbb{R}^n}$-measurable and m_n-integrable, then f_r is $\mathcal{B}_{\mathbb{S}^{n-1}}$-measurable for every $r \in \mathbb{R}^+$ and σ_{n-1}-integrable for m_1-a.e. $r \in \mathbb{R}^+$, and f^y is $\mathcal{B}_{\mathbb{R}^+}$-measurable for every $y \in \mathbb{S}^{n-1}$ and m_1-integrable for σ_{n-1}-a.e. $y \in \mathbb{S}^{n-1}$. Moreover, the m_1-almost everywhere defined function $r \mapsto \int_{\mathbb{S}^{n-1}} f_r(y)\, d\sigma_{n-1}(y)$ is $\mathcal{B}_{\mathbb{R}^+}$-measurable and m_1-integrable, the σ_{n-1}-almost everywhere defined function $y \mapsto \int_0^{+\infty} f^y(r) r^{n-1}\, dr$ is $\mathcal{B}_{\mathbb{S}^{n-1}}$-measurable and σ_{n-1}-integrable, and the equalities in (i) are true.

Proof. (i) If $f = \chi_E$, then the results are the same as the results of Theorem 4.11. Using the linearity of the integrals, we can prove the results in the case of a simple function $\phi : \mathbb{R}^n \to [0, +\infty]$. Finally, applying the Monotone Convergence Theorem to an increasing sequence of simple functions, we can prove the results in the general case $f : \mathbb{R}^n \to [0, +\infty]$.
(ii) We use the results of (i) to pass to the case of functions $f : \mathbb{R}^n \to \overline{\mathbb{R}}$, by writing them as $f = f^+ - f^-$. We next treat the case of $f : \mathbb{R}^n \to \overline{\mathbb{C}}$, by writing $f = \mathrm{Re}(f) + i\,\mathrm{Im}(f)$, after we exclude, in the usual manner, the set $f^{-1}(\{\infty\})$. $\square$

Theorem 4.14 treats integration by polar coordinates for Lebesgue measurable functions.

Theorem 4.12. *(i) If $f : \mathbb{R}^n \to [0, +\infty]$ is $\mathcal{L}_n$-measurable, then f_r is $\mathcal{S}_{n-1}$-measurable for m_1-a.e. $r \in \mathbb{R}^+$ and f^y is $\mathcal{L}_1$-measurable for σ_{n-1}-a.e. $y \in \mathbb{S}^{n-1}$. Moreover, the m_1-almost everywhere defined function $r \mapsto \int_{\mathbb{S}^{n-1}} f_r(y)\, d\sigma_{n-1}(y)$ is $\mathcal{L}_1$-measurable, and the σ_{n-1}-almost everywhere defined function $y \mapsto \int_0^{+\infty} f^y(r) r^{n-1}\, dr$ is $\mathcal{S}_{n-1}$-measurable, and*

$$\begin{aligned}\int_{\mathbb{R}^n} f(x)\, dm_n(x) &= \int_0^{+\infty} \big(\int_{\mathbb{S}^{n-1}} f(ry)\, d\sigma_{n-1}(y) \big) r^{n-1}\, dr \\ &= \int_{\mathbb{S}^{n-1}} \big(\int_0^{+\infty} f(ry) r^{n-1}\, dr \big)\, d\sigma_{n-1}(y).\end{aligned}$$

(ii) If $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ is $\mathcal{L}_n$-measurable and m_n-integrable, then f_r is $\mathcal{S}_{n-1}$-measurable and σ_{n-1}-integrable for m_1-a.e. $r \in \mathbb{R}^+$, and f^y is $\mathcal{L}_1$-measurable and m_1-integrable for σ_{n-1}-a.e. $y \in \mathbb{S}^{n-1}$. Moreover, the m_1-almost everywhere defined function $r \mapsto \int_{\mathbb{S}^{n-1}} f_r(y)\, d\sigma_{n-1}(y)$ is $\mathcal{L}_1$-measurable and m_1-integrable, and the σ_{n-1}-almost everywhere defined function $y \mapsto \int_0^{+\infty} f^y(r) r^{n-1}\, dr$ is $\mathcal{S}_{n-1}$-measurable and σ_{n-1}-integrable, and the equalities in (i) are true.

Proof. We use Theorem 4.12 in the way that we used Theorem 4.11 to prove Theorem 4.13. $\square$

Definition. *A set $E \subseteq \mathbb{R}^n$ is called* ***radial*** *if $x \in E$ implies that $x' \in E$ for all x' with $\|x'\| = \|x\|$. A function $f : \mathbb{R}^n \to Z$ is called* ***radial*** *if $f(x) = f(x')$ for every x, x' with $\|x\| = \|x'\|$.*

It is obvious that E is radial if and only if χ_E is radial.
If the set $E \subseteq \mathbb{R}^n$ is radial, we define the **radial projection** $\widetilde{E}$ of E by

$$\widetilde{E} = \{r \in \mathbb{R}^+ \,|\, x \in E \text{ when } \|x\| = r\} \subseteq \mathbb{R}^+.$$

Also, if $f : \mathbb{R}^n \to Z$ is radial, we define the **radial projection** $\widetilde{f} : \mathbb{R}^+ \to Z$ of f by

$$\widetilde{f}(r) = f(x) \quad \text{for all } x \in \mathbb{R}^n \text{ with } \|x\| = r.$$

It is obvious that a radial set or a radial function is uniquely determined from its radial projection (except from the fact that the radial set may or may not contain the point 0 and that the value of the function at 0 is not determined by its radial projection).

Proposition 4.11. *(i) The radial set $E \subseteq \mathbb{R}^n$ is in $\mathcal{B}_{\mathbb{R}^n}$ or in $\mathcal{L}_n$ if and only if its radial projection is in $\mathcal{B}_{\mathbb{R}^+}$ or, respectively, in $\mathcal{L}_1$. In any case we have*

$$m_n(E) = \sigma_{n-1}(\mathbb{S}^{n-1}) \int_{\widetilde{E}} r^{n-1}\, dr.$$

(ii) If $(Z, \mathcal{S}_Z)$ is a measurable space, then the radial $f : \mathbb{R}^n \to Z$ is $(\mathcal{B}_{\mathbb{R}^n}, \mathcal{S}_Z)$-measurable or $(\mathcal{L}_n, \mathcal{S}_Z)$-measurable if and only if its radial projection is $(\mathcal{B}_{\mathbb{R}^+}, \mathcal{S}_Z)$-measurable or, respectively, $(\mathcal{L}_1, \mathcal{S}_Z)$-measurable.
(iii) If $f : \mathbb{R}^n \to [0, +\infty]$ is $\mathcal{B}_n$ or $\mathcal{L}_n$-measurable or if $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ is $\mathcal{B}_n$ or $\mathcal{L}_n$-measurable and m_n-integrable, then

$$\int_{\mathbb{R}^n} f(x)\, dm_n(x) = \sigma_{n-1}(\mathbb{S}^{n-1}) \int_0^{+\infty} \widetilde{f}(r) r^{n-1}\, dr.$$

Proof. (i) If $E \in \mathcal{B}_{\mathbb{R}^n}$ or $E \in \mathcal{L}_n$ is radial, then for every $y \in \mathbb{S}^{n-1}$ we have $E^y = \widetilde{E}$, and so the result is a consequence of Theorems 4.11 and 4.12.
For the converse we consider the collection of all subsets of $\mathbb{R}^+$ which are radial projections of radial Borel sets in $\mathbb{R}^n$. Then we prove easily that this collection is a σ-algebra which contains all open subsets of $\mathbb{R}^+$, and so it contains all Borel sets in $\mathbb{R}^+$.
Now, if E is radial and $\widetilde{E} \in \mathcal{L}_1$, then there are Borel sets M_1, M_2 in $\mathbb{R}^+$ with $m_1(M_2) = 0$ so that $M_1 \subseteq \widetilde{E}$ and $\widetilde{E} \setminus M_1 \subseteq M_2$. We consider the radial sets $E_1, E_2 \subseteq \mathbb{R}^n$ so that $\widetilde{E_1} = M_1$ and $\widetilde{E_2} = M_2$. By the result of the previous paragraph, E_1, E_2 are Borel sets. Then we have $E_1 \subseteq E$ and $E \setminus E_1 \subseteq E_2$. Since

$$0 = m_n(E_2) = \int_{\mathbb{S}^{n-1}} \big(\int_{(E_2)^y} r^{n-1}\, dr\big)\, d\sigma_{n-1} = \sigma_{n-1}(\mathbb{S}^{n-1}) \int_{\widetilde{E_2}} r^{n-1}\, dr,$$

we have that $\int_{\widetilde{E_2}} r^{n-1}\, dr = 0$, and so $m_1(\widetilde{E_2}) = 0$. This implies that $E \in \mathcal{L}_1$.
(ii) A consequence of the definition of measurability and the result of part (i).
(iii) A consequence of Theorems 4.13 and 4.14. $\square$

Exercises.

4.4.1. Consider, for any $p > 0$, the function $f : \mathbb{R}^n \to [0, +\infty]$, defined by $f(x) = \frac{1}{\|x\|^p}$.
(i) Prove that f is not Lebesgue integrable over $\mathbb{R}^n$.
(ii) For any $\delta > 0$, prove that f is integrable over $\{x \in \mathbb{R}^n \mid \|x\| \geq \delta\}$ if and only if $p > 1$.
(iii) For any $R < +\infty$, prove that f is integrable over $\{x \in \mathbb{R}^n \mid \|x\| \leq R\}$ if and only if $p < 1$.

4.4.2. *The volume of the unit ball in* $\mathbb{R}^n$ *and the surface measure of* $\mathbb{S}^{n-1}$.
(i) If $v_n = m_n(\mathbb{B}_n)$ is the Lebesgue measure of the closed unit ball of $\mathbb{R}^n$ centered at 0, prove that $v_n = 2v_{n-1} \int_0^1 (1-t^2)^{\frac{n-1}{2}}\, dt$.
(ii) Set $J_n = \int_0^1 (1-t^2)^{\frac{n-1}{2}}\, dt$ for $n \geq 0$ and prove the inductive formula $J_n = \frac{n-1}{n} J_{n-2}$, $n \geq 2$.
(iii) Use properties of the gamma-function (see exercise 3.2.15) to prove that

$$m_n(\mathbb{B}_n) = \tfrac{\pi^{n/2}}{\Gamma((n/2)+1)}, \quad \sigma_{n-1}(\mathbb{S}^{n-1}) = \tfrac{2\pi^{n/2}}{\Gamma(n/2)}.$$

4.4.3. *The integral of Gauss and the measures of* $\mathbb{B}_n$ *and of* $\mathbb{S}^{n-1}$.
Define $I_n = \int_{\mathbb{R}^n} e^{-\frac{\|x\|^2}{2}}\, dm_n(x)$.
(i) Prove that $I_n = I_1^n$ for every $n \in \mathbb{N}$.
(ii) Using integration by polar coordinates, prove that $I_2 = 2\pi$ and, hence,

$$\int_{\mathbb{R}^n} e^{-\frac{\|x\|^2}{2}}\, dm_n(x) = (2\pi)^{\frac{n}{2}}.$$

(iii) Using integration by polar coordinates, prove that $(2\pi)^{\frac{n}{2}} = \sigma_{n-1}(\mathbb{S}^{n-1}) \int_0^{+\infty} e^{-\frac{r^2}{2}} r^{n-1}\, dr$ and, hence, $\sigma_{n-1}(\mathbb{S}^{n-1}) = \frac{2\pi^{n/2}}{\Gamma(n/2)}$ and $m_n(\mathbb{B}_n) = \frac{\pi^{n/2}}{\Gamma((n/2)+1)}$.

Chapter 5

Convergence of functions.

5.1 a.e. convergence and uniform a.e. convergence.

The two types of convergence of sequences of functions which are usually studied in elementary courses are the pointwise convergence and the uniform convergence. We recall briefly their definitions and simple properties.

Let A be an arbitrary set and $f, f_n : A \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ for every $n \in \mathbb{N}$. We say that $f_n \to f$ *pointwise* on A if $f_n(x) \to f(x)$ for every $x \in A$. In case $f(x)$ is finite, this means that for every $\epsilon > 0$ there is an $n_0 = n_0(\epsilon, x)$ so that $|f_n(x) - f(x)| \leq \epsilon$ for every $n \geq n_0$.

Let A be an arbitrary set and $f, f_n : A \to \mathbb{C}$ for every $n \in \mathbb{N}$. We say that $f_n \to f$ *uniformly* on A if for every $\epsilon > 0$ there is an $n_0 = n_0(\epsilon)$ so that $|f_n(x) - f(x)| \leq \epsilon$ for every $x \in A$ and every $n \geq n_0$ or, equivalently, $\sup_{x \in A} |f_n(x) - f(x)| \leq \epsilon$ for every $n \geq n_0$. In other words, $f_n \to f$ uniformly on A if and only if $\sup_{x \in A} |f_n(x) - f(x)| \to 0$ as $n \to +\infty$.

It is obvious that uniform convergence on A implies pointwise convergence on A. The converse is not true in general.

Example. Let $f_n = \chi_{(0, \frac{1}{n})} : (0, 1) \to \mathbb{R}$ for every n. Then $f_n \to 0$ pointwise on $(0, 1)$ but not uniformly on $(0, 1)$.

Let us describe some easy properties.

The pointwise limit (if it exists) of a sequence of functions is unique. The same is true for the uniform limit.

Let $f, g, f_n, g_n : A \to \mathbb{C}$ for all $n \in \mathbb{N}$. If $f_n \to f$ and $g_n \to g$ pointwise on A, then $f_n + g_n \to f + g$ and $f_n g_n \to fg$ pointwise on A. The same is true for uniform convergence, provided that in the case of the product the two sequences are uniformly bounded: this means that there is an $M < +\infty$ so that $|f_n(x)| \leq M$ and $|g_n(x)| \leq M$ for every $x \in A$ and every $n \in \mathbb{N}$.

Let $f_n : A \to \mathbb{C}$ for every $n \in \mathbb{N}$. We say that (f_n) is *Cauchy uniformly* on A if for every $\epsilon > 0$ there is an $n_0 = n_0(\epsilon)$ so that $|f_n(x) - f_m(x)| \leq \epsilon$ for every $x \in A$ and every $n, m \geq n_0$ or, equivalently, $\sup_{x \in A} |f_n(x) - f_m(x)| \leq \epsilon$ for every $n, m \geq n_0$. In other words, (f_n) is Cauchy uniformly on A if and only if $\sup_{x \in A} |f_n(x) - f_m(x)| \to 0$ as $n, m \to +\infty$.

If (f_n) is Cauchy uniformly on A then there is an $f : A \to \mathbb{C}$ so that $f_n \to f$ uniformly on A. Indeed, we have that for every $\epsilon > 0$ there is an $n_0 = n_0(\epsilon)$ so that $|f_n(x) - f_m(x)| \leq \epsilon$ for every $x \in A$ and every $n, m \geq n_0$. This implies that for every $x \in A$ the sequence $(f_n(x))$ is a Cauchy sequence of complex numbers, and so it converges to some complex number. Let us define $f : A \to \mathbb{C}$ by $f(x) = \lim_{n \to +\infty} f_n(x)$. Now, if in the above inequality $|f_n(x) - f_m(x)| \leq \epsilon$ we let $m \to +\infty$, we get that $|f_n(x) - f(x)| \leq \epsilon$ for every $x \in A$ and every $n \geq n_0$. Therefore, $f_n \to f$ uniformly on A.

It is almost straightforward to extend these two notions of convergence to measure spaces.

Let $(X, \mathcal{S}, \mu)$ be an arbitrary measure space.

We have already seen the notion of a.e. convergence. If $f, f_n : X \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ for every $n \in \mathbb{N}$, we say that $f_n \to f$ **(pointwise) a.e.** on $A \in \mathcal{S}$ if there is a set $B \in \mathcal{S}$, $B \subseteq A$, so that $\mu(A \setminus B) = 0$ and $f_n \to f$ pointwise on B.

If $f, f_n : X \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ for every $n \in \mathbb{N}$, we say that $f_n \to f$ **uniformly a.e.** on $A \in \mathcal{S}$ if there is a set $B \in \mathcal{S}$, $B \subseteq A$, so that $\mu(A \setminus B) = 0$, f and f_n are finite on B for all n and $f_n \to f$ uniformly on B.

It is clear that uniform a.e. convergence on A implies a.e. convergence on A. The converse is not true in general and the counter-example is the same as above.

If $f_n \to f'$ a.e. on A and $f_n \to f''$ a.e. on A, then $f' = f''$ a.e. on A. Indeed, there are $B, C \in \mathcal{S}$, $B, C \subseteq A$ so that $\mu(A \setminus B) = \mu(A \setminus C) = 0$ and $f_n \to f'$ pointwise on B and $f_n \to f''$ pointwise on C. Therefore, $f_n \to f'$ and $f_n \to f''$ pointwise on $B \cap C$, and so $f' = f''$ on $B \cap C$. Since $\mu(A \setminus (B \cap C)) = 0$, we get that $f' = f''$ a.e. on A. *This is a common feature of almost any notion of convergence in the framework of measure spaces*: the limits may be considered unique only if we agree to identify functions which are equal a.e. on A. This can be made precise by using the tool of equivalence classes in an appropriate manner but we postpone this discussion for later.

We can similarly prove that if $f_n \to f'$ uniformly a.e. on A and $f_n \to f''$ uniformly a.e. on A, then $f' = f''$ a.e. on A.

Moreover, if $f, g, f_n, g_n : A \to \mathbb{C}$ for every n and $f_n \to f$ a.e. on A and $g_n \to g$ a.e. on A, then $f_n + g_n \to f + g$ a.e. on A and $f_n g_n \to fg$ a.e. on A. The same is true for uniform a.e. convergence, provided that in the case of the product the two sequences are *uniformly a.e. bounded*: namely, that there is an $M < +\infty$ so that $|f_n| \le M$ a.e. on A and $|g_n| \le M$ a.e. on A for every n.

Exercises.

Except if specified otherwise, all exercises refer to a measure space $(X, \mathcal{S}, \mu)$, all sets belong to $\mathcal{S}$ and all functions are $\mathcal{S}$-measurable.

5.1.1. Let $f, f_n : A \to \mathbb{C}$ for every $n \in \mathbb{N}$.

(i) If $\phi : \mathbb{C} \to \mathbb{C}$ is continuous and $f_n \to f$ a.e. on A, prove that $\phi \circ f_n \to \phi \circ f$ a.e. on A.

(ii) If $\phi : \mathbb{C} \to \mathbb{C}$ is uniformly continuous and $f_n \to f$ uniformly a.e. on A, prove that $\phi \circ f_n \to \phi \circ f$ uniformly a.e. on A.

5.1.2. If $f_n \to f$ a.e. on A and $|f_n| \le g$ a.e. on A for every $n \in \mathbb{N}$, prove that $|f| \le g$ a.e. on A.

5.1.3. If $E_n \subseteq A$ for every $n \in \mathbb{N}$ and $\chi_{E_n} \to f$ a.e. on A, prove that there exists $E \subseteq A$ so that $f = \chi_E$ a.e. on A. What is the relation between the three sets: E, $\underline{\lim} E_n$, $\overline{\lim} E_n$? (See exercise 1.1.1.)

5.1.4. Let $\mu(A) < +\infty$ and $f_n : A \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ for every $n \in \mathbb{N}$ and every f_n be finite a.e. on A.

(i) Prove that there exists a $g : A \to [0, +\infty)$ and a sequence (r_n) in $\mathbb{R}^+$ so that $|f_n| \le r_n g$ a.e. on A for every n.

Hint. For every $n \in \mathbb{N}$ there is a $k_n \in \mathbb{N}$ so that $\mu(E_n) < \frac{1}{2^n}$, where $E_n = \{x \in A \,|\, |f_n(x)| > k_n\}$. Let $F_n = \bigcup_{k=n}^{+\infty} E_k$ and $F = \bigcap_{n=1}^{+\infty} F_n$. Then $\mu(F) = 0$. Now let $A_n = A \setminus F_n$ so that $A_n \uparrow A \setminus F$. Consider $g = 1$ on A_1 and $g = \max\{1, f_1, \dots, f_{n-1}\}$ on $A_n \setminus A_{n-1}$ for $n \ge 2$ and prove that $|f_n| \le k_n g$ on $A \setminus F$.

(ii) Prove that there is a sequence (λ_n) in $\mathbb{R}^+$ so that $\lambda_n f_n \to 0$ a.e. on A.

5.1.5. Let $\mu(A) < +\infty$ and $f_n : A \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ for every $n \in \mathbb{N}$ and every f_n be finite a.e. on A and $f_n \to 0$ a.e. on A.

(i) Prove that there exists a $g : A \to [0, +\infty)$ and a sequence (ϵ_n) in $\mathbb{R}^+$ so that $\epsilon_n \downarrow 0$ and $|f_n| \le \epsilon_n g$ a.e. on A for every n.

Hint. Look at the hint of exercise 5.1.4.

(ii) Prove that there is a sequence (λ_n) in $\mathbb{R}^+$ so that $\lambda_n \uparrow +\infty$ and $\lambda_n f_n \to 0$ a.e. on A.

5.2 Convergence in the mean.

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

Definition. *Let all $f, f_n : X \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ be $\mathcal{S}$-measurable, and all f, f_n be finite a.e. on A.*
*We say that $f_n \to f$ **in the mean** on $A \in \mathcal{S}$ if $\int_A |f_n - f| \, d\mu \to 0$ as $n \to +\infty$.*
*We say that (f_n) is **Cauchy in the mean** on $A \in \mathcal{S}$ if $\int_A |f_n - f_m| \, d\mu \to 0$ as $m, n \to +\infty$.*

Proposition 5.1. *If $f_n \to f'$ and $f_n \to f''$ in the mean on A, then $f' = f''$ a.e. on A.*

Proof. We have

$$\int_A |f' - f''| \, d\mu \le \int_A |f_n - f'| \, d\mu + \int_A |f_n - f''| \, d\mu \to 0$$

as $n \to +\infty$. Hence, $\int_A |f' - f''| \, d\mu = 0$, and so $f' = f''$ a.e. on A. □

Proposition 5.2. *Let $f_n \to f$ and $g_n \to g$ in the mean on A and $\lambda \in \mathbb{C}$. Then*
(i) $f_n + g_n \to f + g$ in the mean on A.
(ii) $\lambda f_n \to \lambda f$ in the mean on A.

Proof. We have

$$\int_A |(f_n + g_n) - (f + g)| \, d\mu \le \int_A |f_n - f| \, d\mu + \int_A |g_n - g| \, d\mu \to 0,$$

$$\int_A |\lambda f_n - \lambda f| \, d\mu = |\lambda| \int_A |f_n - f| \, d\mu \to 0$$

as $n \to +\infty$. □

It is trivial to prove that, if $f_n \to f$ in the mean on A, then (f_n) is Cauchy in the mean on A. The following basic theorem expresses the converse.

Theorem 5.1. *If (f_n) is Cauchy in the mean on A, then there is $f : X \to \mathbb{C}$ so that $f_n \to f$ in the mean on A. Moreover, there is a subsequence (f_{n_k}) so that $f_{n_k} \to f$ a.e. on A.*
Corollary: if $f_n \to f$ in the mean on A, there is a subsequence (f_{n_k}) so that $f_{n_k} \to f$ a.e. on A.

Proof. For every $k \in \mathbb{N}$ there is $n_k \in \mathbb{N}$ so that $\int_A |f_n - f_m| \, d\mu < \frac{1}{2^k}$ for every $n, m \ge n_k$. Since we may assume that each n_k is as large as we like, we may inductively take (n_k) so that $n_k < n_{k+1}$ for every k. Therefore, (f_{n_k}) is a subsequence of (f_n).
From the construction of n_k and from $n_k < n_{k+1}$, we get $\int_A |f_{n_{k+1}} - f_{n_k}| \, d\mu < \frac{1}{2^k}$ for every k. Then the $\mathcal{S}$-measurable function $G : X \to [0, +\infty]$, defined by $G = \sum_{k=1}^{+\infty} |f_{n_{k+1}} - f_{n_k}|$ on A and by $G = 0$ on A^c, satisfies

$$\int_X G \, d\mu = \sum_{k=1}^{+\infty} \int_A |f_{n_{k+1}} - f_{n_k}| \, d\mu < \sum_{k=1}^{+\infty} \frac{1}{2^k} = 1.$$

Hence, G is finite a.e. on A, and so $\sum_{k=1}^{+\infty} (f_{n_{k+1}}(x) - f_{n_k}(x))$ converges for a.e. $x \in A$. Therefore, there is a $B \in \mathcal{S}$, $B \subseteq A$ so that $\mu(A \setminus B) = 0$ and $\sum_{k=1}^{+\infty} (f_{n_{k+1}}(x) - f_{n_k}(x))$ converges for every $x \in B$. Now, we define the $\mathcal{S}$-measurable function $f : X \to \mathbb{C}$ by

$$f = \begin{cases} f_{n_1} + \sum_{k=1}^{+\infty} (f_{n_{k+1}} - f_{n_k}), & \text{on } B \\ 0, & \text{on } B^c \end{cases}$$

On B we have that

$$f = f_{n_1} + \lim_{K \to +\infty} \sum_{k=1}^{K-1} (f_{n_{k+1}} - f_{n_k}) = \lim_{K \to +\infty} f_{n_K}.$$

and so (f_{n_k}) converges to f a.e. on A.
We also have on B that

$$\begin{aligned}|f_{n_K}-f| &= |f_{n_K}-f_{n_1}-\textstyle\sum_{k=1}^{+\infty}(f_{n_{k+1}}-f_{n_k})|\\ &= |\textstyle\sum_{k=1}^{K-1}(f_{n_{k+1}}-f_{n_k})-\sum_{k=1}^{+\infty}(f_{n_{k+1}}-f_{n_k})| \leq \sum_{k=K}^{+\infty}|f_{n_{k+1}}-f_{n_k}|\end{aligned}$$

for all K. Hence,

$$\textstyle\int_A |f_{n_K}-f|\,d\mu \leq \sum_{k=K}^{+\infty}\int_A |f_{n_{k+1}}-f_{n_k}|\,d\mu < \sum_{k=K}^{+\infty}\frac{1}{2^k}=\frac{1}{2^{K-1}}\to 0$$

as $K\to+\infty$.
Since $n_k\to+\infty$ when $k\to+\infty$, we get

$$\textstyle\int_A |f_k-f|\,d\mu \leq \int_A |f_k-f_{n_k}|\,d\mu+\int_A |f_{n_k}-f|\,d\mu \to 0,$$

and we conclude that $f_n\to f$ in the mean on A. $\square$

Example. We consider the functions $f_1=\chi_{(0,1)}, f_2=\chi_{(0,\frac12)}, f_3=\chi_{(\frac12,1)}, f_4=\chi_{(0,\frac13)}, f_5=\chi_{(\frac13,\frac23)}, f_6=\chi_{(\frac23,1)}, f_7=\chi_{(0,\frac14)}, f_8=\chi_{(\frac14,\frac24)}, f_9=\chi_{(\frac24,\frac34)}, f_{10}=\chi_{(\frac34,1)}$ and so on.
It is clear that $\int_{(0,1)}|f_n|\,dm_1\to 0$, and so $f_n\to 0$ in the mean on $(0,1)$. Theorem 5.1 implies that there exists a subsequence converging to 0 a.e. on $(0,1)$, and it is easy to find many such subsequences: for example, the functions $f_1=\chi_{(0,1)}, f_2=\chi_{(0,\frac12)}, f_4=\chi_{(0,\frac13)}, f_7=\chi_{(0,\frac14)}$ and so on, form one such subsequence.
But it is *not true* that $f_n\to 0$ a.e. on $(0,1)$. In fact, if x is any irrational number in $(0,1)$, then x belongs to infinitely many intervals of the form $(\frac{k-1}{m},\frac{k}{m})$ (for each value of m there is exactly one such value of k), and so $(f_n(x))$ does not converge to 0. It easy to see that $f_n(x)\to 0$ only for every rational $x\in(0,1)$.

We may now complete Proposition 5.2 as follows.

Proposition 5.3. *Let $f_n\to f$ and $g_n\to g$ in the mean on A.*
(i) If there is $M<+\infty$ so that $|f_n|\leq M$ a.e. on A for all n, then $|f|\leq M$ a.e. on A.
(ii) If there is $M<+\infty$ so that $|f_n|\leq M$ a.e. on A and $|g_n|\leq M$ a.e. on A for all n, then $f_ng_n\to fg$ in the mean on A.

Proof. (i) Theorem 5.1 implies that there is a subsequence (f_{n_k}) so that $f_{n_k}\to f$ a.e. on A. Since $|f_{n_k}|\leq M$ a.e. on A for every k, we get that $|f|\leq M$ a.e. on A.
(ii) Using the result of (i), we have

$$\begin{aligned}\textstyle\int_A |f_ng_n-fg|\,d\mu &\leq \textstyle\int_A |f_ng_n-fg_n|\,d\mu+\int_A |fg_n-fg|\,d\mu\\ &\leq M\textstyle\int_A |f_n-f|\,d\mu+M\int_A |g_n-g|\,d\mu\to 0\end{aligned}$$

as $n\to+\infty$. $\square$

Exercises.

Except if specified otherwise, all exercises refer to a measure space $(X,\mathcal{S},\mu)$, all sets belong to $\mathcal{S}$ and all functions are $\mathcal{S}$-measurable.

5.2.1. If $f_n\to f$ in the mean on A and $|f_n|\leq g$ a.e. on A for all n, prove that $|f|\leq g$ a.e. on A.

5.2.2. If $f_n\to f$ a.e. on A and $|f_n|\leq g$ a.e. on A for all n and $\int_A g\,d\mu<+\infty$, prove that $f_n\to f$ in the mean on A.

5.2.3. Look at exercise 5.1.3. If $E_n\subseteq A$ for all n and $\chi_{E_n}\to f$ in the mean on A, prove that there exists $E\subseteq A$ so that $f=\chi_E$ a.e. on A. Prove that $\mu(E_n\triangle E)\to 0$.

5.2.4. Let $E_n\subseteq A$ for all n. If $\mu(E_n\triangle E_m)\to 0$ as $n,m\to+\infty$, prove that there exists $E\subseteq A$ so that $\mu(E_n\triangle E)\to 0$ as $n\to+\infty$. Prove that the metric space $(\mathcal{S}/_\sim,d)$ of exercise 1.2.13 is complete.

5.3 Convergence in measure.

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

Definition. *Let all $f, f_n : X \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ be $\mathcal{S}$-measurable, and all f, f_n be finite a.e. on A.*
*We say that $f_n \to f$ **in measure** on $A \in \mathcal{S}$ if $\mu(\{x \in A \,|\, |f_n(x) - f(x)| \geq \epsilon\}) \to 0$ as $n \to +\infty$ for every $\epsilon > 0$.*
*We say that (f_n) is **Cauchy in measure** on $A \in \mathcal{S}$ if $\mu(\{x \in A \,|\, |f_n(x) - f_m(x)| \geq \epsilon\}) \to 0$ as $n, m \to +\infty$ for every $\epsilon > 0$.*

A useful trick is the inequality

$$\mu(\{x \in A \,|\, |f(x) + g(x)| \geq a + b\}) \leq \mu(\{x \in A \,|\, |f(x)| \geq a\}) + \mu(\{x \in A \,|\, |g(x)| \geq b\}),$$

which is true for every $a, b > 0$. This is due to the set-inclusion

$$\{x \in A \,|\, |f(x) + g(x)| \geq a + b\} \subseteq \{x \in A \,|\, |f(x)| \geq a\} \cup \{x \in A \,|\, |g(x)| \geq b\}.$$

Proposition 5.4. *If $f_n \to f'$ and $f_n \to f''$ in measure on A, then $f' = f''$ a.e. on A.*

Proof. We have

$$\begin{aligned}\mu(\{x \in A \,|\, |f'(x) - f''(x)| \geq \epsilon\}) &\leq \mu\big(\big\{x \in A \,\big|\, |f_n(x) - f'(x)| \geq \tfrac{\epsilon}{2}\big\}\big) \\ &\quad + \mu\big(\big\{x \in A \,\big|\, |f_n(x) - f''(x)| \geq \tfrac{\epsilon}{2}\big\}\big) \to 0.\end{aligned}$$

This implies $\mu(\{x \in A \,|\, |f'(x) - f''(x)| \geq \epsilon\}) = 0$ for every $\epsilon > 0$.
Now we have

$$\{x \in A \,|\, f'(x) \neq f''(x)\} = \textstyle\bigcup_{k=1}^{+\infty} \big\{x \in A \,\big|\, |f'(x) - f''(x)| \geq \tfrac{1}{k}\big\}.$$

Since all terms in the union are null sets, we get that $\{x \in A \,|\, f'(x) \neq f''(x)\}$ is a null set, and we conclude that $f' = f''$ a.e. on A. ☐

Proposition 5.5. *Let $f_n \to f$ and $g_n \to g$ in measure on A and $\lambda \in \mathbb{C}$. Then*
(i) $f_n + g_n \to f + g$ in measure on A.
(ii) $\lambda f_n \to \lambda f$ in measure on A.
(iii) If there is $M < +\infty$ so that $|f_n| \leq M$ a.e. on A for all n, then $|f| \leq M$ a.e. on A.
(iv) If there is $M < +\infty$ so that $|f_n| \leq M$ a.e. on A and $|g_n| \leq M$ a.e. on A for all n, then $f_n g_n \to fg$ in measure on A.

Proof. (i) We apply the usual trick and we have

$$\begin{aligned}\mu(\{x \in A \,|\, |(f_n + g_n)(x) - (f + g)(x)| \geq \epsilon\}) &\leq \mu\big(\big\{x \in A \,\big|\, |f_n(x) - f(x)| \geq \tfrac{\epsilon}{2}\big\}\big) \\ &\quad + \mu\big(\big\{x \in A \,\big|\, |g_n(x) - g(x)| \geq \tfrac{\epsilon}{2}\big\}\big) \to 0\end{aligned}$$

as $n \to +\infty$.
(ii) The case $\lambda = 0$ is trivial. If $\lambda \neq 0$, then

$$\mu(\{x \in A \,|\, |\lambda f_n(x) - \lambda f(x)| \geq \epsilon\}) = \mu\big(\big\{x \in A \,\big|\, |f_n(x) - f(x)| \geq \tfrac{\epsilon}{|\lambda|}\big\}\big) \to 0$$

as $n \to +\infty$.
(iii) We write

$$\begin{aligned}\mu(\{x \in A \,|\, |f(x)| \geq M + \epsilon\}) &\leq \mu\big(\big\{x \in A \,\big|\, |f_n(x)| \geq M + \tfrac{\epsilon}{2}\big\}\big) \\ &\quad + \mu\big(\big\{x \in A \,\big|\, |f_n(x) - f(x)| \geq \tfrac{\epsilon}{2}\big\}\big) \\ &= \mu\big(\big\{x \in A \,\big|\, |f_n(x) - f(x)| \geq \tfrac{\epsilon}{2}\big\}\big) \to 0\end{aligned}$$

as $n \to +\infty$. Hence, $\mu(\{x \in A \,|\, |f(x)| \geq M + \epsilon\}) = 0$ for every $\epsilon > 0$.
We have

$$\{x \in A \,|\, |f(x)| > M\} = \bigcup_{k=1}^{+\infty} \left\{x \in A \,\middle|\, |f(x)| \geq M + \tfrac{1}{k}\right\}$$

and, since all sets of the union are null, we find that $\mu(\{x \in A \,|\, |f(x)| > M\}) = 0$. Hence, $|f| \leq M$ a.e. on A.
(iv) Applying the result of (iii),

$$\begin{aligned}\mu(\{x \in A \,|\, |f_n(x)g_n(x) - f(x)g(x)| \geq \epsilon\}) &\leq \mu\big(\big\{x \in A \,\big|\, |f_n(x)g_n(x) - f_n(x)g(x)| \geq \tfrac{\epsilon}{2}\big\}\big) \\ &\quad + \mu\big(\big\{x \in A \,\big|\, |f_n(x)g(x) - f(x)g(x)| \geq \tfrac{\epsilon}{2}\big\}\big) \\ &\leq \mu\big(\big\{x \in A \,\big|\, |g_n(x) - g(x)| \geq \tfrac{\epsilon}{2M}\big\}\big) \\ &\quad + \mu\big(\big\{x \in A \,\big|\, |f_n(x) - f(x)| \geq \tfrac{\epsilon}{2M}\big\}\big) \to 0\end{aligned}$$

as $n \to +\infty$. □

It is easy to see that if $f_n \to f$ in measure on A, then (f_n) is Cauchy in measure on A. Indeed,

$$\begin{aligned}\mu(\{x \in A \,|\, |f_n(x) - f_m(x)| \geq \epsilon\}) &\leq \mu\big(\big\{x \in A \,\big|\, |f_n(x) - f(x)| \geq \tfrac{\epsilon}{2}\big\}\big) \\ &\quad + \mu\big(\big\{x \in A \,\big|\, |f_m(x) - f(x)| \geq \tfrac{\epsilon}{2}\big\}\big) \to 0\end{aligned}$$

as $n, m \to +\infty$.

Theorem 5.2. *If (f_n) is Cauchy in measure on A, then there is $f : X \to \mathbb{C}$ so that $f_n \to f$ in measure on A. Moreover, there is a subsequence (f_{n_k}) so that $f_{n_k} \to f$ a.e. on A.*
Corollary: if $f_n \to f$ in measure on A, there is a subsequence (f_{n_k}) so that $f_{n_k} \to f$ a.e. on A.

Proof. For every $k \in \mathbb{N}$ we have that $\mu\big(\big\{x \in A \,\big|\, |f_n(x) - f_m(x)| \geq \frac{1}{2^k}\big\}\big) \to 0$ as $n, m \to +\infty$. Hence, there is $n_k \in \mathbb{N}$ so that $\mu\big(\big\{x \in A \,\big|\, |f_n(x) - f_m(x)| \geq \frac{1}{2^k}\big\}\big) < \frac{1}{2^k}$ for every $n, m \geq n_k$. Since we may assume that each n_k is as large as we like, we may inductively take (n_k) so that $n_k < n_{k+1}$ for every k. Thus, (f_{n_k}) is a subsequence of (f_n), and from the construction of n_k and from $n_k < n_{k+1}$ we get $\mu\big(\big\{x \in A \,\big|\, |f_{n_{k+1}}(x) - f_{n_k}(x)| \geq \frac{1}{2^k}\big\}\big) < \frac{1}{2^k}$ for every k. For simplicity we write

$$E_k = \big\{x \in A \,\big|\, |f_{n_{k+1}}(x) - f_{n_k}(x)| \geq \tfrac{1}{2^k}\big\},$$

and so $\mu(E_k) < \frac{1}{2^k}$ for all k. We also define the *subsets* of A:

$$F_m = \bigcup_{k=m}^{+\infty} E_k, \quad F = \bigcap_{m=1}^{+\infty} F_m = \overline{\lim}\, E_k.$$

Then

$$\mu(F) \leq \mu(F_m) \leq \sum_{k=m}^{+\infty} \mu(E_k) < \sum_{k=m}^{+\infty} \tfrac{1}{2^k} = \tfrac{1}{2^{m-1}}$$

for every m. This implies $\mu(F) = 0$.
Now, let $x \in A \setminus F$. Then there is m so that $x \in A \setminus F_m$, which implies that $x \in A \setminus E_k$ for all $k \geq m$. Therefore, $|f_{n_{k+1}}(x) - f_{n_k}(x)| < \frac{1}{2^k}$ for all $k \geq m$, and so

$$\sum_{k=m}^{+\infty} |f_{n_{k+1}}(x) - f_{n_k}(x)| < \sum_{k=m}^{+\infty} \tfrac{1}{2^k} = \tfrac{1}{2^{m-1}}.$$

Thus, the series $\sum_{k=m}^{+\infty}(f_{n_{k+1}}(x) - f_{n_k}(x))$ converges for every $x \in A \setminus F$, and we may define $f : X \to \mathbb{C}$ by

$$f = \begin{cases} f_{n_1} + \sum_{k=1}^{+\infty}(f_{n_{k+1}} - f_{n_k}), & \text{on } A \setminus F \\ 0, & \text{on } A^c \cup F \end{cases}$$

From

$$f = f_{n_1} + \lim_{K \to +\infty} \sum_{k=1}^{K-1}(f_{n_{k+1}} - f_{n_k}) = \lim_{K \to +\infty} f_{n_K}$$

on $A \setminus F$ and from $\mu(F) = 0$, we get that (f_{n_k}) converges to f a.e. on A.
Now, on $A \setminus F_m$ we have

$$\begin{aligned}|f_{n_m} - f| &= |f_{n_m} - f_{n_1} - \sum_{k=1}^{+\infty}(f_{n_{k+1}} - f_{n_k})| \\ &= |\sum_{k=1}^{m-1}(f_{n_{k+1}} - f_{n_k}) - \sum_{k=1}^{+\infty}(f_{n_{k+1}} - f_{n_k})| \leq \sum_{k=m}^{+\infty} |f_{n_{k+1}} - f_{n_k}| < \tfrac{1}{2^{m-1}}.\end{aligned}$$

Therefore, $\{x \in A \,\big|\, |f_{n_m}(x) - f(x)| \geq \frac{1}{2^{m-1}}\} \subseteq F_m$. This implies that

$$\mu\big(\{x \in A \,\big|\, |f_{n_m}(x) - f(x)| \geq \tfrac{1}{2^{m-1}}\}\big) \leq \mu(F_m) < \tfrac{1}{2^{m-1}}.$$

Now, we consider an arbitrary $\epsilon > 0$ and $m_0 \in \mathbb{N}$ large enough so that $\frac{1}{2^{m_0-1}} \leq \epsilon$. Then for every $m \geq m_0$ we have

$$\{x \in A \,|\, |f_{n_m}(x) - f(x)| \geq \epsilon\} \subseteq \{x \in A \,\big|\, |f_{n_m}(x) - f(x)| \geq \tfrac{1}{2^{m-1}}\},$$

and so

$$\mu(\{x \in A \,|\, |f_{n_m}(x) - f(x)| \geq \epsilon\}) < \tfrac{1}{2^{m-1}} \to 0$$

as $m \to +\infty$. Therefore, $f_{n_k} \to f$ in measure on A.
And, finally, since $n_k \to +\infty$ as $k \to +\infty$, we have that

$$\begin{aligned}\mu(\{x \in A \,|\, |f_k(x) - f(x)| \geq \epsilon\}) &\leq \mu\big(\{x \in A \,\big|\, |f_k(x) - f_{n_k}(x)| \geq \tfrac{\epsilon}{2}\}\big) \\ &\quad + \mu\big(\{x \in A \,\big|\, |f_{n_k}(x) - f(x)| \geq \tfrac{\epsilon}{2}\}\big) \to 0\end{aligned}$$

as $k \to +\infty$, and we conclude that $f_n \to f$ in measure on A. □

Example. We consider the example just after Theorem 5.1. If $0 < \epsilon \leq 1$, then the sequence of the values of $m_1(\{x \in (0,1) \,|\, |f_n(x)| \geq \epsilon\})$ is $1, \frac{1}{2}, \frac{1}{2}, \frac{1}{3}, \frac{1}{3}, \frac{1}{3}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \ldots$, and converges to 0. Hence, $f_n \to 0$ in measure on $(0,1)$. But, as we have seen, it is not true that $f_n \to 0$ a.e. on $(0,1)$.

Exercises.

Except if specified otherwise, all exercises refer to a measure space $(X, \mathcal{S}, \mu)$, all sets belong to $\mathcal{S}$ and all functions are $\mathcal{S}$-measurable.

5.3.1. If $\phi : \mathbb{C} \to \mathbb{C}$ is uniformly continuous and $f_n \to f$ in measure on A, prove that $\phi \circ f_n \to \phi \circ f$ in measure on A.

5.3.2. If $f_n \to f$ in measure on A and $|f_n| \leq g$ a.e. on A for all n, prove that $|f| \leq g$ a.e. on A.

5.3.3. Look at exercises 5.1.3 and 5.2.3. If $E_n \subseteq A$ for all n and $\chi_{E_n} \to f$ in measure on A, prove that there exists $E \subseteq A$ so that $f = \chi_E$ a.e. on A. Prove that $\mu(E_n \triangle E) \to 0$.

5.3.4. Let $\sharp$ be the counting measure on $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$. Prove that $f_n \to f$ uniformly on $\mathbb{N}$ if and only if $f_n \to f$ in measure on $\mathbb{N}$.

5.3.5. *A variation of the Lemma of Fatou.*
If $f_n \geq 0$ a.e. on A for all n and $f_n \to f$ in measure on A, prove $\int_A f \, d\mu \leq \underline{\lim}_{n \to +\infty} \int_A f_n \, d\mu$.

5.3.6. (i) If $\mu(A) < +\infty$ and $\sup_{n \in \mathbb{N}} |h_n(x)| < \infty$ for a.e. $x \in A$, prove that for every $\delta > 0$ there is a $B \subseteq A$ so that $\mu(A \setminus B) < \delta$ and $\sup_{x \in B, n \in \mathbb{N}} |h_n(x)| < +\infty$.
(ii) If $\mu(A) < +\infty$, and $f_n \to f$ and $g_n \to g$ in measure on A, prove that $f_n g_n \to fg$ in measure on A.

5.3.7. (i) If $\mu(A) < +\infty$, prove that $f_n \to f$ in measure on A if and only if $\int_A \frac{|f_n - f|}{1 + |f_n - f|} \, d\mu \to 0$.
(ii) Prove that $f_n \to f$ in measure on A if and only if $\inf_{\epsilon > 0} \frac{\epsilon + \mu(\{x \in A \,|\, |f_n(x) - f(x)| \geq \epsilon\})}{1 + \epsilon + \mu(\{x \in A \,|\, |f_n(x) - f(x)| \geq \epsilon\})} \to 0$.

5.3.8. If $f_n \to f$ in measure on A, prove that $\lambda_{f_n}(t) \to \lambda_f(t)$ for every $t \in [0, +\infty)$ which is a point of continuity of λ_f.

5.4 Almost uniform convergence.

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

Definition. *Let all $f, f_n : X \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ be $\mathcal{S}$-measurable, and all f, f_n be finite a.e. on A.*
*We say that $f_n \to f$ **almost uniformly** on $A \in \mathcal{S}$ if for every $\delta > 0$ there is $B \in \mathcal{S}$, $B \subseteq A$, so that $\mu(A \setminus B) < \delta$ and $f_n \to f$ uniformly on B.*
*We say that (f_n) is **Cauchy almost uniformly** on $A \in \mathcal{S}$ if for every $\delta > 0$ there is $B \in \mathcal{S}$, $B \subseteq A$, so that $\mu(A \setminus B) < \delta$ and (f_n) is Cauchy uniformly on B.*

Proposition 5.6. *If $f_n \to f'$ and $f_n \to f''$ almost uniformly on A, then $f' = f''$ a.e. on A.*

Proof. Let us assume that $\mu(E) > 0$, where $E = \{x \in A \,|\, f'(x) \neq f''(x)\}$.
There is $B \in \mathcal{S}$, $B \subseteq A$ so that $\mu(A \setminus B) < \frac{\mu(E)}{2}$ and $f_n \to f'$ uniformly on B. Similarly, there is $C \in \mathcal{S}$, $C \subseteq A$ so that $\mu(A \setminus C) < \frac{\mu(E)}{2}$ and $f_n \to f''$ uniformly on C. We consider $D = B \cap C$, and we have that $\mu(A \setminus D) < \mu(E)$ and $f_n \to f'$ and $f_n \to f''$ uniformly on D. Of course this implies that $f' = f''$ on D, and so $D \cap E = \emptyset$.
But then $E \subseteq A \setminus D$. Therefore, $\mu(E) \leq \mu(A \setminus D) < \mu(E)$, and we arrive at a contradiction. $\square$

Proposition 5.7. *Let $f_n \to f$ and $g_n \to g$ almost uniformly on A. Then*
(i) $f_n + g_n \to f + g$ almost uniformly on A.
(ii) $\lambda f_n \to \lambda f$ almost uniformly on A.
(iii) If there is $M < +\infty$ so that $|f_n| \leq M$ a.e. on A for all n, then $|f| \leq M$ a.e. on A.
(iv) If there is $M < +\infty$ so that $|f_n| \leq M$ a.e. on A and $|g_n| \leq M$ a.e. on A for all n, then $f_n g_n \to fg$ almost uniformly on A.

Proof. (i) For each $\delta > 0$, there is $B' \in \mathcal{S}$, $B' \subseteq A$, so that $\mu(A \setminus B') < \frac{\delta}{2}$ and $f_n \to f$ uniformly on B', and there is $B'' \in \mathcal{S}$, $B'' \subseteq A$, so that $\mu(A \setminus B'') < \frac{\delta}{2}$ and $g_n \to g$ uniformly on B''. We consider $B = B' \cap B''$, and then $\mu(A \setminus B) < \delta$ and $f_n \to f$ and $g_n \to g$ uniformly on B. Then $f_n + g_n \to f + g$ uniformly on B, and, since δ is arbitrary, we conclude that $f_n + g_n \to f + g$ almost uniformly on A.
(ii) This is easier, since, if $f_n \to f$ uniformly on B, then $\lambda f_n \to \lambda f$ uniformly on B.
(iii) Let us assume that $\mu(E) > 0$, where $E = \{x \in A \,|\, |f(x)| > M\}$.
There is $B \in \mathcal{S}$, $B \subseteq A$, so that $\mu(A \setminus B) < \mu(E)$ and $f_n \to f$ uniformly on B. Then we have $|f| \leq M$ a.e. on B, and so $\mu(B \cap E) = 0$. Now, $\mu(E) = \mu(E \setminus B) \leq \mu(A \setminus B) < \mu(E)$, and we arrive at a contradiction.
(iv) Exactly as in the proof of (i), for every $\delta > 0$ there is $B_1 \in \mathcal{S}$, $B_1 \subseteq A$, so that $\mu(A \setminus B_1) < \delta$ and $f_n \to f$ and $g_n \to g$ uniformly on B_1. By the result of (iii), we have $|f| \leq M$ a.e. on A. Hence, there is $B_2 \in \mathcal{S}$, $B_2 \subseteq A$ so that $\mu(A \setminus B_2) = 0$ and $|f_n|, |g_n|, |f| \leq M$ on B_2. We consider $B = B_1 \cap B_2$, and then $\mu(A \setminus B) = \mu(A \setminus B_1) < \delta$. Now, on B we have that

$$|f_n g_n - fg| \leq |f_n g_n - f g_n| + |f g_n - fg| \leq M|f_n - f| + M|g_n - g|,$$

and so $f_n g_n \to fg$ uniformly on B. We conclude that $f_n g_n \to fg$ almost uniformly on A. $\square$

One should notice the difference between the next result and the corresponding Theorems 5.1 and 5.2 for the other two types of convergence: if a sequence converges in the mean or in measure, then a.e. convergence holds for *some* subsequence, while, if it converges almost uniformly, then a.e. convergence holds for the whole sequence (and so for *every* subsequence).

Before the next result let us consider a simple general fact.

Assume that there is a collection of functions $g_i : B_i \to \mathbb{C}$, indexed by the set I of indices, where $B_i \subseteq X$ for every $i \in I$, and that $f_n \to g_i$ pointwise on B_i for every $i \in I$. If $x \in B_i \cap B_j$ for any $i, j \in I$, then by the uniqueness of pointwise limits we have that $g_i(x) = g_j(x)$. Therefore, all limit functions have the same value at each point of the union $B = \bigcup_{i\in I} B_i$ of the domains of definition. Hence, we can define a single function $f : B \to \mathbb{C}$ by $f(x) = g_i(x)$, where $i \in I$ is any index for which $x \in B_i$, and then clearly $f_n \to f$ pointwise on B.

Theorem 5.3. *If (f_n) is Cauchy almost uniformly on A, then there is an $f : X \to \mathbb{C}$ so that $f_n \to f$ almost uniformly on A. Moreover, $f_n \to f$ a.e. on A.*
Corollary: if $f_n \to f$ almost uniformly on A, then $f_n \to f$ a.e. on A.

Proof. For every $k \in \mathbb{N}$ there exists $B_k \in \mathcal{S}$, $B_k \subseteq A$ so that $\mu(A \setminus B_k) < \frac{1}{k}$ and (f_n) is Cauchy uniformly on B_k. Therefore, there is a function $g_k : B_k \to \mathbb{C}$ so that $f_n \to g_k$ uniformly and, hence, pointwise on B_k.
By the general result of the paragraph just before this theorem, there is an $f : B \to \mathbb{C}$, where $B = \bigcup_{k=1}^{+\infty} B_k$, so that $f_n \to f$ pointwise on B. But $\mu(A \setminus B) \leq \mu(A \setminus B_k) < \frac{1}{k}$ for every k, and so $\mu(A \setminus B) = 0$. If we extend $f : X \to \mathbb{C}$ by defining $f = 0$ on B^c, we conclude that $f_n \to f$ a.e. on A.
By the general construction of f, we have that $g_k = f$ on B_k, and so $f_n \to f$ uniformly on B_k. If $\delta > 0$ is arbitrary, we just take k large enough so that $\frac{1}{k} \leq \delta$, and we have that $\mu(A \setminus B_k) < \delta$. Therefore, $f_n \to f$ almost uniformly on A. ☐

Exercises.

Except if specified otherwise, all exercises refer to a measure space $(X, \mathcal{S}, \mu)$, all sets belong to $\mathcal{S}$ and all functions are $\mathcal{S}$-measurable.

5.4.1. If $\phi : \mathbb{C} \to \mathbb{C}$ is uniformly continuous and $f_n \to f$ almost uniformly on A, prove that $\phi \circ f_n \to \phi \circ f$ almost uniformly on A.

5.4.2. If $f_n \to f$ almost uniformly on A and $|f_n| \leq g$ a.e. on A for all n, prove that $|f| \leq g$ a.e. on A.

5.5 Relations between various types of convergence.

In this section we shall see three results describing some relations between the four types of convergence: a.e. convergence, convergence in the mean, convergence in measure, and almost uniform convergence. Many other results are consequences of these.

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

Theorem 5.4. *If $f_n \to f$ almost uniformly on A, then $f_n \to f$ a.e. on A.*
The converse is true under the additional assumption that either
(i) **(Egoroff)** *$\mu(A) < +\infty$ and all f, f_n are finite a.e. on A*
or
(ii) there is a $g : A \to [0, +\infty]$ so that $\int_A g\, d\mu < +\infty$ and $|f_n| \leq g$ a.e. on A for every n.

Proof. The first statement is included in Theorem 5.3.$\mu(A) < +\infty$
(i) Let $\mu(A) < +\infty$ and $f_n \to f$ a.e. on A and all f, f_n be finite a.e. on A.
For each $k, n \in \mathbb{N}$ we consider

$$E_n(k) = \bigcup_{m=n}^{+\infty} \{x \in A \,|\, |f_m(x) - f(x)| > \tfrac{1}{k}\}.$$

If $C = \{x \in A \,|\, f_n(x) \to f(x)\}$, then it is easy to see that $\bigcap_{n=1}^{+\infty} E_n(k) \subseteq A \setminus C$. Since $\mu(A \setminus C) = 0$, we get $\mu(\bigcap_{n=1}^{+\infty} E_n(k)) = 0$ for every k. From $E_n(k) \downarrow \bigcap_{n=1}^{+\infty} E_n(k)$, from $\mu(A) < +\infty$ and from the continuity of μ from above, we get that $\mu(E_n(k)) \to 0$ as $n \to +\infty$.

Hence, for an arbitrary $\delta > 0$ there is $n_k \in \mathbb{N}$ so that $\mu(E_{n_k}(k)) < \frac{\delta}{2^k}$.
We define

$$E = \bigcup_{k=1}^{+\infty} E_{n_k}(k), \quad B = A \setminus E,$$

and we have

$$\mu(E) \leq \sum_{k=1}^{+\infty} \mu(E_{n_k}(k)) < \delta.$$

Also, for every $x \in B$ and for every $k \geq 1$ we have that $|f_m(x) - f(x)| \leq \frac{1}{k}$ for all $m \geq n_k$ or, equivalently, $\sup_{x \in B} |f_m(x) - f(x)| \leq \frac{1}{k}$ for all $m \geq n_k$. This implies, of course, that $f_n \to f$ uniformly on B. Since $\mu(A \setminus B) = \mu(E) < \delta$, we conclude that $f_n \to f$ almost uniformly on A.

(ii) Let $\int_A g\,d\mu < +\infty$ and $f_n \to f$ a.e. on A and $|f_n| \leq g$ a.e. on A for all n.
Then $|f| \leq g$ a.e. on A and, since $\int_A g\,d\mu < +\infty$, all f, f_n are finite a.e. on A. Therefore, $|f_n - f| \leq 2g$ a.e. on A for all n. Using the same notation as in the proof of (i), this implies that there is an $F \subseteq E_n(k)$, $F \in \mathcal{S}$, so that $E_n(k) \setminus F \subseteq \{x \in A \,|\, g(x) > \frac{1}{2k}\}$, and so

$$\mu(E_n(k)) = \mu(E_n(k) \setminus F) \leq \mu\big(\{x \in A \,|\, g(x) > \tfrac{1}{2k}\}\big).$$

Now, it is clear that $\int_A g\,d\mu < +\infty$ implies $\mu\big(\{x \in A \,|\, g(x) > \frac{1}{2k}\}\big) < +\infty$. Therefore, we may apply again the continuity of μ from above to find that $\mu(E_n(k)) \to 0$ as $n \to +\infty$. From this point we repeat the proof of (i) word for word. $\square$

Example. If $f_n = \chi_{(n,n+1)}$ for every $n \geq 1$, then $f_n \to 0$ everywhere on $\mathbb{R}$, but it is not true that $f_n \to 0$ almost uniformly on $\mathbb{R}$.
Indeed, if $0 < \delta \leq 1$, then every Lebesgue measurable $B \subseteq \mathbb{R}$ with $m_1(\mathbb{R} \setminus B) < \delta$ has non-empty intersection with every interval $(n, n+1)$, and so $\sup_{x \in B} |f_n(x)| \geq 1$ for every n.
In this example we have $m_1(\mathbb{R}) = +\infty$, and it is easy to see that there is *no* $g : \mathbb{R} \to [0, +\infty]$ so that $\int_{\mathbb{R}} g(x)\,dm_1(x) < +\infty$ and $f_n \leq g$ a.e. on $\mathbb{R}$ for every n. In fact, if $f_n \leq g$ a.e. on $\mathbb{R}$ for every n, then $g \geq 1$ a.e. on $(1, +\infty)$.

Theorem 5.5. *If $f_n \to f$ almost uniformly on A, then $f_n \to f$ in measure on A.*
Conversely, if $f_n \to f$ in measure on A, then there is a subsequence (f_{n_k}) so that $f_{n_k} \to f$ almost uniformly on A.

Proof. Let $f_n \to f$ almost uniformly on A. We take an arbitrary $\epsilon > 0$.
Then for every $\delta > 0$ there is $B \in \mathcal{S}$, $B \subseteq A$, so that $\mu(A \setminus B) < \delta$ and $f_n \to f$ uniformly on B. Now, there exists an $n_0 \in \mathbb{N}$ so that $|f_n(x) - f(x)| < \epsilon$ for all $n \geq n_0$ and every $x \in B$. Therefore, $\{x \in A \,|\, |f_n(x) - f(x)| \geq \epsilon\} \subseteq A \setminus B$ and, thus, $\mu(\{x \in A \,|\, |f_n(x) - f(x)| \geq \epsilon\}) < \delta$ for all $n \geq n_0$. This implies that $\mu(\{x \in A \,|\, |f_n(x) - f(x)| \geq \epsilon\}) \to 0$ as $n \to +\infty$, and so $f_n \to f$ in measure on A.

The idea for the converse is already in the proof of Theorem 5.2.
Let $f_n \to f$ in measure on A.
Then for all $k \in \mathbb{N}$ we have $\mu\big(\{x \in A \,|\, |f_n(x) - f(x)| \geq \frac{1}{2^k}\}\big) \to 0$ as $n \to +\infty$. Hence, there is $n_k \in \mathbb{N}$ so that $\mu\big(\{x \in A \,|\, |f_n(x) - f(x)| \geq \frac{1}{2^k}\}\big) < \frac{1}{2^k}$ for all $n \geq n_k$, and we may also assume that $n_k < n_{k+1}$ for all k. Therefore, (f_{n_k}) is a subsequence of (f_n) such that

$$\mu\big(\{x \in A \,|\, |f_{n_k}(x) - f(x)| \geq \tfrac{1}{2^k}\}\big) < \tfrac{1}{2^k}$$

for all k. Now, we consider

$$E_k = \{x \in A \,|\, |f_{n_k}(x) - f(x)| \geq \tfrac{1}{2^k}\}, \quad F_m = \bigcup_{k=m}^{+\infty} E_k.$$

Then

$$\mu(F_m) \leq \sum_{k=m}^{+\infty} \mu(E_k) < \sum_{k=m}^{+\infty} \tfrac{1}{2^k} = \tfrac{1}{2^{m-1}}$$

for every m.
If $x \in A \setminus F_m$, then $x \in A \setminus E_k$ for every $k \geq m$, and then $|f_{n_k}(x) - f(x)| < \frac{1}{2^k}$ for every $k \geq m$. This implies that $\sup_{x \in A \setminus F_m} |f_{n_k}(x) - f(x)| \leq \frac{1}{2^k}$ for all $k \geq m$, and so $f_{n_k} \to f$ uniformly on $A \setminus F_m$. Since $\mu(F_m) < \frac{1}{2^{m-1}}$ for all m, we conclude that $f_{n_k} \to f$ almost uniformly on A. $\square$

Example. We consider the example just after Theorem 5.1. There, $f_n \to 0$ in measure on $(0,1)$ but it is not true that $f_n \to 0$ almost uniformly on $(0,1)$. In fact, if we take any δ with $0 < \delta \leq 1$, then every Lebesgue measurable $B \subseteq (0,1)$ with $m_1((0,1) \setminus B) < \delta$ must have non-empty intersection with infinitely many intervals of the form $(\frac{k-1}{m}, \frac{k}{m})$ (at least one for each value of m), and so $\sup_{x \in B} |f_n(x)| \geq 1$ for infinitely many n.

The converse in Theorem 5.6 is a variation of the Dominated Convergence Theorem.

Theorem 5.6. *If $f_n \to f$ in the mean on A, then $f_n \to f$ in measure on A.*
The converse is true under the additional assumption that there exists a $g : X \to [0,+\infty]$ so that $\int_A g\, d\mu < +\infty$ and $|f_n| \leq g$ a.e. on A.

Proof. If $f_n \to f$ in the mean on A, then

$$\mu(\{x \in A \,|\, |f_n(x) - f(x)| \geq \epsilon\}) \leq \tfrac{1}{\epsilon} \int_A |f_n - f|\, d\mu \to 0$$

as $n \to +\infty$. Therefore, $f_n \to f$ in measure on A.
Let us assume that the converse is not true. Then there is some $\epsilon_0 > 0$ and a subsequence (f_{n_k}) of (f_n) so that $\int_A |f_{n_k} - f|\, d\mu \geq \epsilon_0$ for every $k \geq 1$. Since $f_{n_k} \to f$ in measure, Theorem 5.2 implies that there is a subsequence $(f_{n_{k_l}})$ so that $f_{n_{k_l}} \to f$ a.e. on A. From $|f_{n_{k_l}}| \leq g$ a.e. on A we find that $|f| \leq g$ a.e. on A. Now the Dominated Convergence Theorem implies that $\int_A |f_{n_{k_l}} - f|\, d\mu \to 0$ as $l \to +\infty$, and we arrive at a contradiction. $\square$

Example. Let $f_n = n\chi_{(0,\frac{1}{n})}$ for all n. If $0 < \epsilon \leq 1$, then $\mu(\{x \in (0,1) \,|\, |f_n(x)| \geq \epsilon\}) = \frac{1}{n} \to 0$ as $n \to +\infty$, and so $f_n \to 0$ in measure on $(0,1)$. But $\int_0^1 |f_n|\, dm_1 = 1$, and so it is not true that $f_n \to 0$ in the mean on $(0,1)$.
On the other hand, there can be no $g : (0,1) \to [0,+\infty]$ so that $\int_0^1 g\, dm_1 < +\infty$ and $|f_n| \leq g$ a.e. on $(0,1)$ for all n. Otherwise, we would have that $g \geq n$ a.e. on each interval $[\frac{1}{n+1}, \frac{1}{n})$, and so

$$\int_0^1 g\, dm_1 \geq \sum_{n=1}^{+\infty} \int_{1/(n+1)}^{1/n} n\, dm_1 = \sum_{n=1}^{+\infty} n\big(\tfrac{1}{n} - \tfrac{1}{n+1}\big) = \sum_{n=1}^{+\infty} \tfrac{1}{n+1} = +\infty,$$

resulting in a contradiction.

Exercises.

Except if specified otherwise, all exercises refer to a measure space $(X, \mathcal{S}, \mu)$, all sets belong to $\mathcal{S}$ and all functions are $\mathcal{S}$-measurable.

5.5.1. If $f_n \to f'$ with respect to any of the four types of convergence (a.e., in the mean, in measure, almost uniformly) on A and $f_n \to f''$ with respect to any other of the same four types of convergence, prove that $f' = f''$ a.e. on A.

5.5.2. Prove the Dominated Convergence Theorem using the second converse part of Theorem 5.4.

5.5.3. *A variation of the Dominated Convergence Theorem.*
Let $\int_A g\, d\mu < +\infty$ and $|f_n| \leq g$ a.e. on A and $f_n \to f$ in measure on A. Prove that $\int_A f_n\, d\mu \to \int_A f\, d\mu$.
Hint. One can follow three paths. One is to use Theorem 5.6. Another is to reduce to the case of a.e. convergence and use the original version of the theorem. The third path is to use almost uniform convergence.

5.5.4. *A variation of Egoroff's Theorem for continuous parameter.*
Let $\mu(A) < +\infty$ and $f : A \times [0,1] \to \mathbb{C}$ have the properties:
(a) $f(\cdot, y) : A \to \mathbb{C}$ is $\mathcal{S}$-measurable for every $y \in [0,1]$
(b) $f(x, \cdot) : [0,1] \to \mathbb{C}$ is continuous for every $x \in A$.

(i) If $\epsilon, \eta > 0$, prove that $\{x \in A \,|\, |f(x,y) - f(x,0)| \leq \epsilon$ for all $y < \eta\}$ belongs to $\mathcal{S}$.
(ii) Prove that for every $\delta > 0$ there is $B \subseteq A$ so that $\mu(A \backslash B) < \delta$ and $f(\cdot, y) \to f(\cdot, 0)$ uniformly on B as $y \to 0+$.

5.5.5. Prove the converse part of Theorem 5.6 using the converse part of Theorem 5.5.

5.5.6. *The exact relation between convergence in the mean and convergence in measure.*
In all that follows every f_n is integrable over A.
We say that the indefinite integrals of (f_n) are **uniformly absolutely continuous** on A if for every $\epsilon > 0$ there exists $\delta > 0$ so that $|\int_E f_n \, d\mu| < \epsilon$ for all n and all $E \subseteq A$ with $\mu(E) < \delta$.
We say that the indefinite integrals of (f_n) are **equicontinuous from above at** $\emptyset$ on A if for every sequence (E_k) of subsets of A with $E_k \downarrow \emptyset$ and for every $\epsilon > 0$ there exists k_0 so that $|\int_{E_k} f_n \, d\mu| < \epsilon$ for all $k \geq k_0$ and all n.
Prove **Vitali's Theorem**: $f_n \to f$ *in the mean on* A *if and only if* $f_n \to f$ *in measure on* A *and the indefinite integrals of* $(|f_n|)$ *are uniformly absolutely continuous on* A *and equicontinuous from above at* $\emptyset$ *on* A.
How is Theorem 5.6 related to Vitali's Theorem?

5.5.7. If $f : \mathbb{R}^n \to \mathbb{C}$ is continuous in each variable separately, prove that f is Lebesgue measurable.

Chapter 6

Signed measures and complex measures.

6.1 Signed measures.

Let $(X, \mathcal{S})$ be a measurable space.

Definition. *A function $\nu : \mathcal{S} \to \overline{\mathbb{R}}$ is called a* ***signed measure*** *on $(X, \mathcal{S})$ if*
(i) either $\nu(A) \neq -\infty$ for all $A \in \mathcal{S}$ or $\nu(A) \neq +\infty$ for all $A \in \mathcal{S}$,
(ii) $\nu(\emptyset) = 0$,
(iii) $\nu(\bigcup_{j=1}^{+\infty} A_j) = \sum_{j=1}^{+\infty} \nu(A_j)$ for all pairwise disjoint $A_1, A_2, \ldots \in \mathcal{S}$.
If $\nu(A) \in \mathbb{R}$ for every $A \in \mathcal{S}$, then ν is called a ***real measure****.*
It $\nu(A) \geq 0$ for every $A \in \mathcal{S}$, then ν is called a ***non-negative signed measure****. If $\nu(A) \leq 0$ for every $A \in \mathcal{S}$, then ν is called a* ***non-positive signed measure****.*

It is obvious that ν is a non-negative signed measure if and only if it is a measure. Also, ν is a non-negative signed measure if and only if $-\nu$ is a non-positive signed measure.

Proposition 6.1. *Let ν, ν_1, ν_2 be signed measures on $(X, \mathcal{S})$ and $\lambda \in \mathbb{R}$.*
(i) If either $\nu_1(A) \neq -\infty$, $\nu_2(A) \neq -\infty$ for all $A \in \mathcal{S}$ or $\nu_1(A) \neq +\infty$, $\nu_2(A) \neq +\infty$ for all $A \in \mathcal{S}$, then we can define $\nu_1 + \nu_2 : \mathcal{S} \to \overline{\mathbb{R}}$ by

$$(\nu_1 + \nu_2)(A) = \nu_1(A) + \nu_2(A) \quad \text{for all } A \in \mathcal{S}.$$

Then $\nu_1 + \nu_2$ is a signed measure on $(X, \mathcal{S})$.
(ii) We define the function $\lambda\nu : \mathcal{S} \to \overline{\mathbb{R}}$ by

$$(\lambda\nu)(A) = \lambda\, \nu(A), \quad \text{for all } A \in \mathcal{S}$$

(where we follow the convention: $0\,(\pm\infty) = 0$ whenever $\lambda = 0$ and $\nu(A) = \pm\infty$). Then $\lambda\nu$ is a signed measure on $(X, \mathcal{S})$.

Proof. Similar to the proof of Proposition 1.16. □

Definition. *Let ν, ν_1, ν_2 be signed measures on the measurable space $(X, \mathcal{S})$ and $\lambda \in \mathbb{R}$. The signed measures $\nu_1 + \nu_2$ and $\lambda\nu$ on $(X, \mathcal{S})$ which are defined in Proposition 6.1 are called* ***sum*** *of ν_1 and ν_2 and* ***product*** *of ν by λ.*

Example. Let μ_1, μ_2 be two measures on $(X, \mathcal{S})$.
If $\mu_2(X) < +\infty$, then $\mu_2(A) \leq \mu_2(X) < +\infty$ for every $A \in \mathcal{S}$. Then $\nu = \mu_1 - \mu_2$ is well-defined, since $\nu(A) = \mu_1(A) - \mu_2(A) \geq -\mu_2(A) > -\infty$ for all $A \in \mathcal{S}$, and ν is a signed measure on $(X, \mathcal{S})$
Similarly, if $\mu_1(X) < +\infty$, then $\nu = \mu_1 - \mu_2$ is a signed measure on $(X, \mathcal{S})$ with $\nu(A) < +\infty$ for all $A \in \mathcal{S}$.
Thus, *the difference of two measures, at least one of which is finite, is a signed measure*.
Clearly, the difference of two finite measures is a real measure.

Example. Let μ be a measure on $(X, \mathcal{S})$ and $f : X \to \overline{\mathbb{R}}$ be a measurable function such that the integral $\int_X f \, d\mu$ is defined. Lemma 3.10 says that $\int_A f \, d\mu$ is defined for every $A \in \mathcal{S}$. If we consider the function $\lambda : \mathcal{S} \to \overline{\mathbb{R}}$ defined by $\lambda(A) = \int_A f \, d\mu$ for all $A \in \mathcal{S}$, then Proposition 3.6 and Theorem 3.9 imply that λ is a signed measure on $(X, \mathcal{S})$.

Definition. *The signed measure λ which is defined in the last example is called the* ***indefinite integral*** *of f with respect to μ and it is denoted $f\mu$. Thus, the defining relation for $f\mu$ is*

$$(f\mu)(A) = \int_A f \, d\mu, \quad A \in \mathcal{S}.$$

In case $f \geq 0$ a.e. on X the signed measure $f\mu$ is a measure, since $(f\mu)(A) = \int_A f \, d\mu \geq 0$ for every $A \in \mathcal{S}$. Similarly, if $f \leq 0$ a.e. on X, the $f\mu$ is a non-positive signed measure.

Continuing the study of this example, we shall make a few remarks. That the $\int_X f \, d\mu$ is defined means that either $\int_X f^+ \, d\mu < +\infty$ or $\int_X f^- \, d\mu < +\infty$.

Let us consider the case $\int_X f^+ \, d\mu < +\infty$ first. Since $(f^+\mu)(X) = \int_X f^+ \, d\mu < +\infty$, the signed measure $f^+\mu$ is a *finite* measure. The signed measure $f^-\mu$ is a measure (not necessarily finite). Also, for every $A \in \mathcal{S}$ we have $(f^+\mu)(A) - (f^-\mu)(A) = \int_A f^+ \, d\mu - \int_A f^- \, d\mu = \int_A f \, d\mu = (f\mu)(A)$. Therefore, in the case $\int_X f^+ \, d\mu < +\infty$ the signed measure $f\mu$ is the difference of the measures $f^+\mu$ and $f^-\mu$ of which the first is finite:

$$f\mu = f^+\mu - f^-\mu.$$

Similarly, in the case $\int_X f^- \, d\mu < +\infty$ the signed measure $f\mu$ is the difference of the measures $f^+\mu$ and $f^-\mu$ of which the second is finite, since $(f^-\mu)(X) = \int_X f^- \, d\mu < +\infty$.

Property (iii) in the definition of a signed measure ν is called the σ-additivity of ν. It is trivial to see that a signed measure is also finitely additive.

A signed measure is *not*, in general, monotone: if $A, B \in \mathcal{S}$ and $A \subseteq B$, then $B = A \cup (B \setminus A)$ and, hence, $\nu(B) = \nu(A) + \nu(B \setminus A)$, but $\nu(B \setminus A)$ may not be ≥ 0.

Theorem 6.1. *Let ν be a signed measure on $(X, \mathcal{S})$.*
(i) Let $A, B \in \mathcal{S}$ and $A \subseteq B$. If $\nu(B) < +\infty$, then $\nu(A) < +\infty$ and, if $\nu(B) > -\infty$, then $\nu(A) > -\infty$. In particular, if $\nu(B) \in \mathbb{R}$, then $\nu(A) \in \mathbb{R}$.
(ii) If $A, B \in \mathcal{S}$, $A \subseteq B$ and $\nu(A) \in \mathbb{R}$, then $\nu(B \setminus A) = \nu(B) - \nu(A)$.
(iii) If $A_1, A_2, \ldots \in \mathcal{S}$ and $A_n \uparrow A$, then $\nu(A_n) \to \nu(A)$.
(iv) If $A_1, A_2, \ldots \in \mathcal{S}$, $\nu(A_N) \in \mathbb{R}$ for some N and $A_n \downarrow A$, then $\nu(A_n) \to \nu(A)$.

Proof. (i) We have $\nu(B) = \nu(A) + \nu(B \setminus A)$. If $\nu(A) = +\infty$, then $\nu(B \setminus A) > -\infty$, and so $\nu(B) = +\infty$. Similarly, if $\nu(A) = -\infty$, then $\nu(B \setminus A) < +\infty$, and so $\nu(B) = -\infty$.
The proofs of (ii), (iii), (iv) are the same as the proofs of the analogous parts of Theorem 1.4. □

Property (iii) is called **continuity from below** and property (iv) is called **continuity from above**.

6.2 The Hahn and Jordan decompositions.

Let $(X, \mathcal{S})$ be a measurable space.

Definition. *Let ν be a signed measure on $(X, \mathcal{S})$.*
(i) $P \in \mathcal{S}$ is called a ***non-negative set*** *for ν if $\nu(A) \geq 0$ for every $A \in \mathcal{S}$, $A \subseteq P$.*
(ii) $N \in \mathcal{S}$ is called a ***non-positive set*** *for ν if $\nu(A) \leq 0$ for every $A \in \mathcal{S}$, $A \subseteq N$.*
(iii) $Q \in \mathcal{S}$ is called a ***null set*** *for ν if $\nu(A) = 0$ for every $A \in \mathcal{S}$, $A \subseteq Q$.*

It is obvious that an element of $\mathcal{S}$ which is both a non-negative and a non-positive set for ν is a null set for ν. It is also obvious that, if μ is a measure, then every $A \in \mathcal{S}$ is a non-negative set for μ.

Proposition 6.2. *Let ν be a signed measure on $(X, \mathcal{S})$.*
(i) If P is a non-negative set for ν and $P' \in \mathcal{S}$, $P' \subseteq P$, then P' is a non-negative set for ν.
(ii) If $P_1, P_2, \dots$ are non-negative sets for ν, then $\bigcup_{k=1}^{+\infty} P_k$ is a non-negative set for ν.
The same results are also true for non-positive sets and for null sets for ν.

Proof. (i) Trivial.
(ii) Let $A \in \mathcal{S}$, $A \subseteq \bigcup_{k=1}^{+\infty} P_k$. We consider $A_1 = A \cap P_1$ and $A_k = A \cap (P_k \setminus (P_1 \cup \cdots \cup P_{k-1}))$ for $k \geq 2$. Then $A = \bigcup_{k=1}^{+\infty} A_k$, the $A_1, A_2, \ldots \in \mathcal{S}$ are pairwise disjoint, and $A_k \subseteq P_k$ for all k. We then have $\nu(A) = \sum_{k=1}^{+\infty} \nu(A_k) \geq 0$. □

Theorem 6.2. *Let ν be a signed measure on $(X, \mathcal{S})$. Then there exist a non-negative set P and a non-positive set N for ν which form a partition of X, i.e.*

$$P \cup N = X, \quad P \cap N = \emptyset.$$

If P_1 is a non-negative set and N_1 is a non-positive set for ν which form a partition of X, then $P \triangle P_1 = N \triangle N_1$ is a null set for ν.

Proof. We consider the case when $-\infty < \nu(A)$ for every $A \in \mathcal{S}$, and we define the quantity

$$\kappa = \inf\{\nu(N) \mid N \text{ is a non-positive set for } \nu\}.$$

Since $\emptyset$ is a non-positive set with $\nu(\emptyset) = 0$, we have that $\kappa \leq 0$.
Now, we consider a sequence (N_k) of non-positive sets for ν so that $\nu(N_k) \to \kappa$, and we consider the set $N = \bigcup_{k=1}^{+\infty} N_k$. By Proposition 6.1, N is a non-positive set for ν. Thus, $\nu(N \setminus N_k) \leq 0$ for every k, and so $\kappa \leq \nu(N) \leq \nu(N_k)$ for every k. Taking the limit as $k \to +\infty$, we find that $-\infty < \nu(N) = \kappa$.
Therefore, N is *a non-positive set for ν of minimal ν-measure*, and we shall prove that the set $P = X \setminus N$ is a non-negative set for ν.
Let us assume that P is not a non-negative set for ν. Then there is $A_0 \in \mathcal{S}$, $A_0 \subseteq P$, with $-\infty < \nu(A_0) < 0$. The set A_0 is not a non-positive set or, otherwise, the set $N \cup A_0$ would be a non-positive set with $\nu(N \cup A_0) = \nu(N) + \nu(A_0) < \nu(N)$, contradicting the minimality of N. Therefore, there is at least one subset of A_0 in $\mathcal{S}$ having positive ν-measure. This means that

$$\tau_0 := \sup\{\nu(B) \mid B \in \mathcal{S}, B \subseteq A_0\} > 0.$$

Since $0 < \frac{\tau_0}{\tau_0+1} < \tau_0$, there is a $B_1 \in \mathcal{S}$, $B_1 \subseteq A_0$ so that $0 < \frac{\tau_0}{\tau_0+1} < \nu(B_1) \leq \tau_0$. We consider $A_1 = A_0 \setminus B_1$, and we have that $-\infty < \nu(A_1) < \nu(A_1) + \nu(B_1) = \nu(A_0)$. Here we are using Theorem 6.1 to imply $\nu(A_1), \nu(B_1) \in \mathbb{R}$ from $\nu(A_0) \in \mathbb{R}$.
Let us suppose that we have constructed sets $A_0, A_1, \ldots, A_n \in \mathcal{S}$ and $B_1, \ldots, B_n \in \mathcal{S}$ so that

$$\begin{aligned} &A_n \subseteq A_{n-1} \subseteq \cdots \subseteq A_1 \subseteq A_0 \subseteq N, \quad B_n = A_{n-1} \setminus A_n, \ldots, B_1 = A_0 \setminus A_1, \\ &\tau_{k-1} := \sup\{\nu(B) \mid B \in \mathcal{S}, B \subseteq A_{k-1}\} > 0 \quad \text{for all } k = 1, \ldots, n, \\ &0 < \tfrac{\tau_{k-1}}{\tau_{k-1}+1} < \nu(B_k) \leq \tau_{k-1} \quad \text{for all } k = 1, \ldots, n, \\ &-\infty < \nu(A_n) < \nu(A_{n-1}) < \cdots < \nu(A_1) < \nu(A_0) < 0 < +\infty. \end{aligned} \tag{6.1}$$

Now, A_n is not a non-positive set for ν for the same reason that A_0 is not a non-positive set for ν. Therefore, there is at least one subset of A_n in $\mathcal{S}$ having positive ν-measure. This means that

$$\tau_n := \sup\{\nu(B) \mid B \in \mathcal{S}, B \subseteq A_n\} > 0.$$

Then there is $B_{n+1} \in \mathcal{S}$, $B_{n+1} \subseteq A_n$ so that $0 < \frac{\tau_n}{\tau_n+1} < \nu(B_{n+1}) \le \tau_n$. We consider $A_{n+1} = A_n \setminus B_{n+1}$, and we have that $-\infty < \nu(A_{n+1}) < \nu(A_{n+1}) + \nu(B_{n+1}) = \nu(A_n)$.
Thus, we have constructed, inductively, two sequences (A_n), (B_n) satisfying properties (6.1).
Now the sets $B_1, B_2, \ldots$ and $\bigcap_{n=1}^{+\infty} A_n$ are pairwise disjoint and $A_0 = (\bigcap_{n=1}^{+\infty} A_n) \cup (\bigcup_{n=1}^{+\infty} B_n)$.
Therefore,

$$\nu(A_0) = \nu(\textstyle\bigcap_{n=1}^{+\infty} A_n) + \sum_{n=1}^{+\infty} \nu(B_n),$$

from which we get $\sum_{n=1}^{+\infty} \nu(B_n) < +\infty$. This implies that $\nu(B_n) \to 0$, and, by the third property (6.1), we have that $\tau_{n-1} \to 0$.
By continuity from above of ν, the set $A = \bigcap_{n=1}^{+\infty} A_n \in \mathcal{S}$ has $\nu(A) = \lim_{n\to+\infty} \nu(A_n) < 0$.
Moreover, A is not a non-positive set for ν for the same reason that A_0 is not a non-positive set for ν. Therefore, there is a $B \in \mathcal{S}$, $B \subseteq A$ so that $\nu(B) > 0$. But then $B \subseteq A_n$ for all n, and so $0 < \nu(B) \le \tau_n$ for all n. We, thus, arrive at a contradiction with the limit $\tau_n \to 0$.
In the same way we can prove that, in the case when $\nu(A) < +\infty$ for every $A \in \mathcal{S}$, there is a non-negative set P for ν of maximal ν-measure, and then that the set $N = X \setminus P$ is a non-positive set for ν.
Thus, in any case there exist a non-negative set P and a non-positive set N for ν which form a partition of X.

Now, let P_1 be a non-negative set and N_1 be a non-positive set for ν so that $P_1 \cup N_1 = X$ and $P_1 \cap N_1 = \emptyset$. Then, since $P \setminus P_1 = N_1 \setminus N \subseteq P \cap N_1$, the set $P \setminus P_1 = N_1 \setminus N1$ is both a non-negative set and a non-positive set for ν, and so it is a null set for ν. Similarly, $P_1 \setminus P = N \setminus N_1$ is a null set for ν, and we conclude that their union, i.e. $P \triangle P_1 = N \triangle N_1$, is a null set for ν. $\square$

Definition. *Let ν be a signed measure on $(X, \mathcal{S})$. Every partition of X into a non-negative set and a non-positive set for ν is called a **Hahn decomposition** of X for ν.*

Proposition 6.3. *Let ν be a signed measure on $(X, \mathcal{S})$.*
(i) If P and N constitute a Hahn decomposition of X for ν, then

$$\nu(P) = \max\{\nu(A) \mid A \in \mathcal{S}\}, \quad \nu(N) = \min\{\nu(A) \mid A \in \mathcal{S}\}.$$

(ii) If $\nu(A) < +\infty$ for every $A \in \mathcal{S}$, then ν is bounded from above. If $-\infty < \nu(A)$ for every $A \in \mathcal{S}$, then ν is bounded from below.

Proof. (i) If $A \in \mathcal{S}$, then $\nu(P \setminus A) \ge 0$, because $P \setminus A \subseteq P$. This implies

$$\nu(P) = \nu(P \cap A) + \nu(P \setminus A) \ge \nu(P \cap A) \ge 0$$

and, similarly, $\nu(N) \le \nu(N \cap A) \le 0$. Therefore,

$$\nu(A) = \nu(P \cap A) + \nu(N \cap A) \le \nu(P \cap A) \le \nu(P),$$

$$\nu(A) = \nu(P \cap A) + \nu(N \cap A) \ge \nu(N \cap A) \ge \nu(N).$$

(ii) This is a consequence of the result of (i). $\square$

Definition. *Let μ_1, μ_2 be two measures on $(X, \mathcal{S})$. We say that μ_1, μ_2 are **mutually singular** (or that μ_1 is singular to μ_2 or that μ_2 is singular to μ_1) if there exist $A_1, A_2 \in \mathcal{S}$ so that A_1 is null for μ_2 and A_2 is null for μ_1 and $A_1 \cup A_2 = X$, $A_1 \cap A_2 = \emptyset$.*
We use the symbol

$$\mu_1 \perp \mu_2$$

to denote that μ_1, μ_2 are mutually singular.

In other words, *two measures are mutually singular if there is a set in $\mathcal{S}$ which is null for one of them and its complement is null for the other.*

If μ_1, μ_2 are mutually singular and A_1, A_2 are as in the definition, then it is clear that $\mu_1(A) = \mu_1(A \cap A_1)$ and $\mu_2(A) = \mu_2(A \cap A_2)$ for every $A \in \mathcal{S}$. Thus, we may informally say that μ_1 *is concentrated on* A_1 and μ_2 *is concentrated on* A_2.

Theorem 6.3. *Let ν be a signed measure on $(X, \mathcal{S})$. There exist two measures ν^+ and ν^-, at least one of which is finite, so that*

$$\nu = \nu^+ - \nu^-, \quad \nu^+ \perp \nu^-.$$

If ν_1^+, ν_1^- are two measures on $(X, \mathcal{S})$, at least one of which is finite, so that $\nu = \nu_1^+ - \nu_1^-$ and $\nu_1^+ \perp \nu_1^-$, then $\nu_1^+ = \nu^+$ and $\nu_1^- = \nu^-$.

Proof. We consider any Hahn decomposition of X for ν: let P be a non-negative set and N be a non-positive set for ν so that $P \cup N = X$ and $P \cap N = \emptyset$.
We define $\nu^+, \nu^- : \mathcal{S} \to [0, +\infty]$ by

$$\nu^+(A) = \nu(A \cap P), \quad \nu^-(A) = -\nu(A \cap N) \quad \text{for every } A \in \mathcal{S}.$$

It is trivial to see that ν^+, ν^- are *measures* on $(X, \mathcal{S})$. If $\nu(A) < +\infty$ for every $A \in \mathcal{S}$, then $\nu^+(X) = \nu(P) < +\infty$, and so ν^+ is a finite measure. Similarly, if $-\infty < \nu(A)$ for every $A \in \mathcal{S}$, then $\nu^-(X) = -\nu(N) < +\infty$, and so ν^- is a finite measure.
Also,

$$\nu(A) = \nu(A \cap P) + \nu(A \cap N) = \nu^+(A) - \nu^-(A)$$

for all $A \in \mathcal{S}$, and so $\nu = \nu^+ - \nu^-$.
If $A \in \mathcal{S}$ and $A \subseteq N$, then $\nu^+(A) = \nu(A \cap P) = \nu(\emptyset) = 0$. Therefore, N is a null set for ν^+. Similarly, P is a null set for ν^-, and so $\nu^+ \perp \nu^-$.
Now, let ν_1^+, ν_1^- be two *measures* on $(X, \mathcal{S})$, at least one of which is finite, so that $\nu = \nu_1^+ - \nu_1^-$ and $\nu_1^+ \perp \nu_1^-$. Then there exist $P_1 \in \mathcal{S}$ which is null for ν_1^- and $N_1 \in \mathcal{S}$ which is null for ν_1^+ so that $P_1 \cup N_1 = X$ and $P_1 \cap N_1 = \emptyset$.
Then for every $A \in \mathcal{S}$ we have

$$\nu^+(A) = \nu(A \cap P) = \nu_1^+(A \cap P) - \nu_1^-(A \cap P) \le \nu_1^+(A \cap P) \le \nu_1^+(A).$$

Also,

$$\begin{aligned} \nu_1^+(A) &= \nu_1^+(A \cap P_1) + \nu_1^+(A \cap N_1) = \nu_1^+(A \cap P_1) = \nu_1^+(A \cap P_1) - \nu_1^-(A \cap P_1) \\ &= \nu(A \cap P_1) = \nu^+(A \cap P_1) - \nu^-(A \cap P_1) \le \nu^+(A \cap P_1) \le \nu^+(A). \end{aligned}$$

Hence, $\nu_1^+(A) = \nu^+(A)$ for every $A \in \mathcal{S}$, and so $\nu_1^+ = \nu^+$.
The proof of $\nu_1^- = \nu^-$ is similar. □

Definition. *Let ν be a signed measure on $(X, \mathcal{S})$. We say that the mutually singular measures ν^+, ν^-, whose existence and uniqueness is proved in Theorem 6.3, constitute the* ***Jordan decomposition*** *of ν.*
We call ν^+ the ***non-negative variation*** *of ν and ν^- the* ***non-positive variation*** *of ν. We recall from the proof of Theorem 6.3 that ν^+, ν^- are defined by*

$$\nu^+(A) = \nu(A \cap P), \quad \nu^-(A) = -\nu(A \cap N) \quad \textit{for every } A \in \mathcal{S},$$

where P, N constitute any Hahn decomposition of X for ν.
We call the measure $|\nu| = \nu^+ + \nu^-$ the ***absolute variation*** *of ν, and we call the quantity $|\nu|(X)$ the* ***total variation*** *of ν.*

For the measure $|\nu|$ we have that

$$|\nu|(A) = \nu^+(A) + \nu^-(A) = \nu(A \cap P) - \nu(A \cap N) \quad \text{for every } A \in \mathcal{S}.$$

We observe that the total variation of ν is equal to

$$|\nu|(X) = \nu(P) - \nu(N),$$

where the sets P, N constitute a Hahn decomposition of X for ν. Thus, *the total variation of ν is equal to the difference between the largest and the smallest values of ν.*

Moreover, *the total variation is finite if and only if the absolute variation is a finite measure if and only if both the non-negative and the non-positive variations are finite measures if and only if the signed measure takes only finite values.*

Proposition 6.4. *Let ν be a signed measure on $(X, \mathcal{S})$. Then*

$$\nu^+(A) = \max\{\nu(B) \mid B \in \mathcal{S}, B \subseteq A\}, \quad \nu^-(A) = -\min\{\nu(B) \mid B \in \mathcal{S}, B \subseteq A\}$$

for every $A \in \mathcal{S}$.

Proof. Let P, N constitute any Hahn decomposition of X for ν.
Then for every $B \in \mathcal{S}$, $B \subseteq A$ we have

$$\nu(B) = \nu(B \cap P) + \nu(B \cap N) \leq \nu(B \cap P) = \nu^+(B) \leq \nu^+(A).$$

On the other hand, if we consider $B_0 = A \cap P$, then we have $B_0 \in \mathcal{S}$, $B_0 \subseteq A$ and

$$\nu^+(A) = \nu(A \cap P) = \nu(B_0).$$

The proof of $\nu^-(A) = -\min\{\nu(B) \mid B \in \mathcal{S}, B \subseteq A\}$ is similar. □

Proposition 6.5. *Let ν, ν_1, ν_2 be signed measures on $(X, \mathcal{S})$. If $\nu_1 + \nu_2$ is defined, then*

$$(\nu_1 + \nu_2)^+ \leq \nu_1^+ + \nu_2^+, \quad (\nu_1 + \nu_2)^- \leq \nu_1^- + \nu_2^-, \quad |\nu_1 + \nu_2| \leq |\nu_1| + |\nu_2|.$$

Proof. Let P, N constitute any Hahn decomposition of X for $\nu_1 + \nu_2$. Then for every $A \in \mathcal{S}$ we have

$$\begin{aligned}(\nu_1 + \nu_2)^+(A) &= (\nu_1 + \nu_2)(A \cap P) = \nu_1(A \cap P) + \nu_2(A \cap P) \leq \nu_1^+(A \cap P) + \nu_2^+(A \cap P) \\ &\leq \nu_1^+(A) + \nu_2^+(A),\end{aligned}$$

and so $(\nu_1 + \nu_2)^+ \leq \nu_1^+ + \nu_2^+$.
The proof of $(\nu_1 + \nu_2)^- \leq \nu_1^- + \nu_2^-$ is similar, and then, adding the two inequalities, we get $|\nu_1 + \nu_2| \leq |\nu_1| + |\nu_2|$. □

Proposition 6.6. *Let ν be a signed measures on $(X, \mathcal{S})$ and $\kappa \in \mathbb{R}$.*
(i) If $\kappa \geq 0$, then $(\kappa\nu)^+ = \kappa\nu^+$ and $(\kappa\nu)^- = \kappa\nu^-$.
(ii) If $\kappa \leq 0$, then $(\kappa\nu)^+ = -\kappa\nu^-$ and $(\kappa\nu)^- = -\kappa\nu^+$.
(iii) $|\kappa\nu| = |\kappa||\nu|$.

Proof. (i) Let P, N constitute any Hahn decomposition of X for ν. If $\kappa \geq 0$, then P, N constitute a Hahn decomposition of X for $\kappa\nu$ as well. Hence, for every $A \in \mathcal{S}$ we have

$$(\kappa\nu)^+(A) = (\kappa\nu)(A \cap P) = \kappa\nu(A \cap P) = \kappa\nu^+(A),$$

$$(\kappa\nu)^-(A) = (\kappa\nu)(A \cap N) = \kappa\nu(A \cap N) = \kappa\nu^-(A).$$

(ii) The proof is similar: if $\kappa \leq 0$, then N, P constitute a Hahn decomposition of X for $\kappa\nu$.
(iii) A consequence of the results of (i) and (ii). □

Definition. *Let $A \in \mathcal{S}$. If $A_1, \dots, A_n \in \mathcal{S}$ are pairwise disjoint and $A = \bigcup_{k=1}^n A_k$, then $\{A_1, \dots, A_n\}$ is called a (finite)* ***measurable partition*** *of A.*

Theorem 6.4. *Let ν be a signed measure on $(X, \mathcal{S})$ and let $|\nu|$ be the absolute variation of ν. Then*

$$|\nu|(A) = \sup\left\{ \sum_{k=1}^n |\nu(A_k)| \,\middle|\, n \in \mathbb{N}, \{A_1, \dots, A_n\} \text{ is a measurable partition of } A \right\}$$

for every $A \in \mathcal{S}$.

Proof. Let $M = \sup\left\{ \sum_{k=1}^n |\nu(A_k)| \,\middle|\, n \in \mathbb{N}, \{A_1, \dots, A_n\} \text{ is a measurable partition of } A \right\}$, and let P, N constitute a Hahn decomposition of X for ν.
We have

$$|\nu(A)| = |\nu(A \cap P) + \nu(A \cap N)| \le |\nu(A \cap P)| + |\nu(A \cap N)| = \nu(A \cap P) - \nu(A \cap N) = |\nu|(A)$$

for every $A \in \mathcal{S}$. Therefore, if $\{A_1, \dots, A_n\}$ is any measurable partition of $A \in \mathcal{S}$, then we have

$$\sum_{k=1}^n |\nu(A_k)| \le \sum_{k=1}^n |\nu|(A_k) = |\nu|(A).$$

Hence, $M \le |\nu|(A)$.
Now, $\{A \cap P, A \cap N\}$ is a particular measurable partition of A for which we have

$$|\nu(A \cap P)| + |\nu(A \cap N)| = \nu(A \cap P) - \nu(A \cap N) = |\nu|(A).$$

Hence, $|\nu|(A) \le M$. □

It is useful to note something which appeared in the proof of Theorem 6.4, namely that

$$|\nu(A)| \le |\nu|(A) \quad \text{for all } A \in \mathcal{S}.$$

The next two propositions treat the special case of a signed measure which is the indefinite integral of a function with respect to a measure.

Proposition 6.7. *Let μ be a measure on $(X, \mathcal{S})$, $f : X \to \overline{\mathbb{R}}$ be measurable and $\int_X f \, d\mu$ be defined. Then the sets $P = \{x \in X \mid f(x) \ge 0\}$ and $N = \{x \in X \mid f(x) < 0\}$ constitute a Hahn decomposition of X for the signed measure $f\mu$. We also have*

$$(f\mu)^+ = f^+\mu, \quad (f\mu)^- = f^-\mu.$$

Thus, the indefinite integrals $f^+\mu$ and $f^-\mu$ constitute the Jordan decomposition of $f\mu$. Moreover,

$$|f\mu| = |f|\mu.$$

Proof. If $A \in \mathcal{S}$ and $A \subseteq P$, then $(f\mu)(A) = \int_A f \, d\mu \ge 0$, while, if $A \subseteq N$, then $(f\mu)(A) = \int_A f \, d\mu \le 0$. Therefore, P is a non-negative set and N is a non-positive set for $f\mu$. Since $P \cup N = X$ and $P \cap N = \emptyset$, we conclude that P, N constitute a Hahn decomposition of X for $f\mu$.
Now,

$$(f\mu)^+(A) = (f\mu)(A \cap P) = \int_{A \cap P} f \, d\mu = \int_A f \chi_P \, d\mu = \int_A f^+ \, d\mu = (f^+\mu)(A)$$

and, similarly,

$$(f\mu)^-(A) = -(f\mu)(A \cap N) = -\int_{A \cap N} f \, d\mu = -\int_A f \chi_N \, d\mu = \int_A f^- \, d\mu = (f^-\mu)(A)$$

for every $A \in \mathcal{S}$. Therefore, $(f\mu)^+ = f^+\mu$ and $(f\mu)^- = f^-\mu$.
Finally, $|f\mu| = (f\mu)^+ + (f\mu)^- = f^+\mu + f^-\mu = |f|\mu$. □

Clearly, another Hahn decomposition of X for $f\mu$ consists of the sets $P = \{x \in X \,|\, f(x) > 0\}$ and $N = \{x \in X \,|\, f(x) \le 0\}$.

Proposition 6.8. *Let μ be a measure on $(X, \mathcal{S})$, $f : X \to \overline{\mathbb{R}}$ be measurable and $\int_X f\, d\mu$ be defined. Let $E \in \mathcal{S}$.*
(i) E is a non-negative set for $f\mu$ if and only if $f \ge 0$ a.e. on E.
(ii) E is a non-positive set for $f\mu$ if and only if $f \le 0$ a.e. on E.
(iii) E is a null set for $f\mu$ if and only if $f = 0$ a.e. on E.

Proof. (i) Let $f \ge 0$ a.e. on E. If $A \in \mathcal{S}$, $A \subseteq E$, then $f \ge 0$ a.e. on A, and so $(f\mu)(A) = \int_A f\, d\mu \ge 0$. Thus, E is a non-negative set for $f\mu$.
Conversely, let E be a non-negative set for $f\mu$. If $\epsilon > 0$ and $A_\epsilon = \{x \in E \,|\, f(x) \le -\epsilon\}$, then

$$0 \le (f\mu)(A_\epsilon) = \int_{A_\epsilon} f\, d\mu \le -\epsilon\mu(A_\epsilon),$$

and so $\mu(A_\epsilon) = 0$. Now, we have that $\{x \in E \,|\, f(x) < 0\} = \bigcup_{n=1}^{+\infty} A_{1/n}$, and we conclude that $\mu(\{x \in E \,|\, f(x) < 0\}) = 0$. Therefore, $f \ge 0$ a.e. on E.
The proof of (ii) is identical to the proof of (i), and (iii) is a consequence of (i) and (ii). $\square$

Exercises.

6.2.1. Let ν be a signed measure on $(X, \mathcal{S})$ and let μ_1, μ_2 be two measures on $(X, \mathcal{S})$ at least one of which is finite. If $\nu = \mu_1 - \mu_2$, prove that $\nu^+ \le \mu_1$ and $\nu^- \le \mu_2$.

6.2.2. Let f be the Cantor function on $[0, 1]$ extended as 0 on $(-\infty, 0)$ and as 1 on $(1, +\infty)$ and let μ_f be the Lebesgue-Stieltjes measure on $(\mathbb{R}, \mathcal{B}_1)$ induced by f. Prove that $\mu_f \perp m_1$.

6.2.3. (i) Recall that for every $a \in \overline{\mathbb{R}}$ the non-negative part and the non-positive part of a are defined by $a^+ = \max\{a, 0\}$ and $a^- = -\min\{a, 0\}$. Prove that $(a+b)^+ \le a^+ + b^+$ and $(a+b)^- \le a^- + b^-$ for every $a, b \in \overline{\mathbb{R}}$ for which $a + b$ is defined.
(ii) Let ν be a signed measure on $(X, \mathcal{S})$ and let ν^+ and ν^- be the non-negative and the non-positive variation of ν, respectively. Prove

$$\nu^+(A) = \sup \big\{ \textstyle\sum_{k=1}^n \nu(A_k)^+ \,\big|\, n \in \mathbb{N}, \{A_1, \dots, A_n\} \text{ is a measurable partition of } A \big\},$$

$$\nu^-(A) = \sup \big\{ \textstyle\sum_{k=1}^n \nu(A_k)^- \,\big|\, n \in \mathbb{N}, \{A_1, \dots, A_n\} \text{ is a measurable partition of } A \big\}$$

for every $A \in \mathcal{S}$,

6.3 Complex measures.

Let $(X, \mathcal{S})$ be a measurable space.

Definition. *A function $\nu : \mathcal{S} \to \mathbb{C}$ is called a* ***complex measure*** *on $(X, \mathcal{S})$ if*
(i) $\nu(\emptyset) = 0$,
(ii) $\nu(\bigcup_{j=1}^{+\infty} A_j) = \sum_{j=1}^{+\infty} \nu(A_j)$ for every pairwise disjoint $A_1, A_2, \ldots \in \mathcal{S}$.

It is trivial to prove, taking real and imaginary parts, that the functions $\mathrm{Re}(\nu), \mathrm{Im}(\nu) : \mathcal{S} \to \mathbb{R}$, which are defined by $\mathrm{Re}(\nu)(A) = \mathrm{Re}(\nu(A))$ and $\mathrm{Im}(\nu)(A) = \mathrm{Im}(\nu(A))$ for every $A \in \mathcal{S}$, are real measures on $(X, \mathcal{S})$, and so they are bounded. That is, there is an $M < +\infty$ so that $|\mathrm{Re}(\nu)(A)| \le M$ and $|\mathrm{Im}(\nu)(A)| \le M$ for every $A \in \mathcal{S}$. This implies that $|\nu(A)| \le 2M$ for every $A \in \mathcal{S}$, and we have proved the

Proposition 6.9. *Let ν be a complex measure on $(X, \mathcal{S})$. Then ν is bounded, i.e. there is an $M < +\infty$ so that $|\nu(A)| \le M$ for every $A \in \mathcal{S}$.*

Proposition 6.10. *Let ν, ν_1, ν_2 be complex measures on $(X, \mathcal{S})$ and $\lambda \in \mathbb{C}$.*
(i) We define the function $\nu_1 + \nu_2 : \mathcal{S} \to \mathbb{C}$ by

$$(\nu_1 + \nu_2)(A) = \nu_1(A) + \nu_2(A) \quad \textit{for all } A \in \mathcal{S}.$$

Then $\nu_1 + \nu_2$ is a complex measure on $(X, \mathcal{S})$.
(ii) We define the function $\lambda\nu : \mathcal{S} \to \mathbb{C}$ by

$$(\lambda\nu)(A) = \lambda\, \nu(A), \quad \textit{for all } A \in \mathcal{S}.$$

Then $\lambda\nu$ is a complex measure on $(X, \mathcal{S})$.

Proof. Similar to the proof of Proposition 1.16 or of Proposition 6.1. □

Definition. *Let ν, ν_1, ν_2 be complex measures on the measurable space $(X, \mathcal{S})$ and $\lambda \in \mathbb{C}$. The complex measures $\nu_1 + \nu_2$ and $\lambda\nu$ on $(X, \mathcal{S})$ which are defined in Proposition 6.10 are called **sum** of ν_1 and ν_2 and **product** of ν by λ.*

In particular,

$$\nu = \mathrm{Re}(\nu) + i\, \mathrm{Im}(\nu).$$

Lemma 6.1. *Let $K \subseteq \mathbb{C}$ be finite. Then there is $M \subseteq K$, so that $|\sum_{\lambda \in M} \lambda| \geq \frac{1}{6} \sum_{\lambda \in K} |\lambda|$.*

Proof. $\mathbb{C}$ is the union of

$$Q_1 = \{\lambda \mid \mathrm{Re}(\lambda) \geq |\mathrm{Im}(\lambda)|\}, \quad Q_2 = \{\lambda \mid \mathrm{Re}(\lambda) \leq -|\mathrm{Im}(\lambda)|\},$$

$$Q_3 = \{\lambda \mid \mathrm{Im}(\lambda) \geq |\mathrm{Re}(\lambda)|\}, \quad Q_4 = \{\lambda \mid \mathrm{Im}(\lambda) \leq -|\mathrm{Re}(\lambda)|\}.$$

If $\lambda_1, \ldots, \lambda_n \in Q_1$, then

$$|\lambda_1 + \cdots + \lambda_n| \geq \mathrm{Re}(\lambda_1 + \cdots + \lambda_n) = \mathrm{Re}(\lambda_1) + \cdots + \mathrm{Re}(\lambda_n) \geq \tfrac{1}{\sqrt{2}}(|\lambda_1| + \cdots + |\lambda_n|).$$

The same is true if $\lambda_1, \ldots, \lambda_n$ all belong to one of Q_2, Q_3, Q_4.
Now, we split K in four pairwise disjoint subsets K_1, K_2, K_3, K_4, so that each contains elements of K in Q_1, Q_2, Q_3, Q_4, respectively. Then at least one of them, say M, satisfies

$$\textstyle\sum_{\lambda \in M} |\lambda| \geq \frac{1}{4} \sum_{\lambda \in K} |\lambda|.$$

and so

$$\textstyle\big|\sum_{\lambda \in M} \lambda\big| \geq \frac{1}{\sqrt{2}} \sum_{\lambda \in M} |\lambda| \geq \frac{1}{4\sqrt{2}} \sum_{\lambda \in K} |\lambda| \geq \frac{1}{6} \sum_{\lambda \in K} |\lambda|.$$

□

Proposition 6.11. *Let ν be a complex measure on $(X, \mathcal{S})$. If for every $A \in \mathcal{S}$ we define*

$$|\nu|(A) = \sup\big\{ \textstyle\sum_{k=1}^n |\nu(A_k)| \,\big|\, n \in \mathbb{N}, \{A_1, \ldots, A_n\} \textit{ is a measurable partition of } A\big\},$$

then the function $|\nu| : \mathcal{S} \to [0, +\infty]$ is a finite measure on $(X, \mathcal{S})$.

Proof. It is obvious that $|\nu|(\emptyset) = 0$.
Now, let $A^1, A^2, \ldots \in \mathcal{S}$ be pairwise disjoint and $A = \bigcup_{j=1}^{+\infty} A^j$.
If $\{A_1, \ldots, A_n\}$ is an arbitrary measurable partition of A, then, for every j, $\{A_1 \cap A^j, \ldots, A_n \cap A^j\}$ is a measurable partition of A^j. This implies,

$$\begin{aligned}\textstyle\sum_{k=1}^n |\nu(A_k)| &= \textstyle\sum_{k=1}^n |\sum_{j=1}^{+\infty} \nu(A_k \cap A^j)| \leq \sum_{k=1}^n \big(\sum_{j=1}^{+\infty} |\nu(A_k \cap A^j)|\big) \\ &= \textstyle\sum_{j=1}^{+\infty} \big(\sum_{k=1}^n |\nu(A_k \cap A^j)|\big) \leq \sum_{j=1}^{+\infty} |\nu|(A^j)\end{aligned}$$

Taking the supremum of the left side of this inequality, we get $|\nu|(A) \le \sum_{j=1}^{+\infty} |\nu|(A^j)$.
Now, we fix an arbitrary $N \in \mathbb{N}$ and for each $j = 1, \dots, N$ we consider any measurable partition $\{A_1^j, \dots, A_{n_j}^j\}$ of A^j. Then $\{A_1^1, \dots, A_{n_1}^1, \dots, A_1^N, \dots, A_{n_N}^N, \bigcup_{j=N+1}^{+\infty} A^j\}$ is a measurable partition of A, and so

$$|\nu|(A) \ge \sum_{j=1}^N \big(\sum_{k=1}^{n_j} |\nu(A_k^j)|\big) + \big|\nu\big(\bigcup_{j=N+1}^{+\infty} A^j\big)\big| \ge \sum_{j=1}^N \big(\sum_{k=1}^{n_j} |\nu(A_k^j)|\big).$$

Taking the supremum of the right side of this inequality, we get $|\nu|(A) \ge \sum_{j=1}^N |\nu|(A^j)$. Now, taking the limit as $N \to +\infty$, we find $|\nu|(A) \ge \sum_{j=1}^{+\infty} |\nu|(A^j)$.
Hence, $|\nu|(A) = \sum_{j=1}^{+\infty} |\nu|(A^j)$, and so $|\nu|$ is a measure on $(X, \mathcal{S})$.
Finally, we shall prove that $|\nu|$ is finite, i.e. that $|\nu|(X) < +\infty$. One way to prove this is to use the same result for real measures, considering the real measures $\mathrm{Re}(\nu)$ and $\mathrm{Im}(\nu)$. This is done as follows. We consider an arbitrary measurable partition $\{A_1, \dots, A_n\}$ of X, and we have

$$\sum_{k=1}^n |\nu(A_k)| \le \sum_{k=1}^n |\mathrm{Re}(\nu)(A_k)| + \sum_{k=1}^n |\mathrm{Im}(\nu)(A_k)| \le |\mathrm{Re}(\nu)|(X) + |\mathrm{Im}(\nu)|(X).$$

Taking the supremum of the left side of this inequality, we get

$$|\nu|(X) \le |\mathrm{Re}(\nu)|(X) + |\mathrm{Im}(\nu)|(X) < +\infty,$$

since the signed measures $\mathrm{Re}(\nu)$ and $\mathrm{Im}(\nu)$ have finite values.
Another way to prove that $|\nu|(X) < +\infty$ is the following.
We assume that $|\nu|(X) = +\infty$, and we claim that there are $B_1, B_2, \ldots \in \mathcal{S}$ so that

$$B_1 \supseteq B_2 \supseteq B_3 \supseteq \dots, \quad |\nu|(B_k) = +\infty, \quad |\nu(B_k)| \ge k - 1$$

for every k. We take $B_1 = X$ and we assume that we have proven the existence of the first $B_1, \dots, B_k$. Since $|\nu|(B_k) = +\infty$, there is a measurable partition $\{A_1, \dots, A_n\}$ of B_k so that

$$\sum_{m=1}^n |\nu(A_m)| \ge 6(|\nu(B_k)| + k).$$

According to Lemma 6.1, there are some of the $A_1, \dots, A_n$, which we may assume that they are the $A_1, \dots, A_l$, so that

$$\big|\sum_{m=1}^l \nu(A_m)\big| \ge \tfrac{1}{6} \sum_{m=1}^n |\nu(A_m)| \ge |\nu(B_k)| + k.$$

We consider $S = \bigcup_{m=1}^l A_m \subseteq B_k$, and then

$$|\nu(S)| \ge |\nu(B_k)| + k.$$

Since $|\nu|(S) + |\nu|(B_k \setminus S) = |\nu|(B_k) = +\infty$, we have that either $|\nu|(S) = +\infty$ or $|\nu|(B_k \setminus S) = +\infty$. In the first case we consider $B_{k+1} = S \subseteq B_k$, and then $|\nu(B_{k+1})| \ge |\nu(B_k)| + k \ge k$. In the second case we consider $B_{k+1} = B_k \setminus S \subseteq B_k$, and then $|\nu(B_{k+1})| \ge |\nu(S)| - |\nu(B_k)| \ge k$. In any case we have proven the existence of an appropriate B_{k+1}, and so we have proven the claim.
Now we consider the pairwise disjoint $A_1 = B_1 \setminus B_2, A_2 = B_2 \setminus B_3, \dots$ and the $B_\infty = \bigcap_{k=1}^{+\infty} B_k$. Then

$$\begin{aligned}\nu(B_1) - \nu(B_\infty) &= \nu(B_1 \setminus B_\infty) = \nu\big(\bigcup_{m=1}^{+\infty} A_m\big) = \sum_{m=1}^{+\infty} \nu(A_m) \\ &= \lim_{k\to+\infty} \sum_{m=1}^{k-1} \nu(A_m) = \lim_{k\to+\infty}(\nu(B_1) - \nu(B_k)).\end{aligned}$$

Therefore $\lim_{k\to+\infty} \nu(B_k) = \nu(B_\infty)$, i.e. $|\nu(B_\infty)| = +\infty$, and we arrive at a contradiction. □

Definition. *Let ν be a complex measure on $(X, \mathcal{S})$. The measure $|\nu|$ defined in Proposition 6.9 is called the* ***absolute variation*** *of ν and the number $|\nu|(X)$ is called the* ***total variation*** *of ν.*

It is useful to note something which we have already noted for signed measures. If ν is a complex measure, then

$$|\nu(A)| \leq |\nu|(A) \quad \text{for all } A \in \mathcal{S}.$$

Indeed, we may consider $\{A\}$ as a measurable partition of A, and then the definition of $|\nu|(A)$ implies the above inequlity.

Proposition 6.12. *Let ν, ν_1, ν_2 be complex measures on $(X, \mathcal{S})$ and $\lambda \in \mathbb{C}$. Then*
(i) $|\nu_1 + \nu_2| \leq |\nu_1| + |\nu_2|$ and $|\lambda\nu| = |\lambda||\nu|$,
(ii) $|\operatorname{Re}(\nu)| \leq |\nu|$, $|\operatorname{Im}(\nu)| \leq |\nu|$, $|\nu| \leq |\operatorname{Re}(\nu)| + |\operatorname{Im}(\nu)|$.

Proof. (i) We consider an arbitrary measurable partition $\{A_1, \ldots, A_n\}$ of $A \in \mathcal{S}$, and we have

$$\textstyle\sum_{k=1}^n |(\nu_1 + \nu_2)(A_k)| \leq \sum_{k=1}^n |\nu_1(A_k)| + \sum_{k=1}^n |\nu_2(A_k)| \leq |\nu_1|(A) + |\nu_2|(A).$$

Taking the supremum of the left side, we find $|\nu_1 + \nu_2|(A) \leq |\nu_1|(A) + |\nu_2|(A)$.
In the same manner, we have

$$\textstyle\sum_{k=1}^n |(\lambda\nu)(A_k)| = |\lambda| \sum_{k=1}^n |\nu(A_k)| \leq |\lambda||\nu|(A).$$

Taking the supremum of the left side, we find $|\lambda\nu|(A) \leq |\lambda||\nu|(A)$. If $\lambda \neq 0$, we apply the last inequality to the number $\frac{1}{\lambda}$ and to the signed measure $\lambda\nu$, and we get $|\nu|(A) \leq \frac{1}{|\lambda|}|\lambda\nu|(A)$. From the two inequalities we get $|\lambda\nu|(A) = |\lambda||\nu|(A)$ for every $A \in \mathcal{S}$ and every $\lambda \neq 0$. Finally, this last equality is obviously true if $\lambda = 0$.
(ii) In the same manner, if $\{A_1, \ldots, A_n\}$ is any measurable partition of $A \in \mathcal{S}$, we have

$$\textstyle\sum_{k=1}^n |\operatorname{Re}(\nu)(A_k)| \leq \sum_{k=1}^n |\nu(A_k)| \leq |\nu|(A),$$

$$\textstyle\sum_{k=1}^n |\operatorname{Im}(\nu)(A_k)| \leq \sum_{k=1}^n |\nu(A_k)| \leq |\nu|(A).$$

Taking the supremum of the left sides of these two inequalities, we find $|\operatorname{Re}(\nu)|(A) \leq |\nu|(A)$ and $|\operatorname{Im}(\nu)|(A) \leq |\nu|(A)$.
The last inequality is a consequence of the result of (i). $\square$

Example. Let μ be a measure on $(X, \mathcal{S})$ and $f : X \to \overline{\mathbb{C}}$ be a μ-integrable function. Lemma 3.10 implies that $\int_A f \, d\mu$ is a complex number for every $A \in \mathcal{S}$ and Theorem 3.9 implies that the function $\lambda : \mathcal{S} \to \mathbb{C}$, which is defined by $\lambda(A) = \int_A f \, d\mu$ for every $A \in \mathcal{S}$, is a complex measure on $(X, \mathcal{S})$.

Definition. *The complex measure λ which is defined in the last example is called the* ***indefinite integral*** *of f with respect to μ and it is denoted by $f\mu$. Thus,*

$$(f\mu)(A) = \textstyle\int_A f \, d\mu, \quad A \in \mathcal{S}.$$

The next result is the analogue of Proposition 6.6.

Proposition 6.13. *Let μ be a measure on $(X, \mathcal{S})$ and $f : X \to \overline{\mathbb{C}}$ be integrable with respect to μ. Then*

$$|f\mu|(A) = \textstyle\int_A |f| \, d\mu$$

for every $A \in \mathcal{S}$. Hence,

$$|f\mu| = |f|\mu.$$

Proof. If $\{A_1, \ldots, A_n\}$ is an arbitrary measurable partition of $A \in \mathcal{S}$, then

$$\sum_{k=1}^{n} |(f\mu)(A_k)| = \sum_{k=1}^{n} \big| \int_{A_k} f \, d\mu \big| \leq \sum_{k=1}^{n} \int_{A_k} |f| \, d\mu = \int_A |f| \, d\mu.$$

Taking the supremum of the left side of this inequality, we get $|f\mu|(A) \leq \int_A |f| \, d\mu$.
Since f is integrable, it is finite a.e. on X. If $N = \{x \in X \mid f(x) \neq \infty\}$, then $\mu(N^c) = 0$, and Theorem 2.1 implies that there is a sequence (ϕ_m) of measurable simple functions so that

$$\phi_m \to \overline{\text{sign}(f)}, \quad |\phi_m| \uparrow |\overline{\text{sign}(f)}| \leq 1$$

on N. Defining each ϕ_m as 0 on N^c, we have that all these properties hold a.e. on X.
If $\phi_m = \sum_{k=1}^{n_m} \kappa_{m,k} \chi_{E_{m,k}}$ is the standard representation of ϕ_m, then $|\kappa_{m,k}| \leq 1$ for all $k = 1, \ldots, n_m$, and so

$$\big| \int_A f \phi_m \, d\mu \big| = \big| \sum_{k=1}^{n_m} \kappa_{m,k} \int_{A \cap E_{m,k}} f \, d\mu \big| \leq \sum_{k=1}^{n_m} |(f\mu)(A \cap E_{m,k})| \leq |f\mu|(A),$$

where the last inequality is true since $\{A \cap E_{m,1}, \ldots, A \cap E_{m,n_m}\}$ is a measurable partition of A.
By the Dominated Convergence Theorem we get that

$$\int_A |f| \, d\mu = \int_A f \, \overline{\text{sign}(f)} \, d\mu \leq |f\mu|(A).$$

We conclude that $|f\mu|(A) = \int_A |f| \, d\mu$ for every $A \in \mathcal{S}$. $\square$

Exercises.

6.3.1. Let ν be a real or complex measure on $(X, \mathcal{S})$. If $\nu(X) = |\nu|(X)$, prove that $\nu = |\nu|$.

6.3.2. Let ν be a signed or complex measure on $(X, \mathcal{S})$. We say that $\{A_1, A_2, \ldots\}$ is a **countable measurable partition** of $A \in \mathcal{S}$, if $A_k \in \mathcal{S}$ for all k, the sets $A_1, A_2, \ldots$ are pairwise disjoint and $A = A_1 \cup A_2 \cup \cdots$.
Prove that $|\nu|(A) = \sup\{\sum_{k=1}^{+\infty} |\nu(A_k)| \mid \{A_1, A_2, \ldots\}$ is a countable measurable partition of $A\}$ for every $A \in \mathcal{S}$.

6.4 Integration.

Let $(X, \mathcal{S})$ be a measurable space.

The next definition treats only the case when both f and ν have their values in $\overline{\mathbb{R}}$.

Definition. *Let ν be a signed measure on $(X, \mathcal{S})$. If $f : X \to \overline{\mathbb{R}}$ is $\mathcal{S}$-measurable, we say that the **integral** $\int_X f \, d\nu$ of f over X with respect to ν is defined if both $\int_X f \, d\nu^+$ and $\int_X f \, d\nu^-$ are defined and they are neither both $+\infty$ nor both $-\infty$. In such a case we write*

$$\int_X f \, d\nu = \int_X f \, d\nu^+ - \int_X f \, d\nu^-.$$

*Moreover, we say that f is **integrable** over X with respect to ν if $\int_X f \, d\nu$ is finite.*

Proposition 6.14. *Let ν be a signed measure on $(X, \mathcal{S})$ and $f : X \to \overline{\mathbb{R}}$ be measurable. Then f is integrable with respect to ν if and only if f is integrable with respect to both ν^+ and ν^- if and only if f is integrable with respect to $|\nu|$.*

Proof. $\int_X f \, d\nu$ is finite if and only if both $\int_X f \, d\nu^+$ and $\int_X f \, d\nu^-$ are finite or, equivalently, $\int_X |f| \, d\nu^+ < +\infty$ and $\int_X |f| \, d\nu^- < +\infty$ or, equivalently, $\int_X |f| \, d|\nu| < +\infty$ if and only if f is integrable with respect to $|\nu|$. $\square$

Now, let ν be a signed measure or a complex measure on $(X, \mathcal{S})$ and $f : X \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ be $\mathcal{S}$-measurable. If $\int_X |f| \, d|\nu| < +\infty$, then f is finite $|\nu|$-a.e. on X and the $|\nu|$-almost everywhere defined functions $\mathrm{Re}(f), \mathrm{Im}(f)$ satisfy $\int_X |\mathrm{Re}(f)| \, d|\nu| < +\infty$ and $\int_X |\mathrm{Im}(f)| \, d|\nu| < +\infty$. Since $|\mathrm{Re}(\nu)| \leq |\nu|$ and $|\mathrm{Im}(\nu)| \leq |\nu|$, Lemma 6.6 implies that all integrals $\int_X |\mathrm{Re}(f)| \, d|\mathrm{Re}(\nu)|$, $\int_X |\mathrm{Re}(f)| \, d|\mathrm{Im}(\nu)|$, $\int_X |\mathrm{Im}(f)| \, d|\mathrm{Re}(\nu)|$ and $\int_X |\mathrm{Im}(f)| \, d|\mathrm{Im}(\nu)|$ are finite. Proposition 6.12 implies that $\int_X \mathrm{Re}(f) \, d\,\mathrm{Re}(\nu)$, $\int_X \mathrm{Re}(f) \, d\,\mathrm{Im}(\nu)$, $\int_X \mathrm{Im}(f) \, d\,\mathrm{Re}(\nu)$ and $\int_X \mathrm{Im}(f) \, d\,\mathrm{Im}(\nu)$ are all defined and they are real numbers.

Therefore, the following definition is valid.

Definition. *Let ν be a signed measure or a complex measure on $(X, \mathcal{S})$ and $f : X \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ be $\mathcal{S}$-measurable. We say that f is* ***integrable*** *over X with respect to ν if f is integrable with respect to $|\nu|$, and in this case we say that the* ***integral*** *$\int_X f \, d\nu$ of f over X with respect to ν is defined and that its value is given by*

$$\int_X f \, d\nu = \int_X \mathrm{Re}(f) \, d\,\mathrm{Re}(\nu) - \int_X \mathrm{Im}(f) \, d\,\mathrm{Im}(\nu) + i \int_X \mathrm{Re}(f) \, d\,\mathrm{Im}(\nu) + i \int_X \mathrm{Im}(f) \, d\,\mathrm{Re}(\nu).$$

Of course, we have the particular formulas

$$\int_X f \, d\nu = \int_X \mathrm{Re}(f) \, d\nu + i \int_X \mathrm{Im}(f) \, d\nu, \quad \int_X f \, d\nu = \int_X f \, d\,\mathrm{Re}(\nu) + i \int_X f \, d\,\mathrm{Im}(\nu),$$

all under the assumption that $\int_X |f| \, d|\nu| < +\infty$.

Example. Let ν be a signed measure on $(X, \mathcal{S})$ and $E \in \mathcal{S}$ so that $\nu^+(E) < +\infty$ or $\nu^-(E) < +\infty$. Then $\int_X \chi_E \, d\nu^+ < +\infty$ or $\int_X \chi_E \, d\nu^- < +\infty$, respectively, and so $\int_X \chi_E \, d\nu$ is defined and

$$\int_X \chi_E \, d\nu = \int_X \chi_E \, d\nu^+ - \int_X \chi_E \, d\nu^- = \nu^+(E) - \nu^-(E) = \nu(E).$$

Now, let ν be a complex measure on $(X, \mathcal{S})$ and $E \in \mathcal{S}$ so that $|\nu|(E) < +\infty$. Then $\int_X \chi_E \, d|\nu| = |\nu|(E) < +\infty$, and so $\int_X \chi_E \, d\nu$ is defined and, from the previous case,

$$\int_X \chi_E \, d\nu = \int_X \chi_E \, d\,\mathrm{Re}(\nu) + i \int_X \chi_E \, d\,\mathrm{Im}(\nu) = \mathrm{Re}(\nu)(E) + i\,\mathrm{Im}(\nu)(E) = \nu(E).$$

We shall not try to extend all properties of integrals with respect to measures to properties of integrals with respect to signed measures or complex measures. The safe thing to do is to *reduce everything to non-negative and non-positive variations or to real and imaginary parts*.

For completeness, we shall only see a few of the most useful properties, like the linearity properties and the appropriate version of the Dominated Convergence Theorem.

Proposition 6.15. *Let ν, ν_1, ν_2 be signed or complex measures on $(X, \mathcal{S})$ and $f, f_1, f_2 : X \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ be all integrable with respect to these measures. Then, for every $\lambda_1, \lambda_2 \in \mathbb{C}$, we have*

$$\int_X (\lambda_1 f_1 + \lambda_2 f_2) \, d\nu = \lambda_1 \int_X f_1 \, d\nu + \lambda_2 \int_X f_2 \, d\nu,$$

$$\int_X f \, d(\lambda_1 \nu_1 + \lambda_2 \nu_2) = \lambda_1 \int_X f \, d\nu_1 + \lambda_2 \int_X f \, d\nu_2.$$

Proof. We reduce everything to real functions and signed measures. □

Theorem 6.5. *(***Dominated Convergence Theorem***) Let ν be a signed or complex measure on $(X, \mathcal{S})$, and all $f, f_n : X \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ and $g : X \to [0, +\infty]$ be $\mathcal{S}$-measurable. If $f_n \to f$ and $|f_n| \leq g$ on X except on a set which is null for ν, and if $\int_X g \, d|\nu| < +\infty$, then*

$$\int_X f_n \, d\nu \to \int_X f \, d\nu.$$

Proof. A set which is null for ν is, also, null for ν^+ and ν^-, if ν is signed, and null for $\mathrm{Re}(\nu)$ and $\mathrm{Im}(\nu)$, if ν is complex. Moreover, Lemma 6.6 implies that $\int_X g \, d\nu^+ < +\infty$ and $\int_X g \, d\nu^- < +\infty$, if ν is signed, and $\int_X g \, d|\mathrm{Re}(\nu)| < +\infty$ and $\int_X g \, d|\mathrm{Im}(\nu)| < +\infty$, if ν is complex.
Therefore, the proof reduces to the usual Dominated Convergence Theorem for measures. □

Theorem 6.6. *Let ν be a signed or complex measure on $(X,\mathcal{S})$ and $f : X \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ be such that the $\int_X f\, d\nu$ is defined. Then*

$$\big|\int_X f\, d\nu\big| \leq \int_X |f|\, d|\nu|.$$

Proof. We may assume that $\int_X |f|\, d|\nu| < +\infty$ or else the inequality is obvious.
If ϕ is a measurable simple function with standard representation $\phi = \sum_{k=1}^n \kappa_k \chi_{E_k}$ and so that $|\nu|(E_k) < +\infty$ for all k, then we have

$$\begin{aligned}\big|\int_X \phi\, d\nu\big| &= \big|\sum_{k=1}^n \kappa_k \int_X \chi_{E_k}\, d\nu\big| = \big|\sum_{k=1}^n \kappa_k \nu(E_k)\big| \leq \sum_{k=1}^n |\kappa_k||\nu(E_k)| \\ &\leq \sum_{k=1}^n |\kappa_k||\nu|(E_k) = \int_X |\phi|\, d|\nu|.\end{aligned}$$

The proof in the case of a general function f is a standard limiting argument. □

A companion to the previous theorem is

Theorem 6.7. *Let ν be a signed or complex measure on $(X,\mathcal{S})$. Then*

$$|\nu|(A) = \sup\big\{\big|\int_A f\, d\nu\big| \,\big|\, f \text{ is } \mathcal{S}\text{-measurable},\ |f| \leq 1\ \nu\text{-a.e. on } A\big\}$$

for every $A \in \mathcal{S}$, where the functions f have real values, if ν is signed, and complex values, if ν is complex.

Proof. Let $M = \sup\big\{\big|\int_A f\, d\nu\big| \,\big|\, f \text{ is } \mathcal{S}\text{-measurable},\ |f| \leq 1\ \nu\text{-a.e. on } A\big\}$.
If f is $\mathcal{S}$-measurable and $|f| \leq 1$ ν-a.e. on A, then $|f\chi_A| \leq \chi_A$ ν-a.e. on X, and Theorem 6.6 implies

$$\big|\int_A f\, d\nu\big| = \big|\int_X f\chi_A\, d\nu\big| \leq \int_X |f\chi_A|\, d|\nu| \leq \int_X \chi_A\, d|\nu| = |\nu|(A).$$

Hence, $M \leq |\nu|(A)$.
Now, let $\{A_1, \ldots, A_n\}$ be any measurable partition of A.
Then $\sum_{k=1}^n |\nu|(A_k)| = |\nu|(A) < +\infty$, and so $|\nu|(A_k) < +\infty$ for all k. We consider the function $f = \sum_{k=1}^n \kappa_k \chi_{A_k}$, where $\kappa_k = \overline{\operatorname{sign}(\nu(A_k))}$ for all k. Then $|f| \leq 1$ on A, and so

$$M \geq \big|\int_A f\, d\nu\big| = \big|\sum_{k=1}^n \kappa_k \int_A \chi_{A_k}\, d\nu\big| = \big|\sum_{k=1}^n \kappa_k \nu(A_k)\big| = \sum_{k=1}^n |\nu(A_k)|.$$

Hence, $M \geq |\nu|(A)$. □

Finally, we prove a result about integration with respect to an indefinite integral. This is important because, as we shall see in the next section, indefinite integrals are special measures which play an important role among signed or complex measures.

Theorem 6.8. *Let μ be a measure on $(X,\mathcal{S})$ and $f : X \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ be measurable so that $\int_X f\, d\mu$ is defined. A measurable function $g : X \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ is integrable over X with respect to $f\mu$ if and only if gf is integrable over X with respect to μ. In such a case,*

$$\int_X g\, d(f\mu) = \int_X gf\, d\mu.$$

This equality is true in the case of $\mathcal{S}$-measurable $f, g : X \to [0,+\infty]$ without any restriction.

Proof. We consider first the case of $\mathcal{S}$-measurable $f, g : X \to [0,+\infty]$.
If $g = \chi_A$ for some $A \in \mathcal{S}$, then

$$\int_X \chi_A\, d(f\mu) = (f\mu)(A) = \int_A f\, d\mu = \int_X \chi_A f\, d\mu.$$

Thus, the equality $\int_X g\, d(f\mu) = \int_X gf\, d\mu$ is true for $\mathcal{S}$-measurable characteristic functions g. This extends by linearity to $\mathcal{S}$-measurable non-negative simple functions g, and then by the Monotone Convergence Theorem to the general $\mathcal{S}$-measurable non-negative g.
This implies that, in general, $\int_X |g|\, d(|f|\mu) = \int_X |gf|\, d\mu$.
From this we see that g is integrable over X with respect to $f\mu$ if and only if, by definition, g is integrable over X with respect to $|f\mu| = |f|\mu$ if and only if, by the equality we just proved, gf is integrable over X with respect to μ.
The equality $\int_X g\, d(f\mu) = \int_X gf\, d\mu$ can now be established by reducing all functions to non-negative functions and using the special case we proved. □

6.5 Lebesgue decomposition, Radon-Nikodym derivative.

Let $(X, \mathcal{S})$ be a measurable space.

We extend two definitions from section 6.2. We had formulated the first definition only for signed measures and the second definition only for measures.

Definition. *Let ν be a complex measure on $(X, \mathcal{S})$ and $A \in \mathcal{S}$. We say that A is a* ***null set*** *for ν if $\nu(B) = 0$ for every $B \in \mathcal{S}$, $B \subseteq A$.*

Definition. *Let ν_1, ν_2 be two signed or complex measures on $(X, \mathcal{S})$. We say that ν_1, ν_2 are* ***mutually singular*** *if there exist $A_1, A_2 \in \mathcal{S}$ so that A_1 is null for ν_2 and A_2 is null for ν_1 and $A_1 \cup A_2 = X$, $A_1 \cap A_2 = \emptyset$.*
We use the symbol

$$\nu_1 \perp \nu_2$$

to denote that ν_1, ν_2 are mutually singular.

Lemma 6.2. *(i) Let ν be a signed measure on $(X, \mathcal{S})$ and $A \in \mathcal{S}$. Then A is null for ν if and only if it is null for both ν^+, ν^- if and only if it is null for $|\nu|$.*
(ii) Let ν be a complex measure on $(X, \mathcal{S})$ and $A \in \mathcal{S}$. Then A is null for ν if and only if A is null for both $\mathrm{Re}(\nu)$ and $\mathrm{Im}(\nu)$ if and only if A is null for $|\nu|$.

Proof. Let A be null for $|\nu|$. For every $B \in \mathcal{S}$, $B \subseteq A$, we have that $|\nu(B)| \leq |\nu|(B) = 0$, and so A is null for ν.
Conversely, let A be null for ν. If $\{A_1, \ldots, A_n\}$ is any measurable partition of A, then $\nu(A_k) = 0$ for all k, and so $\sum_{k=1}^n |\nu(A_k)| = 0$. Hence, $|\nu|(A) = 0$, and so A is null for $|\nu|$.
If ν is signed, then from $|\nu| = \nu^+ + \nu^-$ we have that A is null for both ν^+, ν^- if and only if it is null for $|\nu|$.
If ν is complex, then from $\nu = \mathrm{Re}(\nu) + i\,\mathrm{Im}(\nu)$ we have that A is null for both $\mathrm{Re}(\nu), \mathrm{Im}(\nu)$ if and only if it is null for ν. □

Lemma 6.3. *(i) Let ν_1 and ν_2 be two signed measures on $(X, \mathcal{S})$. Then ν_1 and ν_2 are mutually singular if and only if each of ν_1^+, ν_1^- and each of ν_2^+, ν_2^- are mutually singular if and only if $|\nu_1|$ and $|\nu_2|$ are mutually singular.*
(ii) Let ν_1 and ν_2 be complex measures on $(X, \mathcal{S})$. Then, ν_1 and ν_2 are mutually singular if and only if each of $\mathrm{Re}(\nu_1), \mathrm{Im}(\nu_1)$ and each of $\mathrm{Re}(\nu_2), \mathrm{Im}(\nu_2)$ are mutually singular if and only if $|\nu_1|$ and $|\nu_2|$ are mutually singular.

Proof. The proof is a trivial consequence of Lemma 6.1. □

Lemma 6.4. *(i) Let ν, ν_1, ν_2 be signed measures on $(X, \mathcal{S})$ and $\lambda_1, \lambda_2 \in \mathbb{R}$. If $\nu_1 \perp \nu$, $\nu_2 \perp \nu$ and $\lambda_1\nu_1 + \lambda_2\nu_2$ is defined, then $(\lambda_1\nu_1 + \lambda_2\nu_2) \perp \nu$.*
(ii) Let ν, ν_1, ν_2 be complex measures on $(X, \mathcal{S})$ and $\lambda_1, \lambda_2 \in \mathbb{C}$. If $\nu_1 \perp \nu$, $\nu_2 \perp \nu$, then $(\lambda_1\nu_1 + \lambda_2\nu_2) \perp \nu$.

Proof. There are $A_1, B_1, A_2, B_2 \in \mathcal{S}$ so that $A_1 \cup B_1 = X = A_2 \cup B_2$, $A_1 \cap B_1 = \emptyset = A_2 \cap B_2$, A_1 is null for ν_1, A_2 is null for ν_2 and B_1, B_2 are both null for ν. Then $B_1 \cup B_2$ is null for ν and $A_1 \cap A_2$ is null for both ν_1 and ν_2 and, hence, for $\lambda_1\nu_1 + \lambda_2\nu_2$. Since $(A_1 \cap A_2) \cup (B_1 \cup B_2) = X$ and $(A_1 \cap A_2) \cap (B_1 \cup B_2) = \emptyset$, we have that $(\lambda_1\nu_1 + \lambda_2\nu_2) \perp \nu$. □

Definition. *Let μ be a measure and ν be a signed or complex measure on $(X, \mathcal{S})$. We say that ν is* ***absolutely continuous*** *with respect to μ if $\nu(A) = 0$ for every $A \in \mathcal{S}$ with $\mu(A) = 0$, and we denote this by*

$$\nu \ll \mu.$$

Example. Let $f : X \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ be measurable so that the $\int_X f\,d\mu$ is defined (recall that in the case of $\overline{\mathbb{C}}$ this means that f is integrable). Then *the indefinite integral $f\mu$ is absolutely continuous with respect to μ.*
This is obvious: if $A \in \mathcal{S}$ has $\mu(A) = 0$, then $(f\mu)(A) = \int_A f\,d\mu = 0$.

Lemma 6.5. *Let μ be a measure and ν, ν_1, ν_2 be signed or complex measures on $(X, \mathcal{S})$.*
(i) If ν is signed, then $\nu \ll \mu$ if and only if $\nu^+ \ll \mu$ and $\nu^- \ll \mu$ if and only if $|\nu| \ll \mu$.
(ii) If ν is complex, then $\nu \ll \mu$ if and only if $\mathrm{Re}(\nu) \ll \mu$ and $\mathrm{Im}(\nu) \ll \mu$ if and only if $|\nu| \ll \mu$.
(iii) If $\nu \ll \mu$ and $\nu \perp \mu$, then $\nu = 0$.
(iv) If ν_1, ν_2 are signed and $\lambda_1, \lambda_2 \in \mathbb{R}$ and $\lambda_1\nu_1 + \lambda_2\nu_2$ is defined and $\nu_1 \ll \mu$, $\nu_2 \ll \mu$, then $\lambda_1\nu_1 + \lambda_2\nu_2 \ll \mu$.
(v) If ν_1, ν_2 are complex and $\lambda_1, \lambda_2 \in \mathbb{C}$ and $\nu_1 \ll \mu$, $\nu_2 \ll \mu$, then $\lambda_1\nu_1 + \lambda_2\nu_2 \ll \mu$.

Proof. (i-ii) Let $|\nu| \ll \mu$. If $A \in \mathcal{S}$, $\mu(A) = 0$, then $|\nu(A)| \le |\nu|(A) = 0$, and so $\nu(A) = 0$. Hence, $\nu \ll \mu$.
Conversely, Let $\nu \ll \mu$, and let $A \in \mathcal{S}$ with $\mu(A) = 0$. If $\{A_1, \dots, A_n\}$ is any measurable partition of A, then $\mu(A_k) = 0$ for all k, and so $\nu(A_k) = 0$ for all k. Hence, $\sum_{k=1}^n |\nu(A_k)| = 0$, and this implies that $|\nu|(A) = 0$. Thus, $|\nu| \ll \mu$.
Since $\nu(A) = 0$ is equivalent to $\mathrm{Re}(\nu)(A) = \mathrm{Im}(\nu)(A) = 0$, the first equivalence is obvious.
If ν is signed, then from $|\nu| = \nu^+ + \nu^-$ we have that $\nu^+ \ll \mu$ and $\nu^- \ll \mu$ if and only if $|\nu| \ll \mu$.
If ν is complex, then from $\nu = \mathrm{Re}(\nu) + i\,\mathrm{Im}(\nu)$ we have that $\mathrm{Re}(\nu) \ll \mu$ and $\mathrm{Im}(\nu) \ll \mu$ if and only if $\nu \ll \mu$.
(iii) We consider sets $M, N \in \mathcal{S}$ so that $M \cup N = X$, $M \cap N = \emptyset$, M is a null set for ν and N is a null set for μ. Then $\mu(N) = 0$ and $\nu \ll \mu$ imply that N is a null set for ν. But then $X = M \cup N$ is a null set for ν, and so $\nu = 0$.
(iv-v) If $A \in \mathcal{S}$ has $\mu(A) = 0$, then $\nu_1(A) = \nu_2(A) = 0$, and so $(\lambda_1\nu_1 + \lambda_2\nu_2)(A) = 0$. ☐

The next result justifies the term *absolutely continuous* at least in the special case of a finite ν.

Proposition 6.16. *Let μ be a measure and ν be a real or a complex measure on $(X, \mathcal{S})$. Then $\nu \ll \mu$ if and only if for every $\epsilon > 0$ there is a $\delta > 0$ so that $|\nu(A)| < \epsilon$ for every $A \in \mathcal{S}$ with $\mu(A) < \delta$.*

Proof. Let us assume that for every $\epsilon > 0$ there is a $\delta > 0$ so that $|\nu(A)| < \epsilon$ for every $A \in \mathcal{S}$ with $\mu(A) < \delta$. If $\mu(A) = 0$, then $\mu(A) < \delta$ for every $\delta > 0$, and so $|\nu(A)| < \epsilon$ for every $\epsilon > 0$. Hence, $\nu(A) = 0$, and so $\nu \ll \mu$.
Conversely, let us assume that there is some $\epsilon_0 > 0$ so that for every $\delta > 0$ there is $A \in \mathcal{S}$ with $\mu(A) < \delta$ and $|\nu(A)| \ge \epsilon_0$. Then for every $k \in \mathbb{N}$ there is $A_k \in \mathcal{S}$ with $\mu(A_k) < \frac{1}{2^k}$ and $|\nu|(A_k) \ge |\nu(A_k)| \ge \epsilon_0$. We consider $B_k = \bigcup_{l=k}^{+\infty} A_l$, and then $\mu(B_k) < \frac{1}{2^{k-1}}$ and $|\nu|(B_k) \ge |\nu|(A_k) \ge \epsilon_0$ for every k. If we set $B = \bigcap_{k=1}^{+\infty} B_k$, then we have $\mu(B) = 0$. Since $B_k \downarrow B$, the continuity of $|\nu|$ from above implies $|\nu|(B) \ge \epsilon_0$. Therefore, $|\nu|$ is not absolutely continuous with respect to μ. Now Lemma 6.4 implies that ν is not absolutely continuous with respect to μ. ☐

Theorem 6.9. *Let μ be a measure on $(X, \mathcal{S})$.*
(i) If $\lambda, \lambda_1, \rho, \rho_1$ are signed or complex measures on $(X, \mathcal{S})$ so that $\lambda \ll \mu$, $\lambda_1 \ll \mu$ and $\rho \perp \mu$, $\rho_1 \perp \mu$ and $\lambda + \rho = \lambda_1 + \rho_1$, then $\lambda = \lambda_1$ and $\rho = \rho_1$.
(ii) If $f, f_1 : X \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ are μ-integrable over X and $f\mu = f_1\mu$, then $f = f_1$ μ-a.e. on X.
(iii) If $f, f_1 : X \to \overline{\mathbb{R}}$ are $\mathcal{S}$-measurable and the $\int_X f\,d\mu$, $\int_X f_1\,d\mu$ are defined and $f\mu = f_1\mu$, then $f = f_1$ μ-a.e. on X, provided that μ restricted on the set $\{x \in X \mid f(x) \ne f_1(x)\}$ is semifinite.

Proof. (i) There exist sets $M, M_1, N, N_1 \in \mathcal{S}$ with $M \cup N = X = M_1 \cup N_1$, $M \cap N = \emptyset = M_1 \cap N_1$ so that N, N_1 are null for μ, M is null for ρ and M_1 is null for ρ_1. If we set $K = N \cup N_1$, then K is null for μ and $K^c = M \cap M_1$ is null for both ρ and ρ_1. Since $\lambda \ll \mu$, $\lambda_1 \ll \mu$, we have

that K is null for both λ and λ_1.
If $A \in \mathcal{S}$, $A \subseteq K$, then

$$\rho(A) = \rho(A) + \lambda(A) = \rho_1(A) + \lambda_1(A) = \rho_1(A).$$

If $A \in \mathcal{S}$, $A \subseteq K^c$, then $\rho(A) = 0 = \rho_1(A)$. Therefore, for every $A \in \mathcal{S}$ we have

$$\rho(A) = \rho(A \cap K) + \rho(A \cap K^c) = \rho_1(A \cap K) + \rho_1(A \cap K^c) = \rho_1(A),$$

and so $\rho = \rho_1$. A symmetric argument implies that $\lambda = \lambda_1$.
(ii) We have

$$\textstyle\int_A (f - f_1)\, d\mu = \int_A f\, d\mu - \int_A f_1\, d\mu = (f\mu)(A) - (f_1\mu)(A) = 0$$

for all $A \in \mathcal{S}$. Now, Theorem 3.3 implies $f = f_1$ μ-a.e. on X.
(iii) Let $t, s \in \mathbb{R}$ with $t < s$, and let $A_{t,s} = \{x \in X \,|\, f(x) \le t < s \le f_1(x)\}$.
If $0 < \mu(A_{t,s}) < +\infty$, we consider $B = A_{t,s}$. If $\mu(A_{t,s}) = +\infty$, we consider any $B \in \mathcal{S}$ so that $B \subseteq A_{t,s}$ and $0 < \mu(B) < +\infty$. In any case, we have

$$(f\mu)(B) = \textstyle\int_B f\, d\mu \le t\mu(B), \quad (f_1\mu)(B) = \int_B f_1\, d\mu \ge s\mu(B),$$

and so $s\mu(B) \le t\mu(B)$. This implies $\mu(B) = 0$, which is false.
The only remaining case is $\mu(A_{t,s}) = 0$. Now we observe that

$$\{x \in X \,|\, f(x) < f_1(x)\} = \textstyle\bigcup_{t,s\in\mathbb{Q},\, t<s} A_{t,s},$$

which implies $\mu(\{x \in X \,|\, f(x) < f_1(x)\}) = 0$. Similarly, $\mu(\{x \in X \,|\, f(x) > f_1(x)\}) = 0$, and we conclude that $f = f_1$ μ-a.e. on X. $\square$

Lemma 6.6. *Let μ, ν be finite measures on $(X, \mathcal{S})$. If μ, ν are not mutually singular, then there is $\epsilon_0 > 0$ and $A_0 \in \mathcal{S}$ with $\mu(A_0) > 0$ so that $\frac{\nu(A)}{\mu(A)} \ge \epsilon_0$ for every $A \in \mathcal{S}$, $A \subseteq A_0$ with $\mu(A) > 0$.*

Proof. For every $n \in \mathbb{N}$ we consider a Hahn decomposition of the signed measure $\nu - \frac{1}{n}\mu$. There are sets $P_n, N_n \in \mathcal{S}$ so that $P_n \cup N_n = X$, $P_n \cap N_n = \emptyset$ and P_n is a non-negative set and N_n is a non-positive set for $\nu - \frac{1}{n}\mu$.
We consider $N = \bigcap_{n=1}^{+\infty} N_n$. Since $N \subseteq N_n$, we get $(\nu - \frac{1}{n}\mu)(N) \le 0$ for all n. Then $\nu(N) \le \frac{1}{n}\mu(N)$ for all n and, since $\mu(N) < +\infty$, we have $\nu(N) = 0$.
We consider $P = \bigcup_{n=1}^{+\infty} P_n$, and then $P \cup N = X$ and $P \cap N = \emptyset$. If $\mu(P) = 0$, then μ and ν are mutually singular. Therefore, $\mu(P) > 0$, and this implies that $\mu(P_N) > 0$ for at least one N. We define $A_0 = P_N$ for such an N and we set $\epsilon_0 = \frac{1}{N}$ for the same N.
Now, $\mu(A_0) > 0$. Since A_0 is a non-negative set for $\nu - \epsilon_0\mu$, for every $A \in \mathcal{S}$, $A \subseteq A_0$ with $\mu(A) > 0$ we get $\nu(A) - \epsilon_0\mu(A) \ge 0$, and so $\frac{\nu(A)}{\mu(A)} \ge \epsilon_0$. $\square$

Lebesgue-Radon-Nikodym Theorem. `The signed case`. *Let ν be a σ-finite signed measure and μ be a σ-finite measure on $(X, \mathcal{S})$. Then there exist unique σ-finite signed measures λ and ρ on $(X, \mathcal{S})$ so that*

$$\nu = \lambda + \rho, \quad \lambda \ll \mu, \quad \rho \perp \mu.$$

Moreover, there exists a $\mathcal{S}$-measurable $f : X \to \overline{\mathbb{R}}$ so that the $\int_X f\, d\mu$ is defined and

$$\lambda = f\mu.$$

If f_1 is another such function, then $f_1 = f$ μ-a.e. on X.
If ν is non-negative, then λ and ρ are non-negative and $f \ge 0$ μ-a.e. on X.
If ν is real, then λ and ρ are real and f is integrable over X with respect to μ.

Proof. The uniqueness part of the statement is a consequence of Theorem 6.9. Observe that μ is σ-finite, and so it is semifinite.
Therefore, we need to prove the existence of λ, ρ and f.
(a) We first consider the special case when both μ, ν are finite measures on $(X, \mathcal{S})$.
We define $\mathcal{C}$ to be the collection of all $\mathcal{S}$-measurable $f : X \to [0, +\infty]$ with the property

$$\int_A f \, d\mu \le \nu(A), \quad A \in \mathcal{S}.$$

The function 0, obviously, belongs to $\mathcal{C}$ and, if $f_1, f_2 \in \mathcal{C}$, then $f = \max\{f_1, f_2\} \in \mathcal{C}$. Indeed, if $A \in \mathcal{S}$, we consider $A_1 = \{x \in A \,|\, f_2(x) \le f_1(x)\}$ and $A_2 = \{x \in A \,|\, f_1(x) < f_2(x)\}$, and we have

$$\int_A f \, d\mu = \int_{A_1} f \, d\mu + \int_{A_2} f \, d\mu = \int_{A_1} f_1 \, d\mu + \int_{A_2} f_2 \, d\mu \le \nu(A_1) + \nu(A_2) = \nu(A).$$

We define

$$\kappa = \sup\big\{ \int_X f \, d\mu \,|\, f \in \mathcal{C}\big\}. \tag{6.2}$$

Since $0 \in \mathcal{C}$ and $\int_X f \, d\mu \le \nu(X)$ for all $f \in \mathcal{C}$, we have $0 \le \kappa \le \nu(X) < +\infty$.
Now, there is a sequence (f_n) in $\mathcal{C}$ so that $\int_X f_n \, d\mu \to \kappa$. We define $g_1 = f_1$ and, inductively, $g_n = \max\{g_{n-1}, f_n\}$ for all $n \ge 2$. Then $g_n \in \mathcal{C}$ for all n. We define $f = \lim_{n\to+\infty} g_n$, and then $g_n \uparrow f$. From $\int_A g_n \, d\mu \le \nu(A)$ for all n and all $A \in \mathcal{S}$ we get, by the Monotone Convergence Theorem, that $\int_A f \, d\mu \le \nu(A)$ for all $A \in \mathcal{S}$. Therefore, $f \in \mathcal{C}$, and so $\int_X f \, d\mu \le \kappa$. On the other hand, we have that $f_n \le g_n \le f$ for all n. Thus, $\int_X f_n \, d\mu \le \int_X f \, d\mu \le \kappa$ for all n and, since $\int_X f_n \, d\mu \to \kappa$, we conclude that

$$\int_X f \, d\mu = \kappa < +\infty.$$

In other words, f is a maximazing element of $\mathcal{C}$ for (6.2).
Since $(\nu - f\mu)(A) = \nu(A) - \int_A f \, d\mu \ge 0$ for all $A \in \mathcal{S}$, the signed measure $\nu - f\mu$ is a finite measure.
If $\nu - f\mu$ and μ are not mutually singular, then by Lemma 6.5 there is $\epsilon_0 > 0$ and $A_0 \in \mathcal{S}$ with $\mu(A_0) > 0$ so that

$$\frac{\nu(A)}{\mu(A)} - \frac{1}{\mu(A)} \int_A f \, d\mu = \frac{(\nu - f\mu)(A)}{\mu(A)} \ge \epsilon_0$$

for all $A \in \mathcal{S}$, $A \subseteq A_0$ with $\mu(A) > 0$. Thus,

$$\int_A (f + \epsilon_0 \chi_{A_0}) \, d\mu \le \nu(A)$$

for all $A \in \mathcal{S}$, $A \subseteq A_0$. Now for any $A \in \mathcal{S}$ we have

$$\begin{aligned}\int_A (f + \epsilon_0 \chi_{A_0}) \, d\mu &= \int_{A\cap A_0} (f + \epsilon_0 \chi_{A_0}) \, d\mu + \int_{A\setminus A_0} (f + \epsilon_0 \chi_{A_0}) \, d\mu \\ &\le \nu(A \cap A_0) + \int_{A\setminus A_0} (f + \epsilon_0 \chi_{A_0}) \, d\mu = \nu(A \cap A_0) + \int_{A\setminus A_0} f \, d\mu \\ &\le \nu(A \cap A_0) + \nu(A \setminus A_0) = \nu(A).\end{aligned}$$

This implies that $f + \epsilon_0 \chi_{A_0}$ belongs to $\mathcal{C}$, and so

$$\kappa + \epsilon_0 \mu(A_0) = \int_X (f + \epsilon_0 \chi_{A_0}) \, d\mu \le \kappa.$$

This is false and we arrived at a contradiction. Therefore, $\nu - f\mu \perp \mu$.
We set $\rho = \nu - f\mu$ and $\lambda = f\mu$ and we have the decomposition $\nu = \lambda + \rho$ with $\lambda \ll \mu$, $\rho \perp \mu$. Both λ and ρ are finite measures and $f : X \to [0, +\infty]$ is integrable with respect to μ, since

$$\lambda(X) = \int_X f \, d\mu = \kappa < +\infty, \quad \rho(X) = \nu(X) - \int_X f \, d\mu = \nu(X) - \kappa < +\infty.$$

(b) We now suppose that both μ, ν are σ-finite measures on $(X, \mathcal{S})$.
Then there are pairwise disjoint $F_1, F_2, \ldots \in \mathcal{S}$ so that $X = \bigcup_{k=1}^{+\infty} F_k$ and $\mu(F_k) < +\infty$ for all k

and pairwise disjoint $G_1, G_2, \ldots \in \mathcal{S}$ so that $X = \bigcup_{l=1}^{+\infty} G_l$ and $\nu(G_l) < +\infty$ for all l. Then the sets $F_k \cap G_l$ are pairwise disjoint, they cover X and $\mu(F_k \cap G_l) < +\infty$, $\nu(F_k \cap G_l) < +\infty$ for all k, l. We enumerate them as $E_1, E_2, \ldots$, and then we have $X = \bigcup_{n=1}^{+\infty} E_n$ and $\mu(E_n) < +\infty$, $\nu(E_n) < +\infty$ for all n.
We consider the restrictions μ_n and ν_n of μ and ν on each E_n. Namely,

$$\mu_n(A) = \mu(A \cap E_n), \quad \nu_n(A) = \nu(A \cap E_n) \quad \text{for all } A \in \mathcal{S}.$$

Then all μ_n, ν_n are finite measures on $(X, \mathcal{S})$, and we also have

$$\mu(A) = \textstyle\sum_{n=1}^{+\infty} \mu_n(A), \quad \nu(A) = \sum_{n=1}^{+\infty} \nu_n(A) \quad \text{for all } A \in \mathcal{S}.$$

Applying the results of part (a), we see that there exist finite measures λ_n, ρ_n on $(X, \mathcal{S})$ and μ_n-integrable $f_n : X \to [0, +\infty]$ so that

$$\nu_n = \lambda_n + \rho_n, \quad \lambda_n \ll \mu_n, \quad \rho_n \perp \mu_n, \quad \lambda_n(A) = \textstyle\int_A f_n \, d\mu_n \quad \text{for all } A \in \mathcal{S}.$$

From $\nu_n(E_n^c) = 0$ we get that $\lambda_n(E_n^c) = \rho_n(E_n^c) = 0$. Now, since $\mu_n(A) = \lambda_n(A) = 0$ for every $A \in \mathcal{S}$, $A \subseteq E_n^c$, the relation $\lambda_n(A) = \int_A f_n \, d\mu_n$ remains true for all $A \in \mathcal{S}$ if we change f_n and make it 0 on E_n^c. Hence, we may assume that

$$f_n = 0 \quad \text{on } E_n^c, \quad \lambda_n(A) = \textstyle\int_{A \cap E_n} f_n \, d\mu_n \quad \text{for all } A \in \mathcal{S}.$$

We define $\lambda, \rho : \mathcal{S} \to [0, +\infty]$ and $f : X \to [0, +\infty]$ by

$$\lambda(A) = \textstyle\sum_{n=1}^{+\infty} \lambda_n(A), \ \ \rho(A) = \sum_{n=1}^{+\infty} \rho_n(A), \ \ f(x) = \sum_{n=1}^{+\infty} f_n(x) \quad \text{for all } A \in \mathcal{S}, x \in X.$$

It is trivial to see that λ and ρ are measures on $(X, \mathcal{S})$ and that f is $\mathcal{S}$-measurable.
Now, the equality $\nu = \lambda + \rho$ is obvious.
If $A \in \mathcal{S}$ has $\mu(A) = 0$, then $\mu_n(A) = \mu(A \cap E_n) = 0$, and so $\lambda_n(A) = 0$ for all n. Hence, $\lambda(A) = 0$, and so $\lambda \ll \mu$.
Since $\rho_n \perp \mu_n$, there is $R_n \in \mathcal{S}$ so that R_n is null for μ_n and R_n^c is null for ρ_n. But, then $R_n' = R_n \cap E_n$ is also null for μ_n and $R_n'^c = R_n^c \cup E_n^c$ is null for ρ_n. Since R_n' is obviously null for all μ_m, $m \neq n$, we have that R_n' is null for μ. Then $R = \bigcup_{n=1}^{+\infty} R_n'$ is null for μ and $R^c = \bigcap_{n=1}^{+\infty} R_n'^c$ is null for all ρ_n and, hence, for ρ. We conclude that $\rho \perp \mu$.
The λ and ρ are σ-finite since $\lambda(E_n) = \lambda_n(E_n) < +\infty$ and $\rho(E_n) = \rho_n(E_n) < +\infty$ for all n.
Finally, for every $A \in \mathcal{S}$,

$$\begin{aligned} \lambda(A) &= \textstyle\sum_{n=1}^{+\infty} \lambda_n(A) = \sum_{n=1}^{+\infty} \int_{A \cap E_n} f_n \, d\mu_n = \sum_{n=1}^{+\infty} \int_{A \cap E_n} f \, d\mu_n \\ &= \textstyle\sum_{n=1}^{+\infty} \int_{A \cap E_n} f \, d\mu = \int_A f \, d\mu. \end{aligned} \tag{6.3}$$

The fourth equality is true because $\int_{E_n} f \, d\mu_n = \int_{E_n} f \, d\mu$ for all $\mathcal{S}$-measurable $f : X \to [0, +\infty]$. This is justified as follows. If $f = \chi_A$ with $A \in \mathcal{S}$, then the equality becomes $\mu_n(A \cap E_n) = \mu(A \cap E_n)$ which is true. Then the equality holds, by linearity, for non-negative $\mathcal{S}$-measurable simple functions. Finally, by the Monotone Convergence Theorem, it holds for all $\mathcal{S}$-measurable $f : X \to [0, +\infty]$.
Now, from (6.3) we conclude that $\lambda = f\mu$ and that $\lambda \ll \mu$.
(c) In the general case we have $\nu = \nu^+ - \nu^-$, and both ν^+, ν^- are σ-finite measures on $(X, \mathcal{S})$. We apply the result of part (b) and we get σ-finite measures $\lambda_1, \lambda_2, \rho_1, \rho_2$ so that $\nu^+ = \lambda_1 + \rho_1$, $\nu^- = \lambda_2 + \rho_2$ and $\lambda_1 \ll \mu$, $\lambda_2 \ll \mu$, $\rho_1 \perp \mu$, $\rho_2 \perp \mu$. Since either ν^+ or ν^- is a finite measure, we have that either λ_1, ρ_1 are finite or λ_2, ρ_2 are finite. Now, we consider $\lambda = \lambda_1 - \lambda_2$ and $\rho = \rho_1 - \rho_2$, and we have that $\nu = \lambda + \rho$ and $\lambda \ll \mu$, $\rho \perp \mu$.
There are also $\mathcal{S}$-measurable $f_1, f_2 : X \to [0, +\infty]$ so that $\lambda_1 = f_1\mu$ and $\lambda_2 = f_2\mu$. Then, either $\int_X f_1 \, d\mu = \lambda_1(X) < +\infty$ or $\int_X f_2 \, d\mu = \lambda_2(X) < +\infty$, and so either $f_1 < +\infty$ μ-a.e. on X or

$f_2 < +\infty$ μ-a.e. on X. Hence, the function $f = f_1 - f_2$ is defined μ-a.e. on X and the integral $\int_X f\,d\mu = \int_X f_1\,d\mu - \int_X f_2\,d\mu$ exists. Now,

$$\lambda(A) = \lambda_1(A) - \lambda_2(A) = \int_A f_1\,d\mu - \int_A f_2\,d\mu = \int_A f\,d\mu$$

for all $A \in \mathcal{S}$, and so $\lambda = f\mu$. $\square$

Lebesgue-Radon-Nikodym Theorem. `The complex case`. *Let ν be a complex measure and μ be a σ-finite measure on $(X, \mathcal{S})$. Then there exist unique complex measures λ and ρ on $(X, \mathcal{S})$ so that*

$$\nu = \lambda + \rho, \quad \lambda \ll \mu, \quad \rho \perp \mu.$$

Moreover, there exists a $\mathcal{S}$-measurable $f : X \to \overline{\mathbb{C}}$ so that f is integrable over X with respect to μ and

$$\lambda = f\mu.$$

If f_1 is another such function, then $f_1 = f$ μ-a.e. on X.
If ν is non-negative, then λ and ρ are non-negative and $f \geq 0$ μ-a.e. on X.
If ν is real, then λ and ρ are real and f is extended-real valued.

Proof. The measures $\mathrm{Re}(\nu)$ and $\mathrm{Im}(\nu)$ are real measures, and, by the previous theorem which deals with the signed case, we have that there exist real measures $\lambda_1, \lambda_2, \rho_1, \rho_2$ on $(X, \mathcal{S})$ so that $\mathrm{Re}(\nu) = \lambda_1 + \rho_1$, $\mathrm{Im}(\nu) = \lambda_2 + \rho_2$ and $\lambda_1 \ll \mu$, $\lambda_2 \ll \mu$ and $\rho_1 \perp \mu$, $\rho_2 \perp \mu$. We define $\lambda = \lambda_1 + i\lambda_2$ and $\rho = \rho_1 + i\rho_2$. Then $\nu = \lambda + \rho$ and $\lambda \ll \mu$ and $\rho \perp \mu$. There are also μ-integrable $f_1, f_2 : X \to \overline{\mathbb{R}}$ so that $\lambda_1 = f_1\mu$ and $\lambda_2 = f_2\mu$. The function $f = f_1 + if_2 : X \to \mathbb{C}$ is μ-a.e. defined, it is μ-integrable, and

$$(f\mu)(A) = \int_A f\,d\mu = \int_A f_1\,d\mu + i\int_A f_2\,d\mu = \lambda_1(A) + i\lambda_2(A) = \lambda(A)$$

for all $A \in \mathcal{S}$. Hence, $\lambda = f\mu$.
The uniqueness is an easy consequence of Theorem 6.11. $\square$

Definition. *(i) Let ν be a signed measure or a complex measure and μ be a measure on $(X, \mathcal{S})$. If there exist, necessarily unique, signed or complex measures λ and ρ on $(X, \mathcal{S})$, so that $\nu = \lambda + \rho$, $\lambda \ll \mu$ and $\rho \perp \mu$, then we say that λ and ρ constitute the* ***Lebesgue decomposition*** *of ν with respect to μ. Also, λ is called the* ***absolutely continuous part*** *and ρ is called the* ***singular part*** *of ν with respect to μ.*
(ii) Let ν be a signed or complex measure and μ be a measure on $(X, \mathcal{S})$ so that $\nu \ll \mu$. If there exists a $\mathcal{S}$-measurable $f : X \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ so that $\int_X f\,d\mu$ is defined and $\nu = f\mu$, then f is called a ***Radon-Nikodym derivative*** *of ν with respect to μ. Any Radon-Nikodym derivative of ν with respect to μ is denoted*

$$\frac{d\nu}{d\mu}.$$

The two Lebesgue-Radon-Nikodym Theorems say that, *if ν and μ are σ-finite, then ν has a unique Lebesgue decomposition with respect to μ.* Moreover, *if ν and μ are σ-finite and $\nu \ll \mu$, then there exists a Radon-Nikodym derivative of ν with respect to μ, which is unique if we disregard μ-null sets.* This is true because $\nu = \nu + 0$ is, necessarily, the Lebesgue decomposition of ν with respect to μ.

We should make some remarks about Radon-Nikodym derivatives.
1. The symbol $\frac{d\nu}{d\mu}$ appears as a fraction of two quantities but it is not. It is like the well known symbol $\frac{dy}{dx}$ of the derivative in elementary calculus.
2. The definition allows *all* Radon-Nikodym derivatives of ν with respect to μ to be denoted by *the same* symbol $\frac{d\nu}{d\mu}$. This is not absolutely strict and it would be more correct to say that $\frac{d\nu}{d\mu}$ is *the collection (or class) of all Radon-Nikodym derivatives of ν with respect to μ.* It is simpler to follow the tradition and use the same symbol for all derivatives. Actually, there is no danger for

confusion in doing this, because the equality $f = \frac{d\nu}{d\mu}$ or its equivalent $\nu = f\mu$ acquires its real meaning through the $\nu(A) = \int_A f\,d\mu$, $A \in \mathcal{S}$.
3. As we just observed, the real meaning of the symbol $\frac{d\nu}{d\mu}$ is through the equality $\nu(A) = \int_A \frac{d\nu}{d\mu}\,d\mu$ for all $A \in \mathcal{S}$, which, after *formally* simplifying the fraction (!!!), changes into the *true* equality $\nu(A) = \int_A d\nu$.
4. Theorem 6.9 implies that the Radon-Nikodym of $\nu \ll \mu$ with respect to μ, if it exists, is unique *when μ is a semifinite measure*, provided we disregard sets of zero μ-measure.

The following propositions give some properties of Radon-Nikodym derivatives of calculus type.

Proposition 6.17. *Let ν_1, ν_2 be complex or σ-finite signed measures and μ be a σ-finite measure on $(X, \mathcal{S})$. If $\nu_1 \ll \mu$, $\nu_2 \ll \mu$ and if $\nu_1 + \nu_2$ is defined, then $\nu_1 + \nu_2 \ll \mu$ and*

$$\tfrac{d(\nu_1+\nu_2)}{d\mu} = \tfrac{d\nu_1}{d\mu} + \tfrac{d\nu_2}{d\mu} \quad \mu\text{-a.e. on } X.$$

Proof. We have $(\nu_1 + \nu_2)(A) = \int_A \frac{d\nu_1}{d\mu}\,d\mu + \int_A \frac{d\nu_2}{d\mu}\,d\mu = \int_A (\frac{d\nu_1}{d\mu} + \frac{d\nu_2}{d\mu})\,d\mu$ for all $A \in \mathcal{S}$. ☐

Proposition 6.18. *Let ν be a complex or a σ-finite signed measure and μ be a σ-finite measure on $(X, \mathcal{S})$. If $\nu \ll \mu$ and $\kappa \in \mathbb{C}$ or $\mathbb{R}$, then $\kappa\nu \ll \mu$ and*

$$\tfrac{d(\kappa\nu)}{d\mu} = \kappa \tfrac{d\nu}{d\mu} \quad \mu\text{-a.e. on } X.$$

Proof. We have $(\kappa\nu)(A) = \kappa \int_A \frac{d\nu}{d\mu}\,d\mu = \int_A (\kappa \frac{d\nu}{d\mu})\,d\mu$ for all $A \in \mathcal{S}$. ☐

The following is the *chain rule*.

Proposition 6.19. *Let ν be a complex or σ-finite signed measure and μ_1, μ_2 be σ-finite measures on $(X, \mathcal{S})$. If $\nu \ll \mu_1$ and $\mu_1 \ll \mu_2$, then $\nu \ll \mu_2$ and*

$$\tfrac{d\nu}{d\mu_2} = \tfrac{d\nu}{d\mu_1} \tfrac{d\mu_1}{d\mu_2} \quad \mu_2\text{-a.e. on } X.$$

Proof. If $A \in \mathcal{S}$ has $\mu_2(A) = 0$, then $\mu_1(A) = 0$, and so $\nu(A) = 0$. Therefore, $\nu \ll \mu_2$. Theorem 6.8 implies that $\nu(A) = \int_A \frac{d\nu}{d\mu_1}\,d\mu_1 = \int_A \frac{d\nu}{d\mu_1} \frac{d\mu_1}{d\mu_2}\,d\mu_2$ for every $A \in \mathcal{S}$. ☐

Proposition 6.20. *Let μ_1 and μ_2 be two σ-finite measures on $(X, \mathcal{S})$. If $\mu_1 \ll \mu_2$ and $\mu_2 \ll \mu_1$, then*

$$\tfrac{d\mu_1}{d\mu_2} \tfrac{d\mu_2}{d\mu_1} = 1 \quad \mu_1\text{-a.e. on } X.$$

Proof. We have $\mu_1(A) = \int_A d\mu_1$ for every $A \in \mathcal{S}$, and so $\frac{d\mu_1}{d\mu_1} = 1$ μ_1-a.e. on X. Now the result is a trivial consequence of Proposition 6.19. ☐

Proposition 6.21. *If ν is a σ-finite measure on $(X, \mathcal{S})$, then $\nu \ll |\nu|$ and*

$$\left|\tfrac{d\nu}{d|\nu|}\right| = 1 \quad \nu\text{-a.e. on } X.$$

Proof. We have $\left|\frac{d\nu}{d|\nu|}\right| |\nu| = \left|\frac{d\nu}{d|\nu|}|\nu|\right| = |\nu|$. Thus, $\left|\frac{d\nu}{d|\nu|}\right| = 1$ $|\nu|$-a.e. on X. ☐

Exercises.

6.5.1. Let $\sharp$ be the counting measure on $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$ and μ be the point-mass distribution on $\mathbb{N}$ induced by the function $a_n = \frac{1}{2^n}$, $n \in \mathbb{N}$. Prove that there is an $\epsilon_0 > 0$ and a sequence (E_k) of subsets of $\mathbb{N}$, so that $\mu(E_k) \to 0$ and $\sharp(E_k) \geq \epsilon_0$ for all k. On the other hand, prove that $\sharp \ll \mu$.

6.5.2. Let ν_1, μ_1 be σ-finite measures on $(X_1, \mathcal{S}_1)$ and ν_2, μ_2 be σ-finite measures on $(X_2, \mathcal{S}_2)$. If $\nu_1 \ll \mu_1$ and $\nu_2 \ll \mu_2$, prove that $\nu_1 \otimes \nu_2 \ll \mu_1 \otimes \mu_2$ and that $\frac{d(\nu_1 \otimes \nu_2)}{d(\mu_1 \otimes \mu_2)}(x_1, x_2) = \frac{d\nu_1}{d\mu_1}(x_1) \frac{d\nu_2}{d\mu_2}(x_2)$ for $(\mu_1 \otimes \mu_2)$-a.e. $(x_1, x_2) \in X_1 \times X_2$.

6.5.3. Let $\sharp$ be the counting measure on $(\mathbb{R}, \mathcal{B}_1)$.

(i) Prove that $m_1 \ll \sharp$. Is there any f so that $m_1 = f\sharp$?

(ii) Is there any Lebesgue decomposition of $\sharp$ with respect to m_1 ?

6.5.4. *Generalization of the Lebesgue-Radon-Nikodym Theorem.*
Let ν be a signed measure and μ be a σ-finite measure on $(X, \mathcal{S})$ so that $\nu \ll \mu$. Prove that there is a measurable $f : X \to \overline{\mathbb{R}}$, so that $\int_X f\, d\mu$ exists and $\nu = f\mu$.

6.5.5. *Generalization of the Lebesgue Decomposition Theorem.*
Let ν be a σ-finite signed measure and μ a measure on $(X, \mathcal{S})$. Prove that there are unique σ-finite signed measures λ, ρ on $(X, \mathcal{S})$ so that $\lambda \ll \mu$, $\rho \perp \mu$ and $\nu = \lambda + \rho$.

6.5.6. Let ν, μ be two measures on $(X, \mathcal{S})$ with $\nu \ll \mu$. If $\lambda = \mu + \nu$, prove that $\nu \ll \lambda$. If $f : X \to [0, +\infty]$ is measurable and $\nu = f\lambda$, prove that $0 \leq f < 1$ μ-a.e. on X and $\nu = \frac{f}{1-f}\mu$.

6.5.7. Let ν be a signed measure on $(X, \mathcal{S})$. Prove that $\nu^+, \nu^- \ll |\nu|$ and find formulas for the Radon-Nikodym derivatives $\frac{d\nu^+}{d|\nu|}$ and $\frac{d\nu^-}{d|\nu|}$.

6.5.8. Let μ be a finite measure on $(X, \mathcal{S})$. We define $d(A, B) = \mu(A \triangle B)$ for all $A, B \in \mathcal{S}$.

(i) Prove that $(\mathcal{S}, d)$ is a complete metric space.

(ii) If ν is a real or a complex measure on $(X, \mathcal{S})$, prove that ν is continuous on $\mathcal{S}$ (with respect to d) if and only if ν is continuous at $\emptyset$ (with respect to d) if and only if $\nu \ll \mu$.

6.5.9. *Conditional Expectation.*
Let $(X, \mathcal{S})$ be a measurable space and $\mathcal{S}_0$ be a σ-algebra with $\mathcal{S}_0 \subseteq \mathcal{S}$. Let μ be a measure on $(X, \mathcal{S})$ which is σ-finite on $(X, \mathcal{S}_0)$ and let us denote by the same symbol μ the restriction of the measure on $(X, \mathcal{S}_0)$.
If $f : X \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ is $\mathcal{S}$-measurable and $\int_X f\, d\mu$ exists, prove that there is a $\mathcal{S}_0$-measurable $f_0 : X \to \overline{\mathbb{R}}$ or, respectively, $\overline{\mathbb{C}}$ so that $\int_X f_0\, d\mu$ exists and $\int_A f_0\, d\mu = \int_A f\, d\mu$ for all $A \in \mathcal{S}_0$.
If h_0 has the same properties as f_0, prove that $h_0 = f_0$ μ-a.e. on X.
Any f_0 with the above properies is called a **conditional expectation** of f with respect to $\mathcal{S}_0$ and it is denoted by $E(f|\mathcal{S}_0)$.
Prove:

(i) $E(f|\mathcal{S}) = f$ μ-a.e. on X.

(ii) $E(f + g|\mathcal{S}_0) = E(f|\mathcal{S}_0) + E(g|\mathcal{S}_0)$ μ-a.e. on X.

(iii) $E(\kappa f|\mathcal{S}_0) = \kappa E(f|\mathcal{S}_0)$ μ-a.e. on X.

(iv) If g is $\mathcal{S}_0$-measurable, then $E(gf|\mathcal{S}_0) = gE(f|\mathcal{S}_0)$ μ-a.e. on X.

(v) If $\mathcal{S}_1 \subseteq \mathcal{S}_0 \subseteq \mathcal{S}$, then $E(f|\mathcal{S}_1) = E(E(f|\mathcal{S}_0)|\mathcal{S}_1)$ μ-a.e. on X.

6.6 Differentiation.

DIFFERENTIATION OF INDEFINITE INTEGRALS OVER $\mathbb{R}^n$.

Let $f : [a, b] \to \mathbb{R}$ be a Riemann integrable function. The Fundamental Theorem of Calculus says that for every $x \in [a, b]$ which is a continuity point of f we have $\frac{d}{dx} \int_a^x f(y)\, dy = f(x)$. Of course, this means that

$$\lim_{r \to 0+} \big(\int_a^{x+r} f(y)\, dy - \int_a^x f(y)\, dy \big)/r = \lim_{r \to 0+} \big(\int_a^x f(y)\, dy - \int_a^{x-r} f(y)\, dy \big)/r = f(x).$$

Adding the two limits, we find

$$\lim_{r \to 0+} \big(\int_{x-r}^{x+r} f(y)\, dy \big)/(2r) = f(x).$$

In this (and the next) section we shall prove a far reaching generalisation of this result: *a fundamental theorem of calculus for indefinite Lebesgue integrals and, more generally, for locally finite Borel measures on* $\mathbb{R}^n$.

Wiener's Lemma. *Let $B_1, \dots, B_m$ be open balls in $\mathbb{R}^n$. There exist pairwise disjoint $B_{i_1}, \dots, B_{i_k}$ so that $m_n(B_{i_1}) + \cdots + m_n(B_{i_k}) \geq \frac{1}{3^n}\, m_n(B_1 \cup \cdots \cup B_m)$.*

Proof. From $B_1, \dots, B_m$ we choose a ball B_{i_1} with largest radius. (There may be more than one balls with the same largest radius and we choose any one of them.) Together with B_{i_1} we collect all other balls, its *satellites*, which intersect it and call their union (B_{i_1} included) C_1. Since each of these balls has radius not larger than the radius of B_{i_1}, we see that $C_1 \subseteq B_{i_1}^*$, where $B_{i_1}^*$ is the ball with the same center as B_{i_1} and radius three times the radius of B_{i_1}.
Therefore, $m_n(C_1) \leq m_n(B_{i_1}^*) = 3^n m_n(B_{i_1})$.
The remaining balls have empty intersection with B_{i_1} and from them we choose a ball B_{i_2} with largest radius. Of course, B_{i_2} does not intersect B_{i_1}. Together with B_{i_2} we collect all other balls (from the remaining ones), its satellites, which intersect it and call their union (B_{i_2} included) C_2. Since each of these balls has radius not larger than the radius of B_{i_2}, we have $C_2 \subseteq B_{i_2}^*$, where $B_{i_2}^*$ is the ball with the same center as B_{i_2} and radius three times the radius of B_{i_2}.
Therefore, $m_n(C_2) \leq m_n(B_{i_2}^*) = 3^n m_n(B_{i_2})$.
We continue this procedure and, since at every step at least one ball is collected (B_{i_1} at the first step, B_{i_2} at the second step and so on), after at most m steps, say at the kth step, the procedure will stop. Namely, after the first $k-1$ steps, the remaining balls have empty intersection with $B_{i_1}, \dots, B_{i_{k-1}}$ and from them we choose a ball B_{i_k} with largest radius. This B_{i_k} does not intersect $B_{i_1}, \dots, B_{i_{k-1}}$. All remaining balls intersect B_{i_k}, they are its satellites, (since this is the step where the procedure stops) and form their union (B_{i_k} included) C_k. Since each of these balls has radius not larger than the radius of B_{i_k}, we have $C_k \subseteq B_{i_k}^*$, where $B_{i_k}^*$ is the ball with the same center as B_{i_k} and radius three times the radius of B_{i_k}.
Therefore, $m_n(C_k) \leq m_n(B_{i_k}^*) = 3^n m_n(B_{i_k})$.
Clearly, each of the original balls $B_1, \dots, B_m$ is either chosen as one of $B_{i_1}, \dots, B_{i_k}$ or is a satellite of one of $B_{i_1}, \dots, B_{i_k}$. Therefore, $B_1 \cup \cdots \cup B_m = C_1 \cup \cdots \cup C_k$, and so

$$\begin{aligned} m_n(B_1 \cup \cdots \cup B_m) &= m_n(C_1 \cup \cdots \cup C_k) \leq m_n(C_1) + \cdots + m_n(C_k) \\ &\leq 3^n\big(m_n(B_{i_1}) + \cdots + m_n(B_{i_k})\big) \end{aligned}$$

and the proof is complete. □

Definition. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ be Lebesgue measurable. We say f is* ***locally Lebesgue integrable*** *if for all $x \in \mathbb{R}^n$ there is an open neighborhood U_x of x so that $\int_{U_x} |f(y)|\, dm_n(y) < +\infty$.*

Lemma 6.7. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ be locally Lebesgue integrable. Then $\int_M |f(y)|\, dm_n(y) < +\infty$ for every bounded set $M \in \mathcal{L}_n$.*

Proof. Let f be locally Lebesgue integrable and $M \in \mathcal{L}_n$ be bounded. We consider any compact $K \subseteq \mathbb{R}^n$ so that $M \subseteq K$. Such a K is the closure of M or just a closed ball or a closed cube including M. For each $x \in K$ there is an open neighborhood U_x of x so that $\int_{U_x} |f(y)|\, dm_n(y) < +\infty$. Since $K \subseteq \bigcup_{x\in K} U_x$, there are finitely many $x_1, \dots, x_m$ so that $M \subseteq K \subseteq U_{x_1} \cup \cdots \cup U_{x_m}$. This implies

$$\int_M |f(y)|\, dm_n(y) \leq \int_{U_{x_1}} |f(y)|\, dm_n(y) + \cdots + \int_{U_{x_m}} |f(y)|\, dm_n(y) < +\infty.$$

If, conversely, $\int_M |f(y)|\, dm_n(y) < +\infty$ for every bounded set $M \in \mathcal{L}_n$, then f is locally Lebesgue integrable since $\int_{B(x;1)} |f(y)|\, dm_n(y) < +\infty$ for every x. □

Proposition 6.22. *Let $f, f_1, f_2 : \mathbb{R}^n \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ be locally Lebesgue integrable and $\kappa \in \mathbb{C}$. Then*
(i) f is finite a.e. on $\mathbb{R}^n$,
(ii) $f_1 + f_2$ is defined a.e. on $\mathbb{R}^n$ and it is locally Lebesgue integrable,
(iii) κf is locally Lebesgue integrable.

Proof. (i) Lemma 6.6 implies $\int_{B(0;k)} |f(y)| \, dm_n(y) < +\infty$, and so f is finite a.e. on $B(0;k)$ for every k. Since $\mathbb{R}^n = \bigcup_{k=1}^{+\infty} B(0;k)$, we have that f is finite a.e. on $\mathbb{R}^n$.
(ii) By the result of (i), both f_1, f_2 are finite a.e. on $\mathbb{R}^n$, and so $f_1 + f_2$ is defined a.e. on $\mathbb{R}^n$. We have

$$\int_M |f_1(y) + f_2(y)| \, dm_n(y) \leq \int_M |f_1(y)| \, dm_n(y) + \int_M |f_2(y)| \, dm_n(y) < +\infty$$

for every bounded $M \in \mathcal{L}_n$, and, by Lemma 6.6, $f_1 + f_2$ is locally Lebesgue integrable.
(iii) Similarly,

$$\int_M |\kappa f(y)| \, dm_n(y) = |\kappa| \int_M |f(y)| \, dm_n(y) < +\infty$$

for all bounded $M \in \mathcal{L}_n$, and so κf is locally Lebesgue integrable. $\square$

The need for local Lebesgue integrability (or for local finiteness of measures) is for definitions like the following one to make sense. Of course, we may restrict to Lebesgue integrability if we like.

Definition. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ be locally Lebesgue integrable. Then $M(f) : \mathbb{R}^n \to [0, +\infty]$, defined by*

$$M(f)(x) = \sup_{B \text{ open ball},\, B \ni x} \frac{1}{m_n(B)} \int_B |f(y)| \, dm_n(y)$$

for all $x \in \mathbb{R}^n$, is called the ***Hardy-Littlewood maximal function*** *of f.*

Proposition 6.23. *Let $f, f_1, f_2 : \mathbb{R}^n \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ be locally Lebesgue integrable and $\kappa \in \mathbb{C}$. Then*
(i) $M(f_1 + f_2) \leq M(f_1) + M(f_2)$,
(ii) $M(\kappa f) = |\kappa| M(f)$.

Proof. (i) For all x and all open balls $B \ni x$,

$$\begin{aligned} \frac{1}{m_n(B)} \int_B |f_1(y) + f_2(y)| \, dm_n(y) &\leq \frac{1}{m_n(B)} \int_B |f_1(y)| \, dm_n(y) + \frac{1}{m_n(B)} \int_B |f_2(y)| \, dm_n(y) \\ &\leq M(f_1)(x) + M(f_2)(x). \end{aligned}$$

Taking the supremum of the left side, we get $M(f_1 + f_2)(x) \leq M(f_1)(x) + M(f_2)(x)$.
(ii) Similarly, for all x and all open balls $B \ni x$,

$$\frac{1}{m_n(B)} \int_B |\kappa f(y)| \, dm_n(y) = |\kappa| \frac{1}{m_n(B)} \int_B |f(y)| \, dm_n(y) \leq |\kappa| M(f)(x)$$

and, taking the supremum of the left side, we get $M(\kappa f)(x) \leq |\kappa| M(f)(x)$. Now, if $\kappa \neq 0$, we apply this inequality to the number $\frac{1}{\kappa}$ and to the function κf, and we get $M(f)(x) \leq \frac{1}{|\kappa|} M(\kappa f)(x)$. The two inequalities imply $M(\kappa f)(x) = |\kappa| M(f)(x)$. On the other hand, if $\kappa = 0$, then the equality is trivial. $\square$

Lemma 6.8. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ be locally Lebesgue integrable. Then for every $t > 0$ the set $\{x \in \mathbb{R}^n \,|\, t < M(f)(x)\}$ is open in $\mathbb{R}^n$.*

Proof. Let $U = \{x \in \mathbb{R}^n \,|\, t < M(f)(x)\}$ and $x \in U$. Then $t < M(f)(x)$, and so there is an open ball $B \ni x$ so that

$$t < \frac{1}{m_n(B)} \int_B |f(y)| \, dm_n(y).$$

If we take an arbitrary $x' \in B$, then

$$t < \frac{1}{m_n(B)} \int_B |f(y)| \, dm_n(y) \leq M(f)(x').$$

Therefore, $B \subseteq U$, and so U is open in $\mathbb{R}^n$. $\square$

Since $\{x \in \mathbb{R}^n \,|\, t < M(f)(x)\}$ is open, it is also a Lebesgue set.

Hardy-Littlewood Theorem. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ be Lebesgue integrable. Then for every $t > 0$ we have*

$$m_n(\{x \in \mathbb{R}^n \,|\, t < M(f)(x)\}) \leq \tfrac{3^n}{t} \int_{\mathbb{R}^n} |f(y)| \, dm_n(y).$$

Proof. We consider an arbitrary compact $K \subseteq U = \{x \in \mathbb{R}^n \,|\, t < M(f)(x)\}$. Then for every $x \in K$ we have $t < M(f)(x)$, and this implies that there is an open ball B_x containing x so that

$$t < \tfrac{1}{m_n(B_x)} \int_{B_x} |f(y)| \, dm_n(y).$$

Since $K \subseteq \bigcup_{x \in K} B_x$, there are $x_1, \dots, x_m$ so that $K \subseteq B_{x_1} \cup \dots \cup B_{x_m}$. Wiener's Lemma implies that there exist *pairwise disjoint* $B_{x_{i_1}}, \dots, B_{x_{i_k}}$ so that

$$m_n(B_{x_1} \cup \dots \cup B_{x_m}) \leq 3^n \big(m_n(B_{x_{i_1}}) + \dots + m_n(B_{x_{i_k}})\big).$$

Then

$$\begin{aligned} m_n(K) &\leq 3^n \big(m_n(B_{x_{i_1}}) + \dots + m_n(B_{x_{i_k}})\big) \\ &\leq \tfrac{3^n}{t} \big(\int_{B_{x_{i_1}}} |f(y)| \, dm_n(y) + \dots + \int_{B_{x_{i_k}}} |f(y)| \, dm_n(y)\big) \\ &= \tfrac{3^n}{t} \int_{B_{x_{i_1}} \cup \dots \cup B_{x_{i_k}}} |f(y)| \, dm_n(y) \leq \tfrac{3^n}{t} \int_{\mathbb{R}^n} |f(y)| \, dm_n(y). \end{aligned}$$

By the regularity of m_n, the supremum of $m_n(K)$ for all compact $K \subseteq U$ is equal to $m_n(U)$, and we conclude that $m_n(U) \leq \frac{3^n}{t} \int_{\mathbb{R}^n} |f(y)| \, dm_n(y)$. $\square$

Observe that $m_n(\{x \in \mathbb{R}^n \,|\, t < M(f)(x)\})$ is nothing but the value at t of the distribution function $\lambda_{M(f)}$ of $M(f)$. Therefore, another way to state the result of the Hardy-Littlewood Theorem is

$$\lambda_{M(f)}(t) \leq \tfrac{3^n}{t} \int_{\mathbb{R}^n} |f(y)| \, dm_n(y).$$

Definition. *Let $(X, \mathcal{S}, \mu)$ be a measure space and $g : X \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ be $\mathcal{S}$-measurable. We say that g is* ***weakly μ-integrable*** *over X if there is a constant $c < +\infty$ so that $\lambda_{|g|}(t) \leq \frac{c}{t}$ for every $t > 0$.*

Another way to state the Hardy-Littlewood Theorem is: *if f is Lebesgue integrable, then $M(f)$ is weakly Lebesgue integrable.*

Proposition 6.24. *Let $(X, \mathcal{S}, \mu)$ be a measure space, $g, g_1, g_2 : X \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ be weakly μ-integrable and $\kappa \in \mathbb{C}$. Then*
(i) g is finite a.e. on X,
(ii) $g_1 + g_2$ is defined a.e. on X and it is weakly μ-integrable,
(iii) κg is weakly μ-integrable.

Proof. (i) $\lambda_{|g|}(t) \leq \frac{c}{t}$ for all $t > 0$ implies that

$$\mu(\{x \in X \,|\, |g(x)| = +\infty\}) \leq \mu(\{x \in X \,|\, n < |g(x)|\}) \leq \tfrac{c}{n}$$

for all n, and so $\mu(\{x \in X \,|\, |g(x)| = +\infty\}) = 0$.
(ii) By (i) both g_1 and g_2 are finite a.e. on X, and so $g_1 + g_2$ is defined a.e. on X.
If $\mu(\{x \in X \,|\, t < |g_1(x)|\}) \leq \frac{c_1}{t}$ and $\mu(\{x \in X \,|\, t < |g_2(x)|\}) \leq \frac{c_2}{t}$ for all $t > 0$, then

$$\begin{aligned} \mu(\{x \in X \,|\, t < |g_1(x) + g_2(x)|\}) &\leq \mu\big(\big\{x \in X \,\big|\, \tfrac{t}{2} < |g_1(x)|\big\}\big) + \mu\big(\big\{x \in X \,\big|\, \tfrac{t}{2} < |g_2(x)|\big\}\big) \\ &\leq \tfrac{2c_1 + 2c_2}{t} \end{aligned}$$

for all $t > 0$.
(iii) If $\mu(\{x \in X \,|\, t < |g(x)|\}) \leq \frac{c}{t}$ for all $t > 0$, then

$$\mu(\{x \in X \,|\, t < |\kappa g(x)|\}) = \mu\big(\big\{x \in X \,\big|\, \tfrac{t}{|\kappa|} < |g(x)|\big\}\big) \leq \tfrac{c|\kappa|}{t}$$

for all $t > 0$. $\square$

Proposition 6.25. *Let $(X, \mathcal{S}, \mu)$ be a measure space and $g : X \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ be μ-integrable. Then g is weakly μ-integrable.*

Proof. We have

$$\lambda_{|g|}(t) = \mu(\{x \in X \mid t < |g(x)|\}) \leq \tfrac{1}{t} \int_{\{x \in X \mid t < |g(x)|\}} |g| \, d\mu \leq \tfrac{1}{t} \int_X |g| \, d\mu$$

for all $t > 0$. Therefore, $\lambda_{|g|}(t) \leq \frac{c}{t}$ for all $t > 0$, where $c = \int_X |g| \, d\mu$. □

Example. The converse of Proposition 6.25 is not true. Consider, for example, the function $g(x) = \frac{1}{|x|^n}$, $x \in \mathbb{R}^n$. Then

$$\int_{\mathbb{R}^n} |g(x)| \, dm_n(x) = \sigma_{n-1}(\mathbb{S}^{n-1}) \int_0^{+\infty} \tfrac{1}{r^n} \, r^{n-1} \, dr = \sigma_{n-1}(\mathbb{S}^{n-1}) \int_0^{+\infty} \tfrac{1}{r} \, dr = +\infty.$$

But $\{x \in \mathbb{R}^n \mid t < |g(x)|\} = B(0; t^{-1/n})$, the open ball with center 0 and radius $t^{-1/n}$. Thus,

$$\lambda_{|g|}(t) = m_n(B(0; t^{-1/n})) = (t^{-1/n})^n m_n(B(0; 1)) = \tfrac{c}{t},$$

where $c = m_n(B(0; 1))$.

The next result says that the Hardy-Littlewood maximal function of any f is *not* Lebesgue integrable, except only when $f = 0$ m_n-a.e. on $\mathbb{R}^n$.

Proposition 6.26. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ be locally Lebesgue integrable. If $M(f)$ is Lebesgue integrable, then $f = 0$ m_n-a.e. on $\mathbb{R}^n$.*

Proof. Let $A = \{x \in \mathbb{R}^n \mid f(x) \neq 0\}$, and let us assume that $m_n(A) > 0$.
Since $A = \bigcup_{k=1}^{+\infty}(A \cap B(0; k))$, we get that $m_n(A \cap B(0; k)) > 0$ for at least one $k \geq 1$. We set $M = A \cap B(0; k)$, and we have got a bounded $M \in \mathcal{L}_n$ so that $m_n(M) > 0$ and $\|x\| \leq k$ for every $x \in M$. Since $f(x) \neq 0$ for every $x \in M$, we have that $\int_M |f(y)| \, dm_n(y) > 0$.
We consider any x with $\|x\| \geq k$, and we observe that there is an open ball B of diameter $\|x\|+k+1$ containing x and including M. Then

$$m_n(B) = \big(\tfrac{\|x\|+k+1}{2}\big)^n m_n(B(0; 1)) \leq \big(\tfrac{3\|x\|}{2}\big)^n m_n(B(0; 1)),$$

and so

$$M(f)(x) \geq \tfrac{1}{m_n(B)} \int_B |f(y)| \, dm_n(y) \geq \tfrac{2^n}{3^n \|x\|^n m_n(B(0;1))} \int_M |f(y)| \, dm_n(y) = \tfrac{c}{\|x\|^n},$$

with $c = \frac{2^n}{3^n m_n(B(0;1))} \int_M |f(y)| \, dm_n(y) > 0$. This implies

$$\int_{\mathbb{R}^n} |M(f)(x)| \, dm_n(x) \geq c \int_{\{x \in \mathbb{R}^n \mid \|x\| \geq k\}} \tfrac{1}{\|x\|^n} \, dm_n(x) = +\infty.$$

Therefore, if $M(f)$ is Lebesgue integrable, then $m_n(A) = 0$. □

The next result is a direct generalization of the Fundamental Theorem of Calculus and the proofs are identical.

Lemma 6.9. *Let $g : \mathbb{R}^n \to \mathbb{C}$ be continuous on $\mathbb{R}^n$. Then*

$$\lim_{r \to 0+} \tfrac{1}{m_n(B(x;r))} \int_{B(x;r)} |g(y) - g(x)| \, dm_n(y) = 0$$

for every $x \in \mathbb{R}^n$.

Proof. Let $\epsilon > 0$ be arbitrary. Then there is $\delta > 0$ so that $|g(y) - g(x)| \leq \epsilon$ for every $y \in \mathbb{R}^n$ with $\|y - x\| < \delta$. Then

$$\tfrac{1}{m_n(B(x;r))} \int_{B(x;r)} |g(y) - g(x)| \, dm_n(y) \leq \tfrac{1}{m_n(B(x;r))} \int_{B(x;r)} \epsilon \, dm_n(y) = \epsilon$$

for every $r < \delta$. □

Lebesgue's Theorem. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ be locally Lebesgue integrable. Then,*

$$\lim_{r\to 0+} \frac{1}{m_n(B(x;r))} \int_{B(x;r)} |f(y) - f(x)|\, dm_n(y) = 0$$

for m_n-a.e. $x \in \mathbb{R}^n$.

Proof. (a) Let f be Lebesgue integrable.
We consider an arbitrary $\epsilon > 0$. Theorem 3.14 implies that there is $g : \mathbb{R}^n \to \mathbb{C}$ continuous on $\mathbb{R}^n$ so that $\int_{\mathbb{R}^n} |g - f|\, dm_n < \epsilon$. For all $x \in \mathbb{R}^n$ and $r > 0$ we get

$$\begin{aligned}
&\frac{1}{m_n(B(x;r))} \int_{B(x;r)} |f(y) - f(x)|\, dm_n(y) \\
&\quad\le \frac{1}{m_n(B(x;r))} \int_{B(x;r)} |f(y) - g(y)| dm_n(y) + \frac{1}{m_n(B(x;r))} \int_{B(x;r)} |g(y) - g(x)| dm_n(y) \\
&\qquad + \frac{1}{m_n(B(x;r))} \int_{B(x;r)} |g(x) - f(x)|\, dm_n(y) \\
&\quad\le M(f-g)(x) + \frac{1}{m_n(B(x;r))} \int_{B(x;r)} |g(y) - g(x)|\, dm_n(y) + |g(x) - f(x)|.
\end{aligned}$$

We set

$$A(f)(x;r) = \frac{1}{m_n(B(x;r))} \int_{B(x;r)} |f(y) - f(x)|\, dm_n(y)$$

and the last inequality together with Lemma 6.8 imply

$$\overline{\lim_{r\to 0+}}\, A(f)(x;r) \le M(f-g)(x) + 0 + |g(x) - f(x)|.$$

Now for every $t > 0$ we get

$$\begin{aligned}
&m_n^*(\{x \in \mathbb{R}^n \,|\, t < \overline{\lim_{r\to 0+}}\, A(f)(x;r)\}) \\
&\qquad\le m_n(\{x \in \mathbb{R}^n \,|\, \tfrac{t}{2} < M(f-g)(x)\}) + m_n(\{x \in \mathbb{R}^n \,|\, \tfrac{t}{2} < |g(x) - f(x)|\}) \\
&\qquad\le \tfrac{2\cdot 3^n}{t} \int_{\mathbb{R}^n} |f - g|\, dm_n + \tfrac{2}{t} \int_{\mathbb{R}^n} |f - g|\, dm_n \le \tfrac{2\cdot 3^n + 2}{t}\, \epsilon,
\end{aligned}$$

where the second inequality is a consequence of the Hardy-Littlewood Theorem. Since ϵ is arbitrary, for all $t > 0$ we have $m_n^*(\{x \in \mathbb{R}^n \,|\, t < \overline{\lim}_{r\to 0+} A(f)(x;r)\}) = 0$. By the subadditivity of m_n^*,

$$\begin{aligned}
&m_n^*(\{x \in \mathbb{R}^n \,|\, 0 < \overline{\lim_{r\to 0+}}\, A(f)(x;r)\}) \\
&\qquad\le \textstyle\sum_{k=1}^{+\infty} m_n^*\big(\{x \in \mathbb{R}^n \,\big|\, \frac{1}{k} < \overline{\lim}_{r\to 0+} A(f)(x;r)\}\big) = 0,
\end{aligned}$$

and so $m_n^*(\{x \in \mathbb{R}^n \,|\, 0 < \overline{\lim}_{r\to 0+} A(f)(x;r)\}) = 0$.
Thus, $\overline{\lim}_{r\to 0+} A(f)(x;r) \le 0$ for m_n-a.e. $x \in \mathbb{R}^n$ and, since $A(f)(x;r) \ge 0$ for every $x \in \mathbb{R}^n$ and $r > 0$, we conclude that $\lim_{r\to 0+} A(f)(x;r) = 0$ for m_n-a.e. $x \in \mathbb{R}^n$.
(b) Now let f be locally Lebesgue integrable. We fix an arbitrary $k \ge 2$ and consider the function $h = f\chi_{B(0;k)}$. Then h is Lebesgue integrable and for every $x \in B(0; k-1)$ and every $r \le 1$ we have $A(f)(x;r) = A(h)(x;r)$. By what we have already proved this implies that $\lim_{r\to 0+} A(f)(x;r) = 0$ for m_n-a.e. $x \in B(0;k-1)$. Since k is arbitrary, we conclude that $\lim_{r\to 0+} A(f)(x;r) = 0$ for m_n-a.e. $x \in \mathbb{R}^n$. □

Definition. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ be locally Lebesgue integrable. The set L_f of all $x \in \mathbb{R}^n$ for which* $\lim_{r\to 0+} \frac{1}{m_n(B(x;r))} \int_{B(x;r)} |f(y) - f(x)|\, dm_n(y) = 0$ *is called the* ***Lebesgue set*** *of f.*

Example. If x is a continuity point of f, then x belongs to the Lebesgue set of f. The proof of this fact is, actually, the proof of Lemma 6.8.

Theorem 6.10. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ be locally Lebesgue integrable. Then for every x in the Lebesgue set of f we have*

$$\lim_{r\to 0+} \frac{1}{m_n(B(x;r))} \int_{B(x;r)} f(y)\, dm_n(y) = f(x).$$

Proof. Indeed, for all $x \in L_f$ we have

$$\left|\frac{1}{m_n(B(x;r))}\int_{B(x;r)} f(y)\,dm_n(y) - f(x)\right| \le \frac{1}{m_n(B(x;r))}\int_{B(x;r)} |f(y)-f(x)|\,dm_n(y) \to 0$$

as $r \to 0+$. $\square$

Definition. *Let $x \in \mathbb{R}^n$ and $\mathcal{C}$ be a collection of sets in $\mathcal{L}_n$ with the property that there is a $c > 0$ so that for every $E \in \mathcal{C}$ there is a ball $B(x;r)$ with $E \subseteq B(x;r)$ and $m_n(E) \ge cm_n(B(x;r))$. Then the collection $\mathcal{C}$ is called a* ***thick family of sets*** *at x.*

Example. Any collection of cubes containing x and any collection of balls containing x is a thick family of sets at x.

Example. Consider any collection $\mathcal{C}$ all elements of which are bounded intervals S containing x. Let A_S be the length of the largest edge and a_S be the length of the smallest edge of S. If there is a constant $c > 0$ so that $\frac{a_S}{A_S} \ge c$ for every $S \in \mathcal{C}$, then $\mathcal{C}$ is a thick family of sets at x.

Theorem 6.11. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ be locally Lebesgue integrable. Then for every x in the Lebesgue set of f and for every thick family $\mathcal{C}$ of sets at x we have*

$$\lim_{E\in\mathcal{C},m_n(E)\to 0+} \frac{1}{m_n(E)}\int_E |f(y)-f(x)|\,dm_n(y) = 0$$

$$\lim_{E\in\mathcal{C},m_n(E)\to 0+} \frac{1}{m_n(E)}\int_E f(y)\,dm_n(y) = f(x).$$

Proof. There is a $c > 0$ so that for every $E \in \mathcal{C}$ there is a ball $B(x;r_E)$ with $E \subseteq B(x;r_E)$ and $m_n(E) \ge cm_n(B(x;r_E))$. If $x \in L_f$, then for every $\epsilon > 0$ there is a $\delta > 0$ so that $r < \delta$ implies

$$\frac{1}{m_n(B(x;r))}\int_{B(x;r)} |f(y)-f(x)|\,dm_n(y) < c\epsilon.$$

If $m_n(E) < c\delta^n m_n(B(0;1))$, then $r_E < \delta$, and so

$$\frac{1}{m_n(E)}\int_E |f(y)-f(x)|\,dm_n(y) \le \frac{1}{cm_n(B(x;r_E))}\int_{B(x;r_E)} |f(y)-f(x)|\,dm_n(y) < \epsilon.$$

Therefore, $\lim_{E\in\mathcal{C},m_n(E)\to 0+} \frac{1}{m_n(E)}\int_E |f(y)-f(x)|\,dm_n(y) = 0$.
The proof of the second limit is now trivial. $\square$

DIFFERENTIATION OF BOREL MEASURES ON $\mathbb{R}^n$.

Definition. *Any signed or complex measure on $(\mathbb{R}^n, \mathcal{B}_n)$ is called a* ***Borel*** *signed or complex measure on $\mathbb{R}^n$.*

Definition. *Let ν be a Borel signed measure on $\mathbb{R}^n$. We say that ν is* ***locally finite*** *if for every $x \in \mathbb{R}^n$ there is an open neighborhood U_x of x so that $\nu(U_x)$ is finite.*

This definition is indifferent for complex measures, since complex measures take only finite values.

Proposition 6.27. *Let ν be a Borel signed measure on $\mathbb{R}^n$. Then ν is locally finite if and only if ν^+ and ν^- are both locally finite if and only if $|\nu|$ is locally finite.*

Proof. Since $|\nu| = \nu^+ + \nu^-$, the second equivalence is trivial to prove. It is also trivial to prove that ν is locally finite if $|\nu|$ is locally finite.
Let ν be locally finite. For an arbitrary $x \in \mathbb{R}^n$ there is an open neighborhood U_x of x so that $\nu(U_x)$ is finite. Since $\nu(U_x) = \nu^+(U_x) - \nu^-(U_x)$, both $\nu^+(U_x)$ and $\nu^-(U_x)$, and so also $|\nu|(U_x)$ are finite. Therefore, $|\nu|$ is locally finite. $\square$

Proposition 6.28. *Let ν be a locally finite Borel signed measure on $\mathbb{R}^n$. Then $\nu(M)$ is finite for all bounded Borel sets $M \subseteq \mathbb{R}^n$.*

Proof. Proposition 6.27 implies that $|\nu|$ is locally finite. Now let $M \in \mathcal{B}_n$ be bounded. We consider any compact $K \subseteq \mathbb{R}^n$ so that $M \subseteq K$. For every $x \in K$ there is an open neighborhood U_x of x so that $|\nu|(U_x) < +\infty$. Since $K \subseteq \bigcup_{x \in K} U_x$, there are finitely many $x_1, \ldots, x_m$ so that $M \subseteq K \subseteq U_{x_1} \cup \cdots \cup U_{x_m}$. This implies

$$|\nu(M)| \leq |\nu|(M) \leq |\nu|(U_{x_1}) + \cdots + |\nu|(U_{x_m}) < +\infty,$$

and so $\nu(M)$ is finite. □

Theorem 6.12. *Let ρ be a locally finite Borel signed measure or a Borel complex measure on $\mathbb{R}^n$ with $\rho \perp m_n$. Then*

$$\lim_{r \to 0+} \tfrac{\rho(B(x;r))}{m_n(B(x;r))} = 0$$

for m_n-a.e. $x \in \mathbb{R}^n$.

Proof. If ρ is complex, then $|\rho|$ is a finite Borel measure on $\mathbb{R}^n$. Proposition 6.27 implies that, if ρ is signed, then $|\rho|$ is a locally finite Borel measure on $\mathbb{R}^n$. Moreover, Lemma 6.22 implies that $|\rho| \perp m_n$. Hence, there exist sets $R, M \in \mathcal{B}_n$ with $M \cup R = \mathbb{R}^n$, $M \cap R = \emptyset$ so that R is null for m_n and M is null for $|\rho|$.
We define

$$A(|\rho|)(x;r) = \tfrac{|\rho|(B(x;r))}{m_n(B(x;r))},$$

we take an arbitrary $t > 0$, and we consider the set

$$M_t = \big\{x \in M \,\big|\, t < \overline{\lim}_{r \to 0+} A(|\rho|)(x;r)\big\}.$$

Since $|\rho|$ is a regular measure and $|\rho|(M) = 0$, there is an open set U so that $M_t \subseteq M \subseteq U$ and $|\rho|(U) < \epsilon$. For each $x \in M_t$, there is a small enough $r_x > 0$ so that

$$t < A(|\rho|)(x;r_x) = \tfrac{|\rho|(B(x;r_x))}{m_n(B(x;r_x))}$$

and $B(x;r_x) \subseteq U$.
We consider the open set $V = \bigcup_{x \in M_t} B(x;r_x)$, and an arbitrary compact set $K \subseteq V$. Now, there exist finitely many $x_1, \ldots, x_m \in M_t$ so that $K \subseteq B(x_1;r_{x_1}) \cup \cdots \cup B(x_m;r_{x_m})$. Wiener's Lemma implies that there exist *pairwise disjoint* $B(x_{i_1};r_{x_{i_1}}), \ldots, B(x_{i_k};r_{x_{i_k}})$ so that

$$m_n(B(x_1;r_{x_1}) \cup \cdots \cup B(x_m;r_{x_m})) \leq 3^n\big(m_n(B(x_{i_1};r_{x_{i_1}})) + \cdots + m_n(B(x_{i_k};r_{x_{i_k}}))\big).$$

All these imply that

$$m_n(K) \leq \tfrac{3^n}{t}\big(|\rho|(B(x_{i_1};r_{x_{i_1}})) + \cdots + |\rho|(B(x_{i_k};r_{x_{i_k}}))\big) \leq \tfrac{3^n}{t}\,|\rho|(U) < \tfrac{3^n}{t}\,\epsilon.$$

By the regularity of m_n and since K is an arbitrary compact subset of V, we get $m_n(V) \leq \frac{3^n}{t}\,\epsilon$. Since $M_t \subseteq V$, we have that $m_n^*(M_t) \leq \frac{3^n}{t}\,\epsilon$. Since ϵ is arbitrary, we conclude that M_t is a Lebesgue set and $m_n(M_t) = 0$.
Finally, since

$$\big\{x \in M \,\big|\, \overline{\lim}_{r \to 0+} A(|\rho|)(x;r) \neq 0\big\} = \textstyle\bigcup_{k=1}^{+\infty} M_{1/k},$$

we get $\overline{\lim}_{r \to 0+} A(|\rho|)(x;r) = 0$ for m_n-a.e. $x \in \mathbb{R}^n$. Since $A(|\rho|)(x;r) \geq 0$ for all $x \in \mathbb{R}^n$ and all $r > 0$, we conclude that $\lim_{r \to 0+} A(|\rho|)(x;r) = 0$ for m_n-a.e. $x \in \mathbb{R}^n$. □

Lemma 6.10. *Let ν be a locally finite Borel signed measure on $\mathbb{R}^n$. Then ν is σ-finite and let $\nu = \lambda + \rho$ be the Lebesgue decomposition of ν with respect to m_n, where $\lambda \ll m_n$ and $\rho \perp m_n$. Then both λ and ρ are locally finite Borel signed measures.*
Moreover, if f is any Radon-Nikodym derivative of λ with respect to m_n, then f is locally Lebesgue integrable.

Proof. Since $\mathbb{R}^n = \bigcup_{k=1}^{+\infty} B(0;k)$ and $\nu(B(0;k))$ is finite for every k, we find that ν is σ-finite and the first of the Lebesgue-Radon-Nikodym Theorems implies the existence of the Lebesgue decomposition of ν.
Since $\rho \perp m_n$, there exist $R, N \in \mathcal{B}_n$ with $R \cup N = X$, $R \cap N = \emptyset$ so that R is null for m_n and N is null for ρ. From $\lambda \ll m_n$, we see that R is null for λ, as well.
Now let $M \in \mathcal{B}_n$ be bounded. Since $\nu(M)$ is finite, Theorem 6.1 implies that $\nu(M \cap N)$ is finite. Now we have

$$\lambda(M) = \lambda(M \cap R) + \lambda(M \cap N) = \lambda(M \cap N) = \lambda(M \cap N) + \rho(M \cap N) = \nu(M \cap N),$$

and so $\lambda(M)$ is finite. From $\nu(M) = \lambda(M) + \rho(M)$ we get that $\rho(M)$ is also finite. We conclude that λ and ρ are locally finite.
Again, let $M \in \mathcal{B}_n$ be bounded. Then $\int_M f(x)\, dm_n(x) = \lambda(M)$ is finite. This implies that f is locally Lebesgue integrable. $\square$

Theorem 6.13. *Let ν be a locally finite Borel signed measure or a Borel complex measure on $\mathbb{R}^n$. If f is any Radon-Nikodym derivative of the absolutely continuous part of ν with respect to m_n, then*

$$\lim_{r\to 0+} \frac{\nu(B(x;r))}{m_n(B(x;r))} = f(x)$$

for m_n-a.e. $x \in \mathbb{R}^n$.

Proof. Let $\nu = \lambda + \rho$ be the Lebesgue decomposition of ν with respect to m_n, where $\lambda \ll m_n$, $\rho \perp m_n$ and $\lambda = f m_n$. If ν is signed, Lemma 6.9 implies that ρ is a locally finite Borel signed measure and f is locally Lebesgue integrable. If ν is complex, then ρ is complex and f is Lebesgue integrable. Lebesgue's Theorem and Theorem 6.12 imply

$$\lim_{r\to 0+} \frac{\nu(B(x;r))}{m_n(B(x;r))} = \lim_{r\to 0+} \frac{1}{m_n(B(x;r))} \int_{B(x;r)} f(y)\, dm_n(y) + \lim_{r\to 0+} \frac{\rho(B(x;r))}{m_n(B(x;r))} = f(x)$$

for m_n-a.e. $x \in \mathbb{R}^n$. $\square$

Theorem 6.14. *Let ν be a locally finite Borel signed measure or a Borel complex measure on $\mathbb{R}^n$. If f is any Radon-Nikodym derivative of the absolutely continuous part of ν with respect to m_n, then, for m_n-a.e. $x \in \mathbb{R}^n$,*

$$\lim_{E\in\mathcal{C}, m_n(E)\to 0+} \frac{\nu(E)}{m_n(E)} = f(x)$$

for every thick family $\mathcal{C}$ of sets at x.

Proof. If ρ is the singular part of ν with respect to m_n, then $|\rho| \perp m_n$, and Theorem 6.12 implies

$$\lim_{r\to 0+} \frac{|\rho|(B(x;r))}{m_n(B(x;r))} = 0$$

for m_n-a.e. $x \in \mathbb{R}^n$.
Now, we consider any x for which $\lim_{r\to 0+} \frac{|\rho|(B(x;r))}{m_n(B(x;r))} = 0$, and any thick family $\mathcal{C}$ of sets at x. Then there is $c > 0$ so that for every $E \in \mathcal{C}$ there is a ball $B(x;r_E)$ with $E \subseteq B(x;r_E)$ and $m_n(E) \geq c\, m_n(B(x;r_E))$. For every $\epsilon > 0$ there is a $\delta > 0$ so that $r < \delta$ implies

$$\frac{|\rho|(B(x;r))}{m_n(B(x;r))} < c\epsilon.$$

If $m_n(E) < c\delta^n m_n(B(0;1))$, then $r_E < \delta$, and so

$$\Big|\frac{\rho(E)}{m_n(E)}\Big| \leq \frac{|\rho|(E)}{m_n(E)} \leq \frac{1}{c} \frac{|\rho|(B(x;r_E))}{m_n(B(x;r_E))} < \epsilon.$$

This means that, for m_n-a.e. $x \in \mathbb{R}^n$,

$$\lim_{E\in\mathcal{C}, m_n(E)\to 0+} \frac{\rho(E)}{m_n(E)} = 0$$

for every thick family $\mathcal{C}$ of sets at x.
We combine this with Theorem 6.11 to complete the proof. $\square$

Exercises.

6.6.1. *A variation of the Hardy-Littlewood maximal function.*

Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ or $\overline{\mathbb{C}}$ be locally Lebesgue integrable.

We define $M^*(f)(x) = \sup_{r>0} \frac{1}{m_n(B(x;r))} \int_{B(x;r)} |f(y)| \, dm_n(y)$ for every $x \in \mathbb{R}^n$.

(i) Prove that the set $\{x \in \mathbb{R}^n \,|\, t < M^*(f)(x)\}$ is open for every $t > 0$.

(ii) Prove that $\frac{1}{2^n} \, M(f)(x) \leq M^*(f)(x) \leq M(f)(x)$ for every $x \in \mathbb{R}^n$.

One may define other variants of the Hardy-Littlewood maximal function by taking the supremum of the mean values of $|f|$ over open cubes containing the point x or open cubes centered at the point x. The results are similar.

6.6.2. Vitali's Covering Theorem.

Let $E \subseteq \mathbb{R}^n$ and let $\mathcal{C}$ be a collection of open balls with the property that for every $x \in E$ and every $\epsilon > 0$ there is a $B \in \mathcal{C}$ so that $x \in B$ and $m_n(B) < \epsilon$. Prove that there are pairwise disjoint $B_1, B_2, \ldots \in \mathcal{C}$ so that $m_n^*(E \setminus \bigcup_{k=1}^{+\infty} B_k) = 0$.

6.6.3. *Points of density.*

Let $E \in \mathcal{L}_n$. If $x \in \mathbb{R}^n$, we set $D_E(x) = \lim_{r\to 0+} \frac{m_n(E\cap B(x;r))}{m_n(B(x;r))}$ whenever the limit exists. Observe that this limit (if it exists) is a number in the interval $[0, 1]$. If $D_E(x) = 1$, we say that x is a **density point** of E.

(i) If x is an interior point of E, prove that it is a density point of E.

(ii) Prove that m_n-a.e. $x \in E$ is a density point of E.

(iii) For any $\alpha \in (0, 1)$ find $x \in \mathbb{R}$ and $E \in \mathcal{L}_1$ so that $D_E(x) = \alpha$. Also find $x \in \mathbb{R}$ and $E \in \mathcal{L}_1$ so that $D_E(x)$ does not exist.

6.7 Functions of bounded variation.

Chapter 7

The classical Banach spaces.

7.1 Some facts from functional analysis.

NORMED SPACES.

Definition. *Let Z be a linear space over the field $F = \mathbb{R}$ or over the field $F = \mathbb{C}$ and let the function $\|\cdot\| : Z \to \mathbb{R}$ have the properties:*
(i) $\|u + v\| \leq \|u\| + \|v\|$, for all $u, v \in Z$,
(ii) $\|\kappa u\| = |\kappa|\|u\|$, for all $u \in Z$ and $\kappa \in F$,
(iii) $\|u\| = 0$ implies $u = 0$, where 0 is the zero element of Z.
Then, $\|\cdot\|$ is called a ***norm*** *on Z and $(Z, \|\cdot\|)$ is called a* ***normed space****.*

If $F = \mathbb{R}$, we say that $(Z, \|\cdot\|)$ is a real normed space and, if $F = \mathbb{C}$, we say that $(Z, \|\cdot\|)$ is a complex normed space.

If it is obvious from the context which $\|\cdot\|$ we are talking about, we shall say that Z is a normed space.

Proposition 7.1. *If $\|\cdot\|$ is a norm on the linear space Z, then*
(i) $\|0\| = 0$, where 0 is the zero element of Z,
(ii) $\|-u\| = \|u\|$, for all $u \in Z$,
(iii) $\|u\| \geq 0$, for all $u \in Z$.

Proof. Exercise. □

Proposition 7.2. *Let $(Z, \|\cdot\|)$ be a normed space. If we define $d : Z \times Z \to \mathbb{R}$ by $d(u, v) = \|u - v\|$ for all $u, v \in Z$, then d is a metric on Z.*

Proof. Exercise. □

Definition. *Let $(Z, \|\cdot\|)$ be a normed space. If d is the metric defined in Proposition 7.2, then d is called the* ***metric induced*** *on Z by $\|\cdot\|$.*

Therefore, if $(Z, \|\cdot\|)$ is a normed space, then (Z, d) is a metric space and we can study all notions related to the notion of a metric space, like convergence of sequences, open and closed sets and so on. Open balls in Z have the form $B(u; r) = \{v \in Z \mid \|v - u\| < r\}$. A sequence (u_n) in Z converges to $u \in Z$ if $\|u_n - u\| \to 0$ as $n \to +\infty$. We denote this by: $u_n \to u$ or $\lim_{n \to +\infty} u_n = u$. A set $U \subseteq Z$ is open if for every $u \in U$ there is an $r > 0$ so that $B(u; r) \subseteq U$. Any union of open sets is open and any finite intersection of open sets is open. The sets $\emptyset$ and Z are open. A set $K \subseteq Z$ is closed if $Z \setminus K$ is open or, equivalently, if the limit of every sequence in K (which has a limit) belongs to K. Any intersection of closed sets is closed and any finite union of closed sets is closed. The sets $\emptyset$ and Z are closed. A set $K \subseteq Z$ is compact if every open cover of K has a finite subcover of K. Equivalently, K is compact if every sequence in K has a convergent subsequence with limit in K. A sequence (u_n) in Z is a Cauchy sequence if $\|u_n - u_m\| \to 0$ as $n, m \to +\infty$. Every convergent sequence is Cauchy. If every Cauchy sequence is convergent, then Z is a complete metric space.

Definition. *If the normed space $(Z, \|\cdot\|)$ is complete as a metric space (with the metric induced by the norm), then it is called a **Banach space**.*

If there is no danger of confusion, we say that Z is a Banach space.

Example. The space $\mathbb{R}^n$ with the Euclidean norm defined by $\|x\| = |x| = (x_1^2 + \cdots + x_n^2)^{1/2}$ is a familiar real Banach space.
The space $\mathbb{C}^n$ with the norm defined by $\|x\| = |x| = (|x_1|^2 + \cdots + |x_n|^2)^{1/2}$ is a complex Banach space.

There are some special results based on the combination of the linear and the metric structure of a normed space. We first define, as in any linear space, $u + A = \{u + v \,|\, v \in A\}$ and $\kappa A = \{\kappa v \,|\, v \in A\}$ for all $A \subseteq Z, u \in Z$ and $\kappa \in F$. We also define for every $u \in Z$ and every $\kappa > 0$ the **translation** $\tau_u : Z \to Z$ and the **dilation** $l_\kappa : Z \to Z$, by $\tau_u(v) = v + u$ and $l_\kappa(v) = \kappa v$ for all $v \in Z$. It is trivial to prove that translations and dilations are one-to-one transformations of Z onto Z and that $\tau_u^{-1} = \tau_{-u}$ and $l_\kappa^{-1} = l_{1/\kappa}$. It is obvious that $u + A = \tau_u(A)$ and $\kappa A = l_\kappa(A)$.

Proposition 7.3. *Let $(Z, \|\cdot\|)$ be a normed space.*
(i) $u + B(v; r) = B(u + v; r)$ for all $u, v \in Z$ and $r > 0$.
(ii) $\kappa B(v; r) = B(\kappa v; |\kappa| r)$ for all $v \in Z$, $\kappa \in F \setminus \{0\}$ and $r > 0$.
(iii) If $u_n \to u$ and $v_n \to v$ in Z, then $u_n + v_n \to u + v$ in Z.
(iv) If $\kappa_n \to \kappa$ in F and $u_n \to u$ in Z, then $\kappa_n u_n \to \kappa u$ in Z.
(v) If $u_n \to u$ in Z, then $\|u_n\| \to \|u\|$.
(vi) Translations and dilations are homeomorphisms. This means that they, together with their inverses, are continuous on Z.
(vii) If A is open or closed or compact in Z and $u \in Z$, then $u + A$ is, respectively, open or closed or compact in Z.
(viii) If A is open or closed or compact in Z and $\kappa \in F \setminus \{0\}$, then κA is, respectively, open or closed or compact in Z.

Proof. Exercise. □

INNER PRODUCT SPACES.

Definition. *Let Z be a linear space over the field $F = \mathbb{R}$ or over the field $F = \mathbb{C}$ and let the function $\langle\cdot,\cdot\rangle : Z \times Z \to F$ have the properties:*
(i) $\langle u_1 + u_2, v\rangle = \langle u_1, v\rangle + \langle u_2, v\rangle$, for all $u_1, u_2, v \in Z$,
(ii) $\langle \kappa u, v\rangle = \kappa\langle u, v\rangle$, for all $u, v \in Z$ and $\kappa \in F$,
(iii) $\langle u, u\rangle \geq 0$ for all $u \in Z$ and, also, $\langle u, u\rangle = 0$ implies $u = 0$.
(iv) $\langle v, u\rangle = \overline{\langle u, v\rangle}$ for all $u, v \in Z$.
*Then, $\langle\cdot,\cdot\rangle$ is called an **inner product** on Z and $(Z, \langle\cdot,\cdot\rangle)$ is called an **inner product space**.*

If $F = \mathbb{R}$, we say that Z is a real inner product space and, if $F = \mathbb{C}$, we say that Z is a complex inner product space. Of coure, if $F = \mathbb{R}$, then property (iv) becomes $\langle v, u\rangle = \langle u, v\rangle$ for all $u, v \in Z$.

Proposition 7.4. *Let $(Z, \langle\cdot,\cdot\rangle)$ be an inner product space.*
(i) $\langle u, v_1 + v_2\rangle = \langle u, v_1\rangle + \langle u, v_2\rangle$, for all $u, v_1, v_2 \in Z$,
(ii) $\langle u, \kappa v\rangle = \overline{\kappa}\langle u, v\rangle$, for all $u, v \in Z$ and $\kappa \in F$,
(iii) $\langle 0, v\rangle = \langle u, 0\rangle = 0$ for all $u, v \in Z$.
(iv) $\langle u + v, u + v\rangle = \langle u, u\rangle + 2\,\mathrm{Re}(\langle u, v\rangle) + \langle v, v\rangle$ for all $u, v \in Z$.
(v) $\langle u + v, u + v\rangle + \langle u - v, u - v\rangle = 2\langle u, u\rangle + 2\langle v, v\rangle$ for all $u, v \in Z$.
(vi) $|\langle u, v\rangle|^2 \leq \langle u, u\rangle\langle v, v\rangle$ for all $u, v \in Z$.

Proof. The proofs of (i), (ii), (iii), (iv) and (v) are left as an exercise.
(vi) If $u = 0$ then the inequality takes the form of an equality: $0 = 0$. Now let $u \neq 0$ and, thus, $\langle u, u\rangle > 0$. Then for every $\kappa \in F$ we get from (iv) that $0 \leq \langle \kappa u + v, \kappa u + v\rangle = |\kappa|^2\langle u, u\rangle + 2\,\mathrm{Re}(\kappa\langle u, v\rangle) + \langle v, v\rangle$. We finish the proof, using $\kappa = -\frac{\overline{\langle u,v\rangle}}{\langle u,u\rangle}$. □

Proposition 7.5. *If $(Z, \langle\cdot,\cdot\rangle)$ is an inner product space, we define $\|u\| = (\langle u, u\rangle)^{1/2}$ for all $u \in Z$. Then $\|\cdot\|$ is a norm on Z.*

Proof. All properties of a norm are trivial to prove. We shall only prove the last property:

$$\begin{aligned}\|u+v\|^2 &= \langle u+v, u+v\rangle = \langle u, u\rangle + 2\,\mathrm{Re}(\langle u, v\rangle) + \langle v, v\rangle \leq \|u\|^2 + 2|\langle u, v\rangle| + \|v\|^2 \\ &\leq \|u\|^2 + 2(\langle u, u\rangle)^{1/2}(\langle v, v\rangle)^{1/2} + \|v\|^2 = \|u\|^2 + 2\|u\|\|v\| + \|v\|^2 = (\|u\| + \|v\|)^2\end{aligned}$$

which implies that $\|u + v\| \leq \|u\| + \|v\|$ for all $u, v \in Z$. □

We see that every inner product space Z becomes a normed space with a norm which is defined using the inner product of Z and whatever properties we prove for normed spaces they hold also for inner product spaces.

Equalities (iv) and (v) of Proposition 7.4 take the forms $\|u+v\|^2 = \|u\|^2+2\,\mathrm{Re}(\langle u, v\rangle)+\|v\|^2$ and $\|u + v\|^2 + \|u - v\|^2 = 2\|u\|^2 + 2\|v\|^2$. Inequality (vi) is called the **Cauchy-Schwartz inequality** and takes the form

$$|\langle u, v\rangle| \leq \|u\|\,\|v\|.$$

Definition. *If the inner product space $(Z, \langle\cdot,\cdot\rangle)$ is complete (as a normed space) then it is called a* ***Hilbert space****.*

Example. $\mathbb{R}^n$ with the Euclidean norm, defined by $\|x\| = |x| = (x_1^2+\cdots+x_n^2)^{1/2}$, is a real Hilbert space. Indeed, the well known inner product defined by $\langle x, y\rangle = x\cdot y = x_1y_1+\cdots+x_ny_n$ induces the norm of the space.
Similarly, $\mathbb{C}^n$ with the norm defined by $\|x\| = |x| = (|x_1|^2 + \cdots + |x_n|^2)^{1/2}$ is a complex Hilbert space. Now, the appropriate inner product is defined by $\langle x, y\rangle = x \cdot y = x_1\,\overline{y_1} + \cdots + x_n\,\overline{y_n}$.

Definition. *Let $(Z, \langle\cdot,\cdot\rangle)$ be an inner product space. If $\langle u, v\rangle = 0$, we say that u, v are* ***orthogonal****. If $\langle u, v\rangle = 0$ for every $v \in B \subseteq Z$, then we say that u, B are orthogonal. If $\langle u, v\rangle = 0$ for every $u \in A \subseteq Z$ and every $v \in B \subseteq Z$, then we say that A, B are orthogonal. In each case we write, respectively, $u \perp v$, $u \perp B$ and $A \perp B$.*

Proposition 7.6. *Let $(Z, \langle\cdot,\cdot\rangle)$ be an inner product space.*
(i) If $u_n \to u$ and $v_n \to v$, then $\langle u_n, v_n\rangle \to \langle u, v\rangle$.
(ii) If $u \perp v$ and $u \perp w$, then $u \perp (v + w)$. Also, if $u \perp v$, then $u \perp (\kappa v)$ for all $\kappa \in F$.
(iii) If $v_n \to v$ and $u \perp v_n$ for all n, then $u \perp v$.
(iv) If $u \perp A$, then $u \perp V$, where V is the closed linear subspace of Z generated by A. Also, if $A \perp B$, then $U \perp V$, where U, V are the closed linear subspaces of Z generated by A, B, respectively.
(v) If $u_1, \ldots, u_n$ are pairwise orthogonal, then $\|u_1 + \cdots + u_n\|^2 = \|u_1\|^2 + \cdots + \|u_n\|^2$.

Proof. Exercise. □

From now on, if there is no danger of confusion, we shall say normed space Z or inner product space Z instead of normed space $(Z, \|\cdot\|)$ or inner product space $(Z, \langle\cdot,\cdot\rangle)$.

F.Riesz Theorem. *Let Z be a Hilbert space, $u \in Z$ and V be a closed linear subspace of Z. Then there is a unique $v_0 \in V$ such that $\|u-v_0\| \leq \|u-v\|$ for all $v \in V$. This v_0 is the unique element of V satisfying $(u - v_0) \perp V$.*

Proof. Let $d = \inf\{\|u - v\| \mid v \in V\}$.
Then $d \geq 0$ and there is a sequence $(v_n) \in V$ so that $\|u - v_n\| \to d$.
Since V is a linear subspace, we have that $\frac{1}{2}(v_n + v_m) \in V$ and hence $\|u - \frac{1}{2}(v_n + v_m)\| \geq d$ for all n, m. Now we apply (v) of Proposition 7.4 to $\frac{1}{2}(u - v_n)$ and $\frac{1}{2}(u - v_m)$ and get

$$\|v_n - v_m\|^2 = 2\|u - v_n\|^2 + 2\|u - v_m\|^2 - 4\|u - \tfrac{1}{2}(v_n + v_m)\|^2 \leq 2\|u - v_n\|^2 + 2\|u - v_m\|^2 - 4d^2.$$

Taking the limit, we find $\|v_n - v_m\| \to 0$. Thus, (v_n) is a Cauchy sequence and since Z is complete, we get $v_n \to v_0$ for some $v_0 \in Z$. Also, since V is closed, $v_0 \in V$. Now, $v_n \to v_0$ implies that $\|u - v_n\| \to \|u - v_0\|$ and, hence, $\|u - v_0\| = d$. I.e. $\|u - v_0\| \leq \|u - v\|$ for all $v \in V$.
Now we take any $\kappa \in F$ and any $v \in V$, $v \neq 0$. Since $v_0 + \kappa v \in V$, we get

$$\|u - v_0\|^2 \leq \|u - (v_0 + \kappa v)\|^2 = \|u - v_0\|^2 + 2\,\mathrm{Re}(\overline{\kappa}\langle u - v_0, v\rangle) + |\kappa|^2\|v\|^2.$$

Using $\kappa = -\frac{\langle u - v_0, v\rangle}{\|v\|^2}$ we find $|\langle u - v_0, v\rangle|^2 \leq 0$. Therefore, $\langle u - v_0, v\rangle = 0$ for all $v \in V$ and we conclude that $u - v_0 \perp V$.
If $u - v_0 \perp V$ and $u - v_1 \perp V$ for some other $v_1 \in V$ we get $(v_1 - v_0) \perp V$ and, since $v_1 - v_0 \in V$, we find $(v_1 - v_0) \perp (v_1 - v_0)$. This implies $v_1 - v_0 = 0$ and so v_0 is unique. □

Definition. *Let Z be a Hilbert space and V be a closed linear subspace of Z. For every $u \in Z$ the unique $v_0 \in V$ which is such that $(u - v_0) \perp V$ is called the* ***projection*** *of u on V and it is denoted $P_V(u)$.*

Definition. *Let Z be an inner product space and $A \subseteq Z$ so that $0 \notin A$. The set A is called* ***orthogonal*** *if $u \perp v$ for all $u, v \in A$, $u \neq v$. The set A is called* ***orthonormal*** *if it is orthogonal and $\|u\| = 1$ for all $u \in A$.*

Every orthogonal set A can become orthonormal when we multiply every element of A by an appropriate number so that its norm becomes 1. More precisely, the set $B = \{\frac{1}{\|u\|}\,u \mid u \in A\}$ is orthonormal.

Bessel's Inequality. *Let Z be an inner product space and $A \subseteq Z$ be an orthonormal set. Then for every $u \in Z$ we have $\sum_{e\in A} |\langle u, e\rangle|^2 \leq \|u\|^2$.*

Proof. Take any finite subset $\{e_1, \dots, e_n\} \subseteq A$ and consider $v_0 = \langle u, e_1\rangle e_1 + \cdots + \langle u, e_n\rangle e_n$. Part (v) of Proposition 7.6 implies $\|v_0\|^2 = |\langle u, e_1\rangle|^2 + \cdots + |\langle u, e_n\rangle|^2$ and, hence,

$$\langle u - v_0, v_0\rangle = \langle u, v_0\rangle - \|v_0\|^2 = \overline{\langle u, e_1\rangle}\,\langle u, e_1\rangle + \cdots + \overline{\langle u, e_n\rangle}\,\langle u, e_n\rangle - \|v_0\|^2 = 0.$$

Therefore,

$$|\langle u, e_1\rangle|^2 + \cdots + |\langle u, e_n\rangle|^2 = \|v_0\|^2 \leq \|v_0\|^2 + \|u - v_0\|^2 = \|v_0 + (u - v_0)\|^2 = \|u\|^2.$$

Since this is true for every finite subset of A, we conclude that $\sum_{e\in A} |\langle u, e\rangle|^2 \leq \|u\|^2$. □

Proposition 7.7. *Let Z be a Hilbert space and $A \subseteq Z$ be an orthonormal set. If $(\kappa_e)_{e\in A}$ is a family in F indexed over A with $\sum_{e\in A} |\kappa_e|^2 < +\infty$, then the sum $\sum_{e\in A} \kappa_e\, e$ can be defined as an element of Z, and this element satisfies $\langle \sum_{e\in A} \kappa_e\, e, v\rangle = \sum_{e\in A} \kappa_e\langle e, v\rangle$ for all $v \in Z$.*
In particular, $\langle \sum_{e\in A} \kappa_e\, e, e'\rangle = \kappa_{e'}$ for all $e' \in A$. Moreover, the sum $\sum_{e\in A} \kappa_e\, e$ belongs to the closed linear subspace generated by A.

Proof. From $\sum_{e\in A} |\kappa_e|^2 < +\infty$ we get that the set of $e \in A$ for which $\kappa_e \neq 0$ is countable. Thus, let $e_1, e_2, \dots$ be any enumeration of the elements of A with $\kappa_e \neq 0$ and then we have $\kappa_e = 0$ when $e \in A \setminus \{e_1, e_2, \dots\}$.
Now, if $e_1, e_2, \dots$ are finite, say $e_1, \dots, e_n$, then we obviously define $\sum_{e\in A} \kappa_e\, e$ to be the element $u = \kappa_{e_1} e_1 + \cdots + \kappa_{e_n} e_n$ and then u belongs to the closed linear subspace generated by A.

Moreover, $\langle u, v\rangle = \kappa_{e_1}\langle e_1, v\rangle + \cdots + \kappa_{e_n}\langle e_n, v\rangle = \sum_{e\in A}\kappa_e\langle e, v\rangle$.
If $e_1, e_2, \dots$ are infinitely many, we consider the partial sums $s_n = \kappa_{e_1} e_1 + \cdots + \kappa_{e_n} e_n$ for all n. Then for $m < n$ we get $\|s_n - s_m\|^2 = \|\kappa_{e_{m+1}} e_{m+1} + \cdots + \kappa_{e_n} e_n\|^2 = |\kappa_{e_{m+1}}|^2 + \cdots + |\kappa_{e_n}|^2$. Now, $\sum_{k=1}^{+\infty}|\kappa_{e_k}|^2 = \sum_{e\in A}|\kappa_e|^2 < +\infty$ and this implies that $\|s_n - s_m\| \to 0$ as $n, m \to +\infty$. Since Z is complete, (s_n) converges to some $u \in Z$ and we define $\sum_{e\in A}\kappa_e\, e = u = \lim_{n\to+\infty} s_n$. Since every s_n belongs to the closed linear subspace generated by A, the same is true for u.
Furthermore, from $s_n \to u$ and from $\langle s_n, v\rangle = \kappa_{e_1}\langle e_1, v\rangle + \cdots + \kappa_{e_n}\langle e_n, v\rangle$ we find $\langle u, v\rangle = \sum_{k=1}^{+\infty}\kappa_{e_k}\langle e_k, v\rangle = \sum_{e\in A}\kappa_e\langle e, v\rangle$. □

Proposition 7.8. *Let Z be a Hilbert space, $A \subseteq Z$ be an orthonormal set and V be the closed linear subspace generated by A. Then for every $u \in Z$ the sum $\sum_{e\in A}\langle u, e\rangle\, e$ can be defined as an element v_0 of Z. This v_0 is equal to the projection $P_V(u)$ of u on V.*

Proof. By Bessel's inequality, we have $\sum_{e\in A}|\langle u, e\rangle|^2 \le \|u\|^2 < +\infty$. Now, Proposition 7.7 implies that the sum $\sum_{e\in A}\langle u, e\rangle\, e$ can be defined as an element v_0 of Z which belongs to V and satisfies $\langle v_0, e\rangle = \langle u, e\rangle$ for all $e \in A$. Thus, $(u - v_0) \perp e$ for all $e \in A$. Therefore, $(u - v_0) \perp V$ and we conclude that $v_0 = P_V(u)$. □

Combining the last result with the F. Riesz Theorem we conclude that

$$\|u - \textstyle\sum_{e\in A}\langle u, e\rangle\, e\| \le \|u - \sum_{e\in A}\kappa_e\, e\|$$

for all $(\kappa_e)_{e\in A}$ with $\sum_{e\in A}|\kappa_e|^2 < +\infty$.

Definition. *Let Z be a Hilbert space and $A \subseteq Z$ be an orthonormal set. We say that A is an* ***orthonormal basis*** *of Z if the closed linear subspace generated by A is Z.*

Proposition 7.9. *Let Z be a Hilbert space and $A \subseteq Z$ be an orthonormal set. Then the following are equivalent:*
(i) A is an orthonormal basis of Z.
(ii) $u = \sum_{e\in A}\langle u, e\rangle\, e$ for all $u \in Z$.
(iii) $\|u\|^2 = \sum_{e\in A}|\langle u, e\rangle|^2$ for all $u \in Z$.

Proof. Let A be an orthonormal basis of Z. We consider the element $v_0 = \sum_{e\in A}\langle u, e\rangle\, e$ and we shall prove that $u = v_0$. Indeed, Proposition 7.8 says that v_0 is the projection of u on the closed linear subspace generated by A, which is Z. But the projection of u on Z is u itself.
If we assume that $u = \sum_{e\in A}\langle u, e\rangle\, e$ for all $u \in Z$, then this implies that $\|u\|^2 = \sum_{e\in A}|\langle u, e\rangle|^2$ for all $u \in Z$.
Finally, let $\|u\|^2 = \sum_{e\in A}|\langle u, e\rangle|^2$ for all $u \in Z$. We assume that A is not an orthonormal basis of Z, i.e. that the closed linear subspace V which is generated by A is a proper subspace of Z. Thus, there is a $u \in Z \setminus V$. By Proposition 7.8, the projection of u on V is $v_0 = \sum_{e\in A}\langle u, e\rangle\, e$ and then $(u - v_0) \perp V$. Thus, $(u - v_0) \perp v_0$ and $\|u\|^2 = \|u - v_0\|^2 + \|v_0\|^2 > \|v_0\|^2 = \sum_{e\in A}|\langle u, e\rangle|^2$. We arrive at a contradiction and, hence, A is an orthonormal basis of Z. □

Proposition 7.10. *Every Hilbert space has an orthonormal basis.*

Proof. We consider the family $\mathcal{F} = \{A \,|\, A \text{ is an orthonormal subset of } Z\}$.
This family is non-empty. Indeed, we may consider any $u \in Z$ with $\|u\| = 1$ and then $A = \{u\}$ is an orthonormal subset of Z.
We define a partial order $\prec$ on $\mathcal{F}$ by: $A_1 \prec A_2$ if $A_1 \subseteq A_2$.
Now assume that $\mathcal{G}$ is a totally ordered subfamily of $\mathcal{F}$. We consider the set $A_0 = \bigcup_{A\in\mathcal{G}} A$ and then it is easy to prove that A_0 is an orthonormal subset of Z. Since $A \prec A_0$ for all $A \in \mathcal{G}$, we get that A_0 is an upper bound of $\mathcal{G}$.
Zorn's Lemma implies that there is a maximal element of $\mathcal{F}$. I.e. there is an orthonormal subset A of Z such that there is no orthonormal subset of Z which is strictly larger than A. We shall prove

that A is an orthonormal basis of Z.
Assume that A is not an orthonormal basis of Z. Then the closed linear subspace V generated by A is strictly smaller that Z. Thus, there is some $u \in Z \setminus V$. We consider the projection $v_0 = P_V(u)$ of u on V and we have that $(u - v_0) \perp V$. Then $u' = u - v_0 \neq 0$ and the element $e = \frac{1}{\|u'\|}\, u'$ satisfies $\|e\| = 1$ and $e \perp V$ and, thus, $e \perp A$. Therefore, $A \cup \{e\}$ is an orthonormal set strictly larger than A and we get a contradiction. $\square$

Definition. *Let Z be a Hilbert space and A be an orthonormal basis of Z. For every $u \in Z$ the series $\sum_{e \in A} \langle u, e\rangle\, e$ is called the **Fourier series** of u with respect to A. Thus, every element of Z is equal to its Fourier series with respect to any orthonormal basis of Z. The numbers $\langle u, e\rangle$, for all $e \in A$, are called the **Fourier coefficients** of u with respect to A.*

If we consider any closed linear subspace V of a Hilbert space Z, then V is also complete and, hence, a Hilbert space. Therefore, *every closed linear subspace of a Hilbert space has an orthonormal basis*.

Proposition 7.11. *Let Z be a Hilbert space. If Z has a countable orthonormal basis then Z is separable. If Z is separable then every orthonormal basis of Z is countable.*

Proof. Let Z be separable and let A be any orthonormal basis of Z. Let also B be a countable dense subset of Z. Then the open balls $B(e; \frac{\sqrt{2}}{2})$ for all $e \in A$ are disjoint and each of them contains at least one element b_e of B. The elements b_e are disjoint and, thus, the mapping $e \mapsto b_e$ from A into B is one-to-one. Therefore, A is countable.
Now let the orthonormal basis A of Z be countable. Then the set of all linear combinations of elements of A with rational coefficients is countable and dense in Z. Therefore, Z is separable. $\square$

BOUNDED LINEAR OPERATORS.

Definition. *Let Z and W be two linear spaces over the same F and a function $T : Z \to W$. Then T is called a **linear transformation** or a **linear operator** from Z to W if $T(u+v) = T(u) + T(v)$ and $T(\kappa u) = \kappa T(u)$ for all $u, v \in Z$ and all $\kappa \in F$.*

The following are familiar from elementary Linear Algebra. Let $T : Z \to W$ be a linear operator. Then T is one-to-one if and only if $T(u) = 0$ (the zero element of W) implies $u = 0$ (the zero element of Z). The subset $N(T) = \{u \in Z \,|\, T(u) = 0\}$ of Z, called the **kernel** of T, is a linear subspace of Z. Similarly, the subset $R(T) = \{T(u) \,|\, u \in Z\}$ of W, called the **range** of T, is a linear subspace of W. Now, T is one-to-one if and only if $N(T) = \{0\}$ and T is onto if and only if $R(T) = W$. If $T : Z \to W$ is one-to-one and onto, then the inverse function $T^{-1} : W \to Z$ is also a linear operator. In this case we say that *the linear spaces Z and W are identified*. By this we mean that we may view the two spaces *as a single space whose elements have two "names"*. I.e. we view the elements u of Z and $T(u)$ of W as a single element with the two names: u and $T(u)$. In fact the linear relations between elements are unaffected by changing their "names": $z = u + v$ if and only if $T(z) = T(u) + T(v)$ and $z = \kappa u$ if and only if $T(z) = \kappa T(u)$. If $T : Z \to W$ is one-to-one but not onto, then we may consider the restriction $T : Z \to R(T)$. This is a linear operator which is one-to-one and onto and we may say that the linear spaces Z and $R(T)$ are identified and that *Z is identified with a linear subspace of W* or that *$R(T)$ is a "copy" of Z inside W*.

Definition. *Let $(Z, \|\cdot\|_Z)$ and $(W, \|\cdot\|_W)$ be two normed spaces and $T : Z \to W$ be a linear operator. We say that T is **bounded** if there is a constant $M < +\infty$ so that $\|T(u)\|_W \leq M\|u\|_Z$ for all $u \in Z$.*

From now on when we have two normed spaces $(Z, \|\cdot\|_Z)$ and $(W, \|\cdot\|_W)$ we shall denote, for simplicity, both norms with the same symbol $\|\cdot\|$. For instance, the relation $\|T(u)\|_W \leq M\|u\|_Z$ will be simplified to $\|T(u)\| \leq M\|u\|$.

Theorem 7.1. *Let Z and W be two normed spaces and $T : Z \to W$ be a linear operator. The following are equivalent.*
(i) T is bounded.
(ii) T is continuous on Z.
(iii) T is continuous at 0.

Proof. Suppose that T is bounded and, hence, there is $M < +\infty$ so that $\|T(u)\| \leq M\|u\|$ for every $u \in Z$. If $u_n \to u$ in Z, then $\|T(u_n) - T(u)\| = \|T(u_n - u)\| \leq M\|u_n - u\| \to 0$ and, thus, $T(u_n) \to T(u)$ in W. Therefore, T is continuous on Z.
If T is continuous on Z, then it is certainly continuous at 0.
Suppose that T is continuous at 0. Then there is $\delta > 0$ so that $\|T(u)\| = \|T(u) - T(0)\| < 1$ for every u with $\|u\| = \|u - 0\| < \delta$. We take any $u \in Z \setminus \{0\}$ and any $t > 1$ and we get $\big\|\frac{\delta}{t\|u\|}\, u\big\| = \frac{\delta}{t} < \delta$. Therefore, $\big\|T\big(\frac{\delta}{t\|u\|}\, u\big)\big\| < 1$ and, hence, $\|T(u)\| \leq \frac{t}{\delta}\, \|u\|$. This is trivially true also for $u = 0$ and we conclude that $\|T(u)\| \leq \frac{t}{\delta}\, \|u\|$ for every $u \in Z$. Letting $t \to 1+$, we get $\|T(u)\| \leq M\|u\|$, where $M = \frac{1}{\delta}$. Therefore, T is bounded. $\square$

Proposition 7.12. *Let Z and W be two normed spaces and $T : Z \to W$ be a bounded linear operator. Then there is a smallest M_0 with the property: $\|T(u)\| \leq M_0\|u\|$ for every $u \in Z$. This M_0 is characterized by the two properties:*
(i) $\|T(u)\| \leq M_0\|u\|$ for every $u \in Z$,
(ii) for every $M < M_0$ there is a $u \in Z$ so that $\|T(u)\| > M\|u\|$.

Proof. We consider $M_0 = \inf\{M \mid \|T(u)\| \leq M\|u\| \text{ for every } u \in Z\}$.
The set $L = \{M \mid \|T(u)\| \leq M\|u\| \text{ for every } u \in Z\}$ is non-empty by assumption and included in $[0, +\infty)$. Therefore M_0 exists and $M_0 \geq 0$. We take a sequence (M_n) in L so that $M_n \to M_0$ and, from $\|T(u)\| \leq M_n\|u\|$ for every $u \in Z$, we get $\|T(u)\| \leq M_0\|u\|$ for every $u \in Z$. Therefore, M_0 is the smallest element of L.
If $M < M_0$, then $M \notin L$ and, hence, there is a $u \in Z$ so that $\|T(u)\| > M\|u\|$. $\square$

Definition. *Let Z and W be two normed spaces and $T : Z \to W$ be a bounded linear operator. The smallest M for which $\|T(u)\| \leq M\|u\|$ for every $u \in Z$ is called the* ***norm*** *of T and it is denoted $\|T\|$.*

The zero linear operator $0 : Z \to W$ is bounded and, since $\|0(u)\| = 0 \leq 0\|u\|$ for every $u \in Z$, we have that $\|0\| = 0$. On the other hand, if T is a bounded linear operator with $\|T\| = 0$, then $\|T(u)\| \leq 0\|u\| = 0$ for every $u \in Z$ and, hence, T is the zero linear operator.

Proposition 7.13. *Let Z and W be two normed spaces and $T : Z \to W$ be a bounded linear operator. Then $\|T\| = \sup_{u \in Z, u \neq 0} \frac{\|T(u)\|}{\|u\|} = \sup_{u \in Z, \|u\|=1} \|T(u)\| = \sup_{u \in Z, \|u\| \leq 1} \|T(u)\|$.*

Proof. It is clear that $\sup_{u \in Z, \|u\|=1} \|T(u)\| \leq \sup_{u \in Z, \|u\| \leq 1} \|T(u)\|$.
Writing $v = \frac{u}{\|u\|}$ for every $u \in Z \setminus \{0\}$, we have that $\|v\| = 1$. Therefore, $\sup_{u \in Z, u \neq 0} \frac{\|T(u)\|}{\|u\|} = \sup_{u \in Z, u \neq 0} \big\|T\big(\frac{u}{\|u\|}\big)\big\| \leq \sup_{u \in Z, \|u\|=1} \|T(u)\|$.
For all u with $\|u\| \leq 1$, we get $\|T(u)\| \leq \|T\|\|u\| \leq \|T\|$ and, thus, $\sup_{u \in Z, \|u\| \leq 1} \|T(u)\| \leq \|T\|$.
If we set $M = \sup_{u \in Z, u \neq 0} \frac{\|T(u)\|}{\|u\|}$, then $\frac{\|T(u)\|}{\|u\|} \leq M$ and, hence, $\|T(u)\| \leq M\|u\|$ for all $u \neq 0$. Since this is obviously true for $u = 0$, we have that $\|T\| \leq M$ and this finishes the proof. $\square$

Definition. *Let Z and W be two normed spaces and $T : Z \to W$ be a bounded linear operator. If T is onto W and $\|T(u)\| = \|u\|$ for every $u \in Z$, then we say that T is an* ***isometry*** *from Z* ***onto*** *W or between Z and W.*
If $\|T(u)\| = \|u\|$ for every $u \in Z$ (but T is not necessarily onto W), we say that T is an ***isometry*** *from Z* ***into*** *W.*

Proposition 7.14. *Let Z and W be two normed spaces.*
(i) If T is an isometry from Z into W, then T is one-to-one.
(ii) If T is an isometry from Z onto W, then T^{-1} is an isometry from W onto Z.

Proof. Exercise. □

If T is an isometry from Z onto W, then it is not only that we may identify Z and W as linear spaces (see the discussion before the definition of a bounded linear operator) but we may also identify them as *metric* spaces: the distances between elements are unaffected by changing their "names": $\|T(u) - T(v)\| = \|T(u - v)\| = \|u - v\|$.

If T is an isometry from Z into W, then T is an isometry from Z onto $R(T)$ and we may identify Z with the subspace $R(T)$ of W or we may view $R(T)$ as a "copy" of Z inside W.

BOUNDED LINEAR FUNCTIONALS.

As in any linear space, we define a **linear functional** on Z to be a function $l : Z \to F$ which satisfies $l(u + v) = l(u) + l(v)$ and $l(\kappa u) = \kappa l(u)$ for every $u, v \in Z$ and $\kappa \in F$.

Since F itself is a linear space, a linear functional is a special case of a linear operator.

Definition. *Let Z be a normed space and l be a linear functional on Z. Then we say that l is* ***bounded*** *if there is an $M < +\infty$ so that $|l(u)| \leq M\|u\|$ for all $u \in Z$.*

Theorem 7.2 and Propositions 7.15 and 7.16 are special cases of Theorem 7.1 and Propositions 7.12 and 7.13. Hence, they do not need new proofs.

Theorem 7.2. *Let Z be a normed space and l be a linear functional on Z. The following are equivalent.*
(i) l is bounded.
(ii) l is continuous on Z.
(iii) l is continuous at 0.

Proposition 7.15. *Let Z be a normed space and l be a bounded linear functional on Z. Then there is a smallest M_0 with the property: $|l(u)| \leq M\|u\|$ for every $u \in Z$. This M_0 is characterized by the two properties:*
(i) $|l(u)| \leq M_0\|u\|$ for every $u \in Z$,
(ii) for every $M < M_0$ there is a $u \in Z$ so that $|l(u)| > M\|u\|$.

Definition. *Let Z be a normed space and l be a bounded linear functional on Z. The smallest M for which $|l(u)| \leq M\|u\|$ for every $u \in Z$ is called the* ***norm*** *of l and it is denoted $\|l\|_*$.*

The zero linear functional $0 : Z \to F$ is bounded and, since $|0(u)| = 0 \leq 0\|u\|$ for every $u \in Z$, we get $\|0\|_* = 0$. Conversely, if $l \in Z^*$ has $\|l\|_* = 0$, then $|l(u)| \leq 0\|u\| = 0$ for every $u \in Z$ and, hence, l is the zero linear functional on Z.

Proposition 7.16. *Let Z be a normed space and $l \in Z^*$. Then $\|l\|_* = \sup_{u\in Z, u\neq 0} \frac{|l(u)|}{\|u\|} = \sup_{u\in Z,\|u\|=1} |l(u)| = \sup_{u\in Z,\|u\|\leq 1} |l(u)|$.*

We define the **sum** $l_1 + l_2 : Z \to F$ of two linear functionals l_1, l_2 on Z by $(l_1 + l_2)(u) = l_1(u) + l_2(u)$ for all $u \in Z$ and the **product** $\kappa l : Z \to F$ of a linear functional l on Z and a $\kappa \in F$ by $(\kappa l)(u) = \kappa l(u)$ for all $u \in Z$. It is trivial to prove that $l_1 + l_2$ and κl are linear functionals on Z and that the set $Z' = \{l \,|\, l \text{ is a linear functional on } Z\}$ becomes a linear space under this sum and product. Z' is called the **algebraic dual** of Z. The **zero element** of Z' is the linear functional $0 : Z \to F$ defined by $0(u) = 0$ for all $u \in Z$ and the **opposite** of a linear functional l on Z is the linear functional $-l : Z \to F$ defined by $(-l)(u) = -l(u)$ for all $u \in Z$.

Proposition 7.17. *Let Z be a normed space, l, l_1, l_2 be bounded linear functionals on Z and $\kappa \in F$. Then $l_1 + l_2$ and κl are bounded linear functionals on Z and $\|l_1 + l_2\|_* \le \|l_1\|_* + \|l_2\|_*$ and $\|\kappa l\|_* = |\kappa| \|l\|_*$.*

Proof. We have $|(l_1 + l_2)(u)| \le |l_1(u)| + |l_2(u)| \le \|l_1\|_* \|u\| + \|l_2\|_* \|u\| = (\|l_1\|_* + \|l_2\|_*)\|u\|$ for all $u \in Z$. This implies that $l_1 + l_2$ is bounded and that $\|l_1 + l_2\|_* \le \|l_1\|_* + \|l_2\|_*$.
Similarly, $|(\kappa l)(u)| = |\kappa||l(u)| \le |\kappa| \|l\|_* \|u\|$ for all $u \in Z$. This implies that κl is bounded and that $\|\kappa l\|_* \le |\kappa| \|l\|_*$. If $\kappa = 0$, then the equality is obvious. If $\kappa \ne 0$, to get the opposite inequality, we write $|\kappa||l(u)| = |(\kappa l)(u)| \le \|\kappa l\|_* \|u\|$ and, hence, $|l(u)| \le \frac{\|\kappa l\|_*}{|\kappa|} \|u\|$ for all $u \in Z$ and, hence, $\|l\|_* \le \frac{\|\kappa l\|_*}{|\kappa|}$. □

Definition. *Let Z be a normed space. The set of all bounded linear functionals on Z or, equivalently, of all continuous linear functionals on Z,*

$$Z^* = \{ l \mid l \text{ is a bounded linear functional on } Z \},$$

is called the ***topological dual*** *of Z or the* ***norm-dual*** *of Z or just the* ***dual*** *of Z.*

Proposition 7.17 together with the remarks about the norm of the zero functional imply that Z^* is a linear subspace of Z' and that $\| \cdot \|_* : Z^* \to \mathbb{R}$ is a norm on Z^*.

Theorem 7.3. *If Z is a normed space, then Z^* is a Banach space.*

Proof. Let (l_n) be a Cauchy sequence in Z^*.
For all $u \in Z$ we have $|l_n(u) - l_m(u)| = |(l_n - l_m)(u)| \le \|l_n - l_m\|_* \|u\| \to 0$ as $n, m \to +\infty$. Thus, $(l_n(u))$ is a Cauchy sequence in F and, hence, converges to some element of F. We define $l : Z \to F$ by

$$l(u) = \lim_{n \to +\infty} l_n(u) \qquad \text{for all } u \in Z.$$

It is easy to show that l is linear, i.e. $l \in Z'$, and we shall show that $l \in Z^*$ and $\|l_n - l\|_* \to 0$.
Now, there is N so that $\|l_n - l_m\|_* \le 1$ for all $n, m \ge N$. Then $|l_n(u) - l_m(u)| \le \|l_n - l_m\|_* \|u\| \le \|u\|$ for all $u \in Z$ and all $n, m \ge N$ and, taking the limit as $n \to +\infty$ and $m = N$, we find $|l(u) - l_N(u)| \le \|u\|$ for all $u \in Z$. Therefore, $|l(u)| \le |l_N(u)| + \|u\| \le (\|l_N\|_* + 1)\|u\|$ for all $u \in Z$. Hence, $l \in Z^*$.
Moreover, for an arbitrary $\epsilon > 0$ there is N so that $\|l_n - l_m\|_* \le \epsilon$ for all $n, m \ge N$. Then $|l_n(u) - l_m(u)| \le \|l_n - l_m\|_* \|u\| \le \epsilon \|u\|$ for all $u \in Z$ and all $n, m \ge N$. Taking the limit as $m \to +\infty$, we find $|l_n(u) - l(u)| \le \epsilon \|u\|$ for all $u \in Z$ and all $n \ge N$. Therefore, $\|l_n - l\|_* \le \epsilon$ for all $n \ge N$ and, thus, $\|l_n - l\|_* \to 0$. □

In case $F = \mathbb{C}$, the linear space Z can also be considered as a linear space over $\mathbb{R}$. Therefore we may distinguish between real-linear and complex-linear functionals on Z. A *complex-linear functional* on Z is the same as a linear functional on Z, i.e. a function $l : Z \to \mathbb{C}$ satisfying: $l(u + v) = l(u) + l(v)$ and $l(\kappa u) = \kappa l(u)$ for $\kappa \in \mathbb{C}$ and $u, v \in Z$. A *real-linear functional* on Z is a function $l : Z \to \mathbb{R}$ satisfying: $l(u + v) = l(u) + l(v)$ and $l(\kappa u) = \kappa l(u)$ for $\kappa \in \mathbb{R}$ and $u, v \in Z$.

Proposition 7.18. *Let Z be a normed space over $\mathbb{C}$. For every bounded linear functional (i.e. complex-linear functional) l on Z the $m = \mathrm{Re}(l)$ is a bounded real-linear functional on Z with $\|m\|_* = \|l\|_*$. Conversely, for every bounded real-linear functional m on Z there is a unique bounded linear functional l on Z so that $\mathrm{Re}(l) = m$.*
The two functionals l, m satisfy the relation $l(u) = m(u) - i m(iu)$ for all $u \in Z$.

Proof. If $l : Z \to \mathbb{C}$ is a bounded linear functional on Z, then it is trivial to show that $m = \mathrm{Re}(l) : Z \to \mathbb{R}$ is a real-linear functional on Z and we leave it as an exercise.
We have $l(u) = \mathrm{Re}(l)(u) + i\,\mathrm{Im}(l)(u)$ and $l(iu) = \mathrm{Re}(l)(iu) + i\,\mathrm{Im}(l)(iu)$. Since $l(iu) = il(u)$,

we get $\mathrm{Re}(l)(iu) + i\,\mathrm{Im}(l)(iu) = i\,\mathrm{Re}(l)(u) - \mathrm{Im}(l)(u)$ for all $u \in Z$. Equating real parts, we find $\mathrm{Im}(l)(u) = -\mathrm{Re}(l)(iu) = -m(iu)$ for all $u \in Z$ and, thus,

$$l(u) = \mathrm{Re}(l)(u) + i\,\mathrm{Im}(l)(u) = m(u) - im(iu) \qquad \text{for all } u \in Z.$$

Now, for all $u \in Z$ we have $|m(u)| = |\mathrm{Re}(l(u))| \le |l(u)| \le \|l\|_*\|u\|$ and, hence, $\|m\|_* \le \|l\|_*$. Also, if $l(u) \neq 0$, we consider $\kappa = \frac{|l(u)|}{l(u)} \in \mathbb{C}$ with $|\kappa| = 1$ and we get

$$|l(u)| = \kappa l(u) = l(\kappa u) = \mathrm{Re}(l)(\kappa u) = m(\kappa u) \le |m(\kappa u)| \le \|m\|_*\|\kappa u\| = \|m\|_*\|u\|.$$

The inequality $|l(u)| \le \|m\|_*\|u\|$ is clearly true if $l(u) = 0$ and, hence, holds for all $u \in Z$. Therefore, $\|l\|_* \le \|m\|_*$.
Conversely, let $m : Z \to \mathbb{R}$ be a bounded real-linear functional on Z. We define $l : Z \to \mathbb{C}$ by $l(u) = m(u) - im(iu)$ for all $u \in Z$. It is obvious that $\mathrm{Re}(l) = m$ and it is easy to show that l is a linear functional on Z. That l is bounded with $\|l\|_* = \|m\|_*$ has already been shown above.
The uniqueness of l with $\mathrm{Re}(l) = m$ has also been shown. Indeed, we proved that, if $\mathrm{Re}(l)(u) = m(u)$ for all $u \in Z$, then $\mathrm{Im}(l)(u) = -m(iu)$ for all $u \in Z$ and, hence, $\mathrm{Im}(l)$ is uniquely determined by m. □

Proposition 7.19. *Let Z be any inner product space. For any $u \in Z$ we define $l_u : Z \to F$ by $l_u(v) = \langle v, u\rangle$ for all $v \in Z$. Then $l_u \in Z^*$ and $\|l_u\|_* = \|u\|$.*

Proof. It is trivial to prove that l_u is linear.
For every $v \in Z$ we have $|l_u(v)| = |\langle v, u\rangle| \le \|v\|\|u\|$ and, thus, $\|l_u\|_* \le \|u\|$.
On the other hand, $|l_u(u)| = \|u\|^2 = \|u\|\|u\|$ and thus $\|l_u\|_* = \|u\|$. □

The following theorem shows the opposite in the case of a Hilbert space.

Theorem 7.4. *Let Z be a Hilbert space. Then for every $l \in Z^*$ there is a unique $u \in Z$ such that $l = l_u$, i.e. such that $l(v) = \langle v, u\rangle$ for all $v \in Z$.*

Proof. If $l = 0$ then we consider $u = 0$ and, clearly, we have $l(u) = 0 = \langle v, u\rangle$ for all $v \in Z$.
Now, let $l \neq 0$. Then the kernel $N(l)$ of l is a proper closed linear subspace of Z. We take any $u_0 \in Z \setminus N(l)$ and the projection $v_0 = P_{N(l)}(u_0)$ of u_0 on $N(l)$. Then $(u_0 - v_0) \perp N(l)$ and we consider the element

$$u = \frac{\overline{l(u_0)}}{\|u_0 - v_0\|^2}\,(u_0 - v_0).$$

Thus, $u \perp N(l)$ and $\|u\| = \frac{|l(u_0)|}{\|u_0 - v_0\|} > 0$ and

$$l(u) = \frac{\overline{l(u_0)}}{\|u_0 - v_0\|^2}\,(l(u_0) - l(v_0)) = \frac{\overline{l(u_0)}}{\|u_0 - v_0\|^2}\,l(u_0) = \frac{|l(u_0)|^2}{\|u_0 - v_0\|^2} = \|u\|^2.$$

Now, for all $v \in Z$ we have that $l(v - \frac{l(v)}{l(u)}\,u) = l(v) - \frac{l(v)}{l(u)}\,l(u) = 0$. Hence, $v - \frac{l(v)}{l(u)}\,u \in N(l)$ and, thus, $\langle v - \frac{l(v)}{l(u)}\,u, u\rangle = 0$. This implies $\langle v, u\rangle = \frac{l(v)}{l(u)}\,\|u\|^2 = l(v)$ for all $v \in Z$. □

Proposition 7.20. *Let Z be a Hilbert space. Then the mapping $T : Z \to Z^*$ defined by $T(u) = l_u$ for all $u \in Z$ is an isometric conjugate-linear operator from Z onto Z^*.*

Proof. We have $T(u_1 + u_2)(v) = l_{u_1+u_2}(v) = \langle v, u_1 + u_2\rangle = \langle v, u_1\rangle + \langle v, u_2\rangle = l_{u_1}(v) + l_{u_2}(v) = T(u_1)(v) + T(u_2)(v)$ for all $v \in Z$ and, hence, $T(u_1 + u_2) = T(u_1) + T(u_2)$. Also, $T(\kappa u)(v) = l_{\kappa u}(v) = \langle v, \kappa u\rangle = \overline{\kappa}\,\langle v, u\rangle = \overline{\kappa}\,l_u(v) = \overline{\kappa}\,T(u)(v)$ for all $v \in Z$ and, hence, $T(\kappa u) = \overline{\kappa}\,T(u)$. Therefore, $T : Z \to Z^*$ is a conjugate-linear operator. Theorem 7.4 implies that T is onto Z^*. Also, $\|T(u)\|_* = \|l_u\|_* = \|u\|$. □

NORMED LATTICES.

This subsection is only about *real* linear spaces.

Definition. *We say that $\leq$ is an* ***order*** *on the real linear space Z if it satisfies*
(i) $u \leq u$ for all $u \in Z$.
(ii) If $u, v \in Z$ and $u \leq v$ and $v \leq u$, then $u = v$.
(iii) If $u, v, w \in Z$ and $u \leq v$ and $v \leq w$, then $u \leq w$.
(iv) If $u, v, w \in Z$ and $u \leq v$, then $u + w \leq v + w$.
(v) If $u, v \in Z$ and $\kappa \in \mathbb{R}^+$ and $u \leq v$, then $\kappa u \leq \kappa v$.
If $\leq$ is an order on the linear space Z, then $(Z, \leq)$ is called an ***ordered linear space****.*

Properties (i), (ii), (iii) define the general *order relation* on any set. Properties (iv) and (v) describe the connection between the order relation and the linear structure of the linear space Z.

For simplicity, from now on we shall say that Z (instead of $(Z, \leq)$) is an ordered linear space.

Definition. *Let Z be an ordered linear space. The set $Z^+ = \{u \in Z \,|\, 0 \leq u\}$ is called the* ***non-negative cone*** *of Z.*

Proposition 7.21. *Let Z be an ordered linear space. Then the non-negative cone Z^+ is closed under addition and under multiplication by positive real numbers. More precisely, (i) if $u, v \in Z^+$, then $u + v \in Z^+$, (ii) if $u \in Z^+$ and $\kappa \in \mathbb{R}^+$, then $\kappa u \in Z^+$.*

Proof. Exercise. □

Definition. *We say that the ordered real linear space Z is a* ***linear lattice*** *if every two elements of Z have a* ***least upper bound*** *or, more precisely, if for every $u, v \in Z$ there is a $w \in Z$ such that*
(i) $u \leq w$ and $v \leq w$,
(ii) if $w' \in Z$ and $u \leq w'$ and $v \leq w'$, then $w \leq w'$.
The least upper bound w of u, v is denoted $u \vee v$.
For every $u \in Z$ we denote $u^+ = u \vee 0$, $u^- = (-u) \vee 0$ and $|u| = u \vee (-u)$ and call them the ***non-negative part****, the* ***non-positive part*** *and the* ***absolute value*** *of u, respectively.*

Example. $\mathbb{R}^n$ is a linear lattice under the order $\leq$ defined by: $x \leq y$ if $x_j \leq y_j$ for all $j = 1, \ldots, n$. We have $x \vee y = (\max\{x_1, y_1\}, \ldots, \max\{x_n, y_n\})$.

Example. The real linear space $\mathbb{R}^X$ of all real valued functions $f : X \to \mathbb{R}$ is a linear lattice under the usual order $\leq$ defined by: $f \leq g$ if $f(x) \leq g(x)$ for all $x \in X$.
We have $(f \vee g)(x) = \max\{f(x), g(x)\}$ for all $x \in X$.
Also $f^+(x) = \max\{f(x), 0\}$, $f^-(x) = \max\{-f(x), 0\}$ and $|f|(x) = \max\{f(x), -f(x)\} = |f(x)|$ for all $x \in X$.

Proposition 7.22. *Let Z be a linear lattice.*
(i) If $u, v \in Z$ then the element $w = u + v - u \vee v$ is the largest lower bound of u, v. More precisely, (a) $w \leq u$ and $w \leq v$ and (b), if $w' \in Z$ and $w' \leq u$ and $w' \leq v$, then $w' \leq w$.
(ii) For all $u \in Z$ we have: $u^+ - u^- = u$ and $u^+ + u^- = |u|$.
(iii) For all $u \in Z$ we have: (a) $0 \leq u$ if and only if $u = u^+$ if and only if $u = |u|$, (b) $u \leq 0$ if and only if $u = -u^-$ if and only if $u = -|u|$.
(iv) For all $u, v \in Z$ we have $u \vee v = \frac{u+v+|u-v|}{2}$.

Proof. Exercise. □

Definition. *Let Z be a linear lattice. If $u, v \in Z$, then the element $w = u + v - u \vee v$ which was defined in (i) of Proposition 7.22, i.e. the* ***largest lower bound*** *of u, v, is denoted $u \wedge v$.*

Therefore, in any linear lattice we have $u \vee v + u \wedge v = u + v$. We also see that (iv) of Proposition 7.22 implies $u \wedge v = \frac{u+v-|u-v|}{2}$.

Definition. *Let Z be a real linear space with a norm $\|\cdot\|$ and an order $\leq$ under which it is a linear lattice. Then we say that $(Z, \|\cdot\|, \leq)$ is a **normed lattice** if $u, v \in Z$ and $|u| \leq |v|$ imply $\|u\| \leq \|v\|$.*
*If also $(Z, \|\cdot\|)$ is a Banach space, then $(Z, \|\cdot\|, \leq)$ is called a **Banach lattice**.*

Example. $\mathbb{R}^n$ with the order defined in one of the previous examples and with the Euclidean norm is a Banach lattice.

Definition. *Let Z be an ordered linear space. A linear functional $l : Z \to \mathbb{R}$ is called **non-negative** if it has non-negative values on the non-negative cone Z^+. This means: $l(u) \geq 0$ for all $u \in Z^+$.*

Proposition 7.23. *Let Z be a normed lattice. For every $l \in Z^*$ there are two non-negative bounded linear functionals $l^+, l^- \in Z^*$ such that $l = l^+ - l^-$. Also, $\|l^+\|_* \leq \|l\|_*$ and $\|l^-\|_* \leq \|l\|_*$.*

Proof. For each $u \in Z^+$, i.e. $u \geq 0$, we define

$$l^+(u) = \sup\{l(v) \mid v \in Z, 0 \leq v \leq u\}.$$

Obviously, $l^+(u) \geq l(0) = 0$ and $l^+(u) \geq l(u)$.
Also, if $0 \leq v \leq u$, then $\|v\| \leq \|u\|$ and, hence, $l(v) \leq |l(v)| \leq \|l\|_* \|v\| \leq \|l\|_* \|u\|$. Therefore,

$$l^+(u) \leq \|l\|_* \|u\| < +\infty. \tag{7.1}$$

For every $\kappa \in \mathbb{R}^+$ and $u \in Z^+$ we have

$$\begin{aligned} l^+(\kappa u) &= \sup\{l(v) \mid v \in Z, 0 \leq v \leq \kappa u\} = \sup\{l(\kappa v) \mid v \in Z, 0 \leq v \leq u\} \\ &= \kappa \sup\{l(v) \mid v \in Z, 0 \leq v \leq u\} = \kappa l^+(u). \end{aligned} \tag{7.2}$$

If $u_1, u_2 \in Z^+$, $0 \leq v_1 \leq u_1$ and $0 \leq v_2 \leq u_2$, then $l(v_1) + l(v_2) = l(v_1 + v_2)$ and, since $0 \leq v_1 + v_2 \leq u_1 + u_2$, it is implied that $l(v_1) + l(v_2) \leq l^+(u_1 + u_2)$. Taking supremum separately over v_1 and over v_2, we find $l^+(u_1) + l^+(u_2) \leq l^+(u_1 + u_2)$.
Now let $0 \leq v \leq u_1 + u_2$. We set $v_1 = u_1 \wedge v$ from which $0 \leq v_1 \leq u_1$ and $v_1 \leq v$. If we set $v_2 = v - v_1$, then it is easy to see that $0 \leq v_2 \leq u_2$ and of course $v = v_1 + v_2$. Hence, $l(v) = l(v_1) + l(v_2) \leq l^+(u_1) + l^+(u_2)$ from which $l^+(u_1 + u_2) \leq l^+(u_1) + l^+(u_2)$.
We conclude that

$$l^+(u_1 + u_2) = l^+(u_1) + l^+(u_2). \tag{7.3}$$

Until now $l^+(u)$ is defined only for $u \in Z^+$. For an arbitrary $u \in Z$ we have $u = u^+ - u^-$, where of course $u^+, u^- \in Z^+$. We then define

$$l^+(u) = l^+(u^+) - l^+(u^-) \qquad \text{for all } u \in Z.$$

Observe that, if $u = v - w$ for any $v, w \in Z^+$, then $u^+ + w = u^- + v$, and from (7.3):

$$l^+(u^+) + l^+(w) = l^+(u^+ + w) = l^+(u^- + v) = l^+(u^-) + l^+(v).$$

Hence, $l^+(u) = l^+(v) - l^+(w)$.
If $u_1, u_2 \in Z$, then $u_1 + u_2 = (u_1^+ + u_2^+) - (u_1^- + u_2^-)$ and from the last identity we get

$$l(u_1+u_2) = l(u_1^+ + u_2^+) - l(u_1^- + u_2^-) = l(u_1^+) + l(u_2^+) - l(u_1^-) - l(u_2^-) = l(u_1) + l(u_2). \tag{7.4}$$

If $u \in Z$ and $\kappa \in \mathbb{R}^+$, then $\kappa u = \kappa u^+ - \kappa u^-$ and (7.2) implies

$$l^+(\kappa u) = l^+(\kappa u^+) - l^+(\kappa u^-) = \kappa l^+(u^+) - \kappa l^+(u^-) = \kappa l^+(u), \tag{7.5}$$

while, if $\kappa \in \mathbb{R}^-$, then $\kappa u = |\kappa| u^- - |\kappa| u^+$ and (7.2) implies, again,

$$l^+(\kappa u) = l^+(|\kappa| u^-) - l^+(|\kappa| u^+) = |\kappa| l^+(u^-) - |\kappa| l^+(u^+) = \kappa l^+(u). \tag{7.6}$$

By (7.4), (7.5) and (7.6), $l^+ : Z \to \mathbb{R}$ is a linear functional.
If $u \in Z$, then from (7.1) we get

$$\begin{aligned}|l^+(u)| &= |l^+(u^+) - l^+(u^-)| \leq \max\{l^+(u^+), l^+(u^-)\} \leq \max\{\|l\|_*\|u^+\|, \|l\|_*\|u^-\|\} \\ &= \|l\|_* \max\{\|u^+\|, \|u^-\|\} \leq \|l\|_*\|u\|,\end{aligned}$$

since $0 \leq u^+ \leq |u|$ and $0 \leq u^- \leq |u|$.
Therefore, l^+ is a non-negative bounded linear functional of Z with $\|l^+\|_* \leq \|l\|_*$.
We also define $l^- = l^+ - l : Z \to \mathbb{R}$. This is, clearly, a linear functional of Z and it is non-negative, since for every $u \in Z^+$ we have $l^-(u) = l^+(u) - l(u) \geq 0$. Also it is obvious that l^- is bounded, since $\|l^-\|_* = \|l^+ - l\|_* \leq \|l^+\|_* + \|l\|_* \leq 2\|l\|_*$. But we can find a better estimate for the norm of l^-, namely $\|l^-\|_* \leq \|l\|_*$.
Indeed, if $0 \leq v \leq u$, then $0 \leq u - v \leq u$ and, hence, $l(v) - l(u) = -l(u - v) \leq |l(u - v)| \leq \|l\|_*\|u - v\| \leq \|l\|_*\|u\|$ and, thus, $l^-(u) = l^+(u) - l(u) \leq \|l\|_*\|u\|$.
Therefore, if $u \in Z$, we have

$$\begin{aligned}|l^-(u)| &= |l^-(u^+) - l^-(u^-)| \leq \max\{l^-(u^+), l^-(u^-)\} \leq \max\{\|l\|_*\|u^+\|, \|l\|_*\|u^-\|\} \\ &= \|l\|_* \max\{\|u^+\|, \|u^-\|\} \leq \|l\|_*\|u\|\end{aligned}$$

and we conclude that $\|l^-\|_* \leq \|l\|_*$. □

EXTENSIONS OF LINEAR FUNCTIONALS.

Assume that Z is a normed space and that Z_0 is a linear subspace of Z with the same norm. If $l_0 : Z_0 \to F$ and $l : Z \to F$ are two bounded linear functionals and l is an extension of l_0, then $\|l_0\|_* \leq \|l\|_*$. Indeed, for every $u \in Z_0$ we have $|l_0(u)| = |l(u)| \leq \|l\|_*\|u\|$ and, thus, $\|l_0\|_* \leq \|l\|_*$. Therefore, *when we extend a linear functional its norm increases* (in the broad sense). The next two very basic facts of Functional Analysis say that we can always extend a linear functional from a subspace to the whole space *keeping its norm fixed*. The first theorem deals with the case $F = \mathbb{R}$ and the second theorem considers the case $F = \mathbb{C}$.

Hahn-Banach Theorem. *Let Z be a normed space over $\mathbb{R}$, Z_0 be a linear subspace of Z and $l_0 \in Z_0^*$. Then there is at least one $l \in Z^*$ which is an extension of l_0 so that $\|l_0\|_* = \|l\|_*$.*

Proof. We consider the collection $\mathcal{F}$ the elements of which are all m with the following properties:
(i) $m : D(m) \to \mathbb{R}$ is a linear functional on $D(m)$ which is a linear subspace of Z,
(ii) m is an extension of l_0, i.e. $Z_0 = D(l_0) \subseteq D(m)$ and $l_0(u) = m(u)$ for all $u \in Z_0$,
(iii) $|m(u)| \leq \|l_0\|_*\|u\|$ for all $u \in D(m)$, i.e. $\|l_0\|_* = \|m\|_*$.
Thus, the elements of $\mathcal{F}$ are all the extensions of l_0 on linear subspaces of Z, which have the same norm as l_0.
Then $\mathcal{F}$ is not empty, since $l_0 \in \mathcal{F}$, and we define the following order relation on $\mathcal{F}$: $m_1 \prec m_2$ if m_2 is an extension of m_1.
Now assume that $\mathcal{G}$ is a totally ordered subcollection of $\mathcal{F}$. We define $Z' = \bigcup\{D(m) \,|\, m \in \mathcal{G}\}$.
Clearly, $Z_0 \subseteq Z' \subseteq Z$. If $u_1, u_2 \in Z'$, there are $m_1, m_2 \in \mathcal{G}$ so that $u_1 \in D(m_1)$ and $u_2 \in D(m_2)$. Since one of m_1, m_2, say m_2, is an extension of the other, we get that $u_1, u_2 \in D(m_2)$ and since $D(m_2)$ is a linear subspace of Z, we have that $u_1 + u_2 \in D(m_2)$ and, thus, $u_1 + u_2 \in Z'$.
Similarly, if $u \in Z'$, there is an $m \in \mathcal{G}$ so that $u \in D(m)$ and, hence, for all $\kappa \in \mathbb{R}$ we have $\kappa u \in D(m)$ and, thus, $\kappa u \in Z'$. Therefore, Z' is a linear subspace of Z. Now take any $u \in Z'$, whence $u \in D(m)$ for some $m \in \mathcal{G}$. If there is another $m' \in \mathcal{G}$ so that $u \in D(m')$, then since one of m, m' is an extension of the other we get $m(u) = m'(u)$. This implies that we can define a function $l' : Z' \to \mathbb{R}$ so that $l'(u) = m(u)$ for any $m \in \mathcal{G}$ with $u \in D(m)$.
We have seen that, if $u_1, u_2 \in Z'$, then there is some $m \in \mathcal{G}$ so that $u_1, u_2 \in D(m)$ and, thus, $l'(u_1 + u_2) = m(u_1 + u_2) = m(u_1) + m(u_2) = l'(u_1) + l'(u_2)$. In the same way we can prove that $l'(\kappa u) = \kappa l'(u)$ for all $u \in Z'$ and $\kappa \in \mathbb{R}$. Therefore, l' is a linear functional on Z'. It is clear

that l' is an extension of l_0 and that $|l'(u)| \leq \|l_0\|_* \|u\|$ for all $u \in Z'$. Thus, $l' \in \mathcal{F}$. It is also clear that l' is an extension of all $m \in \mathcal{G}$ and, hence, l' is an upper bound of $\mathcal{G}$.
Now, Zorn's Lemma implies that $\mathcal{F}$ has at least one maximal element. In other words there is some l with the properties (i), (ii) and (iii) so that there is no m with the same properties which is a *proper* extension of l.
Now it is enough to prove that $D(l) = Z$.
To get a contradiction we assume that $D(l) \neq Z$ and we take any $u_0 \in Z \setminus D(l)$. We consider the linear subspace

$$W = \{u + \kappa u_0 \,|\, u \in D(l), \kappa \in \mathbb{R}\}.$$

Then $D(l)$ is a proper subspace of W and we shall define a linear functional $m : W \to \mathbb{R}$ so that $m(u) = l(u)$ for all $u \in D(l)$ and $|m(u)| \leq \|l\|_* \|u\| = \|l_0\|_* \|u\|$ for all $u \in W$. Then m is a proper extension of l with the properties (i), (ii) and (iii) and we arrive at a contradiction.
To define m we consider an *a priori* arbitrary $\kappa_0 \in \mathbb{R}$ and we consider

$$m(u + \kappa u_0) = l(u) + \kappa\kappa_0, \qquad u \in D(l), \kappa \in \mathbb{R}.$$

Then it is easy to see that m is a linear functional on W and that $m(u) = l(u)$ for every $u \in D(l)$. It remains to choose κ_0 so that $|m(u + \kappa u_0)| \leq \|l\|_* \|u + \kappa u_0\|$ or, equivalently, $|l(u) + \kappa\kappa_0| \leq \|l\|_* \|u + \kappa u_0\|$ for all $u \in D(l)$ and $\kappa \in \mathbb{R}$.
When $\kappa = 0$ what we want takes the form $|l(u)| \leq \|l\|_* \|u\|$ and this is true independently of the choice of κ_0. For $\kappa \neq 0$ what we have to prove takes the following successive equivalent forms:

$$|l(u) + \kappa\kappa_0| \leq \|l\|_* \|u + \kappa u_0\|, \qquad u \in D(l), \kappa \neq 0$$

$$|l(\tfrac{1}{\kappa}\, u) + \kappa_0| \leq \|l\|_* \|\tfrac{1}{\kappa}\, u + u_0\|, \qquad u \in D(l), \kappa \neq 0$$

$$|l(u) + \kappa_0| \leq \|l\|_* \|u + u_0\|, \qquad u \in D(l)$$

$$-l(u) - \|l\|_* \|u + u_0\| \leq \kappa_0 \leq -l(u) + \|l\|_* \|u + u_0\|, \qquad u \in D(l). \tag{7.7}$$

Now, if we prove that

$$-l(u_1) - \|l\|_* \|u_1 + u_0\| \leq -l(u_2) + \|l\|_* \|u_2 + u_0\|, \qquad u_1, u_2 \in D(l), \tag{7.8}$$

then we get $\sup\{-l(u) - \|l\|_* \|u + u_0\| \,|\, u \in D(l)\} \leq \inf\{-l(u) + \|l\|_* \|u + u_0\| \,|\, u \in D(l)\}$ and then we can choose any number κ_0 between the supremum and the infimum and this number certainly satisfies (7.7). But (7.8) is equivalent to $l(u_2) - l(u_1) \leq \|l\|_* (\|u_2 + u_0\| + \|u_1 + u_0\|)$.
But $l(u_2) - l(u_1) = l(u_2 - u_1) \leq \|l\|_* \|u_2 - u_1\| \leq \|l\|_* (\|u_2 + u_0\| + \|u_1 + u_0\|)$. ☐

Bohnenblust-Sobczyk Theorem. *Let Z be a normed space over $\mathbb{C}$, Z_0 be a linear subspace of Z and $l_0 \in Z_0^*$. Then there is at least one $l \in Z^*$ which is an extension of l_0 so that $\|l_0\|_* = \|l\|_*$.*

Proof. We know from Proposition 7.18 that $m_0 = \mathrm{Re}(l_0)$ is a real-linear functional on Z_0 with $\|m_0\|_* = \|l_0\|_*$. By the Hahn-Banach Theorem, there is a real-linear functional m on Z which is an extension of m_0 with $\|m_0\|_* = \|m\|_*$. By Proposition 7.18 again, there is a linear functional l on Z such that $\mathrm{Re}(l) = m$. Then we have $\|m\|_* = \|l\|_*$, hence $\|l_0\|_* = \|l\|_*$, and, also, $l(u) = m(u) - im(iu) = m_0(u) - im_0(iu) = l_0(u)$ for all $u \in Z_0$, which means that l is an extension of l_0. ☐

Proposition 7.24. *Let Z be a normed space. Then $\|u\| = \max\{|l(u)| \,|\, l \in Z^*, \|l\|_* \leq 1\}$ for all $u \in Z$.*

Proof. First we observe that for every $l \in Z^*$ with $\|l\|_* \leq 1$ we have $|l(u)| \leq \|l\|_* \|u\| \leq \|u\|$. Therefore, $\sup\{|l(u)| \,|\, l \in Z^*, \|l\|_* \leq 1\} \leq \|u\|$.
Now, we consider the linear subspace $Z_0 = \{\kappa u \,|\, \kappa \in F\}$ of Z and we define $l_0 : Z_0 \to F$ by $l_0(\kappa u) = \kappa \|u\|$ for all $\kappa \in F$.
It is clear that l_0 is linear and that $|l_0(\kappa u)| = |\kappa| \|u\| = \|\kappa u\|$ for all $\kappa \in F$. Thus, $\|l_0\|_* = 1$.
Then there is an $l \in Z^*$ which is an extension of l_0 with $\|l\|_* = \|l_0\|_* = 1$. Since $|l(u)| = |l_0(u)| = \|u\|$, the proof is finished. ☐

Definition. *Let Z be a normed space and Z^* be its dual space. The dual of Z^* is denoted Z^{**}.*

Proposition 7.25. *Let Z be a normed space. For every $u \in Z$ we define $L_u : Z^* \to F$ by $L_u(l) = l(u)$ for all $l \in Z^*$. Then $L_u \in Z^{**}$ and $\|L_u\|_{**} = \|u\|$.*

Proof. It is clear that L_u is linear.
Also, $\|L_u\|_{**} = \sup\{|L_u(l)| \,|\, l \in Z^*, \|l\|_* \leq 1\} = \sup\{|l(u)| \,|\, l \in Z^*, \|l\|_* \leq 1\} = \|u\|$, where the last equality is due to Proposition 7.24. □

Proposition 7.26. *Let Z be a normed space. The mapping $T : Z \to Z^{**}$ defined by $T(u) = L_u$ is an isometry from Z into Z^{**}.*

Proof. Exercise. □

Through the mapping T we may *identify* each $u \in Z$ with the corresponding $T(u) = L_u \in Z^{**}$ and we may view every $u \in Z$ as a bounded linear functional on Z^*. If we write u instead of L_u then the relation $L_u(l) = l(u)$ becomes $u(l) = l(u)$ for $u \in Z$, $l \in Z^*$. This symmetric relation says that l acts as a function on u and also that u (meaning: L_u) acts as a function on l.

Definition. *If the mapping $T : Z \to Z^{**}$ is onto we say that $(Z, \|\cdot\|)$ is* ***reflexive****.*

Proposition 7.27. *Every Hilbert space is reflexive.*

Proof. Let Z be a Hilbert space and take any $L \in Z^{**}$. We define $l : Z \to F$ by $l(v) = \overline{L(l_v)}$ for all $v \in Z$.
We recall that $l_v \in Z^*$ is such that $l_v(w) = \langle w, v\rangle$ for all $w \in Z$.
We know that $l_{v_1+v_2} = l_{v_1} + l_{v_2}$ and $l_{\kappa v} = \overline{\kappa}\, l_v$ and, hence, l is a linear functional on Z. Also $|l(v)| = |L(l_v)| \leq \|L\|\|l_v\|_* = \|L\|\|v\|$ and, thus, $l \in Z^*$ with $\|l\| \leq \|L\|$.
Now, there is some $u \in Z$ so that $l = l_u$, i.e. $l(v) = \langle v, u\rangle$ for all $v \in Z$. Of course this means that $L(l_v) = \overline{l(v)} = \langle u, v\rangle$ for all $v \in Z$.
On the other hand, $T(u)(l_v) = l_v(u) = \langle u, v\rangle$ for all $v \in Z$.
Therefore, $T(u)(l_v) = L(l_v)$ for all $v \in Z$. Since every element of Z^* is of the form l_v for some $v \in Z$, we get that $T(u) = L$. □

WEAK AND WEAK* CONVERGENCE.

Definition. *Let Z be a normed space and Z^* be its dual space. We say that a sequence (u_n) in Z* ***converges weakly*** *to $u \in Z$ if $l(u_n) \to l(u)$ for every $l \in Z^*$ and we write $u_n \xrightarrow{w} u$.*
Similarly, we say that a sequence (l_n) in Z^* ***converges weakly**** *to $l \in Z^*$ if $l_n(u) \to l(u)$ for every $u \in Z$ and we write $l_n \xrightarrow{w*} l$.*

Proposition 7.28. *If $u_n \to u$ in Z, then $u_n \xrightarrow{w} u$ in Z. If $l_n \to l$ in Z^*, then $l_n \xrightarrow{w*} l$ in Z^*.*

Proof. Exercise. □

Therefore, convergence in Z is stronger than weak convergence in Z and convergence in Z^* is stronger than weak* convergence in Z^*.

Definition. *Let Z be a normed space.*
$F \subseteq Z$ is called ***weakly sequentially closed*** *if (u_n) in F and $u_n \xrightarrow{w} u$ imply $u \in F$.*
$F \subseteq Z^*$ is called ***weakly* sequentially closed*** *if (l_n) in F and $l_n \xrightarrow{w*} l$ imply $l \in F$.*

Proposition 7.29. *Let Z be a normed space. If $F \subseteq Z$ is weakly sequentially closed then it is closed. If $F \subseteq Z^*$ is weakly* sequentially closed then it is closed.*

Proof. Exercise. □

Proposition 7.30. *If* $u_n \xrightarrow{w} u$ *then* $\|u\| \leq \underline{\lim}_{n\to+\infty} \|u_n\|$.
If $l_n \xrightarrow{w*} l$ *then* $\|l\| \leq \underline{\lim}_{n\to+\infty} \|l_n\|$.

Proof. For every $l \in Z^*$ with $\|l\|_* \leq 1$ we have $|l(u_n)| \leq \|u_n\|$ for all n and, hence, $|l(u)| \leq \underline{\lim}_{n\to+\infty} \|u_n\|$. Proposition 7.24 implies $\|u\| \leq \underline{\lim}_{n\to+\infty} \|u_n\|$.
Similarly, for every $u \in Z$ with $\|u\| \leq 1$ we have $|l_n(u)| \leq \|l_n\|_*$ for all n and, hence, $|l(u)| \leq \underline{\lim}_{n\to+\infty} \|l_n\|_*$. Therefore, $\|l\|_* \leq \underline{\lim}_{n\to+\infty} \|l_n\|_*$. □

Proposition 7.31. *Let Z be a normed space. Every closed ball in Z is weakly sequentially closed. Every closed ball in Z^* is weakly* sequentially closed.*

Proof. Exercise. Use Proposition 7.30. □

Uniform Boundedness Principle. *Let Z be a Banach space and (l_n) be a sequence in Z^* so that* $\sup_{n\in\mathbb{N}} |l_n(u)| < +\infty$ *for every $u \in Z$. Then* $\sup_{n\in\mathbb{N}} \|l_n\|_* < +\infty$.

Proof. For each $k \in \mathbb{N}$ we consider $F_k = \{u \in Z \,|\, |l_n(u)| \leq k \text{ for all } n \in \mathbb{N}\} \subseteq Z$. Due to the continuity of each l_n it is easy to show that every F_k is closed in Z. Also, because of the hypothesis that $\sup_{n\in\mathbb{N}} |l_n(u)| < +\infty$ for every $u \in Z$, we get $Z = \bigcup_{k=1}^{+\infty} F_k$. Now, since Z is a *complete* metric space, the classical Theorem of Baire implies that at least one of the sets F_k has non-empty interior. I.e. there is some k_0 and some ball $B(u_0; r_0)$ so that $\mathrm{cl}(B(u_0; r_0)) \subseteq F_{k_0}$. This means that $|l_n(u)| \leq k_0$ for all n and all $u \in Z$ with $\|u - u_0\| \leq r_0$. In particular, $|l_n(u_0)| \leq k_0$ for all n. Now, if $\|u\| \leq 1$, we have that $\|(r_0 u + u_0) - u_0\| \leq r_0$ and, thus,

$$|l_n(u)| = \frac{1}{r_0} |l_n(r_0 u)| = \frac{1}{r_0} |l_n(r_0 u + u_0) - l_n(u_0)| \leq \frac{1}{r_0} (|l_n(r_0 u + u_0)| + |l_n(u_0)|) \leq \frac{2k_0}{r_0}$$

for all n. This implies that $\|l_n\|_* \leq \frac{2k_0}{r_0}$ for all n. □

Proposition 7.32. *Let Z be a Banach space and (l_n) be a sequence in Z^*.*
(i) If (l_n) is weakly convergent then* $\sup_{n\in\mathbb{N}} \|l_n\|_* < +\infty$.
(ii) If $\lim_{n\to+\infty} l_n(u)$ *exists in F for every $u \in Z$, then (l_n) is weakly* convergent.*

Proof. Assume that $l_n \xrightarrow{w*} l$. Then $l_n(u) \to l(u)$ and, hence, $\sup_{n\in\mathbb{N}} |l_n(u)| < +\infty$ for every $u \in Z$. By the Uniform Boundedness Principle we get $\sup_{n\in\mathbb{N}} \|l_n\|_* < +\infty$.
If $\lim_{n\to+\infty} l_n(u)$ exists in F for every $u \in Z$, then again $\sup_{n\in\mathbb{N}} |l_n(u)| < +\infty$ for every $u \in Z$ and, as before, $\sup_{n\in\mathbb{N}} \|l_n\|_* < +\infty$. Now, we define $l : Z \to F$ by $l(u) = \lim_{n\to+\infty} l_n(u)$ for all $u \in Z$ and it is easy to see that the linearity of all l_n implies the linearity of l. Moreover, if we set $M = \sup_{n\in\mathbb{N}} \|l_n\|_*$, then we have $|l_n(u)| \leq \|l_n\|_* \|u\| \leq M\|u\|$ for all $u \in Z$ and, taking the limit, $|l(u)| \leq M\|u\|$. Thus, $l \in Z^*$ and $l_n \xrightarrow{w*} l$. □

Uniform Boundedness Principle. *Let Z be a normed space and (u_n) be a sequence in Z so that* $\sup_{n\in\mathbb{N}} |l(u_n)| < +\infty$ *for every $l \in Z^*$. Then* $\sup_{n\in\mathbb{N}} \|u_n\| < +\infty$.

Proof. For each $k \in \mathbb{N}$ we consider $F_k = \{l \in Z^* \,|\, |l(u_n)| \leq k \text{ for all } n \in \mathbb{N}\} \subseteq Z^*$. It is easy to see that every F_k is closed in Z^* and, because of the hypothesis that $\sup_{n\in\mathbb{N}} |l(u_n)| < +\infty$ for every $l \in Z^*$, we get $Z^* = \bigcup_{k=1}^{+\infty} F_k$. Since Z^* is a complete metric space, at least one of the sets F_k has non-empty interior. I.e. there is some k_0 and some ball $B(l_0; r_0)$ so that $\mathrm{cl}(B(l_0; r_0)) \subseteq F_{k_0}$. This means that $|l(u_n)| \leq k_0$ for all n and all $l \in Z^*$ with $\|l - l_0\|_* \leq r_0$. In particular, $|l_0(u_n)| \leq k_0$ for all n. Now, if $\|l\|_* \leq 1$, we get $\|(r_0 l + l_0) - l_0\|_* \leq r_0$ and, thus,

$$|l(u_n)| = \frac{1}{r_0} |(r_0 l)(u)| = \frac{1}{r_0} |(r_0 l_n + l_0)(u) - l_0(u)| \leq \frac{1}{r_0} (|(r_0 l + l_0)(u)| + |l_0(u)|) \leq \frac{2k_0}{r_0}$$

for all n. Proposition 7.24 implies that $\|u_n\| \leq \frac{2k_0}{r_0}$ for all n. □

Proposition 7.33. *Let Z be a normed space and (u_n) be a sequence in Z.*
(i) If (u_n) is weakly convergent then $\sup_{n\in\mathbb{N}} \|u_n\| < +\infty$.
(ii) If $\lim_{n\to+\infty} l(u_n)$ exists in F for every $l \in Z^*$, then there is an $L \in Z^{**}$ so that $l(u_n) \to L(l)$ for every $l \in Z^*$. If, moreover, Z is reflexive, then (u_n) is weakly convergent.

Proof. (i) Exercise.
(ii) We consider the elements $L_{u_n} \in Z^{**}$. Then $\lim_{n\to+\infty} L_{u_n}(l) = \lim_{n\to+\infty} l(u_n)$ exists in F for every $l \in Z^*$ and Proposition 7.32 (applied to the Banach space Z^*) implies that there is some $L \in Z^{**}$ so that $L_{u_n} \xrightarrow{w*} L$ as elements of the dual Z^{**} of Z^*. This means that $l(u_n) = L_{u_n}(l) \to L(l)$ for all $l \in Z^*$.
If Z is reflexive, then there is some $u \in Z$ so that $L = L_u$ and, hence, $l(u_n) \to L(l) = L_u(l) = l(u)$ for all $l \in Z^*$. I.e. $u_n \xrightarrow{w} u$. □

Definition. *Let Y be a non-empty set, $(Y_i, \mathcal{T}_i)_{i\in I}$ be a family of topological spaces and consider functions $f_i : Y \to Y_i$ for each $i \in I$. The smallest topology $\mathcal{T}_w$ on Y under which all functions f_i are continuous is called the* ***weak topology*** *on Y induced by the family of functions $(f_i)_{i\in I}$.*

Proposition 7.34. *Let Y be a non-empty set, $(Y_i, \mathcal{T}_i)_{i\in I}$ be a family of topological spaces and consider functions $f_i : Y \to Y_i$ for each $i \in I$ and let $\mathcal{T}_w$ be the weak topology on Y induced by $(f_i)_{i\in I}$. Let also $(X, \mathcal{T})$ be a topological space and $f : X \to Y$. Then f is continuous if and only if all $f_i \circ f$ are continuous.*

Proof. It is obvious that, if f is continuous, then all f_i are continuous.
Conversely, assume that all $f_i \circ f$ are continuous. We define $\mathcal{T}' = \{U \subseteq Y \mid f^{-1}(U) \in \mathcal{T}\}$. It is easy to see that $\emptyset \in \mathcal{T}'$, $Y \in \mathcal{T}'$ and that $\mathcal{T}'$ is closed under unions and countable intersections. Thus, $\mathcal{T}'$ is a topology on Y. Now, if $U_i \in \mathcal{T}_i$, then, since $f_i \circ f$ is continuous, we have that $f^{-1}(f_i^{-1}(U_i)) \in \mathcal{T}$ and, hence, $f_i^{-1}(U_i) \in \mathcal{T}'$. Therefore, f_i is continuous under the topology $\mathcal{T}'$ on Y. Since this is true for all $i \in I$, we get that $\mathcal{T}_w$ is smaller that $\mathcal{T}'$. This implies that $f^{-1}(U) \in \mathcal{T}$ for all $U \in \mathcal{T}_w$ and, hence, f is continuous. □

Definition. *Let $(Y_i, \mathcal{T}_i)_{i\in I}$ be a family of topological spaces and let $Y = \prod_{i\in I} Y_i$ be the product space of all Y_i. For each $i \in I$ we consider the projection $\pi_i : Y \to Y_i$ defined by $\pi_i(y) = y_i$ for every $y = (y_i)_{i\in I}$. Then the weak topology on Y induced by the family of projections $(\pi_i)_{i\in I}$ is called the* ***product topology*** *on Y.*

Definition. *Let Z be a normed space.*
The weak topology on Z induced by the family of functions Z^* is called the ***weak topology*** *on Z.*
The weak topology on Z^* induced by the family of functions $(L_u)_{u\in Z}$ is called the ***weak* topology*** *on Z^*.*

Proposition 7.35. *Let Z be a normed space.*
The weak topology on Z is weaker that the metric space topology on Z induced by its norm.
The weak topology on Z^* is weaker than the metric space topology on Z^* induced by its norm.*

Proof. Exercise. □

Thus, if $K \subseteq Z$ is weakly closed, then it is closed. Similarly, if $K \subseteq Z^*$ is weakly* closed, then it is closed.

WEAK AND WEAK* COMPACTNESS.

Definition. *Let Z be a normed space.*
$K \subseteq Z$ is called ***weakly sequentially compact*** *if every sequence in K has a subsequence which is weakly convergent to an element of K.*
$K \subseteq Z^*$ is called ***weakly* sequentially compact*** *if every sequence in K has a subsequence which is weakly* convergent to an element of K.*

Proposition 7.36. *Let Z be a normed space.*
Every weakly sequentially compact $K \subseteq Z$ is weakly sequentially closed and bounded.
Every weakly sequentially compact $K \subseteq Z^*$ is weakly* sequentially closed and -if Z is Banach- bounded.*

Proof. Exercise. $\square$

Theorem 7.5. *Let Z be a separable Banach space. Then a $K \subseteq Z^*$ is weakly* sequentially compact if and only if it is bounded and weakly* sequentially closed. In particular, every closed ball in Z^* is weakly* sequentially compact.*

Proof. Assume that $\|l_n\|_* \leq M < +\infty$ for all n.
We consider a dense countable subset $\{u_1, u_2, \ldots\}$ of Z.
Since $|l_n(u_1)| \leq M\|u_1\|$ for all n and since $\{y \in F \,|\, |y| \leq M\|u_1\|\}$ is compact, there is a subsequence $(l_n^{(1)})$ of (l_n) such that $(l_n^{(1)}(u_1))$ is a convergent sequence in F. I.e.

$$\lim_{n\to+\infty} l_n^{(1)}(u_1) \text{ exists in } F.$$

Since $|l_n^{(1)}(u_2)| \leq M\|u_2\|$ for all n and since $\{y \in F \,|\, |y| \leq M\|u_2\|\}$ is compact, there is a subsequence $(l_n^{(2)})$ of $(l_n^{(1)})$ such that $(l_n^{(2)}(u_2))$ is a convergent sequence in F. I.e.

$$\lim_{n\to+\infty} l_n^{(2)}(u_2) \text{ exists in } F.$$

We continue inductively and for every u_j we construct a sequence $(l_n^{(j)})$ so that

$$\lim_{n\to+\infty} l_n^{(j)}(u_j) \text{ exists in } F$$

and so that $(l_n^{(j)})$ is a subsequence of $(l_n^{(j-1)})$ for all $j \geq 2$ and $(l_n^{(1)})$ is a subsequence of (l_n).
Now we consider the *diagonal* sequence

$$(l_n^{(n)}).$$

This is a subsequence of (l_n). Also for every j the sequence $(l_n^{(n)})$ after its j-th term is a subsequence of $(l_n^{(j)})$ and, thus,

$$\lim_{n\to+\infty} l_n^{(n)}(u_j) \text{ exists in } F.$$

For simplicity we write $l_n^* = l_n^{(n)}$ and, thus, $\lim_{n\to+\infty} l_n^*(u_j)$ exists in F for every u_j. We shall prove that this is true for all $u \in Z$.
We take any $u \in Z$ and any $\epsilon > 0$. Then there is a u_j so that $\|u - u_j\| < \epsilon$ and then there is an n_0 so that $|l_n^*(u_j) - l_m^*(u_j)| < \epsilon$ for all $m, n \geq n_0$. Hence,

$$\begin{aligned} |l_n^*(u) - l_m^*(u)| &\leq |l_n^*(u) - l_n^*(u_j)| + |l_n^*(u_j) - l_m^*(u_j)| + |l_m^*(u_j) - l_m^*(u)| \\ &\leq M\|u - u_j\| + |l_n^*(u_j) - l_m^*(u_j)| + M\|u_j - u\| < (2M+1)\epsilon. \end{aligned}$$

Therefore, the sequence $(l_n^*(u))$ is Cauchy in F and $\lim_{n\to+\infty} l_n^*(u)$ exists in F.
Now we define $l : Z \to F$ by

$$l(u) = \lim_{n\to+\infty} l_n^*(u) \in F, \qquad u \in Z.$$

Since all l_n^* are linear on Z, it is clear that l is a linear functional on Z. We also have that $|l_n^*(u)| \leq M\|u\|$ for all n and all $u \in Z$ and, taking the limit,

$$|l(u)| \leq M\|u\|, \qquad u \in Z.$$

Therefore, $l \in Z^*$ and $l_n^* \xrightarrow{w*} l$. $\square$

Lemma 7.1. *If Z^* is separable, then Z is separable.*

Proof. Let Q be a countable dense subset of Z^*. Clearly, we may assume that $0 \notin Q$. For each $m \in Q$ there is some $u_m \in Z$ so that $\|u_m\| = 1$ and $|m(u_m)| > \frac{1}{2}\|m\|_*$. Then the set $P = \{u_m \,|\, m \in Q\}$ is countable.
Now we consider the closed linear subspace $\tilde{P}$ which is produced by P, i.e. all limits of linear combinations of elements of P. Then $\tilde{P}$ is separable, since every element of $\tilde{P}$ is the limit of linear combinations of elements of P with rational coefficients.
It is enough to prove that $\tilde{P} = Z$.
We assume that $\tilde{P}$ is a proper subspace of Z and then there is some $u_0 \in Z \setminus \tilde{P}$ with $\|u_0 - u\| \geq 1$ for all $u \in \tilde{P}$. We consider the linear subspace $Z_0 = \{u + \kappa u_0 \,|\, u \in \tilde{P}, \kappa \in F\}$ and the linear functional $l_0 : Z_0 \to F$ defined by

$$l_0(u + \kappa u_0) = \kappa, \qquad u \in \tilde{P}, \kappa \in F.$$

We have that

$$|l_0(u + \kappa u_0)| = |\kappa| \leq |\kappa| \|\tfrac{1}{\kappa}\, u + u_0\| = \|u + \kappa u_0\|$$

for all $u \in \tilde{P}$ and $\kappa \in F$. Therefore $l_0 \in Z_0^*$ with $\|l_0\|_* \leq 1$.
Now there is some $l \in Z^*$ which is an extension of l_0 with $\|l_0\|_* = \|l\|_*$.
Since Q is dense in Z^*, there is some $m \in Q$ so that $\|m - l\|_* \leq \frac{1}{2}\|m\|_*$. Then $u_m \in P$ and, thus, $l(u_m) = 0$. This implies that

$$\tfrac{1}{2}\|m\|_* < |m(u_m)| = |m(u_m) - l(u_m)| \leq \|m - l\|_* \|u_m\| \leq \tfrac{1}{2}\|m\|_*$$

and we arrive at a contradiction. □

Lemma 7.2. *If Z is reflexive and W is a closed linear subspace of Z, then W is also reflexive.*

Proof. We define the usual isometry $T_W : W \to W^{**}$ by

$$T_W(u)(l) = l(u), \qquad l \in W^*, u \in W$$

and we want to prove that it is onto W^{**}. We know that the similar isometry $T_Z : Z \to Z^{**}$ defined by

$$T_Z(u)(l) = l(u), \qquad l \in Z^*, u \in Z$$

is onto.
We take an arbitrary $\tilde{L} \in W^{**}$ and consider any $l \in Z^*$. We then define the restriction $\tilde{l}$ of l on W and we have that $\tilde{l} \in W^*$ with $\|\tilde{l}\|_* \leq \|l\|_*$. And finally we define $L : Z^* \to F$ by

$$L(l) = \tilde{L}(\tilde{l}), \qquad l \in Z^*.$$

It is easy to show that L is linear and, since $|L(l)| = |\tilde{L}(\tilde{l})| \leq \|\tilde{L}\|_{**}\|\tilde{l}\|_* \leq \|\tilde{L}\|_{**}\|l\|_*$, we have that L is bounded and $\|L\|_{**} \leq \|\tilde{L}\|_{**}$. Thus, $L \in Z^{**}$.
Now, since T_Z is onto, there is a $u_0 \in Z$ so that $T_Z(u_0) = L$. I.e. $T_Z(u_0)(l) = l(u_0)$ for all $l \in Z^*$.
For the moment we assume that $u_0 \notin W$. Then there is some $c > 0$ so that $\|u_0 - u\| \geq c$ for all $u \in W$. Now, as in the proof of Lemma 8.1, we consider the linear subspace $Z_0 = \{u + \kappa u_0 \,|\, u \in W, \kappa \in F\}$ and we define $l_0 : Z_0 \to F$ by

$$l_0(u + \kappa u_0) = \kappa, \qquad u \in W, \kappa \in F.$$

Then l_0 is linear and we have that

$$|l_0(u + \kappa u_0)| = |\kappa| \leq \tfrac{|\kappa|}{c}\|\tfrac{1}{\kappa}\, u + u_0\| = \tfrac{1}{c}\|u + \kappa u_0\|$$

for all $u \in W$ and $\kappa \in F$. Therefore, $l_0 \in Z_0^*$ and we know that there is an $l \in Z^*$ which is an extension of l_0 with $\|l\|_* = \|l_0\|_*$.
Now, the restriction $\tilde{l}$ of l on W is the same as the restriction of l_0 on W, which is 0. Hence

$$1 = l(u_0) = T_Z(u_0)(l) = L(l) = \tilde{L}(\tilde{l}) = \tilde{L}(0) = 0$$

and we get a contradiction.
Therefore, $u_0 \in W$ and now for every $l \in W^*$ we consider some $l' \in Z^*$ which is an extension of l, i.e. so that $l = \tilde{l}'$ and we get

$$T_W(u_0)(l) = l(u_0) = l'(u_0) = T_Z(u_0)(l') = L(l') = \tilde{L}(l).$$

Thus, $T_W(u_0) = \tilde{L}$ and T_W is onto. □

Theorem 7.6. *Let Z be a reflexive normed space. Then a $K \subseteq Z$ is weakly sequentially compact if and only if it is bounded and weakly sequentially closed. In particular, every closed ball in Z is weakly sequentially compact.*

Proof. Let (u_n) be a sequence in B, i.e. $\|u_n\| \leq 1$ for all n. We consider the closed linear subspace W of Z which is produced by all u_n, i.e. all limits of linear combinations of all the u_n. Then W is separable.
Lemma 7.2 implies that W is reflexive. Since W is separable and W^{**} is isometric to W, we get that W^{**} is also separable. Now Lemma 7.1 implies that W^* is separable too and we apply Theorem 7.5 to the space W^*.
We consider the isometry $T : W \to W^{**}$ and the sequence $(L_n) = (T(u_n))$ in W^{**} which is bounded since $\|L_n\|_{**} = \|u_n\| \leq 1$ for all n. Then there is a subsequence (L_{n_k}) which converges weakly* to some $L \in W^{**}$. I.e.

$$L_{n_k}(l) \to L(l)$$

for all $l \in W^*$. Since T is onto, there is some $u \in W$ so that $T(u) = L$. Now we have that

$$l(u_{n_k}) = T(u_{n_k})(l) = L_{n_k}(l) \to L(l) = T(u)(l) = l(u)$$

for all $l \in W^*$ and hence for all $l \in Z^*$ and, thus, $u_{n_k} \xrightarrow{w} u$. Moreover, we get that

$$\|u\| \leq \varliminf_{k \to +\infty} \|u_{n_k}\| \leq 1$$

and, hence, $u \in B$. □

Banach-Alaoglou Theorem. *Let Z be a normed space and Z^* be its dual. Then the closed unit ball $B^* = \{l \in Z^* \mid \|l\|_* \leq 1\}$ is weak* compact.*

Proof. We consider the case $F = \mathbb{R}$. The case $F = \mathbb{C}$ is similar and we leave as an exercise.
If $l \in B^*$, then we have $|l(u)| \leq \|u\|$ or equivalently $l(u) \in [-\|u\|, \|u\|]$ for all $u \in Z$. We define the product space

$$W = \textstyle\prod_{u \in Z} [-\|u\|, \|u\|]$$

with the product topology (each closed interval has the usual Euclidean topology). By the Theorem of Tychonov W is a compact topological space.
We also define the mapping $T : B^* \to W$ by

$$T(l) = (l(u))_{u \in Z}, \qquad l \in B^*$$

and it is clear that T is one-to-one.
Let $X = T(B^*) \subseteq W$ so that $T^{-1} : X \to B^*$. Then for every $x \in X$ there is an $l \in B^*$ so that $x = T(l) = (l(u))_{u \in Z}$ and, thus,

$$(L_u \circ T^{-1})(x) = L_u(l) = l(u) = \pi_u(x), \qquad u \in Z.$$

Therefore, $L_u \circ T^{-1} = \pi_u$ for all $u \in Z$. Now all π_u are continuous on W under the product topology and, hence, they are continuous on $X \subseteq W$ under the product topology. Therefore, all $L_u \circ T^{-1}$ are continuous on X under the product topology and Proposition 7.34 implies that T^{-1} is continuous from X under the product topology to B^* under the weak* topology. Now it is enough to prove that X is compact under the product topology and, since W is compact under the product topology, it is enough to prove that X is closed under the product topology.
Let $x \in W$ be a limit point of X. Take $u_1, u_2 \in Z$ and consider an arbitrary $\epsilon > 0$ and the open intervals

$$I_{u_1} = (x_{u_1} - \epsilon, x_{u_1} + \epsilon), \quad I_{u_2} = (x_{u_2} - \epsilon, x_{u_2} + \epsilon), \quad I_{u_1+u_2} = (x_{u_1+u_2} - \epsilon, x_{u_1+u_2} + \epsilon).$$

Consider also $I_u = \mathbb{R}$ for all $u \in Z$, $u \neq u_1, u_2, u_1 + u_2$ and take the open neighborhood $N = \prod_{u \in Z} I_u$ of x. Then there is an $x' \in X$ so that $x' \in N$. I.e. there is an $l \in B^*$ so that $T(l) = (l(u))_{u \in Z} \in N$ or equivalently $l(u_1) \in I_{u_1}$, $l(u_2) \in I_{u_3}$, $l(u_1 + u_2) \in I_{u_1+u_2}$. This means that

$$|x_{u_1} - l(u_1)| < \epsilon, \quad |x_{u_2} - l(u_2)| < \epsilon, \quad |x_{u_1+u_2} - l(u_1 + u_2)| < \epsilon.$$

Since $l(u_1)+l(u_2) = l(u_1+u_2)$, we get $|x_{u_1}+x_{u_2}-x_{u_1+u_2}| < 3\epsilon$ and, hence, $x_{u_1}+x_{u_2} = x_{u_1+u_2}$.
Similarly, we can prove that $\kappa x_u = x_{\kappa u}$ for all $u \in Z$, $\kappa \in \mathbb{R}$.
Now, if we define $l : Z \to \mathbb{R}$ by

$$l(u) = x_u, \qquad u \in Z,$$

then l is a linear functional on Z. Moreover, for every $u \in Z$ we have that $|l(u)| = |x_u| \leq \|u\|$ and hence l is a bounded linear functional on Z with $\|l\|_* \leq 1$. I.e. $l \in B^*$. Therefore,

$$x = (x_u)_{u \in Z} = (l(u))_{u \in Z} = T(l)$$

with $l \in B^*$ and, hence, $x \in X$. This implies that X is closed under the product topology. □

7.2 The spaces $B(X)$, $C(X)$, $BC(X)$ and $C_0(X)$.

Definition. *Let X be non-empty and $B(X)$ be the space of all bounded functions $f : X \to F$.*

If there is no danger of confusion we shall use the notation B for $B(X)$.

The sum of two bounded functions and the product of a bounded function with a number are bounded functions. Therefore, B is a linear space over F.

Definition. *We define*

$$\|f\|_u = \sup_{x \in X} |f(x)|$$

for every $f \in B$.

It is easy to see that $\|\cdot\|_u$ is a norm on B. In fact, $\|0\|_u = \sup_{x \in X} 0 = 0$ and, if $\|f\|_u = 0$, then $\sup_{x \in X} |f(x)| = 0$ and, hence, $f(x) = 0$ for all $x \in X$.
Moreover, $\|\kappa f\|_u = \sup_{x \in X} |\kappa f(x)| = |\kappa| \sup_{u \in X} |f(x)| = |\kappa| \|f\|_u$ and, finally, $|f(x)+g(x)| \leq |f(x)| + |g(x)| \leq \|f\|_u + \|g\|_u$ for all $x \in X$ and, hence, $\|f + g\|_u \leq \|f\|_u + \|g\|_u$.

We call $\|\cdot\|_u$ the **uniform norm** on B.

If $F = \mathbb{R}$, then, besides the uniform norm, B is equipped with the natural order defined by: $f \leq g$ if $f(x) \leq g(x)$ for all $x \in X$. Thus, B is a normed lattice, since it is clear that $|f| \leq |g|$ implies $\|f\|_u \leq \|g\|_u$ for all $f, g \in B$.

Theorem 7.7. *B is a Banach space. Hence, if $F = \mathbb{R}$, then B is a Banach lattice.*

Proof. Let (f_n) be a Cauchy sequence in B. Then for any $x \in X$ we have $|f_n(x) - f_m(x)| \leq \|f_n - f_m\|_u \to 0$ as $m, n \to +\infty$. This means that $(f_n(x))$ is a Cauchy sequence in F and, therefore, it converges. We denote $f(x) = \lim_{n\to+\infty} f_n(x)$ and in this way a function $f : X \to F$ is defined.
For $\epsilon = 1$ there is some N so that $\|f_n - f_m\|_u \leq 1$ for all $n, m \geq N$. In particular, $\|f_n - f_N\|_u \leq 1$ for all $n \geq N$ which implies that $|f_n(x) - f_N(x)| \leq 1$ for all $x \in X$ and $n \geq N$. Letting $n \to +\infty$, we find $|f(x) - f_N(x)| \leq 1$ and, hence, $|f(x)| \leq |f_N(x)| + 1 \leq \|f_N\|_u + 1 < +\infty$ for all $x \in X$. Therefore, $f \in B$.
Now for any $\epsilon > 0$ there is some N so that $\|f_n - f_m\|_u \leq \epsilon$ for all $n, m \geq N$. This implies $|f_n(x) - f_m(x)| \leq \epsilon$ for all $x \in X$ and $n, m \geq N$. Letting $m \to +\infty$, we find $|f_n(x) - f(x)| \leq \epsilon$ for all $x \in X$ and $n \geq N$. Thus $\|f_n - f\|_u \leq \epsilon$ for all $n \geq N$ and (f_n) converges to f in B. $\square$

From now on we shall assume that X is a topological space. This is natural, since our main objects of consideration will be continuous functions and Borel measures on X.

Definition. *The space $C(X)$ consists of all continuous functions $f : X \to F$.*

We write C instead of $C(X)$ if there is no danger of confusion.

Since the sum of two continuous functions and the product of a continuous function with a number are continuous functions, the space C is a linear space over F.

Definition. $BC(X) = B(X) \cap C(X)$.

We may write BC for $BC(X)$.

BC is also a linear space and, as a subspace of B, we may (and do) use as norm the restriction of $\|\cdot\|_u$ on it. In other words, we write $\|f\|_u = \sup_{x\in X} |f(x)|$ for every $f \in BC$.

Exactly as in the case of B, if $F = \mathbb{R}$, then BC is a normed lattice.

Theorem 7.8. *BC is a Banach space. Hence, if $F = \mathbb{R}$, then BC is a Banach lattice.*

Proof. In view of Theorem 7.12, it is enough to prove that BC is a closed subset of B.
Let (f_n) in BC converge to some f in B. Take any $x \in X$ and any $\epsilon > 0$. Then there is some N so that $\|f_n - f\|_u \leq \frac{\epsilon}{3}$ for all $n \geq N$ and, in particular, $\|f_N - f\|_u \leq \frac{\epsilon}{3}$. By continuity of f_N there is some open neighborhood U of x so that $|f_N(y) - f_N(x)| \leq \frac{\epsilon}{3}$ for all $y \in U$. Now for all $y \in U$ we have $|f(y) - f(x)| \leq |f(y) - f_N(y)| + |f_N(y) - f_N(x)| + |f_N(x) - f(x)| \leq \|f - f_N\|_u + \frac{\epsilon}{3} + \|f_N - f\|_u \leq \epsilon$. Therefore, f is continuous at x and, since x is arbitrary, f is continuous on X. Thus $f \in BC$. $\square$

We know that, if X is compact, then every continuous function $f : X \to F$ is also bounded on X. Therefore, if X is compact, then $C = BC$.

Definition. *Let $f \in C(X)$. We say that f **vanishes at infinity** if for every $\epsilon > 0$ there is a compact $K \subseteq X$ so that $|f| < \epsilon$ outside K. We define*

$$C_0(X) = \{f \in C(X) \mid f \text{ vanishes at infinity}\}.$$

Again, we may simplify to C_0.

It is clear that $C_0 \subseteq BC$ and, in fact, that C_0 is a linear subspace of BC. We also take the restriction on C_0 of the uniform norm on BC, that is $\|f\|_u = \sup_{x\in X} |f(x)|$ for all $f \in C_0$.

As in the cases of the spaces B and BC, if $F = \mathbb{R}$, then the space C_0 is a normed lattice. If X is compact, then $C_0 = C = BC$.

Theorem 7.9. *C_0 is a Banach space. Hence, if $F = \mathbb{R}$, then C_0 is a Banach lattice.*

Proof. Exercise. $\square$

Lemma 7.3. *The series $1 - \sum_{n=1}^{+\infty} \frac{1\cdot 3\cdots(2n-3)}{2^n n!}(1 - t^2)^n$ converges to $|t|$ uniformly on $[-1, 1]$.*

Proof. Taylor's theorem implies that $\sqrt{1-x} = 1 - \sum_{n=1}^{+\infty} \frac{1\cdot 3\cdots(2n-3)}{2^n n!} x^n$ when $0 \leq x < 1$ and hence we have that $\sum_{n=1}^{+\infty} \frac{1\cdot 3\cdots(2n-3)}{2^n n!} x^n < 1$ when $0 \leq x < 1$. Since every summand of the series is non-negative, we may let $x \to 1-$ and we deduce that $\sum_{n=1}^{+\infty} \frac{1\cdot 3\cdots(2n-3)}{2^n n!} \leq 1$. Therefore, the series $1 - \sum_{n=1}^{+\infty} \frac{1\cdot 3\cdots(2n-3)}{2^n n!} x^n$ converges to some function uniformly on $[0,1]$. The limiting function is continuous on $[0,1]$ and hence $\sqrt{1-x} = 1 - \sum_{n=1}^{+\infty} \frac{1\cdot 3\cdots(2n-3)}{2^n n!} x^n$ uniformly on $[0,1]$. It just remains to set $x = 1 - t^2$ with $t \in [-1,1]$. $\square$

Kakutani-Krein Theorem. *Let $F = \mathbb{R}$ and X be compact. Let Z be a linear subspace of $C(X) = BC(X)$ with the following properties:*
(i) the constant function 1 belongs to Z,
(ii) $f \vee g \in Z$ for all $f, g \in Z$, i.e. Z is a sublattice of $C(X)$,
(ii) for every $a, b \in X$ with $a \neq b$ there is $f \in Z$ so that $f(a) \neq f(b)$.
Then $\mathrm{cl}(Z) = C(X)$.

Proof. Take an arbitrary $f \in C(X)$ and $a, b \in X$ with $a \neq b$. Then there is $h \in Z$ so that $h(a) \neq h(b)$. It is clear that there are $\kappa_1, \kappa_2 \in \mathbb{R}$ such that the function $g_{a,b} = \kappa_1 h + \kappa_2 \in Z$ satisfies $g_{a,b}(a) = f(a)$ and $g_{a,b}(b) = f(b)$. Then there is an open neighborhood U_b of b so that $|g_{a,b}(x) - f(x)| \leq |g_{a,b}(x) - g_{a,b}(b)| + |f(b) - f(x)| < \epsilon$ for all $x \in U_b$. By compactness, there are $b_1, \ldots, b_n \in X$ so that $X = U_{b_1} \cup \cdots \cup U_{b_n}$. The function $g_a = g_{a,b_1} \vee \cdots \vee g_{a,b_n}$ belongs to Z and $g_a(a) = f(a)$ and $g_a(x) > f(x) - \epsilon$ for all $x \in X$.
Now there is an open neighborhood V_a of a so that $|g_a(x) - f(x)| \leq |g_a(x) - g_a(a)| + |f(a) - f(x)| < \epsilon$ for all $x \in V_a$. By compactness, there are $a_1, \ldots, a_m \in X$ so that $X = V_{a_1} \cup \cdots \cup V_{a_m}$. The function $g = g_{a_1} \wedge \cdots \wedge g_{a_m}$ belongs to Z and $f(x) - \epsilon < g(x) < f(x) + \epsilon$ for all $x \in X$. Thus we can approximate f uniformly by elements of Z and hence $f \in \mathrm{cl}(X)$. $\square$

Stone-Weierstrass Theorem. *Let X be compact and Z be a linear subspace of $C(X) = BC(X)$ with the following properties:*
(i) $fg \in Z$ for every $f, g \in Z$,
(ii) the constant function 1 belongs to Z,
(iii) $\overline{f} \in Z$ for every $f \in Z$,
(iv) for every $a, b \in X$ with $a \neq b$ there is $f \in Z$ so that $f(a) \neq f(b)$.
Then $\mathrm{cl}(Z) = C(X)$.

Proof. Let $C_{\mathbb{R}}(X) \subseteq C(X)$ be the space of all real valued functions in $C(X)$ and $Z_{\mathbb{R}} \subseteq Z$ be the space of all real valued functions in Z. Then $Z_{\mathbb{R}}$ is a linear subspace of $C_{\mathbb{R}}(X)$ and has the properties (i) and (ii). If $a, b \in X$ and $a \neq b$, there is $f \in Z$ so that $f(a) \neq f(b)$. Then $\mathrm{Re}(f) = \frac{1}{2}(f + \overline{f})$ and $\mathrm{Im}(f) = \frac{1}{2i}(f - \overline{f})$ belong to $Z_{\mathbb{R}}$ and either $\mathrm{Re}(f)(a) \neq \mathrm{Re}(f)(b)$ or $\mathrm{Im}(f)(a) \neq \mathrm{Im}(f)(b)$. Hence $Z_{\mathbb{R}}$ has also the property (iv). It is easy to see that the linear space $\mathrm{cl}(Z_{\mathbb{R}})$ also has the properties (i),(ii) and (iv).
Now take $f \in \mathrm{cl}(Z_{\mathbb{R}})$ and $\epsilon > 0$ and consider $K > 0$ such that $-K \leq f(x) \leq K$ for all $x \in X$. Lemma 7.3 implies that there is a real valued polynomial $P(t)$ such that $\big||t| - P(t)\big| \leq \frac{\epsilon}{K}$ for all $t \in [-1,1]$. Then $\big||\frac{f(x)}{K}| - P(\frac{f(x)}{K})\big| \leq \frac{\epsilon}{K}$ and hence $\big||f(x)| - KP(\frac{f(x)}{K})\big| \leq \epsilon$ for all $x \in X$. The function $KP(\frac{f}{K})$ is of the form $\kappa_0 + \kappa_1 f + \cdots + \kappa_n f^n$ and hence it belongs to $\mathrm{cl}(Z_{\mathbb{R}})$. Therefore, $|f|$ is approximated by elements of $\mathrm{cl}(Z_{\mathbb{R}})$ and thus $|f| \in \mathrm{cl}(Z_{\mathbb{R}})$. This implies that $f \vee g = \frac{f+g+|f-g|}{2} \in \mathrm{cl}(Z_{\mathbb{R}})$ for all $f, g \in \mathrm{cl}(Z_{\mathbb{R}})$.
We see that the linear subspace $\mathrm{cl}(Z_{\mathbb{R}})$ of $C_{\mathbb{R}}(X)$ satisfies all the hypotheses of the Kakutani-Krein Theorem and we conclude that $\mathrm{cl}(Z_{\mathbb{R}}) = C_{\mathbb{R}}(X)$.
Now if $f \in C(X)$, then $\mathrm{Re}(f), \mathrm{Im}(f) \in C_{\mathbb{R}}(X)$, hence $\mathrm{Re}(f), \mathrm{Im}(f) \in \mathrm{cl}(Z_{\mathbb{R}})$ and we finally get that $f \in \mathrm{cl}(Z)$. $\square$

The next two results are well-known implications of the Stone-Weierstrass Theorem.

Weierstrass Theorem. *If $X \subseteq \mathbb{R}^n$ is compact, then for every continuous $f : X \to F$ and every $\epsilon > 0$ there is a polynomial $P(x_1, \dots, x_n)$ with coefficients from F such that $|f(x) - P(x)| \le \epsilon$ for all $x \in X$.*

Proof. Let $Z \subseteq C(X)$ be the linear space of all polynomials P with coefficients from F. Then Z satisfies all hypotheses of the Stone-Weierstrass Theorem. $\square$

Definition. *Functions which are finite linear combinations with coefficients from F of functions of the form $e^{i2\pi k\cdot x} = e^{i2\pi(k_1x_1+\cdots+k_nx_n)}$, where $k = (k_1, \dots, k_n) \in \mathbb{Z}^n$, are called **exponential polynomials** on $\mathbb{R}^n$.*

Theorem 7.10. *For every continuous $f : \mathbb{R}^n \to F$ which is 1-periodic with respect to every coordinate of $x = (x_1, \dots, x_n)$ and for every $\epsilon > 0$ there is an exponential polynomial P such that $|f(x) - P(x)| \le \epsilon$ for all $x \in \mathbb{R}^n$.*

Proof. Let $\mathbb{T}$ be the unit circle centered at 0 in $\mathbb{R}^2$ and consider the compact $\mathbb{T}^n \subseteq \mathbb{R}^{2n}$.
Every continuous $f : \mathbb{R}^n \to F$ which is 1-periodic with respect to every coordinate defines $\widetilde{f} : \mathbb{T}^n \to F$ through $\widetilde{f}(y_1, \dots, y_n) = f(x_1, \dots, x_n)$, where $y_k = e^{i2\pi x_k}$ for $1 \le k \le n$. Due to the 1-periodicity of f, the function $\widetilde{f}$ is well defined. It is easy to show that $\widetilde{f}$ is also continuous on $\mathbb{T}^n$. Indeed, if $(y_1, \dots, y_n), (y_1', \dots, y_n')$ are close to each other, then the corresponding $(x_1, \dots, x_n), (x_1', \dots, x_n')$ can be chosen so that they are also close to each other and hence $f(x_1, \dots, x_n), f(x_1', \dots, x_n')$ are close to each other.
Conversely, every continuous $\widetilde{f} : \mathbb{T}^n \to F$ defines $f : \mathbb{R}^n \to F$ through $f(x_1, \dots, x_n) = \widetilde{f}(e^{i2\pi x_1}, \dots, e^{i2\pi x_n})$. This f is 1-periodic in every coordinate and continuous on $\mathbb{R}^n$.
Now take any continuous $f : \mathbb{R}^n \to F$ which is 1-periodic with respect to every coordinate and consider the corresponding $\widetilde{f} : \mathbb{T}^n \to F$ which is continuous on the compact $\mathbb{T}^n \subseteq \mathbb{R}^{2n}$. The Weierstrass Theorem implies that there is a polynomial $\widetilde{P}$ such that $|\widetilde{f}(y) - \widetilde{P}(y)| \le \epsilon$ for all $y \in \mathbb{T}^n$. Then the $P : \mathbb{R}^n \to F$ which corresponds to $\widetilde{P}$ is an exponential polynomial such that $|f(x) - P(x)| \le \epsilon$ for all $x \in \mathbb{R}^n$. $\square$

7.3 The spaces $L^p(X, \mathcal{S}, \mu)$ and their duals.

In this section $(X, \mathcal{S}, \mu)$ will be a fixed measure space.

Definition. *If $0 < p < +\infty$, we define the space $\mathcal{L}^p(X, \mathcal{S}, \mu)$ to be the set of all measurable functions $f : X \to \overline{F}$, $F = \mathbb{R}$ or $F = \mathbb{C}$, with*

$$\int_X |f|^p \, d\mu < +\infty.$$

Thus, $\mathcal{L}^1(X, \mathcal{S}, \mu)$ is the set of all functions which are integrable over X with respect to μ.

Whenever any of X, $\mathcal{S}$, μ is uniquely determined by the context of discussion, we may omit it from the symbol of the space. Therefore, we may simply write $\mathcal{L}^p$ or $\mathcal{L}^p(X)$ or $\mathcal{L}^p(\mu)$ etc.

Proposition 7.37. *$\mathcal{L}^p$ is a linear space over F.*

Proof. We shall use the trivial inequality $(a + b)^p \le 2^p(a^p + b^p)$ for all $a, b \ge 0$. This can be proved by $(a + b)^p \le (2\max\{a, b\})^p = 2^p \max\{a^p, b^p\} \le 2^p(a^p + b^p)$.
Now, if $f_1, f_2 \in \mathcal{L}^p$, then $\int_X |f_1 + f_2|^p \, d\mu \le 2^p \int_X |f_1|^p \, d\mu + 2^p \int_X |f_2|^p \, d\mu < +\infty$ and, hence, $f_1 + f_2 \in \mathcal{L}^p$.
Also, if $f \in \mathcal{L}^p$ and $\kappa \in F$, then $\int_X |\kappa f|^p \, d\mu = |\kappa|^p \int_X |f|^p \, d\mu < +\infty$ and, hence, $\kappa f \in \mathcal{L}^p$. $\square$

Definition. *Let $f : X \to \overline{F}$ be measurable. We say that f is **essentially bounded** on X (with respect to μ) if there is $M < +\infty$ so that $|f| \le M$ a.e. on X.*

Proposition 7.38. *Let $f : X \to \overline{F}$ be measurable. If f is essentially bounded on X, then there is a smallest M with the property: $|f| \leq M$ a.e. on X. This smallest M_0 is characterized by:*
(i) $|f| \leq M_0$ a.e. on X,
(ii) $\mu(\{x \in X \,|\, |f(x)| > M\}) > 0$ for every $M < M_0$.

Proof. We consider the set $A = \{M \,|\, |f| \leq M$ a.e. on $X\}$ and $M_0 = \inf A$.
The set A is non-empty and is included in $[0, +\infty)$ and, hence, M_0 exists. We take any sequence (M_n) in A such that $M_n \to M_0$. From $M_n \in A$ we find $\mu(\{x \in X \,|\, |f(x)| > M_n\}) = 0$ for every n and, since $\{x \in X \,|\, |f(x)| > M_0\} = \bigcup_{n=1}^{+\infty} \{x \in X \,|\, |f(x)| > M_n\}$, we have that $\mu(\{x \in X \,|\, |f(x)| > M_0\}) = 0$. Therefore, $|f| \leq M_0$ a.e. on X.
On the other hand, if $M < M_0$, then $M \notin A$ and, hence, $\mu(\{x \in X \,|\, |f(x)| > M\}) > 0$. □

Definition. *Let $f : X \to \overline{F}$ be measurable. If f is essentially bounded, then the smallest M with the property that $|f| \leq M$ a.e. on X is called the **essential supremum** of f over X (with respect to μ) and it is denoted by* ess-sup$_{X,\mu}(f)$.

Again, we may simply write ess-sup(f) or ess-sup$_X(f)$ instead of ess-sup$_{X,\mu}(f)$.

Definition. *We define $\mathcal{L}^\infty(X, \mathcal{S}, \mu)$ to be the set of all measurable functions $f : X \to \overline{F}$ which are essentially bounded on X.*

Proposition 7.39. *$\mathcal{L}^\infty$ is a linear space over F.*

Proof. If $f_1, f_2 \in \mathcal{L}^\infty$, then there are sets $A_1, A_2 \in \mathcal{S}$ so that $\mu(A_1^c) = \mu(A_2^c) = 0$ and $|f_1| \leq$ ess-sup(f_1) on A_1 and $|f_2| \leq$ ess-sup(f_2) on A_2. If we set $A = A_1 \cap A_2$, then we have $\mu(A^c) = 0$ and $|f_1 + f_2| \leq |f_1| + |f_2| \leq$ ess-sup$(f_1) +$ ess-sup(f_2) on A. Hence $f_1 + f_2$ is essentially bounded on X and ess-sup$(f_1 + f_2) \leq$ ess-sup$(f_1) +$ ess-sup(f_2).
If $f \in \mathcal{L}^\infty$ and $\kappa \in F$, then there is $A \in \mathcal{S}$ with $\mu(A^c) = 0$ so that $|f| \leq$ ess-sup(f) on A. We now have $|\kappa f| \leq |\kappa|$ ess-sup(f) on A. Hence κf is essentially bounded on X and ess-sup$(\kappa f) \leq |\kappa|$ ess-sup(f). If $\kappa = 0$, this inequality obviously becomes equality. If $\kappa \neq 0$, we apply the same inequality to $\frac{1}{\kappa}$ and κf and get ess-sup$(f) =$ ess-sup$(\frac{1}{\kappa}(\kappa f)) \leq \frac{1}{|\kappa|}$ ess-sup(κf). Therefore, ess-sup$(\kappa f) = |\kappa|$ ess-sup(f). □

Definition. *Let $1 \leq p \leq +\infty$. We define $p' = \frac{p}{p-1}$, if $1 < p < +\infty$, $p' = +\infty$, if $p = 1$, and $p' = 1$, if $p = +\infty$. We say that p' is the **conjugate** of p or the **dual** of p.*

The definition in the cases $p = 1$ and $p = +\infty$ is justified by $\lim_{p \to 1+} \frac{p}{p-1} = +\infty$ and by $\lim_{p \to +\infty} \frac{p}{p-1} = 1$.
It is easy to see that, if p' is the conjugate of p, then $1 \leq p' \leq +\infty$ and p is the conjugate of p'. Moreover, p, p' are related by the symmetric equality

$$\frac{1}{p} + \frac{1}{p'} = 1.$$

Lemma 7.4. *Let $0 < t < 1$ and $a, b \geq 0$. Then $a^t b^{1-t} \leq ta + (1-t)b$.*

Proof. A simple Calculus exercise. □

Hölder's Inequalities. *Let $1 \leq p, p' \leq +\infty$ and p, p' be conjugate to each other. If $f \in \mathcal{L}^p$ and $g \in \mathcal{L}^{p'}$, then $fg \in \mathcal{L}^1$ and*

$$\int_X |fg|\, d\mu \leq \begin{cases} (\int_X |f|^p\, d\mu)^{1/p} (\int_X |g|^{p'}\, d\mu)^{1/p'}, & \text{if } 1 < p, p' < +\infty \\ \int_X |f|\, d\mu \ \text{ess-sup}(g), & \text{if } p = 1, p' = +\infty \\ \text{ess-sup}(f) \ \int_X |g|\, d\mu, & \text{if } p = +\infty, p' = 1 \end{cases}$$

Proof. We start with the case $1 < p, p' < +\infty$.
If $\int_X |f|^p \, d\mu = 0$ or if $\int_X |g|^{p'} \, d\mu = 0$, then either $f = 0$ a.e. on X or $g = 0$ a.e. on X and the inequality is trivially true in the form of equality: $0 = 0$.
So we assume that $A = \int_X |f|^p \, d\mu > 0$ and $B = \int_X |g|^{p'} \, d\mu > 0$. Applying Lemma 7.3 with $t = \frac{1}{p}, 1 - t = 1 - \frac{1}{p} = \frac{1}{p'}$ and $a = \frac{|f(x)|^p}{A}, b = \frac{|g(x)|^{p'}}{B}$, we get $\frac{|fg|}{A^{1/p}B^{1/p'}} \leq \frac{1}{p}\frac{|f|^p}{A} + \frac{1}{p'}\frac{|g|^{p'}}{B}$ a.e. on X. Integrating, we find $\frac{1}{A^{1/p}B^{1/p'}} \int_X |fg| \, d\mu \leq \frac{1}{p} + \frac{1}{p'} = 1$ and this implies the inequality we wanted to prove.
Now let $p = 1$, $p' = +\infty$. Since $|g| \leq \text{ess-sup}(g)$ a.e. on X, we have that $|fg| \leq |f| \, \text{ess-sup}(g)$ a.e. on X. Integrating, we find the inequality we want to prove.
The proof in the case $p = +\infty, p' = 1$ is the same as in (b). □

Minkowski's inequalities. *Let* $1 \leq p \leq +\infty$. *If* $f_1, f_2 \in \mathcal{L}^p$, *then*

$$\begin{array}{ll} (\int_X |f_1 + f_2|^p \, d\mu)^{1/p} \leq (\int_X |f_1|^p \, d\mu)^{1/p} + (\int_X |f_2|^p \, d\mu)^{1/p}, & \textit{if } 1 \leq p < +\infty \\ \text{ess-sup}(f_1 + f_2) \leq \text{ess-sup}(f_1) + \text{ess-sup}(f_2), & \textit{if } p = +\infty \end{array}$$

Proof. The case $p = +\infty$ is included in the proof of Proposition 7.39. Also the case $p = 1$ is trivial and the result is already known. Hence, we assume $1 < p < +\infty$.
We write $|f_1 + f_2|^p \leq (|f_1| + |f_2|)|f_1 + f_2|^{p-1} = |f_1||f_1 + f_2|^{p-1} + |f_2||f_1 + f_2|^{p-1}$ a.e. on X and, applying Hölder's inequality, we find

$$\begin{aligned} \int_X |f_1 + f_2|^p \, d\mu &\leq \big(\int_X |f_1|^p \, d\mu \big)^{1/p} \big(\int_X |f_1 + f_2|^{(p-1)p'} \, d\mu \big)^{1/p'} \\ &\quad + \big(\int_X |f_2|^p \, d\mu \big)^{1/p} \big(\int_X |f_1 + f_2|^{(p-1)p'} \, d\mu \big)^{1/p'} \\ &= \big(\int_X |f_1|^p \, d\mu \big)^{1/p} \big(\int_X |f_1 + f_2|^p \, d\mu \big)^{1/p'} \\ &\quad + \big(\int_X |f_2|^p \, d\mu \big)^{1/p} \big(\int_X |f_1 + f_2|^p \, d\mu \big)^{1/p'}. \end{aligned}$$

Simplifying, we get the inequality we want to prove. □

Definition. *Let* $1 \leq p \leq +\infty$ *and* (f_n) *be a sequence in* $\mathcal{L}^p$ *and* $f \in \mathcal{L}^p$. *We say that* $f_n \to f$ ***in the*** p***-mean*** *if as* $n \to +\infty$

$$\begin{array}{ll} \int_X |f_n - f|^p \, d\mu \to 0, & \textit{if } 1 \leq p < +\infty \\ \text{ess-sup}(f_n - f) \to 0, & \textit{if } p = +\infty \end{array}$$

We say that (f_n) *is* ***Cauchy in the*** p***-mean*** *if as* $n, m \to +\infty$

$$\begin{array}{ll} \int_X |f_n - f_m|^p \, d\mu \to 0, & \textit{if } 1 \leq p < +\infty \\ \text{ess-sup}(f_n - f_m) \to 0, & \textit{if } p = +\infty \end{array}$$

It is easy to see that, if (f_n) converges to f in the p-mean, then (f_n) is Cauchy in the p-mean. Indeed, if $1 \leq p < +\infty$, then, by Minkowski's inequalities,

$$\big(\int_X |f_n - f_m|^p \, d\mu \big)^{1/p} \leq \big(\int_X |f_n - f|^p \, d\mu \big)^{1/p} + \big(\int_X |f_m - f|^p \, d\mu \big)^{1/p} \to 0$$

as $m, n \to +\infty$. The proof is identical if $p = +\infty$.

The notion of convergence in the 1-mean coincides with the notion of convergence in the mean on X. Theorem 7.7 is an extension of Theorem 5.1.

Theorem 7.11. *If* (f_n) *is Cauchy in the* p*-mean, then there is* $f \in \mathcal{L}^p$ *so that* $f_n \to f$ *in the* p*-mean. Moreover, there is a subsequence* (f_{n_k}) *so that* $f_{n_k} \to f$ *a.e. on* X.
As a corollary: if $f_n \to f$ *in the* p*-mean, there is a subsequence* (f_{n_k}) *so that* $f_{n_k} \to f$ *a.e. on* X.

Proof. We consider first the case $1 \le p < +\infty$.
First proof. We have that for every k there is n_k so that $\int_X |f_n - f_m|^p \, d\mu < \frac{1}{2^{kp}}$ for every $n, m \ge n_k$. Since we may assume that each n_k is as large as we like, we inductively take (n_k) so that $n_k < n_{k+1}$ for every k. Therefore, (f_{n_k}) is a subsequence of (f_n).
From the construction of n_k and from $n_k < n_{k+1}$, we get that $\int_X |f_{n_{k+1}} - f_{n_k}|^p \, d\mu < \frac{1}{2^{kp}}$ for every k. We define the measurable function $G : X \to [0, +\infty]$ by $G = \sum_{k=1}^{+\infty} |f_{n_{k+1}} - f_{n_k}|$. If $G_K = \sum_{k=1}^{K-1} |f_{n_{k+1}} - f_{n_k}|$ then

$$\big(\int_X G_K^p \, d\mu\big)^{1/p} \le \sum_{k=1}^{K-1} \big(\int_X |f_{n_{k+1}} - f_{n_k}|^p \, d\mu\big)^{1/p} < 1,$$

by Minkowski's inequality. Since $G_K \uparrow G$ on X, we find $\int_X G^p \, d\mu \le 1$ and, thus, $G < +\infty$ a.e. on X. This implies that the series $\sum_{k=1}^{+\infty} (f_{n_{k+1}}(x) - f_{n_k}(x))$ converges for a.e. $x \in X$. Therefore, there is a $B \in \mathcal{S}$ so that $\mu(B^c) = 0$ and $\sum_{k=1}^{+\infty} (f_{n_{k+1}}(x) - f_{n_k}(x))$ converges for every $x \in B$. We define the measurable $f : X \to F$ by

$$f = \begin{cases} f_{n_1} + \sum_{k=1}^{+\infty} (f_{n_{k+1}} - f_{n_k}), & \text{on } B \\ 0, & \text{on } B^c \end{cases}$$

On B we have that $f = f_{n_1} + \lim_{K \to +\infty} \sum_{k=1}^{K-1} (f_{n_{k+1}} - f_{n_k}) = \lim_{K \to +\infty} f_{n_K}$ and, hence, $f_{n_k} \to f$ a.e. on X. We also have on B that

$$\begin{aligned} |f_{n_K} - f| &= |f_{n_K} - f_{n_1} - \textstyle\sum_{k=1}^{+\infty} (f_{n_{k+1}} - f_{n_k})| \\ &= |\textstyle\sum_{k=1}^{K-1} (f_{n_{k+1}} - f_{n_k}) - \sum_{k=1}^{+\infty} (f_{n_{k+1}} - f_{n_k})| \le \sum_{k=K}^{+\infty} |f_{n_{k+1}} - f_{n_k}| \le G \end{aligned}$$

for every K and, hence, $|f_{n_K} - f|^p \le G^p$ a.e. on X for every K. Since we have $\int_X G^p \, d\mu < +\infty$ and that $|f_{n_K} - f| \to 0$ a.e. on X, we apply the Dominated Convergence Theorem and we find that $\int_X |f_{n_K} - f|^p \, d\mu \to 0$ as $K \to +\infty$.
From $n_k \to +\infty$ as $k \to +\infty$ and from Minkowski's inequality, we get $(\int_X |f_k - f|^p \, d\mu)^{1/p} \le (\int_X |f_k - f_{n_k}|^p \, d\mu)^{1/p} + (\int_X |f_{n_k} - f|^p \, d\mu)^{1/p} \to 0$ as $k \to +\infty$ and we conclude that $f_n \to f$ in the p-mean.
Second proof. For every $\epsilon > 0$ we have that

$$\mu(\{x \in X \,|\, |f_n(x) - f_m(x)| \ge \epsilon\}) \le \tfrac{1}{\epsilon} \big(\int_X |f_n - f_m|^p \, d\mu \big)^{1/p}$$

and, hence, (f_n) is Cauchy in measure on X. Theorem 5.2 implies that there is a subsequence (f_{n_k}) so that $f_{n_k} \to f$ a.e. on X.
Now, for every $\epsilon > 0$ there is an N so that $\int_X |f_n - f_m|^p \, d\mu \le \epsilon$ for all $n, m \ge N$. Since $n_k \to +\infty$ as $k \to +\infty$, we use $m = n_k$ for large k and apply the Lemma of Fatou to get

$$\int_X |f_n - f|^p \, d\mu \le \underline{\lim}_{k \to +\infty} \int_X |f_n - f_{n_k}|^p \, d\mu \le \epsilon$$

for all $n \ge N$. Of course this says that $f_n \to f$ in the p-mean.
Now let $p = +\infty$.
For each n, m we have a set $A_{n,m} \in \mathcal{S}$ with $\mu(A_{n,m}^c) = 0$ and $|f_n - f_m| \le \text{ess-sup}(f_n - f_m)$ on $A_{n,m}$. We define $A = \bigcap_{1 \le n,m} A_{n,m}$ and get that $\mu(A^c) = 0$ and $|f_n - f_m| \le \text{ess-sup}(f_n - f_m)$ on A for every n, m. This says that (f_n) is Cauchy uniformly on A and, hence, there is an f so that $f_n \to f$ uniformly on A. Now, $\text{ess-sup}(f_n - f) \le \sup_{x \in A} |f_n(x) - f(x)| \to 0$ as $n \to +\infty$. $\square$

If for every $f \in \mathcal{L}^p$ we set

$$N_p(f) = \begin{cases} \big(\int_X |f|^p \, d\mu \big)^{1/p}, & \text{if } 1 \le p < +\infty \\ \text{ess-sup}(f), & \text{if } p = +\infty \end{cases}$$

then (the proofs of) Propositions 7.37 and 7.39 and Minkowski's Inequalities imply that the function $N_p : \mathcal{L}^p \to \mathbb{R}$ satisfies

$$N_p(f_1 + f_2) \le N_p(f_1) + N_p(f_2), \qquad N_p(\kappa f) = |\kappa| N_p(f)$$

for every $f, f_1, f_2 \in \mathcal{L}^p$ and $\kappa \in F$.

The function N_p has the two properties of a norm but not the third. Indeed, $N_p(f) = 0$ if and only if $f = 0$ a.e. on X. The usual practice is to identify every two functions which are equal a.e. on X so that N_p becomes, *informally*, a norm. The precise way to do this is the following.

Definition. *We define the relation $\sim$ on $\mathcal{L}^p$ as follows: we write $f_1 \sim f_2$ if $f_1 = f_2$ a.e. on X.*

Proposition 7.40. *The relation $\sim$ on $\mathcal{L}^p$ is an equivalence relation.*

Proof. Exercise. □

Like any equivalence relation, the relation $\sim$ defines equivalence classes. The equivalence class $[f]$ of any $f \in \mathcal{L}^p$ is the set of all $g \in \mathcal{L}^p$ which are equivalent to f:

$$[f] = \{g \in \mathcal{L}^p \,|\, g \sim f\} = \{g \in \mathcal{L}^p \,|\, g = f \text{ a.e. on } X\}.$$

Proposition 7.41. *Let $f_1, f_2 \in \mathcal{L}^p$. Then*
(i) $[f_1] = [f_2]$ if and only if $f_1 \sim f_2$ if and only if $f_1 = f_2$ a.e. on X.
(ii) If $[f_1] \cap [f_2] \neq \emptyset$, then $[f_1] = [f_2]$.
Moreover, $\mathcal{L}^p = \bigcup_{f \in \mathcal{L}^p} [f]$.

Proof. Exercise. □

Proposition 7.41 says that any two different equivalence classes have empty intersection and that $\mathcal{L}^p$ is the union of all equivalence classes. In other words, the collection of all equivalence classes is a partition of $\mathcal{L}^p$.

Definition. *We define*

$$L^p(X, \mathcal{S}, \mu) = \mathcal{L}^p(X, \mathcal{S}, \mu)/_\sim = \{[f] \,|\, f \in \mathcal{L}^p(X, \mathcal{S}, \mu)\}.$$

Again, we may write L^p or $L^p(X)$ or $L^p(\mu)$ etc.

The first task is to carry addition and multiplication from $\mathcal{L}^p$ over to L^p.

Proposition 7.42. *Let $f, f_1, f_2, g, g_1, g_2 \in \mathcal{L}^p$ and $\kappa \in F$.*
(i) If $f_1 \sim g_1$ and $f_2 \sim g_2$, then $f_1 + f_2 \sim g_1 + g_2$.
(ii) If $f \sim g$, then $\kappa f \sim \kappa g$.

Proof. Exercise. □

Because of Proposition 7.41, another way to state the results of Proposition 7.42 is: (i) $[f_1] = [g_1]$ and $[f_2] = [g_2]$ imply $[f_1 + g_1] = [f_2 + g_2]$ and (ii) $[f] = [g]$ implies $[\kappa f] = [\kappa g]$. These allow the following definition.

Definition. *We define addition and multiplication in L^p as follows:*

$$[f_1] + [f_2] = [f_1 + f_2], \qquad \kappa[f] = [\kappa f].$$

Now it is a matter of routine to prove that the set L^p becomes a linear space under this addition and multiplication. The zero element of L^p is the equivalence class $[0]$ of the function 0 which is identically 0 on X. The opposite of $[f]$ is the equivalence class $[-f]$.

The next task is to define a norm on L^p.

Proposition 7.43. *Let $f_1, f_2 \in \mathcal{L}^p$. If $f_1 \sim f_2$, then $N_p(f_1) = N_p(f_2)$ or, equivalently,*

$$\begin{aligned} &\textstyle\int_X |f_1|^p \, d\mu = \int_X |f_2|^p \, d\mu, \quad \text{if } 1 \le p < +\infty \\ &\text{ess-sup}(f_1) = \text{ess-sup}(f_2), \quad \text{if } p = +\infty \end{aligned}$$

Proof. Exercise. □

An equivalent way to state the result of Proposition 7.43: (i) $[f_1] = [f_2]$ implies $\int_X |f_1|^p \, d\mu = \int_X |f_2|^p \, d\mu$, if $1 \le p < +\infty$, and (ii) $[f_1] = [f_2]$ implies ess-sup$(f_1) =$ ess-sup(f_2), if $p = +\infty$. These allow the:

Definition. *We define for every $[f] \in L^p$*

$$\|[f]\|_p = N_p(f) = \begin{cases} \big(\int_X |f|^p \, d\mu \big)^{1/p}, & \text{if } 1 \le p < +\infty \\ \text{ess-sup}(f), & \text{if } p = +\infty \end{cases}$$

Proposition 7.44. *The function $\| \cdot \|_p$ is a norm on L^p.*

Proof. $\|[f_1] + [f_2]\|_p = \|[f_1 + f_2]\|_p = N_p(f_1 + f_2) \le N_p(f_1) + N_p(f_2) = \|[f_1]\|_p + \|[f_2]\|_p$.
Also, $\|\kappa[f]\|_p = \|[\kappa f]\|_p = N_p(\kappa f) = |\kappa| N_p(f) = |\kappa| \|[f]\|_p$.
If $\|[f]\|_p = 0$, then $N_p(f) = 0$. This implies $f = 0$ a.e. on X and, hence, $f \sim 0$ or, equivalently, $[f]$ is the zero element of L^p. □

In order to simplify things and not have to use the bracket-notation $[f]$ for the elements of L^p, we shall follow the traditional practice and write f instead of $[f]$. When we do this we must have in mind that the element f of L^p (and *not* the element f of $\mathcal{L}^p$) is not the single function f, but the whole collection of functions each of which is equal to f a.e. on X.

For example:
1. When we write $f_1 = f_2$ for the elements f_1, f_2 of L^p, we mean the more correct $[f_1] = [f_2]$ or, equivalently, that $f_1 = f_2$ a.e. on X.
2. When we write $\int_X fh \, d\mu$ for the element $f \in L^p$, we mean the integral $\int_X fh \, d\mu$ for the element-function $f \in \mathcal{L}^p$ *and*, at the same time, all integrals $\int_X gh \, d\mu$ (equal to each other) for all functions $g \in \mathcal{L}^p$ such that $g = f$ a.e. on X.
3. When we write $\|f\|_p$ for the element $f \in L^p$ we mean the more correct $\|[f]\|_p$ or, equivalently, the expression $\big(\int_X |f|^p \, d\mu \big)^{1/p}$, when $1 \le p < +\infty$, and ess-sup(f), when $p = +\infty$, for the element-function $f \in \mathcal{L}^p$ and at the same time all similar expressions (equal to each other) for all functions $g \in \mathcal{L}^p$ such that $g = f$ a.e. on X.

The inequality of Minkowski takes the form

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

for every $f_1, f_2 \in L^p$.

Hölder's inequality takes the form

$$\|fg\|_1 \le \|f\|_p \|g\|_{p'}$$

for every $f \in L^p$ and $g \in L^{p'}$.

Definition. *We define $\langle \cdot, \cdot \rangle : L^2 \times L^2 \to F$ by*

$$\langle f, g \rangle = \textstyle\int_X f \overline{g} \, d\mu, \qquad f, g \in L^2.$$

Proposition 7.45. *The function $\langle \cdot, \cdot \rangle$ is well-defined and it is an inner product on L^2.*

Proof. If $f_1, f_2, g_1, g_2 \in \mathcal{L}^2$ so that $f_1 \sim f_2$ and $g_1 \sim g_2$, then $f_1 \overline{g_1} \sim f_2 \overline{g_2}$ and thus $\int_X f_1 \overline{g_1} \, d\mu = \int_X f_2 \overline{g_2} \, d\mu$. Therefore, $\langle f, g \rangle$ is well defined for any $f, g \in L^2$.
All properties of an inner product are very easy to verify. □

Obviously
$$\langle f, f\rangle = \int_X |f|^2\, d\mu = \|f\|_2^2.$$
This means that the norm induced by the inner product on L^2 is the same as the already defined norm on L^2.

Theorem 7.12. *All L^p are Banach spaces. In particular, L^2 is a Hilbert space.*

Proof. Let (f_n) be a Cauchy sequence in L^p. This means $\|f_n - f_m\|_p \to 0$, which says that $\int_X |f_n - f_m|^p\, d\mu \to 0$, if $1 \le p < +\infty$, and ess-sup$(f_n - f_m) \to 0$, if $p = +\infty$. Now, Theorem 7.7 implies that the sequence (f_n) in $\mathcal{L}^p$ converges to some $f \in \mathcal{L}^p$ in the p-mean. Therefore, $\int_X |f_n - f|^p\, d\mu \to 0$, if $1 \le p < +\infty$, and ess-sup$(f_n - f) \to 0$, if $p = +\infty$. This means that $\|f_n - f\|_p \to 0$ and (f_n) converges to the element f of L^p. $\square$

Definition. *Let I be an index set and $\sharp$ be the counting measure on $(I, \mathcal{P}(I))$. We denote*
$$l^p(I) = L^p(I, \mathcal{P}(I), \sharp).$$
In particular, if $I = \mathbb{N}$, we denote $l^p = l^p(\mathbb{N})$.

If $1 \le p < +\infty$, then the function $b = (b_i)_{i\in I} : I \to \overline{F}$ belongs to $l^p(I)$ if, by definition, $\int_I |b|^p\, d\sharp < +\infty$ or, equivalently,
$$\textstyle\sum_{i\in I} |b_i|^p < +\infty.$$
If $|b_i| = +\infty$ for at least one $i \in I$, then $\sum_{i\in I} |b_i|^p = +\infty$.

Definition. *If $1 \le p < +\infty$, we say that $b = (b_i)_{i\in I}$ is p-**summable** when $\sum_{i\in I} |b_i|^p < +\infty$.*

Hence, $b = (b_i)_{i\in I}$ is p-summable if and only if it belongs to $l^p(I)$. We also have
$$\|b\|_p = \big(\textstyle\sum_{i\in I} |b_i|^p\big)^{1/p}.$$

When $1 \le p < +\infty$, Minkowski's inequality becomes
$$\big(\textstyle\sum_{i\in I} |b_i^{(1)} + b_i^{(2)}|^p\big)^{1/p} \le \big(\sum_{i\in I} |b_i^{(1)}|^p\big)^{1/p} + \big(\sum_{i\in I} |b_i^{(2)}|^p\big)^{1/p}$$
for all $b_1 = (b_i^{(1)})_{i\in I}$ and $b_2 = (b_i^{(2)})_{i\in I}$ which are p-summable. Similarly, when $1 < p, p' < +\infty$ and p, p' are conjugate, Hölder's inequality becomes
$$\textstyle\sum_{i\in I} |b_i c_i| \le \big(\sum_{i\in I} |b_i|^p\big)^{1/p}\big(\sum_{i\in I} |c_i|^{p'}\big)^{1/p'}$$
for all p-summable $b = (b_i)_{i\in I}$ and all p'-summable $c = (c_i)_{i\in I}$.

Since the only subset of I with zero $\sharp$-measure is the $\emptyset$, we easily see that $b = (b_i)_{i\in I}$ is essentially bounded on I with respect to $\sharp$ if and only if there is an $M < +\infty$ so that $|b_i| \le M$ for all $i \in I$. It is obvious that the smallest M with the property that $|b_i| \le M$ for all $i \in I$ is the $M_0 = \sup_{i\in I} |b_i|$.

Definition. *We say that $b = (b_i)_{i\in I}$ is **bounded** if $\sup_{i\in I} |b_i| < +\infty$.*

Therefore, b is essentially bounded on I with respect to $\sharp$ or, equivalently, $b \in l^\infty(I)$ if and only if b is bounded. Also,
$$\|b\|_\infty = \text{ess-sup}(b) = \sup_{i\in I} |b_i|.$$

The inequality of Minkowski takes the form
$$\sup_{i\in I} |b_i^{(1)} + b_i^{(2)}| \le \sup_{i\in I} |b_i^{(1)}| + \sup_{i\in I} |b_i^{(2)}|$$

for all $b_1 = (b_i^{(1)})_{i\in I}$ and $b_2 = (b_i^{(2)})_{i\in I}$ which are bounded. When $p = 1$ and $p' = +\infty$, Hölder's inequality takes the form

$$\sum_{i\in I} |b_i c_i| \leq \sum_{i\in I} |b_i| \cdot \sup_{i\in I} |c_i|$$

for all summable $b = (b_i)_{i\in I}$ and all bounded $c = (c_i)_{i\in I}$.

The spaces $l^p(I)$ are Banach spaces. In particular, the space $l^2(I)$ is a Hilbert space. The inner product on $l^2(I)$ is given by

$$\langle b, c\rangle = \sum_{i\in I} b_i\, \overline{c_i}$$

for all $b = (b_i)_{i\in I} \in l^2(I)$, $c = (c_i)_{i\in I} \in l^2(I)$.

As we have already mentioned, a particular case is when $I = \mathbb{N}$. Then

$$l^p = \big\{x = (x_1, x_2, \ldots) \,\big|\, \textstyle\sum_{k=1}^{+\infty} |x_k|^p < +\infty\big\}, \quad \text{if } 1 \leq p < +\infty$$
$$l^\infty = \big\{x = (x_1, x_2, \ldots) \,\big|\, \sup_{k\geq 1} |x_k| < +\infty\big\}$$

The corresponding norms are

$$\|x\|_p = \big(\textstyle\sum_{k=1}^{+\infty} |x_k|^p\big)^{1/p}, \quad \text{if } 1 \leq p < +\infty$$
$$\|x\|_\infty = \sup_{k\geq 1} |x_k|$$

For l^2 the inner product is

$$\langle x, y\rangle = \sum_{k=1}^{+\infty} x_k\, \overline{y_k}$$

for every $x = (x_1, x_2, \ldots) \in l^2$, $y = (y_1, y_2, \ldots) \in l^2$.

Theorem 7.13. *The set of all functions of the form* $e^{i2\pi k\cdot x} = e^{i2\pi(k_1x_1+\cdots+k_nx_n)}$*, where* $k = (k_1, \ldots, k_n) \in \mathbb{Z}^n$*, is an orthonormal basis of* $L^2([0,1]^n, m_n)$*.*

Proof. It is easy to see that the functions $e^{i2\pi k\cdot x}$, where $k \in \mathbb{Z}^n$, form an orthonormal set in $L^2([0,1]^n, m_n)$.

Now take an arbitrary $f \in L^2([0,1]^n, m_n)$ and any $\epsilon > 0$. There is a continuous $g : [0,1]^n \to F$ whose support is contained in the open cube $(0,1)^n$ so that $\|f - g\|_2^2 = \int_{[0,1]^n} |f - g|^2\, dm_n \leq \epsilon^2$. We extend g to a function $g : \mathbb{R}^n \to F$ which is 1-periodic in each coordinate. The extended g is also continuous on $\mathbb{R}^n$.

Theorem 7.10 implies that there is an exponential polynomial P such that $|g(x) - P(x)| \leq \epsilon$ for all $x \in \mathbb{R}^n$. Thus $\|g - P\|_2^2 = \int_{[0,1]^n} |g - P|^2\, dm_n \leq \epsilon^2$ and by the triangle inequality of the norm we get $\|f - P\|_2 \leq 2\epsilon$. Therefore, every $f \in L^2([0,1]^n, m_n)$ is in the closed linear span of the functions $e^{i2\pi k\cdot x}$, where $k \in \mathbb{Z}^n$, and we conclude that these functions constitute an orthonormal basis of $L^2([0,1]^n, m_n)$. $\square$

Thus, the Fourier series of any $f \in L^2([0,1]^n, m_n)$ with respect to the so-called **exponential orthonormal basis** of all functions $e_k(x) = e^{i2\pi k\cdot x}$ ($k \in \mathbb{Z}^n$) is the series $\sum_{k\in\mathbb{Z}^n} \langle f, e_k\rangle\, e_k$, i.e.

$$\sum_{k\in\mathbb{Z}^n} \int_{[0,1]^n} f(y) e^{-i2\pi k\cdot y}\, dm_n(y)\, e^{i2\pi k\cdot x}.$$

From now on $p, p' \in [1, +\infty]$ are meant to be conjugate.

Theorem 7.14. *Let* $g \in L^{p'}$*. If* $1 < p \leq +\infty$*, then*

$$\|g\|_{p'} = \max\big\{\big|\textstyle\int_X fg\, d\mu\big| \,\big|\, f \in L^p, \|f\|_p \leq 1\big\}.$$

If μ *is semifinite, the same is true when* $p = 1$ *but with* max *replaced by* sup.

Proof. (a) Let $1 < p \leq +\infty$ and, hence, $1 \leq p' < +\infty$.
For any $f \in L^p$ with $\|f\|_p \leq 1$ we get by Hölder's inequality $|\int_X fg\,d\mu| \leq \|f\|_p\|g\|_{p'} \leq \|g\|_{p'}$. Therefore, $\sup\big\{|\int_X fg\,d\mu| \,\big|\, f \in L^p, \|f\|_p \leq 1\big\} \leq \|g\|_{p'}$.
If $\|g\|_{p'} = 0$, then the inequality between the sup and the $\|g\|_{p'}$ obviously becomes equality. Indeed, we have $g = 0$ a.e. on X and this implies $\int_X fg\,d\mu = 0$ for every $f \in L^p$.
Now let $\|g\|_{p'} > 0$. We consider f_0 defined by $f_0(x) = |g(x)|^{p'-1}\overline{\text{sign}(g(x))}/\|g\|_{p'}^{p'-1}$. Then $f_0(x)g(x) = |g(x)|^{p'}/\|g\|_{p'}^{p'-1}$ and, hence, $\int_X f_0 g\,d\mu = \int_X |g|^{p'}\,d\mu/\|g\|_{p'}^{p'-1} = \|g\|_{p'}$.
If $1 < p, p' < +\infty$, then, since $p(p'-1) = p'$, we have $|f_0(x)|^p = |g(x)|^{p'}/\|g\|_{p'}^{p'}$ and, hence, $\|f_0\|_p = \big(\int_X |f_0|^p\,d\mu\big)^{1/p} = 1$.
If $p = +\infty, p' = 1$, then $|f_0(x)| = 1$ and, thus, $\|f_0\|_\infty = \text{ess-sup}(f_0) = 1$.
We conclude that $\|g\|_{p'} = \max\big\{|\int_X fg\,d\mu| \,\big|\, f \in L^p, \|f\|_p \leq 1\big\}$.
(b) Let $p = 1, p' = +\infty$.
For any $f \in L^1$ with $\|f\|_1 \leq 1$ we have $|\int_X fg\,d\mu| \leq \|f\|_1\|g\|_\infty \leq \|g\|_\infty$. Therefore, $\sup\big\{|\int_X fg\,d\mu| \,\big|\, f \in L^1, \|f\|_1 \leq 1\big\} \leq \|g\|_\infty$.
If $\|g\|_\infty = 0$, then $g = 0$ a.e. on X. This implies that $\int_X fg\,d\mu = 0$ for every $f \in L^p$ and the inequality between the sup and the $\|g\|_\infty$ becomes equality.
Let $\|g\|_\infty > 0$. For all ϵ with $0 < \epsilon < \|g\|_\infty$ we get $\mu(\{x \in X \,|\, \|g\|_\infty - \epsilon < |g(x)| \leq \|g\|_\infty\}) > 0$. If μ is semifinite, there exists a $B \in \mathcal{S}$ so that $B \subseteq \{x \in X \,|\, \|g\|_\infty - \epsilon < |g(x)| \leq \|g\|_\infty\}$ and $0 < \mu(B) < +\infty$. We define the function f_0 by $f_0(x) = \overline{\text{sign}(g(x))}\chi_B(x)/\mu(B)$. Then $f_0(x)g(x) = |g(x)|\chi_B(x)/\mu(B)$ and, hence, $\int_X f_0 g\,d\mu = \int_B |g|\,d\mu/\mu(B) \geq \|g\|_\infty - \epsilon$. Also, $|f_0(x)| = \chi_B(x)/\mu(B)$ and, hence, $\|f_0\|_1 = \int_X |f_0|\,d\mu = \int_B d\mu/\mu(B) = 1$.
These imply $\sup\big\{|\int_X fg\,d\mu| \,\big|\, f \in L^1, \|f\|_1 \leq 1\big\} \geq \|g\|_\infty - \epsilon$ for every ϵ with $0 < \epsilon < \|g\|_\infty$ and we conclude that $\|g\|_\infty = \sup\big\{|\int_X fg\,d\mu| \,\big|\, f \in L^1, \|f\|_1 \leq 1\big\}$. □

Definition. *Let $1 \leq p \leq +\infty$. For every $g \in L^{p'}$ we define $l_g : L^p \to F$ by*

$$l_g(f) = \int_X fg\,d\mu, \qquad f \in L^p.$$

Proposition 7.46. *Let $1 \leq p \leq +\infty$. For every $g \in L^{p'}$ the function l_g belongs to $(L^p)^*$. Moreover, if $1 < p \leq +\infty$, then $\|l_g\|_* = \|g\|_{p'}$ and, if $p = 1$, then $\|l_g\|_* \leq \|g\|_\infty$. If $p = 1$ and μ is semifinite, then $\|l_g\|_* = \|g\|_\infty$.*

Proof. We have $l_g(f_1 + f_2) = \int_X (f_1 + f_2)g\,d\mu = \int_X f_1 g\,d\mu + \int_X f_2 g\,d\mu = l_g(f_1) + l_g(f_2)$. Also, $l_g(\kappa f) = \int_X (\kappa f)g\,d\mu = \kappa\int_X fg\,d\mu = \kappa l_g(f)$. These imply that l_g is a linear functional.
Theorem 7.9 together with Proposition 7.16 imply that, if $1 < p \leq +\infty$, then $\|l_g\|_* = \|g\|_{p'}$. If μ is semifinite, the same is true for $p = 1$.
If $p = 1$, for all $f \in L^1$ we have $|l_g(f)| = |\int_X fg\,d\mu \leq \|g\|_\infty\|f\|_1$. Hence, $\|l_g\|_* \leq \|g\|_\infty$. □

Definition. *Let $1 \leq p \leq +\infty$. We define the mapping $J : L^{p'} \to (L^p)^*$ by $J(g) = l_g$ for all $g \in L^{p'}$.*

Proposition 7.47. *The function J is a bounded linear operator. If $1 < p \leq +\infty$, then J is an isometry from $L^{p'}$ into $(L^p)^*$. This is true when $p = 1$, if μ is semifinite.*

Proof. Exercise. □

Lemma 7.5. *Let $l \in (L^p(X, \mathcal{S}, \mu))^*$. If $E \in \mathcal{S}$, $\mathcal{S}\rceil E = \{A \in \mathcal{S} \,|\, A \subseteq E\}$ is the restriction of $\mathcal{S}$ on E and $\mu\rceil E$ is the restricted measure on $(E, \mathcal{S}\rceil E)$, we define $l\rceil E$ by*

$$(l\rceil E)(h) = l(\widetilde{h}), \qquad h \in L^p(E, \mathcal{S}\rceil E, \mu\rceil E),$$

where $\widetilde{h}$ is the extension of h as 0 on $X \setminus E$.
Then, $l\rceil E \in (L^p(E, \mathcal{S}\rceil E, \mu\rceil E))^*$ and $\|l\rceil E\|_* \leq \|l\|_*$. Moreover,

$$l(f\chi_E) = (l\rceil E)(f\rceil E), \qquad f \in L^p(X, \mathcal{S}, \mu),$$

where $f\rceil E$ is the restriction of f on E.

Proof. For all $h, h_1, h_2 \in L^p(E, \mathcal{S}\rceil E, \mu\rceil E)$ we consider the corresponding extensions $\widetilde{h}, \widetilde{h_1}, \widetilde{h_2} \in L^p(X, \mathcal{S}, \mu)$. Since $\widetilde{h_1} + \widetilde{h_2}$ and $\kappa\widetilde{h}$ are the extensions of $h_1 + h_2$ and κh, respectively, we have $(l\rceil E)(h_1+h_2) = l(\widetilde{h_1+h_2}) = l(\widetilde{h_1})+l(\widetilde{h_2}) = (l\rceil E)(h_1)+(l\rceil E)(h_2)$ and $(l\rceil E)(\kappa h) = l(\widetilde{\kappa h}) = \kappa l(\widetilde{h}) = \kappa(l\rceil E)(h)$. This proves that $l\rceil E$ is linear and $|(l\rceil E)(h)| = |l(\widetilde{h})| \leq \|l\|_* \|\widetilde{h}\|_p = \|l\|_* \|h\|_p$ proves that $l\rceil E$ is bounded and that $\|l\rceil E\|_* \leq \|l\|_*$.
If $f \in L^p(X, \mathcal{S}, \mu)$, then $\widetilde{f\rceil E} = f\chi_E$ on X and, hence, $(l\rceil E)(f\rceil E) = l(\widetilde{f\rceil E}) = l(f\chi_E)$. □

Definition. *The $l\rceil E$ defined in Lemma 7.4 is called the* ***restriction*** *of $l \in (L^p(X, \mathcal{S}, \mu))^*$ on $L^p(E, \mathcal{S}\rceil E, \mu\rceil E)$.*

Theorem 7.15. *Let $1 < p < +\infty$.*
(i) For every $l \in (L^p)^*$ there exists a unique $g \in L^{p'}$ so that $l = l_g$ (see the definition before Proposition 7.46) i.e. so that $l(f) = \int_X fg\, d\mu$ for every $f \in L^p$.
(ii) The function J is an isometry from $L^{p'}$ onto $(L^p)^*$.
If μ is σ-finite, then (i) and (ii) are true also when $p = 1$.

Proof. (a) We consider first the case when μ is a finite measure: $\mu(X) < +\infty$.
Let $l \in (L^p)^*$ and $1 \leq p < +\infty$.
Since $\int_A |\chi_A|^p\, d\mu = \mu(A) < +\infty$, we have that $\chi_A \in L^p$ for every $A \in \mathcal{S}$. We define the function $\nu : \mathcal{S} \to F$ by $\nu(A) = l(\chi_A)$ for all $A \in \mathcal{S}$.
We have $\nu(\emptyset) = l(\chi_\emptyset) = l(0) = 0$.
If $A_1, A_2, \ldots \in \mathcal{S}$ are pairwise disjoint and $A = \bigcup_{j=1}^{+\infty} A_j$, then $\chi_A = \sum_{j=1}^{+\infty} \chi_{A_j}$. Therefore,

$$\begin{aligned}\| \textstyle\sum_{j=1}^n \chi_{A_j} - \chi_A\|_p^p &= \textstyle\int_X |\sum_{j=n+1}^{+\infty} \chi_{A_j}|^p\, d\mu = \int_X |\chi_{\bigcup_{j=n+1}^{+\infty} A_j}|^p\, d\mu \\ &= \mu(\textstyle\bigcup_{j=n+1}^{+\infty} A_j) \to \mu(\emptyset) = 0,\end{aligned}$$

by continuity of μ from above. Linearity and continuity of l imply $\sum_{j=1}^n \nu(A_j) = \sum_{j=1}^n l(\chi_{A_j}) = l(\sum_{j=1}^n \chi_{A_j}) \to l(\chi_A) = \nu(A)$ or, equivalently, that $\sum_{j=1}^{+\infty} \nu(A_j) = \nu(A)$.
Hence, ν is a real or complex measure (depending on whether $F = \mathbb{R}$ or $F = \mathbb{C}$) on $(X, \mathcal{S})$.
We observe that, if $A \in \mathcal{S}$ has $\mu(A) = 0$, then $\nu(A) = l(\chi_A) = l(0) = 0$ because the function χ_A is the zero element of L^p. Therefore, $\nu \ll \mu$ and by the Lebesgue-Radon-Nikodym Theorems there exists a function $g : X \to \overline{F}$ which is integrable over X with respect to μ, so that $l(\chi_A) = \nu(A) = \int_A g\, d\mu = \int_X \chi_A g\, d\mu$ for every $A \in \mathcal{S}$. By linearity of l and of the integral this, clearly, implies $l(\phi) = \int_X \phi g\, d\mu$ for every measurable simple function ϕ on X. This extends to all measurable functions in L^p which are bounded a.e. on X. Indeed, let $f \in L^p$ be such that $|f| \leq M$ a.e. on X for some $M < +\infty$. We take any sequence (ϕ_n) of measurable simple functions with $\phi_n \to f$ and $|\phi_n| \uparrow |f|$ on X. Then $\phi_n g \to fg$ and $|\phi_n g| \leq |fg| \leq M|g|$ a.e. on X. Since $\int_X |g|\, d\mu < +\infty$, the Dominated Convergence Theorem implies that $\int_X \phi_n g\, d\mu \to \int_X fg\, d\mu$. On the other hand, $|\phi_n - f|^p \to 0$ on X and $|\phi_n - f|^p \leq (|\phi_n| + |f|)^p \leq 2^p|f|^p$ on X. The Dominated Convergence Theorem again implies that $\int_X |\phi_n - f|^p\, d\mu \to 0$ and, hence, $\phi_n \to f$ in L^p. By continuity of l we get that $\int_X \phi_n g\, d\mu = l(\phi_n) \to l(f)$ and, hence,

$$l(f) = \textstyle\int_X fg\, d\mu \tag{7.9}$$

for every $f \in L^p$ which is bounded a.e. on X.
Now our first task is to prove that $g \in L^{p'}$.
If $1 < p, p' < +\infty$, we consider a sequence (ψ_n) of measurable non-negative simple functions on X so that $\psi_n \uparrow |g|^{p'-1}$ on X. We define $\phi_n(x) = \psi_n(x)\, \overline{\mathrm{sign}(g(x))}$. Then $0 \leq \phi_n g = \psi_n |g| \uparrow |g|^{p'}$ a.e. on X and each ϕ_n is bounded a.e. on X. Hence,

$$\|\psi_n\|_p^p = \textstyle\int_X \psi_n^p\, d\mu \leq \int_X \psi_n |g|\, d\mu = \int_X \phi_n g\, d\mu = l(\phi_n) \leq \|l\|_* \|\phi_n\|_p \leq \|l\|_* \|\psi_n\|_p,$$

where the last equality is justified by (7.9). This implies $\int_X \psi_n^p \, d\mu = \|\psi_n\|_p^p \leq \|l\|_*^{p'}$ and by the Monotone Convergence Theorem we get $\int_X |g|^{p'} \, d\mu = \lim_{n\to+\infty} \int_X \psi_n^p \, d\mu \leq \|l\|_*^{p'}$. Therefore, $g \in L^{p'}$ and $\|g\|_{p'} \leq \|l\|_*$.
If $p = 1$ and $p' = +\infty$, we consider any possible $t > 0$ such that the set $A = \{x \in X \mid t < |g(x)|\}$ has $\mu(A) > 0$. We define the function $f(x) = \chi_A(x) \overline{\mathrm{sign}(g(x))}$. Then

$$t\mu(A) \leq \int_A |g| \, d\mu = \int_X fg \, d\mu = l(f) \leq \|l\|_* \|f\|_1 \leq \|l\|_* \mu(A),$$

where the last equality is justified by (7.9). This implies that $t \leq \|l\|_*$ and, hence, $|g| \leq \|l\|_*$ a.e. on X. Therefore, g is essentially bounded on X with respect to μ and $\|g\|_\infty \leq \|l\|_*$.
We have proved that in all cases $g \in L^{p'}$ and $\|g\|_{p'} \leq \|l\|_*$.
Now consider an arbitrary $f \in L^p$ and take a sequence (ϕ_n) of measurable simple functions on X so that $\phi_n \to f$ and $|\phi_n| \uparrow |f|$ on X. We have already shown by the Dominated Convergence Theorem that $\phi_n \to f$ in L^p and, hence, $l(\phi_n) \to l(f)$. Moreover,

$$\Big| \int_X \phi_n g \, d\mu - \int_X fg \, d\mu \Big| \leq \int_X |\phi_n - f||g| \, d\mu \leq \|\phi_n - f\|_p \|g\|_{p'} \to 0,$$

since $\|g\|_{p'} < +\infty$. From $l(\phi_n) = \int_X \phi_n g \, d\mu$ we conclude that

$$l(f) = \int_X fg \, d\mu, \qquad \text{for all } f \in L^p.$$

Of coure this implies that $l(f) = l_g(f)$ for every $f \in L^p$ and, hence, $l = l_g = J(g)$. Therefore, J is an isometry from $L^{p'}$ *onto* $(L^p)^*$.
Now assume that $h \in L^{p'}$ also satisfies $l = l_h$. Then $J(h) = l = J(g)$ and, since J is an isometry (and, hence, one-to-one), we get that $h = g$ a.e. on X.
(b) We suppose now that μ is σ-finite and consider an increasing sequence (E_k) in $\mathcal{S}$ so that $E_k \uparrow X$ and $\mu(E_k) < +\infty$ for all k.
Let $l \in (L^p(X, \mathcal{S}, \mu))^*$.
For each k we consider the restriction $l \rceil E_k$ of l on $L^p(E_k, \mathcal{S}\rceil E_k, \mu\rceil E_k)$ which is defined in Lemma 7.4. Since $l\rceil E_k \in (L^p(E_k, \mathcal{S}\rceil E_k, \mu\rceil E_k))^*$ and $\|l\rceil E_k\|_* \leq \|l\|_*$ and since $(\mu\rceil E_k)(E_k) = \mu(E_k) < +\infty$, part (a) implies that there is a *unique* $g_k \in L^{p'}(E_k, \mathcal{S}\rceil E_k, \mu\rceil E_k)$ with $\|g_k\|_{p'} \leq \|l\rceil E_k\|_* \leq \|l\|_*$ and

$$(l\rceil E_k)(h) = \int_{E_k} hg_k \, d(\mu\rceil E_k), \qquad \text{for all } h \in L^p(E_k, \mathcal{S}\rceil E_k, \mu\rceil E_k).$$

In particular,

$$l(f\chi_{E_k}) = (l\rceil E_k)(f\rceil E_k) = \int_{E_k} (f\rceil E_k) g_k \, d(\mu\rceil E_k) \qquad \text{for all } f \in L^p(X, \mathcal{S}, \mu).$$

For $h \in L^p(E_k, \mathcal{S}\rceil E_k, \mu\rceil E_k)$ take its extension h_0 on E_{k+1} as 0 on $E_{k+1} \setminus E_k$. Since $\widetilde{h} = \widetilde{h_0}$ on X, we get

$$\begin{aligned} \int_{E_k} hg_k \, d(\mu\rceil E_k) &= (l\rceil E_k)(h) = l(\widetilde{h}) = l(\widetilde{h_0}) = (l\rceil E_{k+1})(h_0) = \int_{E_{k+1}} h_0 g_{k+1} \, d(\mu\rceil E_{k+1}) \\ &= \int_X \widetilde{h_0 g_{k+1}} \, d\mu = \int_{E_k} (\widetilde{h_0 g_{k+1}})\rceil E_k \, d(\mu\rceil E_k) = \int_{E_k} h(g_{k+1}\rceil E_k) \, d(\mu\rceil E_k). \end{aligned}$$

By the uniqueness result of part (a) we have that $g_{k+1}\rceil E_k = g_k$ a.e. on E_k. We may clearly suppose that $g_{k+1}\rceil E_k = g_k$ on E_k for every k by inductively changing g_{k+1} on a subset of E_k of zero measure.
Now we define the measurable function g on X as equal to g_k on each E_k. I.e. $g\rceil E_k = g_k$ on E_k for every k. Therefore, $l(f\chi_{E_k}) = \int_{E_k} (f\rceil E_k)(g\rceil E_k) \, d(\mu\rceil E_k)$ and, thus,

$$l(f\chi_{E_k}) = \int_{E_k} fg \, d\mu \qquad \text{for all } f \in L^p(X, \mathcal{S}, \mu).$$

If $1 < p' < +\infty$, then, since $|\widetilde{g_k}| \uparrow |g|$ on X, by the Monotone Convergence Theorem,

$$\int_X |g|^{p'}\, d\mu = \lim_{k\to+\infty} \int_X |\widetilde{g_k}|^{p'}\, d\mu = \lim_{k\to+\infty} \int_{E_k} |g_k|^{p'}\, d(\mu\rceil E_k) \le \|l\|_*^{p'} < +\infty.$$

Hence, $g \in L^{p'}(X, \mathcal{S}, \mu)$ and $\|g\|_{p'} \le \|l\|_*$.
If $p' = +\infty$, we have that $|g| = |g_k| \le \|g_k\|_\infty \le \|l\rceil E_k\|_* \le \|l\|_*$ a.e. on E_k for every k. This implies $|g| \le \|l\|_*$ a.e. on X and, thus, $g \in L^\infty(X, \mathcal{S}, \mu)$ and $\|g\|_\infty \le \|l\|_*$.
Hence, in all cases, $g \in L^{p'}(X, \mathcal{S}, \mu)$ and $\|g\|_{p'} \le \|l\|_*$.
For an arbitrary $f \in L^p(X, \mathcal{S}, \mu)$ we get $\|f\chi_{E_k} - f\|_p^p = \int_X |f\chi_{E_k} - f|^p\, d\mu = \int_{E_k^c} |f|^p\, d\mu = \int_X \chi_{E_k^c} |f|^p\, d\mu \to 0$ by the Dominated Convergence Theorem. By continuity of l we have that $l(f) = \lim_{k\to+\infty} l(f\chi_{E_k}) = \lim_{k\to+\infty} \int_{E_k} fg\, d\mu = \int_X fg\, d\mu$. The last equality holds since $|\int_{E_k} fg\, d\mu - \int_X fg\, d\mu| = |\int_{E_k^c} fg\, d\mu| \le (\int_{E_k^c} |f|^p\, d\mu)^{\frac{1}{p}} \|g\|_{p'} \to 0$.
We have proved that

$$l(f) = \int_X fg\, d\mu \qquad \text{for all } f \in L^p(X, \mathcal{S}, \mu)$$

and, thus, $l = l_g = J(g)$. Hence, J is an isometry from $L^{p'}(X, \mathcal{S}, \mu)$ *onto* $(L^p(X, \mathcal{S}, \mu))^*$.
Again, if $h \in L^{p'}(X, \mathcal{S}, \mu)$ also satisfies $l = l_h$, then $J(h) = l = J(g)$ and, since J is an isometry, we get that $h = g$ a.e. on X.
(c) Now let $1 < p, p' < +\infty$ and μ be arbitrary.
Let $l \in (L^p(X, \mathcal{S}, \mu))^*$.
We consider any $E \in \mathcal{S}$ of σ-finite measure and the restriction $l\rceil E$ of l on $L^p(E, \mathcal{S}\rceil E, \mu\rceil E)$ defined in Lemma 7.4. Since $l\rceil E \in (L^p(E, \mathcal{S}\rceil E, \mu\rceil E))^*$ and $\|l\rceil E\|_* \le \|l\|_*$, part (b) implies that there is a unique $g_E \in L^{p'}(E, \mathcal{S}\rceil E, \mu\rceil E)$ so that $\|g_E\|_{p'} \le \|l\rceil E\|_* \le \|l\|_*$ and

$$(l\rceil E)(h) = \int_E h g_E\, d(\mu\rceil E) \qquad \text{for all } h \in L^p(E, \mathcal{S}\rceil E, \mu\rceil E).$$

In particular,

$$l(f\chi_E) = (l\rceil E)(f\rceil E) = \int_E (f\rceil E) g_E\, d(\mu\rceil E) \qquad \text{for all } f \in L^p(X, \mathcal{S}, \mu).$$

Now let E, F be two sets of σ-finite measure with $E \subseteq F$. Repeating the argument in the proof of part (b), with which we showed that $g_{k+1}\rceil E_k = g_k$ a.e. on E_k, we may easily show (just replace E_k by E and E_{k+1} by F) that $g_F\rceil E = g_E$ a.e. on E.
Now, we define

$$M = \sup\big\{ \int_E |g_E|^{p'}\, d(\mu\rceil E) \,\big|\, E \text{ of } \sigma\text{-finite measure}\big\}$$

and then, obviously, $M \le \|l\|_*^{p'} < +\infty$. We take a sequence (E_n) in $\mathcal{S}$ where each E_n has σ-finite measure so that

$$\int_{E_n} |g_{E_n}|^{p'}\, d(\mu\rceil E_n) \to M.$$

We define $E = \bigcup_{n=1}^{+\infty} E_n$ and observe that E has σ-finite measure and, hence, $\int_E |g_E|^{p'}\, d(\mu\rceil E) \le M$. Since $E_n \subseteq E$, by the result of the previous paragraph $g_E\rceil E_n = g_{E_n}$ a.e. on E_n and, hence, $\int_{E_n} |g_{E_n}|^{p'}\, d(\mu\rceil E_n) \le \int_E |g_E|^{p'}\, d(\mu\rceil E) \le M$. Taking the limit as $n \to +\infty$, this implies that $\int_E |g_E|^{p'}\, d(\mu\rceil E) = M$. We set $g = \widetilde{g_E}$ and have $\int_X |g|^{p'}\, d\mu = \int_E |g_E|^{p'}\, d(\mu\rceil E) = M \le \|l\|_*^{p'}$.
Now consider an arbitrary $f \in L^p(X, \mathcal{S}, \mu)$. The set $F = E \cup \{x \in X \,|\, f(x) \neq 0\}$ has σ-finite measure. By $g_F\rceil E = g_E$ a.e. on E we get

$$\begin{aligned} M = \int_E |g_E|^{p'}\, d(\mu\rceil E) &= \int_E |g_F|^{p'}\, d(\mu\rceil F) \le \int_E |g_F|^{p'}\, d(\mu\rceil F) + \int_{F\setminus E} |g_F|^{p'}\, d(\mu\rceil F) \\ &= \int_F |g_F|^{p'}\, d(\mu\rceil F) \le M. \end{aligned}$$

Therefore, $\int_{F\setminus E} |g_F|^{p'}\, d(\mu\rceil F) = 0$ and, hence, $g_F = 0$ a.e. on $F \setminus E$. Now

$$\begin{aligned} l(f) = l(f\chi_F) &= \int_F (f\rceil F) g_F\, d(\mu\rceil F) = \int_E (f\rceil F) g_F\, d(\mu\rceil F) = \int_E (f\rceil F) g_E\, d(\mu\rceil F) \\ &= \int_E (f\rceil E) g_E\, d(\mu\rceil E) = \int_X fg\, d\mu. \end{aligned}$$

Thus, $l = l_g = J(g)$ and, just as in parts (a) and (b), J is an isometry from $L^{p'}(X, \mathcal{S}, \mu)$ *onto* $(L^p(X, \mathcal{S}, \mu))^*$.
Finally, if $h \in L^{p'}(X, \mathcal{S}, \mu)$ also satisfies $l = l_h$, then $J(h) = l = J(g)$ and, since J is an isometry, we get that $h = g$ a.e. on X. $\square$

We know that if $1 < p, p' < +\infty$ and $\frac{1}{p} + \frac{1}{p'} = 1$, then there is an isometry from $L^{p'}$ onto $(L^p)^*$. Also, there is an isometry from L^1 into $(L^\infty)^*$, but in general this is not onto. If μ is σ-finite, then there is an isometry from L^∞ onto $(L^1)^*$.

In all these cases we may *identify* every $g \in L^{p'}$ with the corresponding $l_g \in (L^p)^*$ which is defined by $l_g(f) = \int_X fg\, d\mu$ for all $f \in L^p$. Hence, we may view every $g \in L^{p'}$ as a bounded linear functional on L^p and if we write g instead of l_g, then the defining relation of l_g can be written

$$g(f) = \int_X fg\, d\mu, \qquad f \in L^p, g \in L^{p'}.$$

Observe the symmetry $\int_X fg\, d\mu = \int_X gf\, d\mu$ which permits us to write

$$g(f) = f(g) = \int_X fg\, d\mu, \qquad f \in L^p, g \in L^{p'}.$$

Hence, *every $g \in L^{p'}$ acts as a bounded linear functional on all $f \in L^p$ and is, thus, an element of $(L^p)^*$ and, conversely, every $f \in L^p$ acts as a bounded linear functional on all $g \in L^{p'}$ and is, thus, an element of $(L^{p'})^*$.*

Proposition 7.48. *If $1 < p < +\infty$, then L^p is reflexive.*

Proof. We have to prove that the mapping $T : L^p \to (L^p)^{**}$ defined in Proposition 7.26 is onto. We recall that T is defined by $T(f)(l) = l(f)$ for all $l \in (L^p)^*$ and every $f \in L^p$.
We consider $p' = \frac{p}{p-1}$. Then $1 < p' < +\infty$ and $\frac{1}{p} + \frac{1}{p'} = 1$.
Now, we recall the isometry $J : L^{p'} \to (L^p)^*$ defined by $J(g) = l_g$, where $l_g(f) = \int_X fg\, d\mu$ for all $f \in L^p$.
We consider any $L \in (L^p)^{**}$ and we define $\tilde{L} = L \circ J : L^{p'} \to F$. I.e. $\tilde{L}(g) = (L \circ J)(g) = L(l_g)$ for all $g \in L^{p'}$. Both L and J are bounded and linear and, hence, $\tilde{L} \in (L^{p'})^*$.
We also recall the isometry $J' : L^p \to (L^{p'})^*$ defined by $J'(f) = l_f$, where $l_f(g) = \int_X gf\, d\mu$ for all $g \in L^{p'}$. Therefore, there is an $f \in L^p$ so that $\tilde{L} = J'(f) = l_f$. Then for every $l \in (L^p)^*$ there is a $g \in L^{p'}$ so that $l = J(g) = l_g$ and

$$T(f)(l) = l(f) = l_g(f) = \int_X fg\, d\mu = l_f(g) = \tilde{L}(g) = L(l_g) = L(l).$$

Hence, $T(f) = L$ and $T : L^p \to (L^p)^{**}$ is onto. $\square$

Definition. *Let $1 \le p < +\infty$ (in the case $p = 1$ we assume also that μ is σ-finite) and (f_n) be a sequence in L^p. We say that (f_n)* ***converges weakly*** *to $f \in L^p$ if $\int_X f_n g\, d\mu \to \int_X fg\, d\mu$ for all $g \in L^{p'}$. In this case we write $f_n \xrightarrow{w} f$.*
Let $1 \le p \le +\infty$ and (f_n) be a sequence in L^p. We say that (f_n) ***converges weakly**** *to $f \in L^p$ if $\int_X f_n g\, d\mu \to \int_X fg\, d\mu$ for all $g \in L^{p'}$. In this case we write $f_n \xrightarrow{w*} f$.*

Let us see the case of weak convergence. If we identify every $g \in L^{p'}$ with the corresponding $l_g \in (L^p)^*$ then $\int_X f_n g\, d\mu \to \int_X fg\, d\mu$ is equivalent to $l_g(f_n) \to l_g(f)$. Now, since for every $l \in (L^p)^*$ there is a $g \in L^{p'}$ so that $l = l_g$, we conclude that $f_n \xrightarrow{w} f$ is equivalent to $l(f_n) \to l(f)$ for all $l \in (L^p)^*$. Therefore, the definition we gave for $f_n \xrightarrow{w} f$ in L^p is a special case of the definition of weak convergence in the case of the general normed space.

We have a similar comment for the case of weak* convergence. If we identify every $f_n \in L^p$ and $f \in L^p$ with the corresponding $l_{f_n} \in (L^{p'})^*$ and $l_f \in (L^{p'})^*$ then $\int_X f_n g\, d\mu \to \int_X fg\, d\mu$ is equivalent to $l_{f_n}(g) \to l_f(g)$. Therefore, the definition we gave for $f_n \xrightarrow{w*} f$ in L^p is the same as the definition of $l_{f_n} \xrightarrow{w*} l_f$ in $(L^{p'})^*$ which is a special case of the definition of weak* convergence in the case of the general dual space.

We observe that, if $1 < p < +\infty$, then the notions of weak convergence and weak* convergence in L^p coincide. The same is true when $p = 1$ if μ is σ-finite.

The next results are special cases of corresponding results of the previous section.

Proposition 7.49. *Let $1 \le p \le +\infty$. If (f_n) is a sequence in L^p such that $\lim_{n\to+\infty} \int_X f_n g\, d\mu$ exists in F for all $g \in L^{p'}$, then*

$$\|f\|_p \le \underline{\lim}_{n\to+\infty} \|f_n\|_p, \qquad \sup_{n\in\mathbb{N}} \|f_n\|_p < +\infty.$$

Proof. A corollary of Proposition 8.3. □

Proposition 7.50. *Let $1 < p \le +\infty$ (in the case $p = +\infty$ we assume also that μ is σ-finite). If (f_n) is a sequence in L^p such that $\lim_{n\to+\infty} \int_X f_n g\, d\mu$ exists in F for all $g \in L^{p'}$, then there is an $f \in L^p$ so that $\int_X f_n g\, d\mu \to \int_X fg\, d\mu$ for all $g \in L^{p'}$.*

Proof. A corollary of Proposition 8.3. □

Proposition 7.51. *Let $1 < p < +\infty$. If (f_n) is a bounded sequence in L^p, then there is an $f \in L^p$ so that $\int_X f_n g\, d\mu \to \int_X fg\, d\mu$ for all $g \in L^{p'}$.*

Proof. This is a corollary of Theorem 8.2 and Proposition 8.10. □

Proposition 7.52. *Let (X, Σ, μ) be a measure space and a countable $P \subseteq \Sigma$ with the property: for every $E \in \Sigma$ with $\mu(E) < +\infty$ and every $\epsilon > 0$ there is an $A \in P$ so that $\mu(A \triangle E) < \epsilon$. If $1 \le p < +\infty$, then $L^p(X, \Sigma, \mu)$ is separable.*

Proof. (a) Let $E \in \Sigma$ with $\mu(E) < +\infty$ and $\epsilon > 0$. We consider the $A \in P$ so that $\mu(A \triangle E) < \epsilon^p$ and we have that $\{x \in X \mid \chi_A(x) \neq \chi_E(x)\} = A \triangle E$ and, thus,

$$\|\chi_A - \chi_E\|_p < \epsilon.$$

(b) Now we consider a simple function $\phi = \sum_{k=1}^n \kappa_k \chi_{E_k}$ so that $\mu(E_k) < +\infty$ for each k and any ϵ with $0 < \epsilon \le 1$. We also take $M = n + \sum_{k=1}^n |\kappa_k| + \sum_{k=1}^n (\mu(E_k))^{1/p}$. Then for every k we find a rational λ_k so that $|\lambda_k - \kappa_k| < \frac{\epsilon}{M}$ and, by the result of (a), an $A_k \in P$ so that $\|\chi_{A_k} - \chi_{E_k}\|_p < \frac{\epsilon}{M}$. Then we consider $\psi = \sum_{k=1}^n \lambda_k \chi_{A_k}$ and we get

$$\begin{aligned} \|\psi - \phi\|_p &\le \| \textstyle\sum_{k=1}^n \lambda_k(\chi_{A_k} - \chi_{E_k}) + \sum_{k=1}^n (\lambda_k - \kappa_k)\chi_{E_k} \|_p \\ &\le \textstyle\sum_{k=1}^n |\lambda_k| \|\chi_{A_k} - \chi_{E_k}\|_p + \frac{\epsilon}{M} \sum_{k=1}^n (\mu(E_k))^{1/p} \\ &< \tfrac{\epsilon}{M}(n\epsilon + \textstyle\sum_{k=1} |\kappa_k|) + \frac{\epsilon}{M} \sum_{k=1}^n (\mu(E_k))^{1/p} < \frac{\epsilon}{M} M = \epsilon. \end{aligned}$$

We observe that the set B of all functions ψ is countable.
(c) Finally, we take any $f \in L^p(X, \Sigma, \mu)$ and any $\epsilon > 0$. Then there is a simple function ϕ so that $\|\phi - f\|_p < \frac{\epsilon}{2}$. By the result of (b) there is some $\psi \in B$ so that $\|\psi - \phi\|_p < \frac{\epsilon}{2}$. Then of course $\|\psi - f\|_p < \epsilon$. □

Proposition 7.53. *Let μ be σ-finite and assume that there is a countable $P \subseteq \Sigma$ with the property: for every $E \in \Sigma$ with $\mu(E) < +\infty$ and every $\epsilon > 0$ there is an $A \in P$ so that $\mu(A \triangle E) < \epsilon$. If (f_n) is a bounded sequence in L^∞, then there is an $f \in L^\infty$ so that $\int_X f_n g\, d\mu \to \int_X fg\, d\mu$ for all $g \in L^1$.*

Proof. This is a corollary of Theorem 8.1 and Proposition 8.16. □

Exercises.

7.3.1. *Approximation*
(i) Let $f \in L^p(X, \mathcal{S}, \mu)$ and $\epsilon > 0$. Prove that there exists a measurable simple function ϕ on X so that $\|f - \phi\|_p < \epsilon$. If $p < +\infty$, then $\phi = 0$ outside a set of finite measure.
(ii) Let $f \in L^p(\mathbb{R}^n, \mathcal{L}_n, m_n)$ and $\epsilon > 0$. If $p < +\infty$, prove that there exists a function g continuous on $\mathbb{R}^n$ and equal to 0 outside some bounded set so that $\|f - g\|_p < \epsilon$.

7.3.2. Let I be any index set and $0 < p < q \leq +\infty$. Prove that $l^p(I) \subseteq l^q(I)$ and $\|b\|_q \leq \|b\|_p$ for every $b \in l^p(I)$.

7.3.3. Let $\mu(X) < +\infty$ and $0 < p < q \leq +\infty$. Prove that $L^q(X, \mu) \subseteq L^p(X, \mu)$ and that $\|f\|_p \leq \mu(X)^{\frac{1}{p} - \frac{1}{q}} \|f\|_q$ for every $f \in L^q(X, \mu)$.

7.3.4. Let $0 < p < q < r \leq +\infty$ and $f \in L^p \cap L^r$. Prove that $f \in L^q$ and, if $\frac{1}{q} = \frac{t}{p} + \frac{1-t}{r}$, then $\|f\|_q \leq \|f\|_p^t \|f\|_r^{1-t}$. Also prove that $\lim_{q \to p+} \|f\|_q = \|f\|_p$ and $\lim_{q \to r-} \|f\|_q = \|f\|_r$.

7.3.5. Let $1 \leq p < r \leq +\infty$. Set $Z = L^p \cap L^r$ and define $\|f\| = \|f\|_p + \|f\|_r$ for every $f \in Z$.
(i) Prove that $\|\cdot\|$ is a norm on Z and that $(Z, \|\cdot\|)$ is a Banach space.
(ii) If $p < q < r$, consider the linear transformation $T : Z \to L^q$ with $T(f) = f$ for every $f \in Z$ (see exercise 7.2.4). Prove that T is bounded.

7.3.6. Let $0 < p < q < r \leq +\infty$ and $f \in L^q$. If $t > 0$ is arbitrary, consider the functions defined by $g(x) = f(x)$ and $h(x) = 0$, if $|f(x)| > t$, and $g(x) = 0$ and $h(x) = f(x)$, if $|f(x)| \leq t$. Prove that $g \in L^p$ and $h \in L^r$ and that $f = g + h$ on X.

7.3.7. Let $1 \leq p < r \leq +\infty$. We define $W = L^p + L^r = \{g + h \,|\, g \in L^p, h \in L^r\}$ and, also, $\|f\| = \inf\big\{\|g\|_p + \|h\|_r \,|\, g \in L^p, h \in L^r, f = g + h\big\}$ for every $f \in W$.
(i) Prove that $\|\cdot\|$ is a norm on W and that $(W, \|\cdot\|)$ is a Banach space.
(ii) If $p < q < r$, consider the linear transformation $T : L^q \to W$ with $T(f) = f$ for every $f \in L^q$ (see exercise 7.2.6). Prove that T is bounded.

7.3.8. Let $0 < p < q < +\infty$. Prove that $L^p(X) \not\subseteq L^q(X)$ if and only if X includes sets of arbitrarily small positive measure and that $L^q(X) \not\subseteq L^p(X)$ if and only if X includes sets of arbitrarily large finite measure.

7.3.9. Let $1 \leq p < +\infty$ and (f_n) be a sequence in L^p so that $\|f_n - f\|_p \to 0$ for some $f \in L^p$. Prove that $f_n \to f$ in measure.

7.3.10. Let $1 \leq p < +\infty$ and (f_n) be a sequence in L^p so that $|f_n| \leq g$ a.e. for every n for some $g \in L^p$. If $f_n \to f$ a.e. or in measure, prove that $\|f_n - f\|_p \to 0$.

7.3.11. Let $1 \leq p < +\infty$ and $f, f_n \in L^p$ for all n. If $f_n \to f$ a.e., prove that $\|f_n - f\|_p \to 0$ if and only if $\|f_n\|_p \to \|f\|_p$.

7.3.12. Let $1 \leq p \leq +\infty$ and $g \in L^\infty(\mu)$.
We define the linear transformation $T : L^p(\mu) \to L^p(\mu)$ with $T(f) = gf$ for every $f \in L^p(\mu)$. Prove that T is bounded, that $\|T\| \leq \|g\|_\infty$ and that $\|T\| = \|g\|_\infty$ if μ is semifinite.

7.3.13. *The inequality of Chebychev.*
If $0 < p < +\infty$ and $f \in L^p$, prove that $\lambda_{|f|}(t) \leq \|f\|_p^p / t^p$ for $0 < t < +\infty$.

7.3.14. *The general Minkowski's Inequality.*
Let $(X_1, \mathcal{S}_1, \mu_1)$ and $(X_2, \mathcal{S}_2, \mu_2)$ be two σ-finite measure spaces and $1 \leq p < +\infty$.
(i) If $f : X_1 \times X_2 \to [0, +\infty]$ is $\mathcal{S}_1 \otimes \mathcal{S}_2$–measurable, prove that

$$\Big(\int_{X_1} \Big(\int_{X_2} f(x_1, x_2)\, d\mu_2(x_2) \Big)^p d\mu_1(x_1) \Big)^{1/p} \leq \int_{X_2} \Big(\int_{X_1} f(x_1, x_2)^p \, d\mu_1(x_1) \Big)^{1/p} d\mu_2(x_2).$$

(ii) If $f(\cdot, x_2) \in L^p(X_1, \mathcal{S}_1, \mu_1)$ for μ_2-a.e. $x_2 \in X_2$ and the function $x_2 \mapsto \|f(\cdot, x_2)\|_p$ is in $L^1(X_2, \mathcal{S}_2, \mu_2)$, prove that $f(x_1, \cdot) \in L^1(X_2, \mathcal{S}_2, \mu_2)$ for μ_1-a.e. $x_1 \in X_1$, that the function $x_1 \mapsto \int_{X_2} f(x_1, \cdot)\, d\mu_2$ is in $L^p(X_1, \mathcal{S}_1, \mu_1)$ and

$$\Big(\int_{X_1} \Big| \int_{X_2} f(x_1, x_2)\, d\mu_2(x_2) \Big|^p d\mu_1(x_1) \Big)^{1/p} \leq \int_{X_2} \Big(\int_{X_1} |f(x_1, x_2)|^p \, d\mu_1(x_1) \Big)^{1/p} d\mu_2(x_2).$$

7.4 The spaces $M(X)$ and $M_{\mathcal{R}}(X)$.

Definition. *Let $(X, \mathcal{S})$ be a measurable space. The set of all real or complex (depending on whether $F = \mathbb{R}$ or $F = \mathbb{C}$) measures on $(X, \mathcal{S})$ is denoted by $M(X, \mathcal{S})$.*

Therefore, all $\nu \in M(X, \mathcal{S})$ have *only finite* (real or complex) values.

If there is no danger of confusion, we shall use the symbol M instead of $M(X, \mathcal{S})$.

We recall addition and multiplication on these spaces. If $\nu_1, \nu_2 \in M$, we define $\nu_1 + \nu_2 \in M$ by $(\nu_1 + \nu_2)(A) = \nu_1(A) + \nu_2(A)$ for all $A \in \mathcal{S}$. We also define $\kappa\nu \in M$ by $(\kappa\nu)(A) = \kappa\nu(A)$ for all $A \in \mathcal{S}$ and $\kappa \in F$.

It is easy to show that M is a linear space over F. The zero element is the measure 0 defined by $0(A) = 0$ for all $A \in \mathcal{S}$. The opposite to ν is $-\nu$ defined by $(-\nu)(A) = -\nu(A)$ for all $A \in \mathcal{S}$.

Definition. *For every $\nu \in M$ we define*

$$\|\nu\| = |\nu|(X).$$

Thus, $\|\nu\|$ is just the total variation of ν.

Proposition 7.54. *$\|\cdot\|$ is a norm on M.*

Proof. Immediate after Propositions 6.6 and 6.9. $\square$

Theorem 7.16. *M is a Banach space.*

Proof. Let (ν_n) be a Cauchy sequence in M. Then $|\nu_n - \nu_m|(X) = \|\nu_n - \nu_m\| \to 0$ as $n, m \to +\infty$ and, hence, $|\nu_n(A) - \nu_m(A)| = |(\nu_n - \nu_m)(A)| \leq |\nu_n - \nu_m|(A) \leq |\nu_n - \nu_m|(X) \to 0$ as $n, m \to +\infty$. This implies that the sequence $(\nu_n(A))$ of numbers is a Cauchy sequence for every $A \in \mathcal{S}$. Therefore, it converges to a finite number and we define $\nu(A) = \lim_{n\to+\infty} \nu_n(A)$ for every $A \in \mathcal{S}$.
It is clear that $\nu(\emptyset) = \lim_{n\to+\infty} \nu_n(\emptyset) = 0$.
Now, let $A_1, A_2, \ldots \in \mathcal{S}$ be pairwise disjoint and $A = \bigcup_{j=1}^{+\infty} A_j$. We take an arbitrary $\epsilon > 0$ and find N so that $\|\nu_n - \nu_m\| \leq \epsilon$ for all $n, m \geq N$. Since $\sum_{j=1}^{+\infty} |\nu_N|(A_j) = |\nu_N|(A) < +\infty$, there is some J so that $\sum_{j=J+1}^{+\infty} |\nu_N|(A_j) \leq \epsilon$. From $|\nu_n| \leq |\nu_n - \nu_N| + |\nu_N|$ we get that, for every $n \geq N$,

$$\begin{aligned}\sum_{j=J+1}^{+\infty}|\nu_n|(A_j) &\leq \sum_{j=J+1}^{+\infty} |\nu_n - \nu_N|(A_j) + \sum_{j=J+1}^{+\infty} |\nu_N|(A_j) \\ &\leq |\nu_n - \nu_N|(\textstyle\bigcup_{j=J+1}^{+\infty} A_j) + \epsilon \leq |\nu_n - \nu_N|(X) + \epsilon = \|\nu_n - \nu_N\| + \epsilon \leq 2\epsilon.\end{aligned} \tag{7.10}$$

Then for any $K \geq J+1$ and $n \geq N$ we have $\sum_{j=J+1}^{K} |\nu_n(A_j)| \leq \sum_{j=J+1}^{K} |\nu_n|(A_j) \leq 2\epsilon$ and, taking the limit as $n \to +\infty$, $\sum_{j=J+1}^{K} |\nu(A_j)| \leq 2\epsilon$. Finally, taking the limit as $K \to +\infty$, we find

$$\sum_{j=J+1}^{+\infty} |\nu(A_j)| \leq 2\epsilon. \tag{7.11}$$

From (7.10) we get $|\nu_n(A) - \sum_{j=1}^{J} \nu_n(A_j)| = |\sum_{j=J+1}^{+\infty} \nu_n(A_j)| \leq \sum_{j=J+1}^{+\infty} |\nu_n|(A_j) \leq 2\epsilon$ for all $n \geq N$ and, taking the limit as $n \to +\infty$,

$$|\nu(A) - \sum_{j=1}^{J} \nu(A_j)| \leq 2\epsilon. \tag{7.12}$$

Altogether, from (7.11) and (7.12) we have

$$|\nu(A) - \sum_{j=1}^{+\infty} \nu(A_j)| \leq |\nu(A) - \sum_{j=1}^{J} \nu(A_j)| + \sum_{j=J+1}^{+\infty} |\nu(A_j)| \leq 4\epsilon.$$

Since ϵ is arbitrary, we get $\nu(A) = \sum_{j=1}^{+\infty} \nu(A_j)$ and we conclude that $\nu \in M$.
For any measurable partition $\{A_1, \ldots, A_p\}$ of X we get $\sum_{k=1}^{p} |(\nu_n - \nu_m)(A_k)| \leq \|\nu_n - \nu_m\| \leq \epsilon$ for every $n, m \geq N$. Taking the limit as $m \to +\infty$, we find $\sum_{k=1}^{p} |(\nu_n - \nu)(A_k)| \leq \epsilon$ for every $n \geq N$ and, taking the supremum of the left side over all measurable partitions $\{A_1, \ldots, A_p\}$ of X, we get $\|\nu_n - \nu\| = |\nu_n - \nu|(X) \leq \epsilon$. Hence, $\|\nu_n - \nu\| \to 0$. $\square$

Lemma 7.6. *Let μ be a real or complex (depending on whether $F = \mathbb{R}$ or $F = \mathbb{C}$) Borel measure on X. For every $f \in BC$ we have*

$$\big| \int_X f\, d\mu \big| \leq \int_X |f|\, d|\mu| \leq \|f\|_u \|\mu\|.$$

Proof. A consequence of Theorem 6.8. □

Let μ be a Borel measure on X. We recall that μ is called regular if for every Borel set E we have (i) $\mu(E) = \inf\{\mu(U) \,|\, U \text{ open } \supseteq E\}$ and (ii) $\mu(E) = \sup\{\mu(K) \,|\, K \text{ compact } \subseteq E\}$.

Definition. *If μ is a real Borel measure on X, then μ is called regular if μ^+ and μ^- are regular. If μ is a complex Borel measure on X, then μ is called regular if* $\mathrm{Re}(\mu)$ *and* $\mathrm{Im}(\mu)$ *are regular. The space of all regular real or complex Borel measures on X is denoted by*

$$M_{\mathcal{R}}(X, \mathcal{B}_X).$$

We write $M_{\mathcal{R}}$ instead of $M_{\mathcal{R}}(X, \mathcal{B}_X)$ if there is no danger of confusion.

It is clear that, if μ is a Borel measure and $\mu(E) < +\infty$, then (i) and (ii) in the definition of regularity are equivalent to the following: for every $\epsilon > 0$ there is an open $U \supseteq E$ and a compact $K \subseteq E$ so that $\mu(U \setminus K) < \epsilon$.

Proposition 7.55. *Let μ be a real or complex Borel measure on X. Then μ is regular if and only if $|\mu|$ is regular.*

Proof. Let μ be real. If μ is regular, then μ^+ and μ^- are regular and, thus, for every Borel set E and $\epsilon > 0$ there are open $U^+, U^- \supseteq E$ and compact $K^+, K^- \subseteq E$ so that $\mu^+(U^+ \setminus K^+) < \epsilon$ and $\mu^-(U^- \setminus K^-) < \epsilon$. We set $K = K^+ \cup K^- \subseteq A$ and $U = U^+ \cap U^- \supseteq A$ and then $\mu^+(U \setminus K) < \epsilon$ and $\mu^-(U \setminus K) < \epsilon$. We add and find $|\mu|(U \setminus K) < 2\epsilon$ and, hence, $|\mu|$ is regular.
Now let $|\mu|$ be regular. Then for every Borel set E and $\epsilon > 0$ there is an open $U \supseteq E$ and a compact $K \subseteq E$ with $|\mu|(U \setminus K) < \epsilon$ and, since $\mu^+, \mu^- \leq |\mu|$, we get the same inequalities for μ^+ and μ^-. Therefore, μ^+ and μ^- are regular and so μ is regular.
If μ is complex, the proof is similar and uses the inequalities $|\mathrm{Re}(\mu)|, |\mathrm{Im}(\mu)| \leq |\mu|$ and $|\mu| \leq |\mathrm{Re}(\mu)| + |\mathrm{Im}(\mu)|$. □

Theorem 7.17. *$M_{\mathcal{R}}$ is a closed linear subspace of M and, hence, a Banach space.*

Proof. If μ_1 and μ_2 are regular Borel measures on X, then $|\mu_1|$ and $|\mu_2|$ are regular. Therefore, for every Borel set E and $\epsilon > 0$ there are open $U_1, U_2 \supseteq E$ and compact $K_1, K_2 \subseteq E$ so that $|\mu_1|(U_1 \setminus K_1) < \epsilon$ and $|\mu_2|(U_2 \setminus K_2) < \epsilon$. We set $K = K_1 \cup K_2 \subseteq E$ and $U = U_1 \cap U_2 \supseteq E$, and thus we find the same inequalities for K and O. We add, using $|\mu_1 + \mu_2| \leq |\mu_1| + |\mu_2|$, and we find $|\mu_1 + \mu_2|(U \setminus K) < 2\epsilon$. Hence, $|\mu_1 + \mu_2|$ is regular and so $\mu_1 + \mu_2$ is regular.
It is even simpler to prove that, if μ is regular and $\kappa \in F$, then $\kappa\mu$ is regular.
Therefore $M_{\mathcal{R}}$ is a linear subspace of M.
Now let (μ_n) be a sequence in $M_{\mathcal{R}}$ converging to μ in M. We consider any Borel set E and $\epsilon > 0$ and find N so that $\|\mu_N - \mu\| < \epsilon$ and then, since $|\mu_N|$ is regular, we find an open $U \supseteq E$ and a compact $K \subseteq E$ so that $|\mu_N|(U \setminus K) < \epsilon$. Then $|\mu|(U \setminus K) \leq |\mu_N|(U \setminus K) + \|\mu_N - \mu\| < 2\epsilon$ and, thus, μ is regular. Therefore, $M_{\mathcal{R}}$ is closed in M. □

We recall Theorem 1.23 which says that, if for every open subset O of X there is an increasing sequence of compact sets whose interiors cover O, then every locally finite Borel measure is regular and, hence, $M_{\mathcal{R}} = M$.

We also recall Theorem 2.2 which says that, *if X is locally compact and Hausdorff, $K \subseteq X$ is compact, $U \subseteq X$ is open and $K \subseteq U$, then there is an $f \prec U$ so that $f = 1$ on K.* Lemma 7.6 is a generalization of this fact to more than one open sets.

Lemma 7.7. *Let X be locally compact and Hausdorff. If $K \subseteq X$ is compact and $U_1, \dots, U_n \subseteq X$ are open so that $K \subseteq U_1 \cup \cdots \cup U_n$, then there exist $f_1 \prec U_1, \dots, f_n \prec U_n$ so that $f_1 + \cdots + f_n = 1$ on K.*

Proof. From the hypothesis, $K \setminus (U_2 \cup \cdots \cup U_n) \subseteq U_1$ so there is an open V_1 so that $\mathrm{cl}(V_1)$ is compact and $K \setminus (U_2 \cup \cdots \cup U_n) \subseteq V_1 \subseteq \mathrm{cl}(V_1) \subseteq U_1$.
Then $K \subseteq V_1 \cup U_2 \cup \cdots \cup U_n$ and, hence, $K \setminus (V_1 \cup U_3 \cup \cdots \cup U_n) \subseteq U_2$. So there is an open V_2 so that $\mathrm{cl}(V_2)$ is compact and $K \setminus (V_1 \cup U_3 \cup \cdots \cup U_n) \subseteq V_2 \subseteq \mathrm{cl}(V_2) \subseteq U_2$.
Then $K \subseteq V_1 \cup V_2 \cup U_3 \cup \cdots \cup U_n$. Continuing inductively, we replace one after the other the $U_1, \dots, U_n$ with open $V_1, \dots, V_n$ so that $\mathrm{cl}(V_1), \dots, \mathrm{cl}(V_n)$ are compact and $K \subseteq V_1 \cup \cdots \cup V_n$ and $\mathrm{cl}(V_j) \subseteq U_j$ for all j.
By Theorem 2.2, there are $g_1, \dots, g_n$ so that $g_j \prec U_j$ and $g_j = 1$ on $\mathrm{cl}(V_j)$ for all j. Also there exists $g_0 : X \to [0,1]$ so that $g_0 = 1$ on K and $g_0 = 0$ out of $V_1 \cup \cdots \cup V_n$.
We define $f_j = \frac{g_j}{1 - g_0 + g_1 + \cdots + g_n}$ for every $j = 1, \dots, n$.
If for any $x \in X$ the $g_0(x) = 0$ is not true, then $x \in V_1 \cup \cdots \cup V_n$ and then $g_j(x) = 1$ for some $j = 1, \dots, n$. Therefore, $1 - g_0 + g_1 + \cdots + g_n \geq 1$ on X and, hence, $f_1, \dots, f_n : X \to [0,1]$ are all continuous on X.
Clearly, $\mathrm{supp}(f_j) \subseteq \mathrm{supp}(g_j)$ and thus $f_j \prec U_j$ for all j. Also, $f_1 + \cdots + f_n = \frac{g_1 + \cdots + g_n}{1 - g_0 + g_1 + \cdots + g_n} = 1$ on K because $g_0 = 1$ on K. □

Definition. *Let K be compact and $U_1, \dots, U_n$ be open subsets of X and $K \subseteq U_1 \cup \cdots \cup U_n$. If $f_1 \prec U_1, \dots, f_n \prec U_n$ and $f_1 + \cdots + f_n = 1$ on K, then the collection $\{f_1, \dots, f_n\}$ is called a* ***partition of unity*** *for K relative to its open cover $\{U_1, \dots, U_n\}$.*

Theorem 7.18. *Let X be locally compact and Hausdorff and $\mu \in M_{\mathcal{R}}$. Then*

$$\|\mu\| = \sup\big\{ \big| \textstyle\int_X f \, d\mu \big| \, \big| \, f \in C_0, \|f\|_u \leq 1 \big\}.$$

Proof. For all $f \in C_0$ with $\|f\|_u \leq 1$, Lemma 7.5 implies that $|\int_X f \, d\mu| \leq \|f\|_u \|\mu\| \leq \|\mu\|$. Therefore, $\sup\big\{ \big| \int_X f \, d\mu \big| \, \big| \, f \in C_0, \|f\|_u \leq 1 \big\} \leq \|\mu\|$.
By the definition of $\|\mu\|$, there are pairwise disjoint Borel sets $A_1, \dots, A_n \subseteq X$ so that $\|\mu\| - \epsilon < |\mu(A_1)| + \cdots + |\mu(A_n)|$. Since μ is regular, for every j there is a compact $K_j \subseteq A_j$ so that $|\mu|(A_j \setminus K_j) < \frac{1}{n} \epsilon$. Therefore, $\|\mu\| - 2\epsilon < |\mu(K_1)| + \cdots + |\mu(K_n)|$. Since $K_1, \dots, K_n$ are pairwise disjoint, it is easy to prove that there are pairwise disjoint open $U_1, \dots, U_n$ so that $K_j \subseteq U_j$ for all j and, taking them smaller if we need to, we may assume that $|\mu|(U_j \setminus K_j) < \frac{1}{n} \epsilon$ for all j. Then for every j there is $f_j \prec U_j$ so that $f_j = 1$ on K_j.
Finally, we define $\kappa_j = \overline{\mathrm{sign}\big(\int_{U_j} f_j \, d\mu \big)}$ for each j and $f = \kappa_1 f_1 + \cdots + \kappa_n f_n$.
It is easy to see that $\|f\|_u \leq 1$. Therefore,

$$\begin{aligned} \big| \textstyle\int_X f \, d\mu \big| &= \big| \textstyle\sum_{j=1}^n \kappa_j \int_{U_j} f_j \, d\mu \big| = \textstyle\sum_{j=1}^n \big| \int_{U_j} f_j \, d\mu \big| \\ &\geq \textstyle\sum_{j=1}^n |\mu(K_j)| - \sum_{j=1}^n \big| \int_{U_j \setminus K_j} f_j \, d\mu \big| \\ &> \|\mu\| - 2\epsilon - \textstyle\sum_{j=1}^n |\mu|(U_j \setminus K_j) > \|\mu\| - 3\epsilon. \end{aligned}$$

Since $\epsilon > 0$ is arbitrary, we conclude that $\sup\big\{ \big| \int_X f \, d\mu \big| \, \big| \, f \in C_0, \|f\|_u \leq 1 \big\} \geq \|\mu\|$ and the proof is complete. □

Definition. *Let X be locally compact and Hausdorff. For every $\mu \in M_{\mathcal{R}}$ we define $l_\mu : C_0 \to F$ by*

$$l_\mu(f) = \textstyle\int_X f \, d\mu \qquad \textit{for all } f \in C_0.$$

Proposition 7.56. *Let X be locally compact and Hausdorff. For every $\mu \in M_{\mathcal{R}}$ the function l_μ belongs to $(C_0)^*$. Moreover, $\|l_\mu\|_* = \|\mu\|$.*

Proof. We have $l_\mu(f_1+f_2)=\int_X(f_1+f_2)\,d\mu=\int_X f_1\,d\mu+\int_X f_2\,d\mu=l_\mu(f_1)+l_\mu(f_2)$. Also, $l_\mu(\kappa f)=\int_X(\kappa f)\,d\mu=\kappa\int_X f\,d\mu=\kappa l_\mu(f)$. These imply that l_μ is a linear functional.
Theorem 7.16 together with Proposition 7.16 imply that $\|l_\mu\|_*=\|\mu\|$. □

Definition. *Let X be locally compact and Hausdorff. We define $J: M_{\mathcal{R}}\to(C_0)^*$ by*

$$J(\mu)=l_\mu \qquad \textit{for all } \mu\in M_{\mathcal{R}}.$$

Proposition 7.57. *The function J is an isometry from $M_{\mathcal{R}}$ into $(C_0)^*$*

Proof. Exercise. □

We recall that, if $F=\mathbb{R}$, then C_0 is a Banach lattice and that a linear functional $l: C_0\to\mathbb{R}$ is called *non-negative* if $l(f)\geq 0$ for every $f\in C_0$ such that $f\geq 0$ (i.e. $f(x)\geq 0$ for all $x\in X$).

F.Riesz-Radon-Banach-Kakutani Theorem. `The real case`. *Let $F=\mathbb{R}$ and X be locally compact and Hausdorff.*
(i) For every $l\in(C_0)^*$ there exists a unique regular real Borel measure μ on X so that $l=l_\mu$, i.e. so that $l(f)=\int_X f\,d\mu$ for all $f\in C_0$.
If l is non-negative, then μ is non-negative.
(ii) The function J is an isometry from $M_{\mathcal{R}}$ onto $(C_0)^*$.

Proof. (i) We consider first the case of a non-negative $l\in(C_0)^*$.
For each open $O\subseteq X$ we define

$$\mu(O)=\sup\{l(f)\,|\,f\prec O\}$$

and then for each $E\subseteq X$ we define

$$\mu^*(E)=\inf\{\mu(O)\,|\,O\text{ open }\supseteq E\}.$$

If O_1,O_2 are open and $O_1\subseteq O_2$, then $f\prec O_1$ implies $f\prec O_2$ and, thus, $\mu(O_1)\leq\mu(O_2)$. Hence, $\mu^*(O)=\mu(O)$ for each open O.
If $f\prec O$, then $l(f)\leq\|l\|_*\|f\|_u\leq\|l\|_*$. Therefore, $\mu(O)\leq\|l\|_*$ and, thus, $\mu^*(E)\leq\|l\|_*$ for every $E\subseteq X$.
It is obvious that $\mu^*(\emptyset)=\mu(\emptyset)=0$ and also that $\mu^*(E_1)\leq\mu^*(E_2)$ for all E_1,E_2 with $E_1\subseteq E_2$.
Let now $E=E_1\cup E_2\cup\cdots$. For each j we take an open $O_j\supseteq E_j$ so that $\mu(O_j)<\mu^*(E_j)+\frac{\epsilon}{2^j}$ and set $O=O_1\cup O_2\cup\cdots$. Let $f\prec O$ and then set $K=\mathrm{supp}(f)\subseteq O$. Then there is N so that $K\subseteq O_1\cup\cdots\cup O_N$ and we consider a partition of unity $\{f_1,\ldots,f_N\}$ for K relative to $\{O_1,\ldots,O_N\}$. Then $f=ff_1+\cdots+ff_N$ and $ff_j\prec O_j$ for each j and, hence,

$$l(f)=l(ff_1)+\cdots+l(ff_N)\leq\mu(O_1)+\cdots+\mu(O_N)\leq\mu(O_1)+\mu(O_N)+\cdots.$$

This implies that $\mu(O)\leq\mu(O_1)+\mu(O_N)+\cdots\leq\mu^*(E_1)+\mu^*(E_2)+\cdots+\epsilon$ and, since $E\subseteq O$, we get $\mu^*(E)\leq\mu^*(E_1)+\mu^*(E_2)+\cdots+\epsilon$ and, finally, $\mu^*(E)\leq\mu^*(E_1)+\mu^*(E_2)+\cdots$. We conclude that μ^* is an outer measure on X.
By the Caratheodory process we define the σ-algebra of μ^*-measurable subsets of X on which the restriction of μ^* is a measure.
Consider any open O and any E. We take an open $O'\supseteq E$ with $\mu(O')<\mu^*(E)+\epsilon$ and $f\prec O'\cap O$ so that $l(f)>\mu(O'\cap O)-\epsilon$. The set $O'\setminus\mathrm{supp}(f)$ is open and we take $g\prec O'\setminus\mathrm{supp}(f)$ so that $l(g)>\mu(O'\setminus\mathrm{supp}(f))-\epsilon$. We observe that $f+g\prec O'$, whence

$$\begin{aligned}\mu^*(E)+\epsilon>\mu(O')\geq l(f+g)=l(f)+l(g)&>\mu(O'\cap O)+\mu(O'\setminus\mathrm{supp}(f))-2\epsilon\\&\geq\mu^*(E\cap O)+\mu^*(E\setminus O)-2\epsilon.\end{aligned}$$

Hence $\mu^*(E) \geq \mu^*(E \cap O) + \mu^*(E \setminus O)$ and this means that O is μ^*-measurable. Therefore, the σ-algebra of μ^*-measurable sets contains all open sets and, thus, includes $\mathcal{B}_X$. We define μ to be the restriction of μ^* on $\mathcal{B}_X$. So μ is a non-negative Borel measure on X. Observe that μ is identical to the already defined μ on the open sets, since we proved that $\mu^*(O) = \mu(O)$ for each open O. We shall now prove that

$$\mu(K) = \inf\{l(f) \mid f \in C_0 \text{ and } \chi_K \leq f \text{ on } X\} \tag{7.13}$$

for all compact $K \subseteq X$. We take any $f \in C_0$ with $f \geq \chi_K$ (e.g. $f \geq 0$ on X and, in particular, $f \geq 1$ on K) and consider the open set $O = \{x \in X \mid f(x) > 1 - \epsilon\} \supseteq K$. If $g \prec O$, then $g \leq \frac{1}{1-\epsilon} f$ on X and then $l(g) \leq \frac{1}{1-\epsilon} l(f)$, since l is non-negative. Therefore, $\mu(O) \leq \frac{1}{1-\epsilon} l(f)$, whence $\mu(K) \leq \frac{1}{1-\epsilon} l(f)$. Since $\epsilon > 0$ is arbitrary, this implies that $\mu(K) \leq l(f)$ and, thus, $\mu(K) \leq \inf\{l(f) \mid f \in C_0 \text{ and } \chi_K \leq f \text{ on } X\}$. We now take an open $O \supseteq K$ with $\mu(O) < \mu(K) + \epsilon$ and then an $f \prec O$ so that $f = 1$ on K. Then $f \geq \chi_K$ and $l(f) \leq \mu(O) < \mu(K) + \epsilon$. Since ϵ is arbitrary, $\inf\{l(f) \mid f \in C_0 \text{ and } \chi_K \leq f \text{ on } X\} \leq \mu(K)$.
We shall next prove the regularity of μ.
For each Borel set E we have $\mu(E) = \mu^*(E) = \inf\{\mu(O) \mid O \text{ open } \supseteq E\}$ and this is the first regularity condition.
We take any Borel set E and find an open $O \supseteq E$ so that $\mu(O) < \mu(E) + \epsilon$. We then find $g \prec O$ so that $l(g) > \mu(O) - \epsilon$ and set $K = \text{supp}(g) \subseteq O$. For each $f \in C_0$ with $f \geq \chi_K$ we get that $f \geq g$ and then $l(f) \geq l(g)$. From (7.4) it is implied that $\mu(K) \geq l(g)$. Therefore, we have a compact $K \subseteq O$ with $\mu(K) > \mu(O) - \epsilon$. Since $\mu(O \setminus E) = \mu(O) - \mu(E) < \epsilon$, there is an open $O' \supseteq O \setminus E$ so that $\mu(O') < 2\epsilon$. We now define $L = K \setminus O'$ and observe that L is a compact subset of E and that $E \setminus L \subseteq (O \setminus K) \cup O'$. Thus, $\mu(E) - \mu(L) \leq \mu(O \setminus K) + \mu(O') < 3\epsilon$ and, hence, $\mu(E) = \sup\{\mu(L) \mid L \text{ compact } \subseteq E\}$. This is the second regularity condition.
Finally we shall prove that $l(f) = \int_X f \, d\mu$ for every $f \in C_0$.
If f is real, we write $f = f^+ - f^-$, where $f^+ \geq 0$ and $f^- \geq 0$ are the non-negative and non-positive parts of f. Therefore, due to the linearity of l and of the integral, it is enough to consider $f \geq 0$ and, multiplying with an appropriate positive constant, we may assume that $f \in C_0$ and $0 \leq f \leq 1$ on X.
We take an arbitrary $N \in \mathbb{N}$ and define $K_k = \{x \in X \mid f(x) \geq \frac{k}{N}\}$ for $0 \leq k \leq N$. For each $k = 1, \ldots, N$ we have that K_k is compact and, obviously, $K_0 = X$. Also for each $j = 0, \ldots, N-1$ we define $f_j = \min\big\{\max\big\{f, \frac{j}{N}\big\}, \frac{j+1}{N}\big\} - \frac{j}{N}$. We have that $f_j \in C_0$ and $\frac{1}{N}\chi_{K_{j+1}} \leq f_j \leq \frac{1}{N}\chi_{K_j}$ for each $j = 0, \ldots, N-1$ and also $f = f_0 + f_1 + \cdots + f_{N-1}$. Adding the last inequalities and integrating, we find

$$\tfrac{1}{N}(\mu(K_1) + \cdots + \mu(K_N)) \leq \int_X f \, d\mu \leq \tfrac{1}{N}(\mu(K_0) + \cdots + \mu(K_{N-1})). \tag{7.14}$$

From $\chi_{K_{j+1}} \leq N f_j$ and (7.13) it is implied that $\mu(K_{j+1}) \leq l(N f_j) = N l(f_j)$. From $N f_j \leq \chi_{K_j}$ it is implied that $N f_j \prec O$ and, thus, $N l(f_j) \leq \mu(O)$ for every open $O \supseteq K_j$. Hence, from the definition of $\mu(K_j) = \mu^*(K_j)$ we get that $N l(f_j) \leq \mu(K_j)$. Therefore, $\frac{1}{N}\mu(K_{j+1}) \leq l(f_j) \leq \frac{1}{N}\mu(K_j)$ and, adding,

$$\tfrac{1}{N}(\mu(K_1) + \cdots + \mu(K_N)) \leq l(f) \leq \tfrac{1}{N}(\mu(K_0) + \cdots + \mu(K_{N-1})).$$

Thi and (7.14) imply

$$\begin{aligned} \left| \int_X f \, d\mu - l(f) \right| &\leq \tfrac{1}{N}(\mu(K_0) + \cdots + \mu(K_{N-1})) - \tfrac{1}{N}(\mu(K_1) + \cdots + \mu(K_N)) \\ &= \tfrac{1}{N}\mu(K_0 \setminus K_N) \leq \tfrac{1}{N}\mu(X) \leq \tfrac{1}{N}\|l\|_* \end{aligned}$$

and, since N is arbitrary, $l(f) = \int_X f \, d\mu$.
That μ is finite (and, hence, a real measure) is clear from the beginning of the proof. In fact, for every $f \prec X$ we have $l(f) \leq \|l\|_* \|f\|_u \leq \|l\|_*$ and, thus, $\mu(X) = \sup\{l(f) \mid f \prec X\} \leq \|l\|_*$.

Now we consider the case of a general $l \in (C_0)^*$.
Proposition 7.13 implies that there are non-negative $l^+, l^- \in (C_0)^*$ so that $l = l^+ - l^-$ and $\|l^+\|_* \leq \|l\|_*$, $\|l^-\|_* \leq \|l\|_*$. Now from the previous theorem we know that there are regular finite Borel measures μ^+ and μ^- on X so that $l^+(f) = \int_X f\, d\mu^+$ and $l^-(f) = \int_X f\, d\mu^-$ for every $f \in C_0$. Therefore, for the regular real Borel measure $\mu = \mu^+ - \mu^-$ we have $l(f) = l^+(f) - l^-(f) = \int_X f\, d\mu^+ - \int_X f\, d\mu^- = \int_X f\, d\mu$ for every $f \in C_0$.
To prove the uniqueness of μ, we assume that there are regular real Borel measures μ_1, μ_2 so that $l(f) = \int_X f\, d\mu_1 = \int_X f\, d\mu_2$ for all $f \in C_0$. We consider the regular real Borel measure $\mu = \mu_1 - \mu_2$ and then we have $\int_X f\, d\mu = 0$ for all $f \in C_0$. Theorem 7.16 implies that $\|\mu\| = 0$ and, hence, $\mu = 0$.
(ii) Clear after Proposition 7.51. □

F.Riesz-Radon-Banach-Kakutani Theorem. `The complex case.` *Let $F = \mathbb{C}$ and X be locally compact and Hausdorff.*
(i) For every $l \in (C_0)^*$ there exists a unique regular complex Borel measure μ on X so that $l = l_\mu$, i.e. so that $l(f) = \int_X f\, d\mu$ for all $f \in C_0$.
If l is non-negative (in other words if $l(f) \geq 0$ for every non-negative $f \in C_0$), then μ is non-negative.
If l is real (in other words if $l(f) \in \mathbb{R}$ for every real $f \in C_0$), then μ is real.
(ii) The function J is an isometry from $M_{\mathcal{R}}$ onto $(C_0)^*$.

Proof. (i) For the general $l \in (C_0)^*$ Proposition 7.18 implies that $\mathrm{Re}(l)$ is a bounded real-linear functional on C_0 with $\|\mathrm{Re}(l)\|_* = \|l\|_*$.
If we apply this to $-il \in (C_0)^*$ we get that also $\mathrm{Im}(l) = \mathrm{Re}(-il)$ is a bounded real-linear functional on C_0 with $\|\mathrm{Im}(l)\|_* = \|-il\|_* = \|l\|_*$.
Now from the previous theorem we know that there are regular real Borel measures μ_1, μ_2 on X so that $\mathrm{Re}(l)(f) = \int_X f\, d\mu_1$ and $\mathrm{Im}(l)(f) = \int_X f\, d\mu_2$ for every real $f \in C_0$. Therefore, if we define $\mu = \mu_1 + i\mu_2$, then μ is a regular complex Borel measure on X and for every real $f \in C_0$ we have $l(f) = \mathrm{Re}(l)(f) + i\,\mathrm{Im}(l)(f) = \int_X f\, d\mu_1 + i \int_X f\, d\mu_2 = \int_X f\, d\mu$. Therefore, for every $f \in C_0$, we get $l(f) = l(\mathrm{Re}(f)) + il(\mathrm{Im}(f)) = \int_X \mathrm{Re}(f)\, d\mu + i \int_X \mathrm{Im}(f)\, d\mu = \int_X f\, d\mu$.
If $l \in (C_0)^*$ is real, then $\mathrm{Im}(l)(f) = 0$ for all real $f \in C_0$. This implies that $\mu_2 = 0$ and, thus, $\mu = \mu_1$ is a real measure.
If $l \in (C_0)^*$ is non-negative, then for every real $f \in C_0$ we can write $f = f^+ - f^-$ with $f^+, f^- \geq 0$ on C_0. Since $l(f^+), l(f^-) \geq 0$, we get that $l(f) = l(f^+) - l(f^-)$ is real for every real $f \in C_0$. From the previous case we conclude that μ is a real measure and that $l(f) = \int_X f\, d\mu$ for every real $f \in C_0$. By the results of the previous theorem (including the uniqueness) we get that μ is a (non-negative) finite measure.
Again the uniqueness of μ is a consequence of Theorem 7.16.
(ii) Clear after Proposition 7.51. □

Finally, if X is locally compact and Hausdorff, then there is an isometry from $M_{\mathcal{R}}$ onto $(C_0)^*$. Now we may *identify* every $\mu \in M_{\mathcal{R}}$ with the corresponding $l_\mu \in (C_0)^*$ which is defined by $l_\mu(f) = \int_X f\, d\mu$ for all $f \in C_0$. We may view every $\mu \in M_{\mathcal{R}}$ as a bounded linear functional on C_0 and if we write μ instead of l_μ, then the defining relation of l_μ can be written

$$\mu(f) = \int_X f\, d\mu, \qquad f \in C_0, g \in M_{\mathcal{R}}.$$

Definition. *Let X be locally compact and Hausdorff and (μ_n) be a sequence in $M_{\mathcal{R}}$. We say that (μ_n)* ***converges weakly**** *to $\mu \in M_{\mathcal{R}}$ if $\int_X f\, d\mu_n \to \int_X f\, d\mu$ for all $f \in C_0$. In this case we write*

$$\mu_n \xrightarrow{w*} \mu.$$

If we identify every $\mu_n \in M_{\mathcal{R}}$ and $\mu \in M_{\mathcal{R}}$ with the corresponding $l_{\mu_n} \in (C_0)^*$ and $l_\mu \in (C_0)^*$ then $\int_X f\, d\mu_n \to \int_X f\, d\mu$ is equivalent to $l_{\mu_n}(f) \to l_\mu(f)$. Therefore, the definition we gave for $\mu_n \xrightarrow{w*} \mu$ in $M_{\mathcal{R}}$ is the same as the definition of $l_{\mu_n} \xrightarrow{w*} l_\mu$ in $(C_0)^*$ which is a special case of the definition of weak* convergence in the case of the general dual space.

Proposition 7.58. *Let X be locally compact and Hausdorff and (μ_n) be a sequence in $M_{\mathcal{R}}$ such that $\lim_{n\to+\infty}\int_X f\,d\mu_n$ exists in F for all $f\in C_0$. Then*

$$\sup_{n\in\mathbb{N}}\|\mu_n\|<+\infty.$$

Also there is a $\mu\in M_{\mathcal{R}}$ so that $\int_X f\,d\mu_n\to\int_X f\,d\mu$ for all $f\in C_0$ and

$$\|\mu\|\leq\underline{\lim}_{n\in\mathbb{N}}\|\mu_n\|.$$

Proof. A corollary of Proposition 8.3. □

Proposition 7.59. *Let X be locally compact and Hausdorff, ν be a non-negative element of $M_{\mathcal{R}}$ and (f_n) be a sequence in $L^1(\nu)$ such that $\lim_{n\to+\infty}\int_X f_n g\,d\nu$ exists in F for all $g\in C_0$. Then there is a $\mu\in M_{\mathcal{R}}$ so that $\int_X f_n g\,d\nu\to\int_X g\,d\mu$ for all $g\in C_0$. Also*

$$\|\mu\|\leq\underline{\lim}_{n\to+\infty}\int_X|f_n|\,d\nu,\qquad \int_X|f_n|\,d\nu<+\infty.$$

Proof. We consider the (real or complex) measures $\mu_n=f_n\nu$ on X. Then $\mu_n\in M_{\mathcal{R}}$ for all n and

$$\int_X f_n g\,d\nu=\int_X g\,d\mu_n,\qquad g\in C_0.$$

The rest is an application of Proposition 8.11. □

Proposition 7.60. *Let X be locally compact and Hausdorff and assume that there is a countable family P of open sets with the property: for every x and every open U with $x\in U$ there is a $W\in P$ so that $x\in W\subseteq \mathrm{cl}(W)\subseteq U$ and $\mathrm{cl}(W)$ is compact. Then C_0 is separable.*

Proof. Following the proofs of Lemma 2.2 and Theorem 2.2 we may easily prove that there is a countable set A of continuous functions with the property: for every compact K and every open U with $K\subseteq U$ there is an $f\in A$ so that $f\prec U$ and $f=1$ on K.
Now we take any $g\in C_0$ so that $0\leq g\leq 1$ on X and any $\epsilon>0$. We consider $N\in\mathbb{N}$ so that $\frac{1}{N}<\epsilon$.
We consider the sets $K_j=\{x\in X\mid \frac{j}{N}\leq g(x)\leq 1\}$ for $j=1,\dots,N$ and $U_j=\{x\in X\mid \frac{j}{N}< g(x)\leq 1\}$ for $j=1,\dots,N-1$. Let also $U_0=X$. Then every K_j is compact and every U_j is open and

$$K_N\subseteq U_{N-1}\subseteq K_{N-1}\subseteq U_{N-2}\subseteq\cdots\subseteq K_2\subseteq U_1\subseteq K_1\subseteq U_0.$$

Now we consider functions $f_1,f_2,\dots,f_N\in A$ so that

$$f_j\prec U_{j-1}\text{ and }f_j=1\text{ on }K_j\quad\text{for each }j=1,\dots,N.$$

Now it is easy to show that the function

$$f=\frac{1}{N}\sum_{j=1}^N f_j$$

satisfies $\|f-g\|_u\leq\frac{1}{N}<\epsilon$.
Finally it is straightforward to extend this result to all $g\in C_0$ and we leave this as an exercise. □

Proposition 7.61. *Let X be locally compact and Hausdorff and assume that there is a countable family P of open sets with the property: for every x and every open U with $x\in U$ there is a $W\in P$ so that $x\in W\subseteq\mathrm{cl}(W)\subseteq U$ and $\mathrm{cl}(W)$ is compact. If (μ_n) is a bounded sequence in $M_{\mathcal{R}}$, then there is a $\mu\in M_{\mathcal{R}}$ so that $\int_X f\,d\mu_n\to\int_X f\,d\mu$ for all $f\in C_0$.*

Proof. This is a corollary of Theorem 8.1 and Proposition 8.18. □

7.5 The spaces $B(X, \mathcal{S})$ and $L^\infty(X, \mathcal{S}, \mu)$ and their duals.

Definition. *Let $(X, \mathcal{S})$ be a measurable space. Then $B(X, \mathcal{S})$ is the space of all bounded measurable functions $f : X \to F$. We define*

$$\|f\|_u = \sup_{x \in X} |f(x)|, \qquad f \in B(X, \mathcal{S}).$$

It is clear that $B(X, \mathcal{S})$ is a linear space over F and that $\|\cdot\|_u$ is a norm on $B(X, \mathcal{S})$.

Example. If $\mathcal{S} = \mathcal{P}(X)$, then $B(X, \mathcal{S}) = B(X)$, i.e. the space of all bounded $f : X \to F$.

Proposition 7.62. *$B(X, \mathcal{S})$ is a Banach space. If $F = \mathbb{R}$, then $B(X, \mathcal{S})$ is a Banach lattice.*

Proof. Exercise. $\square$

Definition. *We denote by*

$$M_f(X, \mathcal{S})$$

the space of all finitely additive real or complex (depending on whether $F = \mathbb{R}$ or $F = \mathbb{C}$) measures on $(X, \mathcal{S})$.

1. Daniell integral.
2. Functions of bounded variation (in the chapter about signed and complex measures).
3. More exercises, especially for the last three chapters.
4. Probability. Probably not as a separate chapter. For example the notion of a probability measure, and more things (Kolmogorov's theorem etc) as exercises.
5. The Hilbert space structure of L^2, orthonormal bases (like the e^{inx}) etc.
6. Haar measure.