NOTES ON

MEASURE THEORY

M. Papadimitrakis
Department of Mathematics
University of Crete

Autumn of 2004

Contents

Chapter 1

σ-algebras

1.1 σ-algebras.

Definition 1.1 *Let X be a non-empty set and Σ a collection of subsets of X. We call Σ a* ***σ-algebra of subsets of*** *X if it is non-empty, closed under complements and closed under countable unions. This means:*
(i) there exists at least one $A \subseteq X$ so that $A \in \Sigma$,
(ii) if $A \in \Sigma$, then $A^c \in \Sigma$, where $A^c = X \setminus A$, and
(iii) if $A_n \in \Sigma$ for all $n \in \boldsymbol{N}$, then $\cup_{n=1}^{+\infty} A_n \in \Sigma$.

The pair (X, Σ) of a non-empty set X and a σ-algebra Σ of subsets of X is called ***a measurable space****.*

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

Proof: Let Σ be any σ-algebra of subsets of X.
(a) Take any $A \in \Sigma$ and consider the sets $A_1 = A$ and $A_n = A^c$ for all $n \geq 2$. Then $X = A \cup A^c = \cup_{n=1}^{+\infty} A_n \in \Sigma$ and also $\emptyset = X^c \in \Sigma$.
(b) Let $A_1, \ldots, A_N \in \Sigma$. Consider $A_n = A_N$ for all $n > N$ and get that $\cup_{n=1}^{N} A_n = \cup_{n=1}^{+\infty} A_n \in \Sigma$.
(c) Let $A_n \in \Sigma$ for all n. Then $\cap_{n=1}^{+\infty} A_n = (\cup_{n=1}^{+\infty} A_n^c)^c \in \Sigma$.
(d) Let $A_1, \ldots, A_N \in \Sigma$. Using the result of (b), we get that $\cap_{n=1}^{N} A_n = (\cup_{n=1}^{N} A_n^c)^c \in \Sigma$.
(e) Finally, let $A, B \in \Sigma$. Using the result of (d), we get that $A \setminus B = A \cap B^c \in \Sigma$.

Here are some simple examples.

Examples
1. The collection $\{\emptyset, X\}$ is a σ-algebra of subsets of X.
2. If $E \subseteq X$ is non-empty and different from X, then the collection $\{\emptyset, E, E^c, X\}$ is a σ-algebra of subsets of X.

3. $\mathcal{P}(X)$, the collection of all subsets of X, is a σ-algebra of subsets of X.
4. Let X be uncountable. The $\{A \subseteq X \,|\, A \text{ is countable or } A^c \text{ is countable}\}$ is a σ-algebra of subsets of X. Firstly, $\emptyset$ is countable and, hence, the collection is non-empty. If A is in the collection, then, considering cases, we see that A^c is also in the collection. Finally, let A_n be in the collection for all $n \in \mathbf{N}$. If all A_n's are countable, then $\cup_{n=1}^{+\infty} A_n$ is also countable. If at least one of the A_n^c's, say $A_{n_0}^c$, is countable, then $(\cup_{n=1}^{+\infty} A_n)^c \subseteq A_{n_0}^c$ is also countable. In any case, $\cup_{n=1}^{+\infty} A_n$ belongs to the collection.

The following result is useful.

Proposition 1.2 *Let Σ be a σ-algebra of subsets of X and consider a finite sequence $\{A_n\}_{n=1}^N$ or an infinite sequence $\{A_n\}$ in Σ. Then there exists a finite sequence $\{B_n\}_{n=1}^N$ or, respectively, an infinite sequence $\{B_n\}$ in Σ with the properties:*
(i) $B_n \subseteq A_n$ for all $n = 1, \dots, N$ or, respectively, all $n \in \mathbf{N}$.
(ii) $\cup_{n=1}^N B_n = \cup_{n=1}^N A_n$ or, respectively, $\cup_{n=1}^{+\infty} B_n = \cup_{n=1}^{+\infty} A_n$.
(iii) the B_n's are pairwise disjoint.

Proof: Trivial, by taking $B_1 = A_1$ and $B_k = A_k \setminus (A_1 \cup \cdots \cup A_{k-1})$ for all $k = 2, \dots, N$ or, respectively, all $k = 2, 3, \dots$.

1.2 Generated σ-algebras.

Proposition 1.3 *The intersection of any σ-algebras of subsets of the same X is a σ-algebra of subsets of X.*

Proof: Let $\{\Sigma_i\}_{i \in I}$ be any collection of σ-algebras of subsets of X, indexed by an arbitrary non-empty set I of indices, and consider the intersection $\Sigma = \cap_{i \in I} \Sigma_i$.

Since $\emptyset \in \Sigma_i$ for all $i \in I$, we get $\emptyset \in \Sigma$ and, hence, Σ is non-empty.

Let $A \in \Sigma$. Then $A \in \Sigma_i$ for all $i \in I$ and, since all Σ_i's are σ-algebras, $A^c \in \Sigma_i$ for all $i \in I$. Therefore $A^c \in \Sigma$.

Let $A_n \in \Sigma$ for all $n \in \mathbf{N}$. Then $A_n \in \Sigma_i$ for all $i \in I$ and all $n \in \mathbf{N}$ and, since all Σ_i's are σ-algebras, we get $\cup_{n=1}^{+\infty} A_n \in \Sigma_i$ for all $i \in I$. Thus, $\cup_{n=1}^{+\infty} A_n \in \Sigma$.

Definition 1.2 *Let X be a non-empty set and $\mathcal{E}$ be an arbitrary collection of subsets of X. The intersection of all σ-algebras of subsets of X which include $\mathcal{E}$ is called* ***the σ-algebra generated by*** *$\mathcal{E}$ and it is denoted by $\Sigma(\mathcal{E})$. Namely*

$$\Sigma(\mathcal{E}) = \cap \{\Sigma \,|\, \Sigma \textit{ is a } \sigma\textit{-algebra of subsets of } X \textit{ and } \mathcal{E} \subseteq \Sigma\}\,.$$

Note that there is at least one σ-algebra of subsets of X which includes $\mathcal{E}$ and this is $\mathcal{P}(X)$. Note also that the term *σ-algebra* used in the name of $\Sigma(\mathcal{E})$ is justified by its definition and by Proposition 1.3.

Proposition 1.4 *Let $\mathcal{E}$ be any collection of subsets of the non-empty X. Then $\Sigma(\mathcal{E})$ is the smallest σ-algebra of subsets of X which includes $\mathcal{E}$. Namely, if Σ is any σ-algebra of subsets of X such that $\mathcal{E} \subseteq \Sigma$, then $\Sigma(\mathcal{E}) \subseteq \Sigma$.*

Proof: If Σ is any σ-algebra of subsets of X such that $\mathcal{E} \subseteq \Sigma$, then Σ is one of the σ-algebras whose intersection is denoted $\Sigma(\mathcal{E})$. Therefore $\Sigma(\mathcal{E}) \subseteq \Sigma$.

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

Examples.
1. Let $E \subseteq X$ and E be non-empty and different from X and consider $\mathcal{E} = \{E\}$. Then $\Sigma(\mathcal{E}) = \{\emptyset, E, E^c, X\}$. To see this just observe that $\{\emptyset, E, E^c, X\}$ is a σ-algebra of subsets of X which contains E and that there can be no smaller σ-algebra of subsets of X containing E, since such a σ-algebra must necessarily contain $\emptyset, X$ and E^c besides E.
2. Let X be an uncountable set and consider $\mathcal{E} = \{A \subseteq X \mid A \text{ is countable}\}$. Then $\Sigma(\mathcal{E}) = \{A \subseteq X | A \text{ is countable or } A^c \text{ is countable}\}$. The argument is the same as before. $\{A \subseteq X | A \text{ is countable or } A^c \text{ is countable}\}$ is a σ-algebra of subsets of X which contains all countable subsets of X and there is no smaller σ-algebra of subsets of X containing all countable subsets of X, since any such σ-algebra must contain all the complements of countable subsets of X.

1.3 Algebras and monotone classes.

Definition 1.3 *Let X be non-empty and $\mathcal{A}$ 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 $A^c \in \mathcal{A}$ and
(iii) if $A, B \in \mathcal{A}$, then $A \cup B \in \mathcal{A}$.

Proposition 1.5 *Every algebra of subsets of X contains at least the sets $\emptyset$ and X, it is closed under finite unions, under finite intersections and under set-theoretic differences.*

Proof: Let $\mathcal{A}$ be any algebra of subsets of X.
(a) Take any $A \in \mathcal{A}$ and consider the sets A and A^c. Then $X = A \cup A^c \in \mathcal{A}$ and then $\emptyset = X^c \in \mathcal{A}$.
(b) It is trivial to prove by induction that for any $n \in \mathbf{N}$ and any $A_1, \ldots, A_n \in \mathcal{A}$ it follows $A_1 \cup \cdots \cup A_n \in \mathcal{A}$.
(c) By the result of (b), if $A_1, \ldots, A_n \in \mathcal{A}$, then $\cap_{k=1}^n A_k = (\cup_{k=1}^n A_k^c)^c \in \mathcal{A}$.
(d) If $A, B \in \mathcal{A}$, using the result of (c), we get that $A \setminus B = A \cap B^c \in \mathcal{A}$.

Examples.
1. Every σ-algebra is also an algebra.
2. If X is an infinite set then the collection $\{A \subseteq X \mid A \text{ is finite or } A^c \text{ is finite}\}$ is an algebra of subsets of X.

If (A_n) is a sequence of subsets of a set X and $A_n \subseteq A_{n+1}$ for all n, we say that the sequence is *increasing*. In this case, if $A = \cup_{n=1}^{+\infty} A_n$, we write

$$A_n \uparrow A.$$

If $A_{n+1} \subseteq A_n$ for all n, we say that the sequence (A_n) is *decreasing* and, if also $A = \cap_{n=1}^{+\infty} A_n$, we write

$$A_n \downarrow A.$$

Definition 1.4 *Let X be a non-empty set and $\mathcal{M}$ 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 closed under countable decreasing intersections. That is, if $A_1, A_2, \ldots \in \mathcal{M}$ and $A_n \uparrow A$, then $A \in \mathcal{M}$ and, if $A_1, A_2, \ldots \in \mathcal{M}$ and $A_n \downarrow A$, then $A \in \mathcal{M}$.*

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

Proposition 1.6 *The intersection of any monotone classes of subsets of the same set X is a monotone class of subsets of X.*

Proof: Let $\{\mathcal{M}_i\}_{i \in I}$ be any collection of monotone classes of subsets of X, indexed by an arbitrary non-empty set I of indices, and consider the intersection $\mathcal{M} = \cap_{i \in I} \mathcal{M}_i$.

Let $A_1, A_2, \ldots \in \mathcal{M}$ with $A_n \uparrow A$. Then $A_n \in \mathcal{M}_i$ for all $i \in I$ and all $n \in \mathbf{N}$ and, since all $\mathcal{M}_i$'s are monotone classes, we get that $A \in \mathcal{M}_i$ for all $i \in I$. Therefore $A \in \mathcal{M}$.

The proof in the case of a countable decreasing intersection is identical.

Definition 1.5 *Let X be a non-empty set and $\mathcal{E}$ be an arbitrary collection of subsets of X. Then the intersection of all monotone classes of subsets of X which include $\mathcal{E}$ is called* ***the monotone class generated by*** *$\mathcal{E}$ and it is denoted by $\mathcal{M}(\mathcal{E})$. Namely*

$$\mathcal{M}(\mathcal{E}) = \cap\{\mathcal{M} \,|\, \mathcal{M} \text{ is a monotone class of subsets of } X \text{ and } \mathcal{E} \subseteq \mathcal{M}\}.$$

There is at least one monotone class including $\mathcal{E}$ and this is $\mathcal{P}(X)$. Also note that the term monotone class, used for $\mathcal{M}(\mathcal{E})$, is justified by Proposition 1.6.

Proposition 1.7 *Let $\mathcal{E}$ be any collection of subsets of the non-empty X. Then $\mathcal{M}(\mathcal{E})$ is the smallest monotone class of subsets of X which includes $\mathcal{E}$. Namely, if $\mathcal{M}$ is any monotone class of subsets of X such that $\mathcal{E} \subseteq \mathcal{M}$, then $\mathcal{M}(\mathcal{E}) \subseteq \mathcal{M}$.*

Proof: If $\mathcal{M}$ is any monotone class of subsets of X such that $\mathcal{E} \subseteq \mathcal{M}$, then $\mathcal{M}$ is one of the monotone classes whose intersection is $\mathcal{M}(\mathcal{E})$. Thus, $\mathcal{M}(\mathcal{E}) \subseteq \mathcal{M}$.

Theorem 1.1 *Let X be a non-empty set and $\mathcal{A}$ an algebra of subsets of X. Then $\mathcal{M}(\mathcal{A}) = \Sigma(\mathcal{A})$.*

Proof: $\Sigma(\mathcal{A})$ is a σ-algebra and, hence, a monotone class. Since $\mathcal{A} \subseteq \Sigma(\mathcal{A})$, Proposition 1.7 implies $\mathcal{M}(\mathcal{A}) \subseteq \Sigma(\mathcal{A})$.

Now it is enough to prove that $\mathcal{M}(\mathcal{A})$ is a σ-algebra. Since $\mathcal{A} \subseteq \mathcal{M}(\mathcal{A})$, Proposition 1.4 will immediately imply that $\Sigma(\mathcal{A}) \subseteq \mathcal{M}(\mathcal{A})$ and this will conclude the proof.

$\mathcal{M}(\mathcal{A})$ is non-empty because $\emptyset \in \mathcal{A} \subseteq \mathcal{M}(\mathcal{A})$.

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 $\mathcal{M}_A$ includes $\mathcal{A}$ and that it is a monotone class of subsets of X. In fact, if $B \in \mathcal{A}$ then $A \cup B \in \mathcal{A}$ and thus $B \in \mathcal{M}_A$. Also, if $B_1, B_2, \ldots \in \mathcal{M}_A$ and $B_n \uparrow B$, then $A \cup B_1, A \cup B_2, \ldots \in \mathcal{M}(\mathcal{A})$ and $A \cup B_n \uparrow A \cup B$. Since $\mathcal{M}(\mathcal{A})$ is a monotone class, we find that $A \cup B \in \mathcal{M}(\mathcal{A})$. Thus, $B \in \mathcal{M}_A$ and $\mathcal{M}_A$ is closed under countable increasing unions. In a similar way we can prove that $\mathcal{M}_A$ is closed under countable decreasing intersections and we conclude that it is a monotone class.

Proposition 1.7 implies that $\mathcal{M}(\mathcal{A}) \subseteq \mathcal{M}_A$. This means that:

i. $A \cup B \in \mathcal{M}(\mathcal{A})$ for all $A \in \mathcal{A}$ and all $B \in \mathcal{M}(\mathcal{A})$.

Now fix any $B \in \mathcal{M}(\mathcal{A})$ and consider $\mathcal{M}_B = \{A \subseteq X \mid A \cup B \in \mathcal{M}(\mathcal{A})\}$. As before, $\mathcal{M}_B$ is a monotone class of subsets of X and, by i, it includes $\mathcal{A}$. Again, Proposition 1.7 implies $\mathcal{M}(\mathcal{A}) \subseteq \mathcal{M}_B$, which means:

ii. $A \cup B \in \mathcal{M}(\mathcal{A})$ for all $A \in \mathcal{M}(\mathcal{A})$ and all $B \in \mathcal{M}(\mathcal{A})$.

We consider the collection $\mathcal{M} = \{A \subseteq X \mid A^c \in \mathcal{M}(\mathcal{A})\}$. As before, we can show that $\mathcal{M}$ is a monotone class of subsets of X and that it includes $\mathcal{A}$. Therefore, $\mathcal{M}(\mathcal{A}) \subseteq \mathcal{M}$, which means:

iii. $A^c \in \mathcal{M}(\mathcal{A})$ for all $A \in \mathcal{M}(\mathcal{A})$.

It is implied by ii and iii that $\mathcal{M}(\mathcal{A})$ is closed under finite unions and under complements.

Now take $A_1, A_2, \ldots \in \mathcal{M}(\mathcal{A})$ and define $B_n = A_1 \cup \cdots \cup A_n$ for all n. From ii we have that $B_n \in \mathcal{M}(\mathcal{A})$ for all n and it is clear that $B_n \subseteq B_{n+1}$ for all n. Since $\mathcal{M}(\mathcal{A})$ is a monotone class, $\cup_{n=1}^{+\infty} A_n = \cup_{n=1}^{+\infty} B_n \in \mathcal{M}(\mathcal{A})$.

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

1.4 Restriction of a σ-algebra.

Proposition 1.8 *Let Σ be a σ-algebra of subsets of X and $Y \subseteq X$ be non-empty. If we denote*

$$\Sigma \rceil Y = \{A \cap Y \mid A \in \Sigma\},$$

then $\Sigma \rceil Y$ is a σ-algebra of subsets of Y.

In case $Y \in \Sigma$, we have $\Sigma \rceil Y = \{A \subseteq Y \mid A \in \Sigma\}$.

Proof: Since $\emptyset \in \Sigma$, we have that $\emptyset = \emptyset \cap Y \in \Sigma \rceil Y$.

If $B \in \Sigma \rceil Y$, then $B = A \cap Y$ for some $A \in \Sigma$. Since $X \setminus A \in \Sigma$, we get that $Y \setminus B = (X \setminus A) \cap Y \in \Sigma \rceil Y$.

If $B_1, B_2, \ldots \in \Sigma \rceil Y$, then, for each k, $B_k = A_k \cap Y$ for some $A_k \in \Sigma$. Since $\cup_{k=1}^{+\infty} A_k \in \Sigma$, we find that $\cup_{k=1}^{+\infty} B_k = (\cup_{k=1}^{+\infty} A_k) \cap Y \in \Sigma \rceil Y$.

Now let $Y \in \Sigma$. If $B \in \Sigma\rceil Y$, then $B = A \cap Y$ for some $A \in \Sigma$ and, hence, $B \subseteq Y$ and $B \in \Sigma$. Therefore $B \in \{C \subseteq Y \,|\, C \in \Sigma\}$. Conversely, if $B \in \{C \subseteq Y \,|\, C \in \Sigma\}$, then $B \subseteq Y$ and $B \in \Sigma$. We set $A = B$ and, thus, $B = A \cap Y$ and $A \in \Sigma$. Therefore $B \in \Sigma\rceil Y$.

Definition 1.6 *Let Σ be a σ-algebra of subsets of X and let $Y \subseteq X$ be non-empty. The σ-algebra $\Sigma\rceil Y$, defined in Proposition 1.8, is called **the restriction of Σ on Y**.*

In general, if $\mathcal{E}$ is any collection of subsets of X and $Y \subseteq X$, we denote

$$\mathcal{E}\rceil Y = \{A \cap Y \,|\, A \in \mathcal{E}\}$$

and call $\mathcal{E}\rceil Y$ **the restriction of $\mathcal{E}$ on** Y.

Theorem 1.2 *Let $\mathcal{E}$ be a collection of subsets of X and $Y \subseteq X$ be non-empty. Then*

$$\Sigma(\mathcal{E}\rceil Y) = \Sigma(\mathcal{E})\rceil Y,$$

where $\Sigma(\mathcal{E}\rceil Y)$ is the σ-algebra of subsets of Y generated by $\mathcal{E}\rceil Y$.

Proof: If $B \in \mathcal{E}\rceil Y$, then $B = A \cap Y$ for some $A \in \mathcal{E} \subseteq \Sigma(\mathcal{E})$ and, thus, $B \in \Sigma(\mathcal{E})\rceil Y$. Hence, $\mathcal{E}\rceil Y \subseteq \Sigma(\mathcal{E})\rceil Y$ and, since, by Proposition 1.8, $\Sigma(\mathcal{E})\rceil Y$ is a σ-algebra of subsets of Y, Proposition 1.4 implies $\Sigma(\mathcal{E}\rceil Y) \subseteq \Sigma(\mathcal{E})\rceil Y$.

Now, define the collection

$$\Sigma = \{A \subseteq X \,|\, A \cap Y \in \Sigma(\mathcal{E}\rceil Y)\}.$$

We have that $\emptyset \in \Sigma$, because $\emptyset \cap Y = \emptyset \in \Sigma(\mathcal{E}\rceil Y)$.

If $A \in \Sigma$, then $A \cap Y \in \Sigma(\mathcal{E}\rceil Y)$. Therefore, $X \setminus A \in \Sigma$, because $(X \setminus A) \cap Y = Y \setminus (A \cap Y) \in \Sigma(\mathcal{E}\rceil Y)$.

If $A_1, A_2, \ldots \in \Sigma$, then $A_1 \cap Y, A_2 \cap Y, \ldots \in \Sigma(\mathcal{E}\rceil Y)$. This implies that $(\cup_{k=1}^{+\infty} A_k) \cap Y = \cup_{k=1}^{+\infty}(A_k \cap Y) \in \Sigma(\mathcal{E}\rceil Y)$ and, thus, $\cup_{k=1}^{+\infty} A_k \in \Sigma$.

We conclude that Σ is a σ-algebra of subsets of X.

If $A \in \mathcal{E}$, then $A \cap Y \in \mathcal{E}\rceil Y \subseteq \Sigma(\mathcal{E}\rceil Y)$ and, hence, $A \in \Sigma$. Therefore, $\mathcal{E} \subseteq \Sigma$ and, by Proposition 1.4, $\Sigma(\mathcal{E}) \subseteq \Sigma$. Now, for an arbitrary $B \in \Sigma(\mathcal{E})\rceil Y$, we have that $B = A \cap Y$ for some $A \in \Sigma(\mathcal{E}) \subseteq \Sigma$ and, thus, $B \in \Sigma(\mathcal{E}\rceil Y)$. This implies that $\Sigma(\mathcal{E})\rceil Y \subseteq \Sigma(\mathcal{E}\rceil Y)$.

1.5 Borel σ-algebras.

Definition 1.7 *Let X be a topological space and $\mathcal{T}$ 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 of subsets of X containing all open subsets of X, is called **the Borel σ-algebra of** X and we denote it $\mathcal{B}_X$:*

$$\mathcal{B}_X = \Sigma(\mathcal{T}), \qquad \mathcal{T} \text{ the topology of } X.$$

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

By definition, all open subsets of X are Borel sets in X and, since $\mathcal{B}_X$ is a σ-algebra, all closed subsets of X (which are the complements of open subsets) are also Borel sets in X. A subset of X is called a G_δ**-set** if it is a countable intersection of open subsets of X. Also, a subset of X is called an F_σ**-set** if it is a countable union of closed subsets of X. It is obvious that all G_δ-sets and all F_σ-sets are Borel sets in X.

Proposition 1.9 *If X is a topological space and $\mathcal{F}$ is the collection of all closed subsets of X, then $\mathcal{B}_X = \Sigma(\mathcal{F})$.*

Proof: Every closed set is contained in $\Sigma(\mathcal{T})$. This is true because $\Sigma(\mathcal{T})$ contains all open sets and hence, being a σ-algebra, contains all closed sets. Therefore, $\mathcal{F} \subseteq \Sigma(\mathcal{T})$. Since $\Sigma(\mathcal{T})$ is a σ-algebra, Proposition 1.4 implies $\Sigma(\mathcal{F}) \subseteq \Sigma(\mathcal{T})$.

Symmetrically, every open set is contained in $\Sigma(\mathcal{F})$. This is because $\Sigma(\mathcal{F})$ contains all closed sets and hence, being a σ-algebra, contains all open sets (the complements of closed sets). Therefore, $\mathcal{T} \subseteq \Sigma(\mathcal{F})$. Since $\Sigma(\mathcal{F})$ is a σ-algebra, Proposition 1.4 implies $\Sigma(\mathcal{T}) \subseteq \Sigma(\mathcal{F})$.

Therefore, $\Sigma(\mathcal{F}) = \Sigma(\mathcal{T}) = \mathcal{B}_X$.

If X is a topological space with the 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.

Theorem 1.3 *Let X be a topological space and let the non-empty $Y \subseteq X$ have the subspace topology. Then*

$$\mathcal{B}_Y = \mathcal{B}_X \rceil Y.$$

Proof: If $\mathcal{T}$ is the topology of X, then $\mathcal{T}\rceil Y$ is the subspace topology of Y. Theorem 1.2 implies that $\mathcal{B}_Y = \Sigma(\mathcal{T}\rceil Y) = \Sigma(\mathcal{T})\rceil Y = \mathcal{B}_X\rceil Y$.

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

Examples of topological spaces are all metric spaces of which the most familiar is the euclidean space $X = \mathbf{R}^n$ with the usual euclidean metric or even any subset X of $\mathbf{R}^n$ with the restriction on X of the euclidean metric. Because of the importance of $\mathbf{R}^n$ we shall pay particular attention on $\mathcal{B}_{\mathbf{R}^n}$.

The typical closed *orthogonal parallelepiped with axis-parallel edges* is a set of the form $Q = [a_1, b_1] \times \cdots \times [a_n, b_n]$, the typical open orthogonal parallelepiped with axis-parallel edges is a set of the form $R = (a_1, b_1) \times \cdots \times (a_n, b_n)$, the typical open-closed orthogonal parallelepiped with axis-parallel edges is a set of the form $P = (a_1, b_1] \times \cdots \times (a_n, b_n]$ and the typical closed-open orthogonal parallelepiped with axis-parallel edges is a set of the form $T = [a_1, b_1) \times \cdots \times [a_n, b_n)$. More generally, the typical orthogonal parallelepiped with axis-parallel edges is a set S, a cartesian product of n bounded intervals of any possible type. In all cases we consider $-\infty < a_j \leq b_j < +\infty$ for all $j = 1, \ldots, n$ and, hence, all orthogonal parallelepipeds with axis-parallel edges are bounded sets in $\mathbf{R}^n$.

If $n = 1$, then the orthogonal parallelepipeds with axis-parallel edges are just the bounded intervals of all possible types in the real line $\mathbf{R}$. If $n = 2$, then the orthogonal parallelepipeds with axis-parallel edges are the usual orthogonal parallelograms of all possible types with axis-parallel sides.

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 $\mathbf{R}^n$**.

Lemma 1.1 *All n-dimensional intervals are Borel sets in $\mathbf{R}^n$.*

Proof: For any $j = 1, \ldots, n$, a half-space of the form $\{x = (x_1, \ldots, x_n) \mid x_j < b_j\}$ or of the form $\{x = (x_1, \ldots, x_n) \mid x_j \leq b_j\}$ is a Borel set in $\mathbf{R}^n$, since it is an open set in the first case and a closed set in the second case. Similarly, a half-space of the form $\{x = (x_1, \ldots, x_n) \mid a_j < x_j\}$ or of the form $\{x = (x_1, \ldots, x_n) \mid a_j \leq x_j\}$ is a Borel set in $\mathbf{R}^n$. Now, every interval S is an intersection of $2n$ of these half-spaces and, therefore, it is also a Borel set in $\mathbf{R}^n$.

Proposition 1.10 *If $\mathcal{E}$ is the collection of all closed or of all open or of all open-closed or of all closed-open or of all intervals in $\mathbf{R}^n$, then $\mathcal{B}_{\mathbf{R}^n} = \Sigma(\mathcal{E})$.*

Proof: By Lemma 1.1 we have that, in all cases, $\mathcal{E} \subseteq \mathcal{B}_{\mathbf{R}^n}$. Proposition 1.4 implies that $\Sigma(\mathcal{E}) \subseteq \mathcal{B}_{\mathbf{R}^n}$.

To show the opposite inclusion consider any open subset U of $\mathbf{R}^n$. For every $x \in U$ find a small open ball B_x centered at x which is included in U. Now, considering the case of $\mathcal{E}$ being the collection of all closed intervals, take an arbitrary $Q_x = [a_1, b_1] \times \cdots \times [a_n, b_n]$ containing x, 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 $U = \cup_{x \in U} Q_x$. But the collection of all possible Q_x's is countable (!) and, thus, the general open subset U of $\mathbf{R}^n$ can be written as a countable union of sets in the collection $\mathcal{E}$. Hence every open U belongs to $\Sigma(\mathcal{E})$ and, since $\Sigma(\mathcal{E})$ is a σ-algebra of subsets of $\mathbf{R}^n$ and $\mathcal{B}_{\mathbf{R}^n}$ is generated by the collection of all open subsets of $\mathbf{R}^n$, Proposition 1.4 implies that $\mathcal{B}_{\mathbf{R}^n} \subseteq \Sigma(\mathcal{E})$.

Of course, the proof of the last inclusion works in the same way with all other types of intervals.

As we said, the intervals in $\mathbf{R}^n$ are cartesian products of n *bounded* intervals in $\mathbf{R}$. If we allow these intervals in $\mathbf{R}$ to become unbounded, we get the so-called **generalized intervals** in $\mathbf{R}^n$, namely all sets of the form $I_1 \times \cdots \times I_n$, where each I_j is any, even unbounded, interval in $\mathbf{R}$. Again, we have the subcollections of all open or all closed or all open-closed or all closed-open generalized intervals. For example, the typical open-closed generalized interval in $\mathbf{R}^n$ is of the form $P = (a_1, b_1] \times \cdots \times (a_n, b_n]$, where $-\infty \leq a_j \leq b_j \leq +\infty$ for all j. The whole space $\mathbf{R}^n$ is an open-closed generalized interval, as well as any of the half spaces $\{x = (x_1, \ldots, x_n) \mid x_j \leq b_j\}$ and $\{x = (x_1, \ldots, x_n) \mid a_j < x_j\}$. In fact, every open-closed generalized interval is, obviously, the intersection of $2n$ such half-spaces.

Proposition 1.11 *The collection* $\mathcal{A} = \{P_1 \cup \cdots \cup P_k \,|\, k \in \mathbf{N}, P_1, \ldots, P_k$ *are pairwise disjoint open-closed generalized intervals in* $\mathbf{R}^n\}$ *is an algebra of subsets of* $\mathbf{R}^n$.
In particular, the following are true:
(i) The intersection of two open-closed generalized intervals is an open-closed generalized interval.
(ii) For all open-closed generalized intervals $P, P_1, \ldots, P_m$ *there are pairwise disjoint open-closed generalized intervals* $P'_1, \ldots, P'_k$ *so that* $P \setminus (P_1 \cup \cdots \cup P_m) = P'_1 \cup \cdots \cup P'_k$.
(iii) For all open-closed generalized intervals $P_1, \ldots, P_m$ *there are pairwise disjoint open-closed generalized intervals* $P'_1, \ldots, P'_k$ *so that* $P_1 \cup \cdots \cup P_m = P'_1 \cup \cdots \cup P'_k$.

Proof: The intervals $(a, b]$ and $(a', b']$ are not disjoint if and only if $a'' < b''$, where $a'' = \max\{a, a'\}$ and $b'' = \min\{b, b'\}$. In case $a'' < b''$, then $(a, b] \cap (a', b'] = (a'', b'']$. Now if $P = (a_1, b_1] \times \cdots \times (a_n, b_n]$ and $P' = (a'_1, b'_1] \times \cdots \times (a'_n, b'_n]$, then P and P' are not disjoint if and only if for all $j = 1, \ldots, n$ we have that $(a_j, b_j]$ and $(a'_j, b'_j]$ are not disjoint. Hence if P, P' 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 $P \cap P' = P''$, where $P'' = (a''_1, b''_1] \times \cdots \times (a''_n, b''_n]$. This proves (i).

If $A = \cup_{i=1}^k P_i$, where the $P_1, \ldots, P_k$ are pairwise disjoint, and $A' = \cup_{j=1}^l P'_j$, where the $P'_1, \ldots, P'_l$ are also pairwise disjoint, are two elements of $\mathcal{A}$, 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 and they all are open-closed generalized intervals, as we have just seen.

Hence, $\mathcal{A}$ is closed under finite intersections.

Consider the open-closed generalized interval $P = (a_1, b_1] \times \cdots \times (a_n, b_n]$. It is easy to see that P^c can be written as the union of $2n$ (some may be empty) pairwise disjoint open-closed generalized 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 $\mathbf{R}$ (they may be empty). If we write $P = I_1 \times \cdots \times I_n$, then P^c is equal to

$$
\begin{array}{ccc}
I_1^{(l)} \times \mathbf{R} \times \cdots \times \mathbf{R} & \cup & I_1^{(r)} \times \mathbf{R} \times \cdots \times \mathbf{R} \quad \cup \\
I_1 \times I_2^{(l)} \times \mathbf{R} \times \cdots \times \mathbf{R} & \cup & I_1 \times I_2^{(r)} \times \mathbf{R} \times \cdots \times \mathbf{R} \quad \cup \\
 & \cdots & \\
I_1 \times \cdots \times I_{n-2} \times I_{n-1}^{(l)} \times \mathbf{R} & \cup & I_1 \times \cdots \times I_{n-2} \times I_{n-1}^{(r)} \times \mathbf{R} \quad \cup \\
I_1 \times \cdots \times I_{n-1} \times I_n^{(l)} & \cup & I_1 \times \cdots \times I_{n-1} \times I_n^{(r)}.
\end{array}
$$

Hence, for every open-closed generalized interval P the complement P^c is an element of $\mathcal{A}$.

Now, if $A = \cup_{i=1}^k P_i$, where the $P_1, \ldots, P_k$ are pairwise disjoint, is any element of $\mathcal{A}$, then $A^c = \cap_{i=1}^k P_i^c$ is a finite intersection of elements of $\mathcal{A}$. Since $\mathcal{A}$ is closed under finite intersections, $A^c \in \mathcal{A}$ and $\mathcal{A}$ 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 $\mathcal{A}$ is closed under finite unions.

Therefore $\mathcal{A}$ is an algebra and (ii) and (iii) are immediate.

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

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

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

Proposition 1.12 *We have*

$$\mathcal{B}_{\mathbf{R}} = \mathcal{B}_{\overline{\mathbf{R}}}\rceil\mathbf{R}$$

and

$$\mathcal{B}_{\overline{\mathbf{R}}} = \big\{A, A \cup \{+\infty\}, A \cup \{-\infty\}, A \cup \{+\infty, -\infty\} \,|\, A \in \mathcal{B}_{\mathbf{R}}\big\}.$$

Also, if $\mathcal{E}$ is the collection containing $\{+\infty\}$ or $\{-\infty\}$ and all closed or all open or all open-closed or all closed-open or all intervals in $\mathbf{R}$, then $\mathcal{B}_{\overline{\mathbf{R}}} = \Sigma(\mathcal{E})$.

Proof: The first equality is immediate from Theorem 1.3.

Now, $\mathbf{R}$ is open in $\overline{\mathbf{R}}$ and, thus, $\mathbf{R} \in \mathcal{B}_{\overline{\mathbf{R}}}$. Therefore, from the first equality and the last statement in Proposition 1.8, we get that

$$\mathcal{B}_{\mathbf{R}} = \{A \subseteq \mathbf{R} \,|\, A \in \mathcal{B}_{\overline{\mathbf{R}}}\}.$$

Therefore, if $A \in \mathcal{B}_{\mathbf{R}}$, then $A \in \mathcal{B}_{\overline{\mathbf{R}}}$. Also, $[-\infty, +\infty)$ is open in $\overline{\mathbf{R}}$ and, hence, $\{+\infty\} \in \mathcal{B}_{\overline{\mathbf{R}}}$. Similarly, $\{-\infty\} \in \mathcal{B}_{\overline{\mathbf{R}}}$ and $\{+\infty, -\infty\} \in \mathcal{B}_{\overline{\mathbf{R}}}$ and we conclude that $\big\{A, A \cup \{+\infty\}, A \cup \{-\infty\}, A \cup \{+\infty, -\infty\} \,|\, A \in \mathcal{B}_{\mathbf{R}}\big\} \subseteq \mathcal{B}_{\overline{\mathbf{R}}}$. Conversely, let $B \in \mathcal{B}_{\overline{\mathbf{R}}}$ and consider $A = B \cap \mathbf{R} \in \mathcal{B}_{\mathbf{R}}$. Then $B = A$ or $B = A \cup \{+\infty\}$ or $B = A \cup \{-\infty\}$ or $B = A \cup \{+\infty, -\infty\}$ and we conclude that $\mathcal{B}_{\overline{\mathbf{R}}} \subseteq \big\{A, A \cup \{+\infty\}, A \cup \{-\infty\}, A \cup \{+\infty, -\infty\} \,|\, A \in \mathcal{B}_{\mathbf{R}}\big\}$.

Let $\mathcal{E} = \big\{\{+\infty\}, (a, b] \,|\, -\infty < a \leq b < +\infty\big\}$.

From all the above, we get that $\mathcal{E} \subseteq \mathcal{B}_{\overline{\mathbf{R}}}$ and, by Proposition 1.4, $\Sigma(\mathcal{E}) \subseteq \mathcal{B}_{\overline{\mathbf{R}}}$. From Proposition 1.10, if $A \in \mathcal{B}_{\mathbf{R}}$, then $A \in \Sigma(\mathcal{E})$. In particular, $\mathbf{R} \in \Sigma(\mathcal{E})$ and, hence, $(-\infty, +\infty] = \mathbf{R} \cup \{+\infty\} \in \Sigma(\mathcal{E})$. Thus, also $\{-\infty\} = \overline{\mathbf{R}} \setminus (-\infty, +\infty] \in \Sigma(\mathcal{E})$ and $\{+\infty, -\infty\} = \{+\infty\} \cup \{-\infty\} \in \Sigma(\mathcal{E})$. From all these, we conclude that $\mathcal{B}_{\overline{\mathbf{R}}} = \big\{A, A \cup \{+\infty\}, A \cup \{-\infty\}, A \cup \{+\infty, -\infty\} \,|\, A \in \mathcal{B}_{\mathbf{R}}\big\} \subseteq \Sigma(\mathcal{E})$.

This concludes the proof of the last statement for this particular choice of $\mathcal{E}$ and the proof is similar for all other choices.

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

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

Proposition 1.13 *We have*

$$\mathcal{B}_{\mathbf{C}} = \mathcal{B}_{\overline{\mathbf{C}}} \rceil \mathbf{C}$$

and

$$\mathcal{B}_{\overline{\mathbf{C}}} = \big\{A, A \cup \{\infty\} \,|\, A \in \mathcal{B}_{\mathbf{C}}\big\}.$$

Also, if $\mathcal{E}$ *is the collection of all closed or all open or all open-closed or all closed-open or all intervals in* $\mathbf{C} = \mathbf{R}^2$*, then* $\mathcal{B}_{\overline{\mathbf{C}}} = \Sigma(\mathcal{E})$.

Proof: The proof is very similar to (and slightly simpler than) the proof of Proposition 1.12. The steps are the same and only minor modifications are needed.

1.6 Exercises.

1. Let X be a non-empty set and $A_1, A_2, \ldots \subseteq X$. We define

$$\limsup_{n\to+\infty} A_n = \cap_{k=1}^{+\infty}\big(\cup_{j=k}^{+\infty} A_j\big), \qquad \liminf_{n\to+\infty} A_n = \cup_{k=1}^{+\infty}\big(\cap_{j=k}^{+\infty} A_j\big).$$

Only in case $\liminf_{n\to+\infty} A_n = \limsup_{n\to+\infty} A_n$, we define

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

Prove the following.
(i) $\limsup_{n\to+\infty} A_n = \{x \in X \,|\, x \in A_n \text{ for infinitely many } n\}$.
(ii) $\liminf_{n\to+\infty} A_n = \{x \in X \,|\, x \in A_n \text{ for all large enough } n\}$.
(iii) $(\liminf_{n\to+\infty} A_n)^c = \limsup_{n\to+\infty} A_n^c$ and $(\limsup_{n\to+\infty} A_n)^c = \liminf_{n\to+\infty} A_n^c$.

(iv) $\liminf_{n\to+\infty} A_n \subseteq \limsup_{n\to+\infty} A_n$.
(v) If $A_n \subseteq A_{n+1}$ for all n, then $\lim_{n\to+\infty} A_n = \cup_{n=1}^{+\infty} A_n$.
(vi) If $A_{n+1} \subseteq A_n$ for all n, then $\lim_{n\to+\infty} A_n = \cap_{n=1}^{+\infty} A_n$.
(vii) Find an example where $\liminf_{n\to+\infty} A_n \neq \limsup_{n\to+\infty} A_n$.
(viii) If $A_n \subseteq B_n$ for all n, then $\limsup_{n\to+\infty} A_n \subseteq \limsup_{n\to+\infty} B_n$ and $\liminf_{n\to+\infty} A_n \subseteq \liminf_{n\to+\infty} B_n$.
(ix) If $A_n = B_n \cup C_n$ for all n, then $\limsup_{n\to+\infty} A_n \subseteq \limsup_{n\to+\infty} B_n \cup \limsup_{n\to+\infty} C_n$, $\liminf_{n\to+\infty} B_n \cup \liminf_{n\to+\infty} C_n \subseteq \liminf_{n\to+\infty} A_n$.

2. 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 increasing countable unions.

3. Let X be non-empty. In the next three cases find $\Sigma(\mathcal{E})$ and $\mathcal{M}(\mathcal{E})$.
(i) $\mathcal{E} = \emptyset$.
(ii) Fix $E \subseteq X$ and let $\mathcal{E} = \{F \mid E \subseteq F \subseteq X\}$.
(iii) Let $\mathcal{E} = \{F \mid F$ is a two-point subset of $X\}$.

4. Let $\mathcal{E}_1, \mathcal{E}_2$ be two collections of subsets of the non-empty X. If $\mathcal{E}_1 \subseteq \mathcal{E}_2 \subseteq \Sigma(\mathcal{E}_1)$, prove that $\Sigma(\mathcal{E}_1) = \Sigma(\mathcal{E}_2)$.

5. Let Y be a non-empty subset of X.
(i) If $\mathcal{A}$ is an algebra of subsets of X, prove that $\mathcal{A}\rceil Y$ is an algebra of subsets of Y.
(ii) If $\mathcal{M}$ is a monotone class of subsets of X, prove that $\mathcal{M}\rceil Y$ is a monotone class of subsets of Y.
(iii) If $\mathcal{T}$ is a topology of X, prove that $\mathcal{T}\rceil Y$ is a topology of Y.

6. Let X be a topological space and Y be a non-empty Borel set in X. Prove that $\mathcal{B}_Y = \{A \subseteq Y \mid A \in \mathcal{B}_X\}$.

7. *Push-forward of a σ-algebra.*

   Let Σ be a σ-algebra of subsets of X and let $f : X \to Y$. Then the collection
$$\{B \subseteq Y \mid f^{-1}(B) \in \Sigma\}$$
is called **the push-forward of Σ by f on Y**.
(i) Prove that the collection $\{B \subseteq Y \mid f^{-1}(B) \in \Sigma\}$ is a σ-algebra of subsets of Y.

   Consider also a σ-algebra Σ' of subsets of Y and a collection $\mathcal{E}$ of subsets of Y so that $\Sigma(\mathcal{E}) = \Sigma'$.
(ii) Prove that, if $f^{-1}(B) \in \Sigma$ for all $B \in \mathcal{E}$, then $f^{-1}(B) \in \Sigma$ for all $B \in \Sigma'$.
(iii) If X, Y are two topological spaces and $f : X \to Y$ is continuous, prove that $f^{-1}(B)$ is a Borel set in X for every Borel set B in Y.

8. *The pull-back of a σ-algebra.*

Let Σ' be a σ-algebra of subsets of Y and let $f : X \to Y$. Then the collection

$$\{f^{-1}(B) \,|\, B \in \Sigma'\}$$

is called **the pull-back of Σ' by f on X.**

Prove that $\{f^{-1}(B) \,|\, B \in \Sigma'\}$ is a σ-algebra of subsets of X.

9. (i) Prove that $\mathcal{B}_{\mathbf{R}^n}$ is generated by the collection of all half-spaces in $\mathbf{R}^n$ of the form $\{x = (x_1, \ldots, x_n) \,|\, a_j < x_j\}$, where $j = 1, \ldots, n$ and $a_j \in \mathbf{R}$.
(ii) Prove that $\mathcal{B}_{\mathbf{R}^n}$ is generated by the collection of all open balls $B(x; r)$ or of all closed balls $\overline{B(x; r)}$, where $x \in \mathbf{R}^n$ and $r \in \mathbf{R}_+$.

10. (i) Prove that $\mathcal{B}_{\overline{\mathbf{R}}}$ is generated by the collection of all $(a, +\infty]$, where $a \in \mathbf{R}$.
(ii) Prove that $\mathcal{B}_{\overline{\mathbf{C}}}$ is generated by the collection of all open discs $B(x; r)$ or of all closed discs $\overline{B(x; r)}$, where $x \in \mathbf{C}$ and $r \in \mathbf{R}_+$.

11. Let X be a metric space with metric d. Prove that every closed $F \subseteq X$ is a G_δ-set by considering the sets $U_n = \{x \in X \,|\, d(x, y) < \frac{1}{n}$ for some $y \in F\}$. Prove, also, that every open $U \subseteq X$ is an F_σ-set.

12. (i) Suppose that $f : \mathbf{R}^n \to \mathbf{R}$. Prove that $\{x \in \mathbf{R}^n \,|\, f$ is continuous at $x\}$ is a G_δ-set in $\mathbf{R}^n$.
(ii) Suppose that $f_k : \mathbf{R}^n \to \mathbf{R}$ is continuous in $\mathbf{R}^n$ for every $k \in \mathbf{N}$. Prove that $\{x \in \mathbf{R}^n \,|\, (f_k(x))$ converges$\}$ is an $F_{\sigma\delta}$-set, i.e. a countable intersection of F_σ-sets.

13. Let $\mathcal{E}$ be an arbitrary collection of subsets of the non-empty X. Prove that for every $A \in \Sigma(\mathcal{E})$ there is some *countable* subcollection $\mathcal{D} \subseteq \mathcal{E}$ so that $A \in \Sigma(\mathcal{D})$.

Chapter 2

Measures

2.1 General measures.

Definition 2.1 *Let* (X, Σ) *be a measurable space. A function* $\mu : \Sigma \to [0, +\infty]$ *is called* ***a measure on*** (X, Σ) *if*
(i) $\mu(\emptyset) = 0$,
(ii) $\mu(\cup_{n=1}^{+\infty} A_n) = \sum_{n=1}^{+\infty} \mu(A_n)$ *for all sequences* (A_n) *of pairwise disjoint sets which are contained in* Σ.

The triple (X, Σ, μ) *of a non-empty set* X*, a* σ*-algebra of subsets of* X *and a measure* μ *on* Σ *is called* ***a measure space****.*

For simplicity and if there is no danger of confusion, we shall say that μ is **a measure on** Σ or **a measure on** X.

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

Property (ii) of a measure is called **σ-additivity** and sometimes a measure is also called **σ-additive measure** to distinguish from a so-called **finitely additive measure** μ which is defined to satisfy $\mu(\emptyset) = 0$ and $\mu(\cup_{n=1}^{N} A_n) = \sum_{n=1}^{N} \mu(A_n)$ for all $N \in \mathbf{N}$ and all pairwise disjoint $A_1, \dots, A_N \in \Sigma$.

Proposition 2.1 *Every measure is finitely additive.*

Proof: Let μ be a measure on the σ-algebra Σ. If $A_1, \dots, A_N \in \Sigma$ are pairwise disjoint, we consider $A_n = \emptyset$ for all $n > N$ and we get $\mu(\cup_{n=1}^{N} A_n) = \mu(\cup_{n=1}^{+\infty} A_n) = \sum_{n=1}^{+\infty} \mu(A_n) = \sum_{n=1}^{N} \mu(A_n)$.

Examples.
1. The simplest measure is the *zero measure* which is denoted o and is defined by $o(A) = 0$ for every $A \in \Sigma$.
2. Let X be an uncountable set and consider $\Sigma = \{A \subseteq X \mid A$ is countable or A^c 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 $A_1, A_2, \ldots \in \Sigma$ be pairwise disjoint. If all of them are countable, then $\cup_{n=1}^{+\infty} A_n$ is also countable and we get

$\mu(\cup_{n=1}^{+\infty} A_n) = 0 = \sum_{n=1}^{+\infty} \mu(A_n)$. Observe that if one of the A_n's, say A_{n_0}, is uncountable, then for all $n \neq n_0$ we have $A_n \subseteq A_{n_0}^c$ which is countable. Therefore $\mu(A_{n_0}) = 1$ and $\mu(A_n) = 0$ for all $n \neq n_0$. Since $(\cup_{n=1}^{+\infty} A_n)^c (\subseteq A_{n_0}^c)$ is countable, we get $\mu(\cup_{n=1}^{+\infty} A_n) = 1 = \sum_{n=1}^{+\infty} \mu(A_n)$.

Theorem 2.1 *Let (X, Σ, μ) be a measure space.*
(i) (Monotonicity) If $A, B \in \Sigma$ and $A \subseteq B$, then $\mu(A) \leq \mu(B)$.
(ii) If $A, B \in \Sigma$, $A \subseteq B$ and $\mu(A) < +\infty$, then $\mu(B \setminus A) = \mu(B) - \mu(A)$.
(iii) (σ-subadditivity) If $A_1, A_2, \ldots \in \Sigma$, then $\mu(\cup_{n=1}^{+\infty} A_n) \leq \sum_{n=1}^{+\infty} \mu(A_n)$.
(iv) (Continuity from below) If $A_1, A_2, \ldots \in \Sigma$ and $A_n \uparrow A$, then $\mu(A_n) \uparrow \mu(A)$.
(v) (Continuity from above) If $A_1, A_2, \ldots \in \Sigma$, $\mu(A_N) < +\infty$ for some N and $A_n \downarrow A$, then $\mu(A_n) \downarrow \mu(A)$.

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) Using Proposition 1.2 we find $B_1, B_2, \ldots \in \Sigma$ which are pairwise disjoint and satisfy $B_n \subseteq A_n$ for all n and $\cup_{n=1}^{+\infty} B_n = \cup_{n=1}^{+\infty} A_n$. By σ-additivity and monotonicity of μ we get $\mu(\cup_{n=1}^{+\infty} A_n) = \mu(\cup_{n=1}^{+\infty} B_n) = \sum_{n=1}^{+\infty} \mu(B_n) \leq \sum_{n=1}^{+\infty} \mu(A_n)$.
(iv) We write $A = A_1 \cup \cup_{k=1}^{+\infty} (A_{k+1} \setminus A_k)$, where all sets whose union is taken in the right side are pairwise disjoint. Applying σ-additivity (and finite additivity), $\mu(A) = \mu(A_1) + \sum_{k=1}^{+\infty} \mu(A_{k+1} \setminus A_k) = \lim_{n \to +\infty} [\mu(A_1) + \sum_{k=1}^{n-1} \mu(A_{k+1} \setminus A_k)] = \lim_{n \to +\infty} \mu\big(A_1 \cup \cup_{k=1}^{n-1} (A_{k+1} \setminus A_k)\big) = \lim_{n \to +\infty} \mu(A_n)$.
(v) We observe that $A_N \setminus A_n \uparrow A_N \setminus A$ and continuity from below implies $\mu(A_N \setminus A_n) \uparrow \mu(A_N \setminus A)$. Now, $\mu(A_N) < +\infty$ implies $\mu(A_n) < +\infty$ for all $n \geq N$ and $\mu(A) < +\infty$. Applying (ii), we get $\mu(A_N) - \mu(A_n) \uparrow \mu(A_N) - \mu(A)$ and, since $\mu(A_N) < +\infty$, we find $\mu(A_n) \downarrow \mu(A)$.

Definition 2.2 *Let (X, Σ, μ) be a measure space.*
*(i) μ is called **finite** if $\mu(X) < +\infty$.*
*(ii) μ is called **σ-finite** if there exist $X_1, X_2, \ldots \in \Sigma$ so that $X = \cup_{n=1}^{+\infty} X_n$ and $\mu(X_n) < +\infty$ for all $n \in \mathbf{N}$.*
*(iii) μ is called **semifinite** if for every $E \in \Sigma$ with $\mu(E) = +\infty$ there is an $F \in \Sigma$ so that $F \subseteq E$ and $0 < \mu(F) < +\infty$.*
*(iv) A set $E \in \Sigma$ is called **of finite μ-measure** if $\mu(E) < +\infty$.*
*(v) A set $E \in \Sigma$ is called **of σ-finite μ-measure** if there exist $E_1, E_2, \ldots \in \Sigma$ so that $E \subseteq \cup_{n=1}^{+\infty} E_n$ and $\mu(E_n) < +\infty$ for all n.*

For simplicity and 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 Σ are of finite measure. More generally, if $E \in \Sigma$ is of finite measure, then all subsets of it in Σ are of finite measure.
2. If μ is σ-finite then all sets in Σ are of σ-finite measure. More generally, if $E \in \Sigma$ is of σ-finite measure, then all subsets of it in Σ 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 σ-finite, applying Proposition 1.2, we see that there exist *pairwise disjoint* $X_1, X_2, \ldots \in \Sigma$ so that $X = \cup_{n=1}^{+\infty} X_n$ and $\mu(X_n) < +\infty$ for all n. Similarly, by taking successive unions, we see that there exist $X_1, X_2, \ldots \in \Sigma$ so that $X_n \uparrow X$ and $\mu(X_n) < +\infty$ for all n. We shall use these two observations freely whenever σ-finiteness appears in the sequel.
6. If μ is finite, then it is also σ-finite. The next result is not so obvious.

Proposition 2.2 *Let (X, Σ, μ) be a measure space. If μ is σ-finite, then it is semifinite.*

Proof: Take $X_1, X_2, \ldots \in \Sigma$ so that $X_n \uparrow X$ and $\mu(X_n) < +\infty$ for all n. Let $E \in \Sigma$ have $\mu(E) = +\infty$. From $E \cap X_n \uparrow E$ and continuity of μ from below, we get $\mu(E \cap X_n) \uparrow +\infty$. Therefore, $\mu(E \cap X_{n_0}) > 0$ for some n_0 and we observe that $\mu(E \cap X_{n_0}) \leq \mu(X_{n_0}) < +\infty$.

Definition 2.3 *Let (X, Σ, μ) be a measure space. $E \in \Sigma$ is called* ***μ-null*** *if $\mu(E) = 0$.*

For simplicity and if there is no danger of confusion, we may say that E is *null* instead of μ-null.

The following is trivial but basic.

Theorem 2.2 *Let (X, Σ, μ) be a measure space.*
(i) If $E \in \Sigma$ is null, then every subset of it in Σ is also null.
(ii) If $E_1, E_2, \ldots \in \Sigma$ are all null, then their union $\cup_{n=1}^{+\infty} E_n$ is also null.

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

2.2 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 2.4 *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*

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

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

Of course, if F is a non-empty finite set, then $\sum_{i \in F} a_i$ is just equal to the sum $\sum_{k=1}^{N} a_{i_k}$, 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 2.3 *If I is countable and $I = \{i_1, i_2, \ldots\}$ is an arbitrary enumeration of it, 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 $\sum_{k=1}^{N} a_{i_k} = \sum_{i\in F} a_i \leq \sum_{i\in I} a_i$. Since N is arbitrary, we find $\sum_{k=1}^{+\infty} a_{i_k} \leq \sum_{i\in I} a_i$.

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 take the maximal, say N, of them. This means that $F \subseteq \{i_1, i_2, \ldots, i_N\}$. Therefore $\sum_{i\in F} a_i \leq \sum_{k=1}^{N} a_{i_k} \leq \sum_{k=1}^{+\infty} a_{i_k}$ and, since F is arbitrary, we find, by the definition of $\sum_{i\in I} a_i$, that $\sum_{i\in I} a_i \leq \sum_{k=1}^{+\infty} a_{i_k}$.

Proposition 2.4 *Let $a : I \to [0, +\infty]$. If $\sum_{i\in I} a_i < +\infty$, then $a_i < +\infty$ for all i and the set $\{i \in I \mid a_i > 0\}$ is countable.*

Proof: Let $\sum_{i\in I} a_i < +\infty$. It is clear that $a_i < +\infty$ for all i (take $F = \{i\}$) and, for arbitrary n, consider the set $I_n = \{i \in I \mid a_i \geq \frac{1}{n}\}$. If F is an arbitrary finite subset of I_n, then $\frac{1}{n}\,\mathrm{card}(F) \leq \sum_{i\in F} a_i \leq \sum_{i\in I} a_i$. Therefore, the cardinality of the arbitrary finite subset of I_n is not larger than the number $n \sum_{i\in I} a_i$ and, hence, the set I_n is finite. But then, $\{i \in I \mid a_i > 0\} = \cup_{n=1}^{+\infty} I_n$ is countable.

Proposition 2.5 *(i) If $a, b : I \to [0, +\infty]$ and $a_i \leq b_i$ for all $i \in I$, then $\sum_{i\in I} a_i \leq \sum_{i\in I} b_i$.*
(ii) If $a : I \to [0, +\infty]$ and $J \subseteq I$, then $\sum_{i\in J} a_i \leq \sum_{i\in I} a_i$.

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

Proposition 2.6 *Let $I = \cup_{k\in K} J_k$, where K is a non-empty set of indices and the J_k's 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: Take an arbitrary finite $F \subseteq I$ and consider the finite sets $F_k = F \cap J_k$. Observe that the set $L = \{k \in K \mid F_k \neq \emptyset\}$ 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 imply that $\sum_{i\in F} a_i \leq \sum_{k\in L} \big(\sum_{i\in J_k} a_i\big) \leq \sum_{k\in K} \big(\sum_{i\in J_k} a_i\big)$. Taking supremum over the finite subsets F of I we find $\sum_{i\in I} a_i \leq \sum_{k\in K} \big(\sum_{i\in J_k} a_i\big)$.

Now take an arbitrary finite $L \subseteq K$ and 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 $\cup_{k\in L} F_k$ of I. Hence $\sum_{k\in L} \big(\sum_{i\in F_k} a_i\big) \leq \sum_{i\in I} a_i$. Taking supremum over the finite subsets F_k of J_k for each $k \in L$, one at a time, we get that $\sum_{k\in L} \big(\sum_{i\in J_k} a_i\big) \leq \sum_{i\in I} a_i$. Finally, taking supremum over the finite subsets L of K, we find $\sum_{k\in K} \big(\sum_{i\in J_k} a_i\big) \leq \sum_{i\in I} a_i$ and conclude the proof.

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

Proposition 2.7 *Let X be non-empty and consider $a : X \to [0,+\infty]$. We define $\mu : \mathcal{P}(X) \to [0,+\infty]$ by*

$$\mu(E) = \sum_{x \in E} a_x , \qquad E \subseteq X.$$

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 $E_1, E_2, \ldots$ are pairwise disjoint and $E = \cup_{n=1}^{+\infty} E_n$, we apply Propositions 2.3 and 2.6 to find $\mu(E) = \sum_{x \in E} a_x = \sum_{n \in \mathbf{N}} \left(\sum_{x \in E_n} a_x \right) = \sum_{n \in \mathbf{N}} \mu(E_n) = \sum_{n=1}^{+\infty} \mu(E_n)$.

Definition 2.5 *The measure on $(X, \mathcal{P}(X))$ defined in the statement of the previous proposition is called* ***the point-mass distribution on X induced by the function a****. The value a_x is called* ***the point-mass at x****.*

Examples.
1. Consider the function which puts point-mass $a_x = 1$ at every $x \in X$. It is then obvious that the induced point-mass distribution is

$$\sharp(E) = \begin{cases} \text{card}(E), & \text{if } E \text{ is a finite} \subseteq X, \\ +\infty, & \text{if } E \text{ is an infinite} \subseteq X. \end{cases}$$

This $\sharp$ is called **the counting measure on X**.
2. Take a particular $x_0 \in X$ and the function which puts 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

$$\delta_{x_0}(E) = \begin{cases} 1, & \text{if } x_0 \in E \subseteq X, \\ 0, & \text{if } x_0 \notin E \subseteq X. \end{cases}$$

This δ_{x_0} is called **the Dirac measure at x_0** or **the Dirac mass at x_0**.

Of course, it is very easy to show directly, without using Proposition 2.7, that these two examples, $\sharp$ and δ_{x_0}, constitute measures.

2.3 Complete measures.

Theorem 2.2(i) 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 2.6 *Let (X, Σ, μ) be a measure space. Suppose that for every $E \in \Sigma$ with $\mu(E) = 0$ and every $F \subseteq E$ it is implied that $F \in \Sigma$ (and hence $\mu(F) = 0$, also). Then μ is called* ***complete*** *and (X, Σ, μ) is* ***a complete measure space****.*

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

Definition 2.7 *If* (X, Σ_1, μ_1) *and* (X, Σ_2, μ_2) *are two measure spaces on the same set* X*, we say that* (X, Σ_2, μ_2) *is an* ***extension*** *of* (X, Σ_1, μ_1) *if* $\Sigma_1 \subseteq \Sigma_2$ *and* $\mu_1(E) = \mu_2(E)$ *for all* $E \in \Sigma_1$.

Theorem 2.3 *Let* (X, Σ, μ) *be a measure space. Then there is a unique smallest complete extension* $(X, \overline{\Sigma}, \overline{\mu})$ *of* (X, Σ, μ)*. Namely, there is a unique measure space* $(X, \overline{\Sigma}, \overline{\mu})$ *so that*
(i) $(X, \overline{\Sigma}, \overline{\mu})$ *is an extension of* (X, Σ, μ)*,*
(ii) $(X, \overline{\Sigma}, \overline{\mu})$ *is complete,*
(iii) if $(X, \overline{\overline{\Sigma}}, \overline{\overline{\mu}})$ *is another complete extension of* (X, Σ, μ)*, then it is an extension also of* $(X, \overline{\Sigma}, \overline{\mu})$*.*

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

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

We prove that $\overline{\Sigma}$ is a σ-algebra. We write $\emptyset = \emptyset \cup \emptyset$, where the first $\emptyset$ belongs to Σ and the second $\emptyset$ is a subset of $\emptyset \in \Sigma$ with $\mu(\emptyset) = 0$. Therefore $\emptyset \in \overline{\Sigma}$.

Let $B \in \overline{\Sigma}$. Then $B = A \cup F$ for $A \in \Sigma$ and $F \subseteq$ of some $E \in \Sigma$ with $\mu(E) = 0$. Write $B^c = A_1 \cup F_1$, where $A_1 = (A \cup E)^c$ and $F_1 = E \setminus (A \cup F)$. Then $A_1 \in \Sigma$ and $F_1 \subseteq E$. Hence $B^c \in \overline{\Sigma}$.

Let $B_1, B_2, \ldots \in \overline{\Sigma}$. Then for every n, $B_n = A_n \cup F_n$ for $A_n \in \Sigma$ and $F_n \subseteq$ of some $E_n \in \Sigma$ with $\mu(E_n) = 0$. Now $\cup_{n=1}^{+\infty} B_n = (\cup_{n=1}^{+\infty} A_n) \cup (\cup_{n=1}^{+\infty} F_n)$, where $\cup_{n=1}^{+\infty} A_n \in \Sigma$ and $\cup_{n=1}^{+\infty} F_n \subseteq \cup_{n=1}^{+\infty} E_n \in \Sigma$ with $\mu(\cup_{n=1}^{+\infty} E_n) = 0$. Therefore $\cup_{n=1}^{+\infty} B_n \in \overline{\Sigma}$.

We now construct $\overline{\mu}$. For every $B \in \overline{\Sigma}$ we write $B = A \cup F$ for $A \in \Sigma$ and $F \subseteq$ of some $E \in \Sigma$ with $\mu(E) = 0$ and define

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

To prove that $\overline{\mu}(B)$ is well defined we consider that we may also have $B = A' \cup F'$ for $A' \in \Sigma$ and $F' \subseteq$ of some $E' \in \Sigma$ with $\mu(E') = 0$ and we must 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 $(X, \overline{\Sigma})$ let $\emptyset = \emptyset \cup \emptyset$ as above and get $\overline{\mu}(\emptyset) = \mu(\emptyset) = 0$. Let also $B_1, B_2, \ldots \in \overline{\Sigma}$ be pairwise disjoint. Then $B_n = A_n \cup F_n$ for $A_n \in \Sigma$ and $F_n \subseteq E_n \in \Sigma$ with $\mu(E_n) = 0$. Observe that the A_n's are pairwise disjoint. Then $\cup_{n=1}^{+\infty} B_n = (\cup_{n=1}^{+\infty} A_n) \cup (\cup_{n=1}^{+\infty} F_n)$ and $\cup_{n=1}^{+\infty} F_n \subseteq \cup_{n=1}^{+\infty} E_n \in \Sigma$ with $\mu(\cup_{n=1}^{+\infty} E_n) = 0$. Therefore $\overline{\mu}(\cup_{n=1}^{+\infty} B_n) = \mu(\cup_{n=1}^{+\infty} A_n) = \sum_{n=1}^{+\infty} \mu(A_n) = \sum_{n=1}^{+\infty} \overline{\mu}(B_n)$.

We now prove that $\overline{\mu}$ is complete. Let $B \in \overline{\Sigma}$ with $\overline{\mu}(B) = 0$ and let $B' \subseteq B$. Write $B = A \cup F$ for $A \in \Sigma$ and $F \subseteq E \in \Sigma$ with $\mu(E) = 0$ and observe that $\mu(A) = \overline{\mu}(B) = 0$. Then write $B' = A' \cup F'$, where $A' = \emptyset \in \Sigma$ and $F' = B' \subseteq E'$ where $E' = A \cup E \in \Sigma$ with $\mu(E') \leq \mu(A) + \mu(E) = 0$. Hence $B' \in \overline{\Sigma}$.

To prove that $(X, \overline{\Sigma}, \overline{\mu})$ is an extension of (X, Σ, μ) we take any $A \in \Sigma$ and write $A = A \cup \emptyset$, where $\emptyset \subseteq \emptyset \in \Sigma$ with $\mu(\emptyset) = 0$. This implies that $A \in \overline{\Sigma}$ and $\overline{\mu}(A) = \mu(A)$.

Now suppose that $(X, \overline{\overline{\Sigma}}, \overline{\overline{\mu}})$ is another complete extension of (X, Σ, μ). Take any $B \in \overline{\Sigma}$ and thus $B = A \cup F$ for $A \in \Sigma$ and $F \subseteq E \in \Sigma$ with $\mu(E) = 0$. But then $A, E \in \overline{\overline{\Sigma}}$ and $\overline{\overline{\mu}}(E) = \mu(E) = 0$. Since $\overline{\overline{\mu}}$ is complete, we get that also $F \in \overline{\overline{\Sigma}}$ and hence $B = A \cup F \in \overline{\overline{\Sigma}}$.

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, Σ, μ). This is obvious, since two *smallest* complete extensions of (X, Σ, μ) must both be extensions of each other and, hence, identical.

Definition 2.8 *If (X, Σ, μ) is a measure space, then its smallest complete extension is called* ***the completion of*** *(X, Σ, μ).*

2.4 Restriction of a measure.

Proposition 2.8 *Let (X, Σ, μ) be a measure space and let $Y \in \Sigma$. If we define $\mu_Y : \Sigma \to [0, +\infty]$ by*

$$\mu_Y(A) = \mu(A \cap Y), \qquad A \in \Sigma,$$

then μ_Y is a measure on (X, Σ) with the properties that $\mu_Y(A) = \mu(A)$ for every $A \in \Sigma$, $A \subseteq Y$, and that $\mu_Y(A) = 0$ for every $A \in \Sigma$, $A \cap Y = \emptyset$.

Proof: We have $\mu_Y(\emptyset) = \mu(\emptyset \cap Y) = \mu(\emptyset) = 0$.

If $A_1, A_2, \ldots \in \Sigma$ are pairwise disjoint, $\mu_Y(\cup_{j=1}^{+\infty} A_j) = \mu\big((\cup_{j=1}^{+\infty} A_j) \cap Y\big) = \mu\big(\cup_{j=1}^{+\infty} (A_j \cap Y)\big) = \sum_{j=1}^{+\infty} \mu(A_j \cap Y) = \sum_{j=1}^{+\infty} \mu_Y(A_j)$.

Therefore, μ_Y is a measure on (X, Σ) and its two properties are trivial to prove.

Definition 2.9 *Let (X, Σ, μ) be a measure space and let $Y \in \Sigma$. The measure μ_Y on (X, Σ) of Proposition 2.8 is called* ***the Y-restriction of μ.***

There is a second kind of restriction of a measure. To define it we recall that, if $Y \in \Sigma$, the restriction $\Sigma\rceil Y$ of the σ-algebra Σ of subsets of X on the non-empty $Y \subseteq X$ is $\Sigma\rceil Y = \{A \subseteq Y \mid A \in \Sigma\}$.

Proposition 2.9 *Let (X, Σ, μ) be a measure space and let $Y \in \Sigma$ be non-empty. We consider $\Sigma\rceil Y = \{A \subseteq Y \mid A \in \Sigma\}$ and define $\mu\rceil Y : \Sigma\rceil Y \to [0, +\infty]$ by*

$$(\mu\rceil Y)(A) = \mu(A), \qquad A \in \Sigma\rceil Y.$$

Then $\mu\rceil Y$ is a measure on $(Y, \Sigma\rceil Y)$.

Proof: Obvious.

Definition 2.10 *Let (X, Σ, μ) be a measure space and let $Y \in \Sigma$ be non-empty. The measure $\mu\rceil Y$ on $(Y, \Sigma\rceil Y)$ of Proposition 2.9 is called* ***the restriction of μ on*** *$\Sigma\rceil Y$.*

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 Σ 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 Σ which are included in Y.

2.5 Uniqueness of measures.

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

Theorem 2.4 *Let $\mathcal{A}$ be an algebra of subsets of X and let μ, ν be two measures on $(X, \Sigma(\mathcal{A}))$. Suppose there exist $A_1, A_2, \ldots \in \mathcal{A}$ with $A_n \uparrow X$ and $\mu(A_k), \nu(A_k) < +\infty$ for all k.*

If μ, ν are equal on $\mathcal{A}$, then they are equal also on $\Sigma(\mathcal{A})$.

Proof: (a) Suppose that $\mu(X), \nu(X) < +\infty$.

We define the collection $\mathcal{M} = \{E \in \Sigma(\mathcal{A}) \,|\, \mu(E) = \nu(E)\}$. It is easy to see that $\mathcal{M}$ is a monotone class. Indeed, let $E_1, E_2, \ldots \in \mathcal{M}$ with $E_n \uparrow E$. By continuity of measures from below, we get $\mu(E) = \lim_{n\to+\infty} \mu(E_n) = \lim_{n\to+\infty} \nu(E_n) = \nu(E)$ and thus $E \in \mathcal{M}$. We do exactly the same when $E_n \downarrow E$, using the continuity of measures from above and the extra assumption $\mu(X), \nu(X) < +\infty$.

Since $\mathcal{M}$ is a monotone class including $\mathcal{A}$, Proposition 1.7 implies that $\mathcal{M}(\mathcal{A}) \subseteq \mathcal{M}$. Now, Theorem 1.1 implies that $\Sigma(\mathcal{A}) \subseteq \mathcal{M}$ and, thus, $\mu(E) = \nu(E)$ for all $E \in \Sigma(\mathcal{A})$.

(b) The general case.

For each k, we consider the A_k-restrictions of μ, ν. Namely,

$$\mu_{A_k}(E) = \mu(E \cap A_k), \qquad \nu_{A_k}(E) = \nu(E \cap A_k)$$

for all $E \in \Sigma(\mathcal{A})$. All μ_{A_k} and ν_{A_k} are finite measures on (X, Σ), because $\mu_{A_k}(X) = \mu(A_k) < +\infty$ and $\nu_{A_k}(X) = \nu(A_k) < +\infty$. We, clearly, have that μ_{A_k}, ν_{A_k} are equal on $\mathcal{A}$ and, by the result of (a), they are equal also on $\Sigma(\mathcal{A})$.

For every $E \in \Sigma(\mathcal{A})$, using the $E \cap A_k \uparrow E$ and the continuity of measures from below, we can write $\mu(E) = \lim_{k\to+\infty} \mu(E \cap A_k) = \lim_{k\to+\infty} \mu_{A_k}(E) = \lim_{k\to+\infty} \nu_{A_k}(E) = \lim_{k\to+\infty} \nu(E \cap A_k) = \nu(E)$.

Thus, μ, ν are equal on $\Sigma(\mathcal{A})$.

2.6 Exercises.

1. Let (X, Σ, μ) be a measure space and $Y \in \Sigma$ be non-empty. Prove that μ_Y is the only measure on (X, Σ) with the properties:
   (i) $\mu_Y(E) = \mu(E)$ for all $E \in \Sigma$ with $E \subseteq Y$,
   (ii) $\mu_Y(E) = 0$ for all $E \in \Sigma$ with $E \subseteq Y^c$.

2. *Positive linear combinations of measures.*

   Let $\mu, \mu_1 \mu_2$ be measures on the measurable space (X, Σ) and $\kappa \in [0, +\infty)$.
   (i) Prove that $\kappa\mu : \Sigma \to [0, +\infty]$, which is defined by

   $$(\kappa\mu)(E) = \kappa \cdot \mu(E)\,, \qquad E \in \Sigma,$$

   (consider $0 \cdot (+\infty) = 0$) is a measure on (X, Σ). This $\kappa\mu$ is called **the product of μ by κ**.
   (ii) Prove that $\mu_1 + \mu_2 : \Sigma \to [0, +\infty]$, which is defined by

   $$(\mu_1 + \mu_2)(E) = \mu_1(E) + \mu_2(E)\,, \qquad E \in \Sigma,$$

   is a measure on (X, Σ). This $\mu_1 + \mu_2$ is called **the sum of μ_1 and μ_2**.

   Thus, we define **positive linear combinations** $\kappa_1\mu_1 + \cdots + \kappa_n\mu_n$.

3. 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 positive linear combination (see Exercise 2.6.2) of Dirac measures:

   $$\mu = \kappa_1\delta_{x_1} + \cdots + \kappa_k\delta_{x_k}.$$

4. Let X be infinite and define for all $E \subseteq X$

   $$\mu(E) = \begin{cases} 0, & \text{if } E \text{ is finite,} \\ +\infty, & \text{if } E \text{ is infinite.} \end{cases}$$

   Prove that μ is a finitely additive measure on $(X, \mathcal{P}(X))$ which is not a measure.

5. Let (X, Σ, μ) be a measure space and $E \in \Sigma$ be of σ-finite measure. If $\{D_i\}_{i \in I}$ is a collection of pairwise disjoint sets in Σ, prove that the set $\{i \in I \,|\, \mu(E \cap D_i) > 0\}$ is countable.

6. Let X be uncountable and define for all $E \subseteq X$

   $$\mu(E) = \begin{cases} 0, & \text{if } E \text{ is countable,} \\ +\infty, & \text{if } E \text{ is uncountable.} \end{cases}$$

   Prove that μ is a measure on $(X, \mathcal{P}(X))$ which is not semifinite.

7. Let (X, Σ, μ) be a complete measure space. If $A \in \Sigma$, $B \subseteq X$ and $\mu(A \triangle B) = 0$, prove that $B \in \Sigma$ and $\mu(B) = \mu(A)$.

8. Let μ be a finitely additive measure on the measurable space (X, Σ).
   (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.

9. Let (X, Σ, μ) be a measure space and $A_1, A_2, \ldots \in \Sigma$. Prove that (see Exercise 1.6.1)
(i) $\mu(\liminf_{n\to+\infty} A_n) \le \liminf_{n\to+\infty} \mu(A_n)$,
(ii) $\limsup_{n\to+\infty} \mu(A_n) \le \mu(\limsup_{n\to+\infty} A_n)$, if $\mu(\cup_{n=1}^{+\infty} A_n) < +\infty$,
(iii) $\mu(\limsup_{n\to+\infty} A_n) = 0$, if $\sum_{n=1}^{+\infty} \mu(A_n) < +\infty$.

10. *Increasing limits of measures are measures.*

Let (μ_n) be a sequence of measures on (X, Σ) which is increasing. Namely, $\mu_n(E) \le \mu_{n+1}(E)$ for all $E \in \Sigma$ and all n. We define

$$\mu(E) = \lim_{n\to+\infty} \mu_n(E)\,, \qquad E \in \Sigma.$$

Prove that μ is a measure on (X, Σ).

11. *The inclusion-exclusion formula.*

Let (X, Σ, μ) be a measure space. Prove that for all n and $A_1, \ldots, A_n \in \Sigma$

$$\mu(\cup_{j=1}^{n} A_j) + \sum_{k \text{ even}} \sum_{1\le i_1<\cdots<i_k\le n} \mu(A_{i_1} \cap \cdots \cap A_{i_k}) = \sum_{k \text{ odd}} \sum_{1\le i_1<\cdots<i_k\le n} \mu(A_{i_1} \cap \cdots \cap A_{i_k}).$$

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

$$\sum_{i\in I} \kappa a_i = \kappa \sum_{i\in I} a_i\,.$$

(consider $0 \cdot (+\infty) = 0$).
(iv) Prove that

$$\sum_{i\in I} (a_i + b_i) = \sum_{i\in I} a_i + \sum_{i\in I} b_i\,.$$

13. *Tonelli's Theorem for sums.*

Let I, J be two sets of indices and $a : I \times J \to [0, +\infty]$. Using Proposition 2.6, 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 the

$$\sum_{i\in I} (a_i + b_i) = \sum_{i\in I} a_i + \sum_{i\in I} b_i$$

for every $a, b : I \to [0, +\infty]$ (see Exercise 2.6.12).

14. 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.

15. *Characterisation of point-mass distributions.*

    Let $X \neq \emptyset$. Prove that every measure μ on $(X, \mathcal{P}(X))$ is a point-mass distribution.

16. *The push-forward of a measure.*

    Let (X, Σ, μ) be a measure space and $f : X \to Y$. We consider the σ-algebra $\Sigma' = \{B \subseteq Y \mid f^{-1}(B) \in \Sigma\}$, the push-forward of Σ by f on Y (see Exercise 1.6.7). We define

    $$\mu'(B) = \mu(f^{-1}(B)), \qquad B \in \Sigma'.$$

    Prove that μ' is a measure on (Y, Σ'). It is called **the push-forward of μ by f on Y**.

17. *The pull-back of a measure.*

    Let (Y, Σ', μ') be a measure space and $f : X \to Y$ be one-to-one and onto Y. We consider the σ-algebra $\Sigma = \{f^{-1}(B) \mid B \in \Sigma'\}$, the pull-back of Σ' by f on X (see Exercise 1.6.8). We define

    $$\mu(A) = \mu'(f(A)), \qquad A \in \Sigma.$$

    Prove that μ is a measure on (X, Σ). It is called **the pull-back of μ' by f on X**.

18. Let (X, Σ, μ) be a measure space.
(i) If $A, B \in \Sigma$ and $\mu(A \triangle B) = 0$, prove that $\mu(A) = \mu(B)$.
(ii) We define $A \sim B$ if $A, B \in \Sigma$ and $\mu(A \triangle B) = 0$. Prove that $\sim$ is an equivalence relation on Σ.

    We assume that $\mu(X) < +\infty$ and define

    $$\overline{d}(A, B) = \mu(A \triangle B), \qquad A, B \in \Sigma.$$

    (iii) Prove that $\overline{d}$ is a pseudometric on Σ. 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 $A, B, C \in \Sigma$.
(iv) On the set $\Sigma/\sim$ of all equivalence classes we define

    $$d([A], [B]) = \overline{d}(A, B) = \mu(A \triangle B), \qquad [A], [B] \in \Sigma/\sim.$$

    Prove that $d([A], [B])$ is well defined and that d is a metric on $\Sigma/\sim$.

19. Let μ be a semifinite measure on the measurable space (X, Σ). Prove that for every $E \in \Sigma$ with $\mu(E) = +\infty$ and every $M > 0$ there is an $F \in \Sigma$ so that $F \subseteq E$ and $M < \mu(F) < +\infty$.

20. *The saturation of a measure space.*

Let (X, Σ, μ) be a measure space. We say that $E \subseteq X$ **belongs locally to** Σ if $E \cap A \in \Sigma$ for all $A \in \Sigma$ with $\mu(A) < +\infty$. We define

$$\widetilde{\Sigma} = \{E \subseteq X \,|\, E \text{ belongs locally to } \Sigma\}.$$

(i) Prove that $\Sigma \subseteq \widetilde{\Sigma}$ and that $\widetilde{\Sigma}$ is a σ-algebra. If $\Sigma = \widetilde{\Sigma}$, then (X, Σ, μ) is called **saturated**.
(ii) If μ is σ-finite, prove that (X, Σ, μ) is saturated.

We define

$$\widetilde{\mu}(E) = \begin{cases} \mu(E), & \text{if } E \in \Sigma, \\ +\infty, & \text{if } E \in \widetilde{\Sigma} \setminus \Sigma. \end{cases}$$

(iii) Prove that $\widetilde{\mu}$ is a measure on $(X, \widetilde{\Sigma})$, and, hence, $(X, \widetilde{\Sigma}, \widetilde{\mu})$ is an extension of (X, Σ, μ).
(iv) If (X, Σ, μ) is complete, prove that $(X, \widetilde{\Sigma}, \widetilde{\mu})$ is also complete.
(v) Prove that $(X, \widetilde{\Sigma}, \widetilde{\mu})$ is a saturated measure space.

$(X, \widetilde{\Sigma}, \widetilde{\mu})$ is called **the saturation of** (X, Σ, μ).

21. *The direct sum of measure spaces.*

Let $\{(X_i, \Sigma_i, \mu_i)\}_{i \in I}$ be a collection of measure spaces, where the X_i's are pairwise disjoint. We define

$$X = \cup_{i \in I} X_i, \qquad \Sigma = \{E \subseteq X \,|\, E \cap X_i \in \Sigma_i \text{ for all } i \in I\}$$

and

$$\mu(E) = \sum_{i \in I} \mu_i(E \cap X_i)$$

for all $E \in \Sigma$.
(i) Prove that (X, Σ, μ) is a measure space. It is called **the direct sum of** $\{(X_i, \Sigma_i, \mu_i)\}_{i \in I}$ and it is denoted

$$\oplus_{i \in I}(X_i, \Sigma_i, \mu_i).$$

(ii) Prove that μ is σ-finite if and only if the set $J = \{i \in I \,|\, \mu_i \neq o\}$ is countable and μ_i is σ-finite for all $i \in J$.

Chapter 3

Outer measures

3.1 Outer measures.

Definition 3.1 *Let X be a non-empty set. A function $\mu^* : \mathcal{P}(X) \to [0, +\infty]$ is called* ***outer measure on X*** *if*
(i) $\mu^*(\emptyset) = 0$,
(ii) $\mu^*(A) \leq \mu^*(B)$ *if* $A \subseteq B \subseteq X$,
(iii) $\mu^*(\cup_{n=1}^{+\infty} A_n) \leq \sum_{n=1}^{+\infty} \mu^*(A_n)$ *for all sequences* (A_n) *of subsets of* X.

Thus, an outer measure on X is defined for *all* subsets of X, it is monotone and σ-subadditive. An outer measure is also finitely subadditive, because for every $A_1, \ldots, A_N \subseteq X$ we set $A_n = \emptyset$ for all $n > N$ and get $\mu^*(\cup_{n=1}^{N} A_n) = \mu^*(\cup_{n=1}^{+\infty} A_n) \leq \sum_{n=1}^{+\infty} \mu^*(A_n) = \sum_{n=1}^{N} \mu^*(A_n)$.

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

Definition 3.2 *Let μ^* be an outer measure on the non-empty set X. We say that the set $A \subseteq X$ is* ***μ^*-measurable*** *if*

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

for all $E \subseteq X$.

We denote Σ_{μ^*} the collection of all μ^*-measurable subsets of X.

Thus, a set is μ^*-measurable if and only if it decomposes every subset of X into two disjoint pieces whose outer measures add to give the outer measure of the subset.

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

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

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

Theorem 3.1 *(Caratheodory) If μ^* is an outer measure on X, then the collection Σ_{μ^*} of all μ^*-measurable subsets of X is a σ-algebra. If we denote μ the restriction of μ^* on Σ_{μ^*}, then (X, Σ_{μ^*}, μ) is a complete measure space.*

Proof: $\mu^*(E \cap \emptyset) + \mu^*(E \cap \emptyset^c) = \mu^*(\emptyset) + \mu^*(E) = \mu^*(E)$ and, thus, $\emptyset \in \Sigma_{\mu^*}$.

If $A \in \Sigma_{\mu^*}$, then $\mu^*(E \cap A^c) + \mu^*(E \cap (A^c)^c) = \mu^*(E \cap A) + \mu^*(E \cap A^c) = \mu^*(E)$ for all $E \subseteq X$. Therefore, $A^c \in \Sigma_{\mu^*}$ and Σ_{μ^*} is closed under complements.

Let now $A, B \in \Sigma_{\mu^*}$ and take an arbitrary $E \subseteq X$. To check $A \cup B \in \Sigma_{\mu^*}$ write $\mu^*(E \cap (A \cup B)) + \mu^*(E \cap (A \cup B)^c) = \mu^*(E \cap (A \cup B)) + \mu^*(E \cap (A^c \cap B^c))$ and use the subadditivity of μ^* for the first term to get $\leq \mu^*(E \cap (A \cap B^c)) + \mu^*(E \cap (B \cap A^c)) + \mu^*(E \cap (A \cap B)) + \mu^*(E \cap (A^c \cap B^c))$. Now combine the first and third term and also the second and fourth term with the μ^*-measurability of B to get $= \mu^*(E \cap A) + \mu^*(E \cap A^c)$, which is $= \mu^*(E)$ by the μ^*-measurability of A.

This proves that $A \cup B \in \Sigma_{\mu^*}$ and by induction we get that Σ_{μ^*} is closed under finite unions. Since it is closed under complements, Σ_{μ^*} is an algebra of subsets of X and, hence, it is also closed under finite intersections and under set-theoretic differences.

Let $A, B \in \Sigma_{\mu^*}$ with $A \cap B = \emptyset$ and get for all $E \subseteq X$ that $\mu^*(E \cap (A \cup B)) = \mu^*([E \cap (A \cup B)] \cap A) + \mu^*([E \cap (A \cup B)] \cap A^c) = \mu^*(E \cap A) + \mu^*(E \cap B)$. By an obvious induction we find that, if $A_1, \ldots, A_N \in \Sigma_{\mu^*}$ are pairwise disjoint and $E \subseteq X$ is arbitrary, then $\mu^*(E \cap (A_1 \cup \cdots \cup A_N)) = \mu^*(E \cap A_1) + \cdots + \mu^*(E \cap A_N)$. If now $A_1, A_2, \ldots \in \Sigma_{\mu^*}$ are pairwise disjoint and $E \subseteq X$ is arbitrary, then, for all N, $\mu^*(E \cap A_1) + \cdots + \mu^*(E \cap A_N) = \mu^*(E \cap (A_1 \cup \cdots \cup A_N)) \leq \mu^*(E \cap (\cup_{n=1}^{+\infty} A_n))$ by the monotonicity of μ^*. Hence $\sum_{n=1}^{+\infty} \mu^*(E \cap A_n) \leq \mu^*(E \cap (\cup_{n=1}^{+\infty} A_n))$. Since the opposite inequality is immediate after the σ-subadditivity of μ^*, we conclude with the *basic equality*

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

for all pairwise disjoint $A_1, A_2, \ldots \in \Sigma_{\mu^*}$ and all $E \subseteq X$.

If $A_1, A_2, \ldots \in \Sigma_{\mu^*}$ are pairwise disjoint and $E \subseteq X$ is arbitrary, then, since Σ_{μ^*} is closed under finite unions, $\cup_{n=1}^{N} A_n \in \Sigma_{\mu^*}$ for all N. Hence $\mu^*(E) = \mu^*(E \cap (\cup_{n=1}^{N} A_n)) + \mu^*(E \cap (\cup_{n=1}^{N} A_n)^c) \geq \sum_{n=1}^{N} \mu^*(E \cap A_n) + \mu^*(E \cap (\cup_{n=1}^{+\infty} A_n)^c)$, where we used the *basic equality* for the first term and the monotonicity of μ^* for the second. Since N is arbitrary, $\mu^*(E) \geq \sum_{n=1}^{+\infty} \mu^*(E \cap A_n) + \mu^*(E \cap (\cup_{n=1}^{+\infty} A_n)^c) = \mu^*(E \cap (\cup_{n=1}^{+\infty} A_n)) + \mu^*(E \cap (\cup_{n=1}^{+\infty} A_n)^c)$ by the *basic equality.* This means that $\cup_{n=1}^{+\infty} A_n \in \Sigma_{\mu^*}$.

If $A_1, A_2, \ldots \in \Sigma_{\mu^*}$ are not necessarily pairwise disjoint, 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 Σ_{μ^*} is an algebra, all B_n's belong to Σ_{μ^*} and they are pairwise disjoint. By the last paragraph, $\cup_{n=1}^{+\infty} A_n = \cup_{n=1}^{+\infty} B_n \in \Sigma_{\mu^*}$. We conclude that Σ_{μ^*} is a σ-algebra.

We now define $\mu : \Sigma_{\mu^*} \to [0, +\infty]$ as the restriction of μ^*, namely

$$\mu(A) = \mu^*(A), \qquad A \in \Sigma_{\mu^*}.$$

Using $E = X$ in the *basic equality*, we get that for all pairwise disjoint $A_1, A_2, \ldots \in \Sigma_{\mu^*}$,

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

Since $\mu(\emptyset) = \mu^*(\emptyset) = 0$, we see that (X, Σ_{μ^*}, μ) is a measure space.

Let $A \in \Sigma_{\mu^*}$ with $\mu(A) = 0$ and $B \subseteq A$. Then $\mu^*(B) \le \mu^*(A) = \mu(A) = 0$ and for all $E \subseteq X$ we get $\mu^*(E \cap B) + \mu^*(E \cap B^c) \le \mu^*(B) + \mu^*(E) = \mu^*(E)$. Therefore, $B \in \Sigma_{\mu^*}$ and μ is complete.

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

Proposition 3.1 *Let μ^* be an outer measure on X.*
(i) If $B \subseteq X$ has $\mu^*(B) = 0$, then B is μ^*-measurable.
(ii) For all pairwise disjoint μ^*-measurable $A_1, A_2, \ldots$ and all $E \subseteq X$

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

Proof: The proof of (i) is in the last part of the proof of the theorem of Caratheodory and (ii) is the *basic equality* in the same proof.

The most widely used method of producing measures is based on the Theorem of Caratheodory. One starts with an outer measure μ^* on X and produces the measure space (X, Σ_{μ^*}, μ). There are mainly two methods of constructing outer measures. One method starts with a (more or less) arbitrary collection $\mathcal{C}$ of subsets of X and a function τ on this collection and it will be described in the next section. The second method will be studied much later and its starting point is a continuous linear functional on a space of continuous functions. The central result related to this method is the important *F. Riesz Representation Theorem.*

There is another method of producing measures. This is the so-called *Daniell method* which we shall describe also later.

3.2 Construction of outer measures.

Theorem 3.2 *Let $\mathcal{C}$ be a collection of subsets of X, containing at least the $\emptyset$, and $\tau : \mathcal{C} \to [0, +\infty]$ be an arbitrary function with $\tau(\emptyset) = 0$. We define*

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

for all $E \subseteq X$, where we agree that $\inf \emptyset = +\infty$.

Then, μ^* is an outer measure on X.

It should be 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 $E\subseteq\cup_{j=1}^{+\infty}C_n\}$ is non-empty. If there is no countable covering of E with elements of $\mathcal{C}$, then this set is empty and we take $\mu^*(E)=\inf\emptyset=+\infty$.

Proof: For $\emptyset$ the covering $\emptyset\subseteq\emptyset\cup\emptyset\cup\cdots$ implies $\mu^*(\emptyset)\leq\tau(\emptyset)+\tau(\emptyset)+\cdots=0$ and, hence, $\mu^*(\emptyset)=0$.

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 $\mu^*(A)\leq\mu^*(B)$ is obviously true. Otherwise, we take an arbitrary covering $B\subseteq\cup_{j=1}^{+\infty}C_n$ with $C_1,\ldots\in\mathcal{C}$. Then we also have $A\subseteq\cup_{j=1}^{+\infty}C_n$ and, by the definition of $\mu^*(A)$, we get $\mu^*(A)\leq\sum_{j=1}^{+\infty}\tau(C_j)$. Taking the infimum of the right side, we find $\mu^*(A)\leq\mu^*(B)$.

Finally, let's prove $\mu^*(\cup_{n=1}^{+\infty}A_n)\leq\sum_{n=1}^{+\infty}\mu^*(A_n)$ for all $A_1,A_2,\ldots\subseteq X$. If the right side is $=+\infty$, the inequality is clear. Therefore we assume that the right side is $<+\infty$ and, hence, that $\mu^*(A_n)<+\infty$ for all n. By the definition of each $\mu^*(A_n)$, for every $\epsilon>0$ there exist $C_{n,1},C_{n,2},\ldots\in\mathcal{C}$ so that $A_n\subseteq\cup_{j=1}^{+\infty}C_{n,j}$ and $\sum_{j=1}^{+\infty}\tau(C_{n,j})<\mu^*(A_n)+\frac{\epsilon}{2^n}$.

Then $\cup_{n=1}^{+\infty}A_n\subseteq\cup_{(n,j)\in\mathbf{N}\times\mathbf{N}}C_{n,j}$ and, using an arbitrary enumeration of $\mathbf{N}\times\mathbf{N}$ and Proposition 2.3, we get by the definition of $\mu^*(\cup_{n=1}^{+\infty}A_n)$ that $\mu^*(\cup_{n=1}^{+\infty}A_n)\leq\sum_{(n,j)\in\mathbf{N}\times\mathbf{N}}\tau(C_{n,j})$. Proposition 2.6 implies $\mu^*(\cup_{n=1}^{+\infty}A_n)\leq\sum_{n=1}^{+\infty}(\sum_{j=1}^{+\infty}\tau(C_{n,j}))<\sum_{n=1}^{+\infty}(\mu^*(A_n)+\frac{\epsilon}{2^n})=\sum_{n=1}^{+\infty}\mu^*(A_n)+\epsilon$. Since ϵ is arbitrary, we conclude that $\mu^*(\cup_{n=1}^{+\infty}A_n)\leq\sum_{n=1}^{+\infty}\mu^*(A_n)$.

3.3 Exercises.

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

2. 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

$$(\kappa\mu^*)(E)=\kappa\cdot\mu^*(E),\qquad(\mu^*_1+\mu^*_2)(E)=\mu^*_1(E)+\mu^*_2(E)$$

(consider $0\cdot(+\infty)=0$) and

$$\max\{\mu^*_1,\mu^*_2\}(E)=\max\{\mu^*_1(E),\mu^*_2(E)\}$$

for all $E\subseteq X$.

3. Let X be a non-empty set and consider $\mu^*(\emptyset)=0$ and $\mu^*(E)=1$ if $\emptyset\neq E\subseteq X$. Prove that μ^* is an outer measure on X and find all the μ^*-measurable subsets of X.

4. For every $E\subseteq\mathbf{N}$ define $\kappa(E)=\limsup_{n\to+\infty}\frac{1}{n}\,\mathrm{card}(E\cap\{1,2,\ldots,n\})$. Is κ an outer measure on $\mathbf{N}$?

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

6. Let μ^* be an outer measure on X. If $A_1, A_2, \ldots \in \Sigma_{\mu^*}$ and $A_n \uparrow A$, prove that $\mu^*(A_n \cap E) \uparrow \mu^*(A \cap E)$ for every $E \subseteq X$.

7. *Extension of a measure, I.*

   Let (X, Σ_0, μ_0) be a measure space. For every $E \subseteq X$ we define

$$\mu^*(E) = \inf\Big\{ \sum_{j=1}^{+\infty} \mu_0(A_j) \,|\, A_1, A_2, \ldots \in \Sigma_0, E \subseteq \cup_{j=1}^{+\infty} A_j \Big\},$$

   (i) Prove that μ^* is an outer measure on X. We say that μ^* **is induced by the measure** μ_0.
   (ii) Prove that $\mu^*(E) = \min\big\{\mu_0(A) \,|\, A \in \Sigma_0, E \subseteq A\big\}$.
   (iii) If (X, Σ_{μ^*}, μ) is the complete measure space which results from μ^* by the theorem of Caratheodory (i.e. μ is the restriction of μ^* on Σ_{μ^*}), prove that (X, Σ_{μ^*}, μ) is an extension of (X, Σ_0, μ_0).
   (iv) Assume that $E \subseteq X$ and $A_1, A_2, \ldots \in \Sigma_0$ with $E \subseteq \cup_{j=1}^{+\infty} A_j$ and $\mu(A_j) < +\infty$ for all j. Prove that $E \in \Sigma_{\mu^*}$ if and only if there is some $A \in \Sigma_0$ so that $E \subseteq A$ and $\mu^*(A \setminus E) = 0$.
   (v) If μ is σ-finite, prove that (X, Σ_{μ^*}, μ) is the completion of (X, Σ_0, μ_0).
   (vi) Let X be an uncountable set, $\Sigma_0 = \{A \subseteq X \,|\, A$ is countable or A^c is countable$\}$ and $\mu_0(A) = \sharp(A)$ for every $A \in \Sigma_0$. Prove that (X, Σ_0, μ_0) is a complete measure space and that $\Sigma_{\mu^*} = \mathcal{P}(X)$. Thus, the result of (v) does not hold in general.
   (vii) Prove that (X, Σ_{μ^*}, μ) is always the saturation (see Exercise 2.6.20) of the completion of (X, Σ_0, μ_0).

8. *Measures on algebras.*

   Let $\mathcal{A}$ be an algebra of subsets of X. We say that $\mu : \mathcal{A} \to [0, +\infty]$ is **a measure on** $(X, \mathcal{A})$ if
   (i) $\mu(\emptyset) = 0$ and
   (ii) $\mu(\cup_{j=1}^{+\infty} A_j) = \sum_{j=1}^{+\infty} \mu(A_j)$ for all pairwise disjoint $A_1, A_2, \ldots \in \mathcal{A}$ with $\cup_{j=1}^{+\infty} A_j \in \mathcal{A}$.

   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 $\mathcal{A}$).

9. *Extension of a measure, II.*

   Let $\mathcal{A}_0$ be an algebra of subsets of the non-empty X and μ_0 be a measure on $(X, \mathcal{A}_0)$ (see Exercise 3.3.8). For every $E \subseteq X$ we define

$$\mu^*(E) = \inf\Big\{ \sum_{j=1}^{+\infty} \mu_0(A_j) \,|\, A_1, A_2, \ldots \in \mathcal{A}_0, E \subseteq \cup_{j=1}^{+\infty} A_j \Big\},$$

(i) Prove that μ^* is an outer measure on X. We say that μ^* **is induced by the measure** μ_0.
(ii) Prove that $\mu^*(A) = \mu_0(A)$ for every $A \in \mathcal{A}_0$.
(iii) Prove that every $A \in \mathcal{A}_0$ is μ^*-measurable and hence $\Sigma(\mathcal{A}_0) \subseteq \Sigma_{\mu^*}$.

Thus, if (after Caratheodory's theorem) μ is the restriction of μ^* on Σ_{μ^*}, the measure space (X, Σ_{μ^*}, μ) is a complete measure space which extends $(X, \mathcal{A}_0, \mu_0)$.
If we consider the restriction $(X, \Sigma(\mathcal{A}_0), \mu)$, then this is also a measure space (perhaps not complete) which extends $(X, \mathcal{A}_0, \mu_0)$.

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

10. *Regular outer measures.*

    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 $A \in \Sigma_{\mu^*}$ so that $E \subseteq A$ and $\mu^*(E) = \mu(A)$ (where μ is the usual restriction of μ^* on Σ_{μ^*}).

    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 (see Exercise 3.3.9).

11. *Measurable covers.*

    Let μ^* be an outer measure on X and μ be the induced measure (the restriction of μ^*) on Σ_{μ^*}. If $E, G \subseteq X$ we say that G **is a μ^*-measurable cover of** E if $E \subseteq G$, $G \in \Sigma_{\mu^*}$ and for all $A \in \Sigma_{\mu^*}$ 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 $E \subseteq G$, $G \in \Sigma_{\mu^*}$ and $\mu^*(E) = \mu(G)$. If $\mu^*(E) < +\infty$, prove that G is a μ^*-measurable cover of E.

12. We say $E \subseteq \mathbf{R}$ **has an infinite condensation point** if E has uncountably many points outside every bounded interval. Define $\mu^*(E) = 0$ if E is countable, $\mu^*(E) = 1$ if E is uncountable and does not have an infinite condensation point and $\mu^*(E) = +\infty$ if E has an infinite condensation point. Prove that μ^* is an outer measure on $\mathbf{R}$ and that $A \subseteq \mathbf{R}$ is μ^*-measurable if and only if either A or A^c is countable. Does every $E \subseteq \mathbf{R}$ have a μ^*-measurable cover? Is μ^* a regular outer measure? (See exercises 3.3.10 and 3.3.11).

13. Consider the collection $\mathcal{C}$ of subsets of $\mathbf{N}$ which contains $\emptyset$ and all the two-point subsets of $\mathbf{N}$. Define $\tau(\emptyset) = 0$ and $\tau(C) = 2$ for all other $C \in \mathcal{C}$. Calculate $\mu^*(E)$ for all $E \subseteq \mathbf{N}$, where μ^* is the outer measure defined as in Theorem 3.2, and find all the μ^*-measurable subsets of $\mathbf{N}$.

Chapter 4

Lebesgue measure on $\mathbf{R}^n$

4.1 Volume of intervals.

We consider the function $\mathrm{vol}_n(S)$ defined for intervals S in $\mathbf{R}^n$, which is just the product of the lengths of the edges of S: the so-called **(n-dimensional) volume** of S. In this section we shall investigate some properties of the volume of intervals.

Lemma 4.1 *Let $P = (a_1, b_1] \times \cdots \times (a_n, b_n]$ and, for each $k = 1, \ldots, n$, let $a_k = c_k^{(0)} < c_k^{(1)} < \cdots < c_k^{(m_k)} = b_k$. If we 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 $1 \le i_1 \le m_1, \ldots, 1 \le i_n \le m_n$, then*

$$vol_n(P) = \sum_{1 \le i_1 \le m_1, \ldots, 1 \le i_n \le m_n} vol_n(P_{i_1,\ldots,i_n}).$$

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

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

Referring to the situation described by Lemma 4.1 we shall use the expression: *the intervals $P_{i_1,\ldots,i_n}$ result from P by subdivision of its edges.*

Lemma 4.2 *Let $P, P_1, \ldots, P_l$ be open-closed intervals and $P_1, \ldots, P_l$ be pairwise disjoint. If $P = P_1 \cup \cdots \cup P_l$, then $vol_n(P) = vol_n(P_1) + \cdots + vol_n(P_l)$.*

Proof: 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 every $j = 1, \ldots, l$.

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

i. 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 $P_{i_1,\ldots,i_n}$'s result from P by subdivision of its edges. It is also almost clear that

ii. the intervals among the $P_{i_1,\ldots,i_n}$ which belong to a P_j result from it by subdivision of its edges (this is due to i).

iii. 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).

We now calculate, using Lemma 4.1 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:

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

Lemma 4.3 *Let $P, P_1, \ldots, P_l$ be open-closed intervals and $P_1, \ldots, P_l$ be pairwise disjoint. If $P_1 \cup \cdots \cup P_l \subseteq P$, then $vol_n(P_1) + \cdots + vol_n(P_l) \leq vol_n(P)$.*

Proof: We know from Proposition 1.11 that $P \setminus (P_1 \cup \cdots \cup P_l) = P_1' \cup \cdots \cup P_k'$ for some pairwise disjoint open-closed intervals $P_1', \ldots, P_k'$. Then $P = P_1 \cup \cdots \cup P_l \cup P_1' \cup \cdots \cup P_k'$ and Lemma 4.2 now implies that $\mathrm{vol}_n(P) = \mathrm{vol}_n(P_1) + \cdots + \mathrm{vol}_n(P_l) + \mathrm{vol}_n(P_1') + \cdots + \mathrm{vol}_n(P_k') \geq \mathrm{vol}_n(P_1) + \cdots + \mathrm{vol}_n(P_l)$.

Lemma 4.4 *Let $P, P_1, \dots, P_l$ be open-closed intervals. If $P \subseteq P_1 \cup \dots \cup P_l$, then $vol_n(P) \le vol_n(P_1) + \dots + vol_n(P_l)$.*

Proof: We first write $P = P_1' \cup \dots \cup P_l'$ where $P_j' = P_j \cap P$ are open-closed intervals included in P. We then write $P = P_1' \cup (P_2' \setminus P_1') \cup \dots \cup \big(P_l' \setminus (P_1' \cup \dots \cup P_{l-1}')\big)$. Each of these l pairwise disjoint sets can, by Proposition 1.11, be written as a finite union of pairwise disjoint open-closed intervals: $P_1' = P_1'$ and

$$P_j' \setminus (P_1' \cup \dots \cup P_{j-1}') = P_1^{(j)} \cup \dots \cup P_{m_j}^{(j)}$$

for $2 \le j \le l$.

Lemma 4.2 for the equality and Lemma 4.3 for the two inequalities imply

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

Lemma 4.5 *Let Q be a closed interval and $R_1, \dots, R_l$ be open intervals so that $Q \subseteq R_1 \cup \dots \cup R_l$. Then $vol_n(Q) \le vol_n(R_1) + \dots + vol_n(R_l)$.*

Proof: Let P and P_j be the open-closed intervals with the same edges as Q and, respectively, R_j. Then $P \subseteq Q \subseteq R_1 \cup \dots \cup R_l \subseteq P_1 \cup \dots \cup P_l$ and by Lemma 4.4, $\mathrm{vol}_n(Q) = \mathrm{vol}_n(P) \le \mathrm{vol}_n(P_1) + \dots + \mathrm{vol}_n(P_l) = \mathrm{vol}_n(R_1) + \dots + \mathrm{vol}_n(R_l)$.

4.2 Lebesgue measure in $\mathbf{R}^n$.

Consider the collection $\mathcal{C}$ of all open intervals in $\mathbf{R}^n$ 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 $R = (a_1, b_1) \times \dots \times (a_n, b_n) \in \mathcal{C}$.

If we define

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

for all $E \subseteq \mathbf{R}^n$, then Theorem 3.2 implies that m_n^* is an outer measure on $\mathbf{R}^n$. We observe that, since $\mathbf{R}^n = \cup_{k=1}^{+\infty} (-k, k) \times \dots \times (-k, k)$, for every $E \subseteq \mathbf{R}^n$ there is a countable covering of E by elements of $\mathcal{C}$.

Now Theorem 3.1 implies that the collection

$$\mathcal{L}_n = \Sigma_{m_n^*}$$

of m_n^*-measurable sets is a σ-algebra of subsets of $\mathbf{R}^n$ and, if m_n is defined as the restriction of m_n^* on $\mathcal{L}_n$, then m_n is a complete measure on $(X, \mathcal{L}_n)$.

Definition 4.1 *(i)* $\mathcal{L}_n$ *is called* ***the*** σ***-algebra of Lebesgue sets in*** $\mathbf{R}^n$*,*
(ii) m_n^* *is called* ***the (n-dimensional) Lebesgue outer measure on*** $\mathbf{R}^n$ *and*
(iipai) m_n *is called* ***the (n-dimensional) Lebesgue measure on*** $\mathbf{R}^n$*.*

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

Theorem 4.1 *Every interval* S *in* $\mathbf{R}^n$ *is a Lebesgue set and*

$$m_n(S) = vol_n(S).$$

Proof: Let $Q = [a_1, b_1] \times \cdots \times [a_n, b_n]$.

Since $Q \subseteq (a_1 - \epsilon, b_1 + \epsilon) \times \cdots \times (a_n - \epsilon, b_n + \epsilon)$, we get by the definition of m_n^* that $m_n^*(Q) \leq \text{vol}_n((a_1 - \epsilon, b_1 + \epsilon) \times \cdots \times (a_n - \epsilon, b_n + \epsilon)) = (b_1 - a_1 + 2\epsilon) \cdots (b_n - a_n + 2\epsilon)$. Since $\epsilon > 0$ is arbitrary, we find $m_n^*(Q) \leq \text{vol}_n(Q)$.

Now take any covering, $Q \subseteq R_1 \cup R_2 \cup \cdots$ of Q by open intervals. Since Q is compact, there is l so that $Q \subseteq R_1 \cup \cdots \cup R_l$ and Lemma 4.5 implies that $\text{vol}_n(Q) \leq \text{vol}_n(R_1) + \cdots + \text{vol}_n(R_l) \leq \sum_{k=1}^{+\infty} \text{vol}_n(R_k)$. Taking the infimum of the right side, we get $\text{vol}_n(Q) \leq m_n^*(Q)$ and, hence,

$$m_n^*(Q) = \text{vol}_n(Q).$$

Now take any general interval S and let $a_1, b_1, \ldots, a_n, b_n$ be the end-points of its edges. Then $Q' \subseteq S \subseteq Q''$, where $Q' = [a_1+\epsilon, b_1-\epsilon] \times \cdots \times [a_n+\epsilon, b_n-\epsilon]$ and $Q'' = [a_1 - \epsilon, b_1 + \epsilon] \times \cdots \times [a_n - \epsilon, b_n + \epsilon]$. Hence $m_n^*(Q') \leq m_n^*(S) \leq m_n^*(Q'')$, namely $(b_1-a_1-2\epsilon)\cdots(b_n-a_n-2\epsilon) \leq m_n^*(S) \leq (b_1-a_1+2\epsilon)\cdots(b_n-a_n+2\epsilon)$. Since $\epsilon > 0$ is arbitrary, we find

$$m_n^*(S) = \text{vol}_n(S).$$

Consider an open-closed interval P and an open interval R. Take the open-closed interval P_R with the same edges as R. Then $m_n^*(R \cap P) \leq m_n^*(P_R \cap P) = \text{vol}_n(P_R \cap P)$ and $m_n^*(R \cap P^c) \leq m_n^*(P_R \cap P^c)$. Now Proposition 1.11 implies $P_R \cap P^c = P_R \setminus P = P_1' \cup \cdots \cup P_k'$ for some pairwise disjoint open-closed intervals $P_1', \ldots, P_k'$. Hence $m_n^*(R \cap P^c) \leq m_n^*(P_1') + \cdots + m_n^*(P_k') = \text{vol}_n(P_1') + \cdots + \text{vol}_n(P_k')$. Altogether, $m_n^*(R \cap P) + m_n^*(R \cap P^c) \leq \text{vol}_n(P_R \cap P) + \text{vol}_n(P_1') + \cdots + \text{vol}_n(P_k')$ and, by Lemma 4.2, this is $= \text{vol}_n(P_R) = \text{vol}_n(R)$. We have just proved that

$$m_n^*(R \cap P) + m_n^*(R \cap P^c) \leq \text{vol}_n(R).$$

Consider any open-closed interval P and any $E \subseteq \mathbf{R}^n$ with $m_n^*(E) < +\infty$. Take, for arbitrary $\epsilon > 0$, a covering $E \subseteq \cup_{j=1}^{+\infty} R_j$ of E by open intervals so that $\sum_{j=1}^{+\infty} \text{vol}_n(R_j) < m_n^*(E) + \epsilon$. This implies $m_n^*(E \cap P) + m_n^*(E \cap P^c) \leq \sum_{j=1}^{+\infty} m_n^*(R_j \cap P) + \sum_{j=1}^{+\infty} m_n^*(R_j \cap P^c) = \sum_{j=1}^{+\infty} [m_n^*(R_j \cap P) + m_n^*(R_j \cap P^c)]$ which, by the last result, is $\leq \sum_{j=1}^{+\infty} \text{vol}_n(R_j) < m_n^*(E) + \epsilon$. This implies that

$$m_n^*(E \cap P) + m_n^*(E \cap P^c) \leq m_n^*(E)$$

and P is a Lebesgue set.

If T is any interval at least one of whose edges is a single point, then $m_n^*(T) = \mathrm{vol}_n(T) = 0$ and, by Proposition 3.1, T is a Lebesgue set. Now, any interval S differs from the open-closed interval P, which has the same sides as S, by finitely many (at most $2n$) T's, and hence S is also a Lebesgue set.

Theorem 4.2 *Lebesgue measure is σ-finite but not finite.*

Proof: We write $\mathbf{R}^n = \cup_{k=1}^{+\infty} Q_k$ with $Q_k = [-k,k] \times \cdots \times [-k,k]$, where $m_n(Q_k) = \mathrm{vol}_n(Q_k) < +\infty$ for all k. On the other hand, for all k, $m_n(\mathbf{R}^n) \geq m_n(Q_k) = (2k)^n$ and, hence, $m_n(\mathbf{R}^n) = +\infty$.

Theorem 4.3 *All Borel sets in $\mathbf{R}^n$ are Lebesgue sets.*

Proof: Theorem 4.1 says that, if $\mathcal{E}$ is the collection of all intervals in $\mathbf{R}^n$, then $\mathcal{E} \subseteq \mathcal{L}_n$. But then $\mathcal{B}_{\mathbf{R}^n} = \Sigma(\mathcal{E}) \subseteq \mathcal{L}_n$.

Therefore all open and all closed subsets of $\mathbf{R}^n$ are Lebesgue sets.

Theorem 4.4 *Let $E \subseteq \mathbf{R}^n$. Then*
(i) $E \in \mathcal{L}_n$ if and only if there is A, a countable intersection of open sets, so that $E \subseteq A$ and $m_n^*(A \setminus E) = 0$.
(ii) $E \in \mathcal{L}_n$ if and only if there is B, a countable union of compact sets, so that $B \subseteq E$ and $m_n^*(E \setminus B) = 0$.

Proof: (i) One direction is easy. If there is A, a countable intersection of open sets, so that $E \subseteq A$ and $m_n^*(A \setminus E) = 0$, then, by Proposition 3.1, $A \setminus E \in \mathcal{L}_n$ and, thus, $E = A \setminus (A \setminus E) \in \mathcal{L}_n$.

To prove the other direction consider, after Theorem 4.2, $Y_1, Y_2, \ldots \in \mathcal{L}_n$ so that $\mathbf{R}^n = \cup_{k=1}^{+\infty} Y_k$ and $m_n(Y_k) < +\infty$ for all k. Define $E_k = E \cap Y_k$ so that $E = \cup_{k=1}^{+\infty} E_k$ and $m_n(E_k) < +\infty$ for all k.

For all k and arbitrary $l \in \mathbf{N}$ find a covering $E_k \subseteq \cup_{j=1}^{+\infty} R_j^{(k,l)}$ by open intervals so that $\sum_{j=1}^{+\infty} \mathrm{vol}_n(R_j^{(k,l)}) < m_n(E_k) + \frac{1}{l2^k}$ and set $U^{(k,l)} = \cup_{j=1}^{+\infty} R_j^{(k,l)}$. Then $E_k \subseteq U^{(k,l)}$ and $m_n(U^{(k,l)}) < m_n(E_k) + \frac{1}{l2^k}$, from which

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

Now set $U^{(l)} = \cup_{k=1}^{+\infty} U^{(k,l)}$. Then $U^{(l)}$ is open and $E \subseteq U^{(l)}$ and it is trivial to see that $U^{(l)} \setminus E \subseteq \cup_{k=1}^{+\infty} (U^{(k,l)} \setminus E_k)$, from which we get

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

Finally, define $A = \cap_{l=1}^{+\infty} U^{(l)}$ to get $E \subseteq A$ and $m_n(A \setminus E) \leq m_n(U^{(l)} \setminus E) < \frac{1}{l}$ for all l and, thus,

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

(ii) If B is a countable union of compact sets so that $B \subseteq E$ and $m_n^*(E \setminus B) = 0$, then, by Proposition 3.1, $E \setminus B \in \mathcal{L}_n$ and 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 an 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 $B = \cup_{j=1}^{+\infty} F_j$, where each F_j is closed. We then write $F_j = \cup_{k=1}^{+\infty} F_{j,k}$, 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: $B = \cup_{(j,k) \in \mathbf{N} \times \mathbf{N}} F_{j,k}$.

Theorem 4.4 says that *every Lebesgue set in* $\mathbf{R}^n$ *is, except from a null set, equal to a Borel set.*

Theorem 4.5 *(i)* m_n *is the only measure on* $(\mathbf{R}^n, \mathcal{B}_{\mathbf{R}^n})$ *with* $m_n(P) = vol_n(P)$ *for every open-closed interval* P.
(ii) $(\mathbf{R}^n, \mathcal{L}_n, m_n)$ *is the completion of* $(\mathbf{R}^n, \mathcal{B}_{\mathbf{R}^n}, m_n)$.

Proof: (i) If μ is any measure on $(\mathbf{R}^n, \mathcal{B}_{\mathbf{R}^n})$ with $\mu(P) = \mathrm{vol}_n(P)$ for all open-closed intervals P, then it is trivial to see that $\mu(P) = +\infty$ for any unbounded generalised open-closed interval P: just take any increasing sequence of open-closed intervals having union P. Therefore $\mu(\cup_{j=1}^m P_j) = \sum_{j=1}^m \mu(P_j) = \sum_{j=1}^m m_n(P_j) = m_n(\cup_{j=1}^m P_j)$ for all pairwise disjoint open-closed generalised intervals $P_1, \ldots, P_m$. Therefore the measures μ and m_n are equal on the algebra $\mathcal{A} = \{\cup_{j=1}^m P_j \,|\, m \in \mathbf{N}, P_1, \ldots, P_m$ pairwise disjoint open-closed generalised intervals$\}$. By Theorem 2.4, the two measures are equal also on $\Sigma(\mathcal{A}) = \mathcal{B}_{\mathbf{R}^n}$.
(ii) Let $(\mathbf{R}^n, \overline{\mathcal{B}_{\mathbf{R}^n}}, \overline{m_n})$ be the completion of $(\mathbf{R}^n, \mathcal{B}_{\mathbf{R}^n}, m_n)$.

By Theorem 4.3, $(\mathbf{R}^n, \mathcal{L}_n, m_n)$ is a complete extension of $(\mathbf{R}^n, \mathcal{B}_{\mathbf{R}^n}, m_n)$. Hence, $\overline{\mathcal{B}_{\mathbf{R}^n}} \subseteq \mathcal{L}_n$ and $\overline{m_n}(E) = m_n(E)$ for every $E \in \overline{\mathcal{B}_{\mathbf{R}^n}}$.

Take any $E \in \mathcal{L}_n$ and, using Theorem 4.4, find a Borel set B so that $B \subseteq E$ and $m_n(E \setminus B) = 0$. Using Theorem 4.4 once more, find a Borel set A so that $(E \setminus B) \subseteq A$ and $m_n(A \setminus (E \setminus B)) = 0$. Therefore, $m_n(A) = m_n(A \setminus (E \setminus B)) + m_n(E \setminus B) = 0$.

Hence, we can write $E = B \cup L$, where $B \in \mathcal{B}_{\mathbf{R}^n}$ and $L = E \setminus B \subseteq A \in \mathcal{B}_{\mathbf{R}^n}$ with $m_n(A) = 0$. After Theorem 2.3, we see that E has the form of the typical element of $\overline{\mathcal{B}_{\mathbf{R}^n}}$ and, thus, $\mathcal{L}_n \subseteq \overline{\mathcal{B}_{\mathbf{R}^n}}$. This concludes the proof.

Theorem 4.6 *Suppose* $E \in \mathcal{L}_n$ *with* $m_n(E) < +\infty$. *For arbitrary* $\epsilon > 0$, *there are pairwise disjoint open intervals* $R_1, \ldots, R_l$ *so that* $m_n(E \triangle (R_1 \cup \cdots \cup R_l)) < \epsilon$.

Proof: We consider a covering $E \subseteq \cup_{j=1}^{+\infty} R_j'$ by open intervals such that $\sum_{j=1}^{+\infty} \mathrm{vol}_n(R_j') < m_n(E) + \frac{\epsilon}{2}$.

Now we consider the open-closed interval P_j' which has the same edges as R_j', and then $E \subseteq \cup_{j=1}^{+\infty} P_j'$ and $\sum_{j=1}^{+\infty} \mathrm{vol}_n(P_j') < m_n(E) + \frac{\epsilon}{2}$.

We take m so that $\sum_{j=m+1}^{+\infty} \mathrm{vol}_n(P_j') < \frac{\epsilon}{2}$ and we observe the inclusions $E \setminus (P_1' \cup \cdots \cup P_m') \subseteq \cup_{j=m+1}^{+\infty} P_j'$ and $(P_1' \cup \cdots \cup P_m') \setminus E \subseteq \left(\cup_{j=1}^{+\infty} P_j'\right) \setminus E$. Thus, $m_n(E \setminus (P_1' \cup \cdots \cup P_m')) \leq \sum_{j=m+1}^{+\infty} \mathrm{vol}_n(P_j') < \frac{\epsilon}{2}$ and $m_n((P_1' \cup \cdots \cup P_m') \setminus E) \leq$

$m_n(\cup_{j=1}^{+\infty} P'_j) - m_n(E) < \frac{\epsilon}{2}$. Adding, we find

$$m_n(E \triangle (P'_1 \cup \cdots \cup P'_m)) < \epsilon.$$

Proposition 1.11 implies that $P'_1 \cup \cdots \cup P'_m = P_1 \cup \cdots \cup P_l$ for some pairwise disjoint open-closed intervals $P_1, \cdots, P_l$ and, thus,

$$m_n(E \triangle (P_1 \cup \cdots \cup P_l)) < \epsilon.$$

We consider R_k to be the open interval with the same edges as P_k so that $\cup_{k=1}^{l} R_k \subseteq \cup_{k=1}^{l} P_k$ and $m_n((\cup_{k=1}^{l} P_k) \setminus (\cup_{k=1}^{l} R_k)) \leq \sum_{k=1}^{l} m_n(P_k \setminus R_k) = 0$. This, easily, implies that

$$m_n(E \triangle (R_1 \cup \cdots \cup R_l)) = m_n(E \triangle (P_1 \cup \cdots \cup P_l)) < \epsilon.$$

4.3 Lebesgue measure and simple transformations.

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

Every $y \in \mathbf{R}^n$ defines the **translation** $\tau_y : \mathbf{R}^n \to \mathbf{R}^n$ by the formula

$$\tau_y(x) = x + y, \quad x \in \mathbf{R}^n.$$

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

$$y + E = \{y + x \,|\, x \in E\} \,(= \tau_y(E)).$$

Every $\lambda > 0$ defines the **dilation** $l_\lambda : \mathbf{R}^n \to \mathbf{R}^n$ by the formula

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

Then l_λ is an one-to-one transformation of $\mathbf{R}^n$ onto $\mathbf{R}^n$ and its inverse transformation is $l_{\frac{1}{\lambda}}$. For every $E \subseteq \mathbf{R}^n$ we define

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

If S is any interval in $\mathbf{R}$, then any translation transforms it onto another interval (of the same type) with the same volume. In fact, if $a_1, b_1, \ldots, a_n, b_n$ are the end-points of the edges of S, then the translated $y + S$ has $y_1 + a_1, y_1 + b_1, \ldots, y_n + a_n, y_n + b_n$ as end-points of its edges. Therefore $\mathrm{vol}_n(y + S) = \big((y_1+b_1)-(y_1+a_1)\big) \cdots \big((y_n+b_n)-(y_n+a_n)\big) = (b_1-a_1) \cdots (b_n-a_n) = \mathrm{vol}_n(S)$.

If we dilate the interval S with $a_1, b_1, \ldots, a_n, b_n$ as end-points of its edges by the number $\lambda > 0$, then we get the interval λS with $\lambda a_1, \lambda b_1, \ldots, \lambda a_n, \lambda b_n$ as end-points of its edges. Therefore, $\mathrm{vol}_n(\lambda S) = (\lambda b_1 - \lambda a_1) \cdots (\lambda b_n - \lambda a_n) = \lambda^n (b_1 - a_1) \cdots (b_n - a_n) = \lambda^n \mathrm{vol}_n(S)$.

Another transformation is r, **reflection through** 0, with the formula

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

This is one-to-one onto $\mathbf{R}^n$ and it is the inverse of itself. We define

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

for all $E \subseteq \mathbf{R}^n$. If S is any interval with $a_1, b_1, \ldots, a_n, b_n$ as end-points of its edges, then $-S$ is an interval with $-b_1, -a_1, \ldots, -b_n, -a_n$ as end-points of its edges and $\mathrm{vol}_n(-S) = (-a_1 + b_1) \cdots (-a_n + b_n) = \mathrm{vol}_n(S)$.

After all these, we may say that *n-dimensional volume of intervals is invariant under translations and reflection* and *it is positive-homogeneous of degree n under dilations.*

We shall see that the same are true for n-dimensional Lebesgue measure of Lebesgue sets in $\mathbf{R}^n$.

Theorem 4.7 *(i) $\mathcal{L}_n$ is invariant under translations, reflection and dilations. That is, for all $A \in \mathcal{L}_n$ we have that $y+A, -A, \lambda A \in \mathcal{L}_n$ for every $y \in \mathbf{R}^n, \lambda > 0$.*
(ii) m_n is invariant under translations and reflection and positive-homogeneous of degree n under dilations. That is, for all $A \in \mathcal{L}_n$ we have that

$$m_n(y + A) = m_n(A), \quad m_n(-A) = m_n(A), \quad m_n(\lambda A) = \lambda^n m_n(A)$$

for every $y \in \mathbf{R}^n, \lambda > 0$.

Proof: Let $E \subseteq \mathbf{R}^n$ and $y \in \mathbf{R}^n$. Then for all coverings $E \subseteq \cup_{j=1}^{+\infty} R_j$ by open intervals we get $y + E \subseteq \cup_{j=1}^{+\infty}(y + R_j)$. Therefore, $m_n^*(y + E) \leq \sum_{j=1}^{+\infty} \mathrm{vol}_n(y + R_j) = \sum_{j=1}^{+\infty} \mathrm{vol}_n(R_j)$. Taking the infimum of the right side, we find that $m_n^*(y + E) \leq m_n^*(E)$. Now, applying this to $y + E$ translated by $-y$, we get $m_n^*(E) = m_n^*(-y + (y + E)) \leq m_n^*(y + E)$. Hence

$$m_n^*(y + E) = m_n^*(E)$$

for all $E \subseteq \mathbf{R}^n$ and $y \in \mathbf{R}^n$.

Similarly, $-E \subseteq \cup_{j=1}^{+\infty}(-R_j)$, which implies $m_n^*(-E) \leq \sum_{j=1}^{+\infty} \mathrm{vol}_n(-R_j) = \sum_{j=1}^{+\infty} \mathrm{vol}_n(R_j)$. Hence $m_n^*(-E) \leq m_n^*(E)$. Applying this to $-E$, we also get $m_n^*(E) = m_n^*(-(-E)) \leq m_n^*(-E)$ and, thus,

$$m_n^*(-E) = m_n^*(E)$$

for all $E \subseteq \mathbf{R}^n$.

Also, $\lambda E \subseteq \cup_{j=1}^{+\infty}(\lambda R_j)$, from which we get $m_n^*(\lambda E) \leq \sum_{j=1}^{+\infty} \mathrm{vol}_n(\lambda R_j) = \lambda^n \sum_{j=1}^{+\infty} \mathrm{vol}_n(R_j)$ and hence $m_n^*(\lambda E) \leq \lambda^n m_n^*(E)$. Applying to $\frac{1}{\lambda}$ and to λE, we find $m_n^*(E) = m_n^*(\frac{1}{\lambda}(\lambda E)) \leq (\frac{1}{\lambda})^n m_n^*(\lambda E)$, which gives

$$m_n^*(\lambda E) = \lambda^n m_n^*(E).$$

Suppose now that $A \in \mathcal{L}_n$ and $E \subseteq \mathbf{R}^n$.

We have $m_n^*(E \cap (y+A)) + m_n^*(E \cap (y+A)^c) = m_n^*\big(y + [(-y+E) \cap A]\big) + m_n^*\big(y + [(-y+E) \cap A^c]\big) = m_n^*\big((-y+E) \cap A\big) + m_n^*\big((-y+E) \cap A^c\big) = m_n^*(-y+E) = m_n^*(E)$. Therefore, $y + A \in \mathcal{L}_n$.

In the same way, $m_n^*(E \cap (-A)) + m_n^*(E \cap (-A)^c) = m_n^*\big(-[(-E) \cap A]\big) + m_n^*\big(-[(-E) \cap A^c]\big) = m_n^*\big((-E) \cap A\big) + m_n^*\big((-E) \cap A^c\big) = m_n^*(-E) = m_n^*(E)$. Therefore, $-A \in \mathcal{L}_n$.

We, finally, have $m_n^*(E \cap (\lambda A)) + m_n^*(E \cap (\lambda A)^c) = m_n^*\big(\lambda[(\frac{1}{\lambda}E) \cap A]\big) + m_n^*\big(\lambda[(\frac{1}{\lambda}E) \cap A^c]\big) = \lambda^n m_n^*\big((\frac{1}{\lambda}E) \cap A\big) + \lambda^n m_n^*\big((\frac{1}{\lambda}E) \cap A^c\big) = \lambda^n m_n^*(\frac{1}{\lambda}E) = m_n^*(E)$. Therefore, $\lambda A \in \mathcal{L}_n$.

If $A \in \mathcal{L}_n$, then $m_n(y+A) = m_n^*(y+A) = m_n^*(A) = m_n(A)$, $m_n(-A) = m_n^*(-A) = m_n^*(A) = m_n(A)$ and $m_n(\lambda A) = m_n^*(\lambda A) = \lambda^n m_n^*(A) = \lambda^n m_n(A)$.

Reflection and dilations are special cases of linear transformations of $\mathbf{R}^n$. As is well known, a *linear transformation of* $\mathbf{R}^n$ is a function $T : \mathbf{R}^n \to \mathbf{R}^n$ such that

$$T(x+y) = T(x) + T(y)\,, \quad T(\kappa x) = \kappa T(x)\,, \qquad x, y \in \mathbf{R}^n, \kappa \in \mathbf{R},$$

and every such T has a *determinant*, $\det(T) \in \mathbf{R}$. In particular, $\det(r) = (-1)^n$ and $\det(l_\lambda) = \lambda^n$.

We recall that a linear transfomation T of $\mathbf{R}^n$ is one-to-one and onto $\mathbf{R}^n$ if and only if $\det(T) \neq 0$. Moreover, if $\det(T) \neq 0$, then T^{-1} is also a linear transformation of $\mathbf{R}^n$ and $\det(T^{-1}) = (\det(T))^{-1}$. Finally, if T, T_1, T_2 are linear transformations of $\mathbf{R}^n$ and $T = T_1 \circ T_2$, then $\det(T) = \det(T_1)\det(T_2)$.

Theorem 4.8 *Let $T : \mathbf{R}^n \to \mathbf{R}^n$ 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 $0 \cdot (+\infty) = 0$.

Proof: At first we assume that $\det(T) \neq 0$.

If T has the form $T(x_1, x_2, \ldots, x_n) = (\lambda x_1, x_2, \ldots, x_n)$ for a certain $\lambda \in \mathbf{R} \setminus \{0\}$, then $\det(T) = \lambda$ and, if $P = (a_1, b_1] \times (a_2, b_2] \times \cdots \times (a_n, b_n]$, then $T(P) = (\lambda a_1, \lambda b_1] \times (a_2, b_2] \times \cdots \times (a_n, b_n]$ or $T(P) = [\lambda b_1, \lambda a_1) \times (a_2, b_2] \times \cdots \times (a_n, b_n]$, depending on whether $\lambda > 0$ or $\lambda < 0$. Thus $T(P)$ is an interval and $m_n(T(P)) = |\lambda| m_n(P) = |\det(T)| m_n(P)$.

If $T(x_1, x_2, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_n) = (x_i, x_2, \ldots, x_{i-1}, x_1, x_{i+1}, \ldots, x_n)$ for a certain $i \neq 1$, then $\det(T) = -1$ and, if $P = (a_1, b_1] \times (a_2, b_2] \times \cdots \times (a_{i-1}, b_{i-1}] \times (a_i, b_i] \times (a_{i+1}, b_{i+1}] \times \cdots \times (a_n, b_n]$, then $T(P) = (a_i, b_i] \times (a_2, b_2] \times \cdots \times (a_{i-1}, b_{i-1}] \times (a_1, b_1] \times (a_{i+1}, b_{i+1}] \times \cdots \times (a_n, b_n]$. Thus $T(P)$ is an interval and $m_n(T(P)) = m_n(P) = |\det(T)| m_n(P)$.

If $T(x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_n) = (x_1, \ldots, x_{i-1}, x_i + x_1, x_{i+1}, \ldots, x_n)$ for a certain $i \neq 1$, then $\det(T) = 1$ and, if $P = (a_1, b_1] \times \cdots \times (a_{i-1}, b_{i-1}] \times (a_i, b_i] \times (a_{i+1}, b_{i+1}] \times \cdots \times (a_n, b_n]$, then $T(P)$ is not an interval any more but $T(P) = \{(y_1, \ldots, y_n) \,|\, y_j \in (a_j, b_j] \text{ for } j \neq i, y_i - y_1 \in (a_i, b_i]\}$ is a Borel set and hence it is in $\mathcal{L}_n$. We define the following three auxilliary sets: $L = (a_1, b_1] \times \cdots \times (a_{i-1}, b_{i-1}] \times (a_i + a_1, b_i + b_1] \times (a_{i+1}, b_{i+1}] \times \cdots \times (a_n, b_n]$, $M = \{(y_1, \ldots, y_n) \,|\, y_j \in (a_j, b_j] \text{ for } j \neq i, a_i + a_1 < y_i \leq a_i + y_1\}$ and $N = \{(y_1, \ldots, y_n) \,|\, y_j \in (a_j, b_j] \text{ for } j \neq i, b_i + a_1 < y_i \leq b_i + y_1\}$. It is easy to see that all four sets, $T(P), L, M, N$, are Borel sets and $T(P) \cap M = \emptyset$, $L \cap N = \emptyset$, $T(P) \cup M = L \cup N$ and that $N = x_0 + M$, where $x_0 = (0, \ldots, 0, b_i - a_i, 0, \ldots, 0)$. Then $m_n(T(P)) + m_n(M) = m_n(L) + m_n(N)$ and $m_n(M) = m_n(N)$, implying that $m_n(T(P)) = m_n(L) = m_n(P) = |\det(T)| m_n(P)$, because L is an interval.

Now, let T be any linear transformation of the above three types. We have shown that

$$m_n(T(P)) = |\det(T)| m_n(P)$$

for every open-closed interval P. If $R = (a_1, b_1) \times \cdots \times (a_n, b_n)$ it is easy to see, just as in the case of open-closed intervals, that $T(R)$ is a Borel set. We consider $P_1 = (a_1, b_1] \times \cdots \times (a_n, b_n]$ and $P_2 = (a_1, b_1 - \epsilon] \times \cdots \times (a_n, b_n - \epsilon]$ and, from $P_2 \subseteq R \subseteq P_1$ we get $T(P_2) \subseteq T(R) \subseteq T(P_1)$. Hence $|\det(T)| m_n(P_2) \leq m_n(T(R)) \leq |\det(T)| m_n(P_1) = |\det(T)| m_n(R)$ and, taking the limit as $\epsilon \to 0+$, we find

$$m_n(T(R)) = |\det(T)| m_n(R)$$

for every open interval R.

Let, again, T be any linear transformation of one of the above three types. Take any $E \subseteq \mathbf{R}^n$ and consider an arbitrary covering $E \subset \cup_{j=1}^{+\infty} R_j$ by open intervals. Then $T(E) \subseteq \cup_{j=1}^{+\infty} T(R_j)$ and hence $m_n^*(T(E)) \leq \sum_{j=1}^{+\infty} m_n(T(R_j)) = |\det(T)| \sum_{j=1}^{+\infty} m_n(R_j)$. Taking the infimum over all coverings, we conclude

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

If T is any linear transformation with $\det(T) \neq 0$, by a well-known result of Linear Algebra, there are linear transformations $T_1, \ldots, T_N$, 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 $E \subseteq \mathbf{R}^n$. In this inequality, use now the set $T(E)$ in the place of E and T^{-1} in the place of T, and get $m_n^*(E) \leq |\det(T^{-1})| m_n^*(T(E)) = |\det(T)|^{-1} m_n^*(T(E))$. Combining the 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 $E \subseteq \mathbf{R}^n$.

Let $A \in \mathcal{L}_n$. For all $E \subseteq \mathbf{R}^n$ we get $m_n^*(E \cap T(A)) + m_n^*(E \cap (T(A))^c) = m_n^*\big(T(T^{-1}(E) \cap A)\big) + m_n^*\big(T(T^{-1}(E) \cap A^c)\big) = |\det(T)| [m_n^*(T^{-1}(E) \cap A) + m_n^*(T^{-1}(E) \cap A^c)] = |\det(T)| m_n^*(T^{-1}(E)) = m_n^*(E)$. This says that $T(A) \in \mathcal{L}_n$. Moreover,

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

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

Let $\{f_1, \ldots, f_m\}$ be a base of V (with $m \leq n-1$) and complete it to a base $\{f_1, \ldots, f_m, f_{m+1}, \ldots, f_n\}$ of $\mathbf{R}^n$. Take the linear transformation $S : \mathbf{R}^n \to \mathbf{R}^n$ given by

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

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

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

We have $S(V) = \cup_{k=1}^{+\infty} Q_k$, where $Q_k = [-k, k] \times \cdots \times [-k, k] \times \{0\} \times \cdots \times \{0\}$. Each Q_k is a closed interval in $\mathbf{R}^n$ with $m_n(Q_k) = 0$. Hence, $m_n(S(V)) = 0$ and, then, $m_n(V) = |\det(S)|^{-1} m_n(S(V)) = 0$.

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

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

is the typical *closed parallelepiped* in $\mathbf{R}^n$. One of the vertices of M is b and $b_1, \ldots, b_n$ (interpreted as vectors) are the edges of M which start from b. For such an M we define the linear transformation $T : \mathbf{R}^n \to \mathbf{R}^n$ by $T(x) = T(x_1, \ldots, x_n) = x_1 b_1 + \cdots + x_n b_n$ for every $x = (x_1, \ldots, x_n) \in \mathbf{R}^n$. We also consider the translation τ_b and observe that

$$M = \tau_b\big(T(Q_0)\big),$$

where $Q_0 = [0, 1]^n$ is the unit qube in $\mathbf{R}^n$. Theorems 4.7 and 4.8 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 matrix of T with respect to the standard basis $\{e_1, \ldots, e_n\}$ of $\mathbf{R}^n$ has as columns the vectors $T(e_1) = b_1, \ldots, T(e_n) = b_n$. We conclude with the rule that *the Lebesgue measure of a closed parallelepiped is given by 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 parallelepiped.

4.4 Cantor set.

Since $\{x\}$ is a degenerate interval, we see that $m_n(\{x\}) = \mathrm{vol}_n(\{x\}) = 0$. In fact, every countable subset of $\mathbf{R}^n$ 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 section is to provide an uncountable set in $\mathbf{R}$ whose Lebesgue measure is zero.

We start with the interval

$$I_0 = [0, 1],$$

then take

$$I_1 = \left[0, \frac{1}{3}\right] \cup \left[\frac{2}{3}, 1\right],$$

next

$$I_2 = \left[0, \frac{1}{9}\right] \cup \left[\frac{2}{9}, \frac{1}{3}\right] \cup \left[\frac{2}{3}, \frac{7}{9}\right] \cup \left[\frac{8}{9}, 1\right],$$

and so on, each time dividing each of the intervals we get at the previous stage into three subintervals of equal length and keeping only the two closed subintervals on the sides.

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

$$C = \cap_{n=1}^{+\infty} I_n$$

and call it **the Cantor set**.

C is a compact subset of $[0, 1]$ with $m_1(C) = 0$. To see this observe that for every n, $m_1(C) \leq m_1(I_n) = 2^n \cdot \frac{1}{3^n}$ which tends to 0 as $n \to +\infty$.

We shall prove by contradiction that C is uncountable. Namely, assume that $C = \{x_1, x_2, \ldots\}$. We shall describe an inductive process of picking one from the subintervals constituting each I_n.

It is obvious that every x_n belongs to I_n, since it belongs to C. At the first step 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 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, just take the left one.) And continue inductively: if you have already chosen $I^{(n-1)}$ from the subintervals of I_{n-1}, then this includes two subintervals of I_n. Choose as $I^{(n)}$ whichever of these two subintervals of $I^{(n-1)}$ does *not* contain x_n. (If both do not contain x_n, just take the left one.)

This produces a sequence $(I^{(n)})$ of intervals with the following properties:
(i) $I^{(n)} \subseteq I_n$ for all n,
(ii) $I^{(n)} \subseteq I^{(n-1)}$ for all n,
(iii) $\text{vol}_1(I^{(n)}) = \frac{1}{3^n} \to 0$ and
(iv) $x_n \notin I^{(n)}$ for all n.
From (ii) and (iii) we conclude that the intersection of all $I^{(n)}$'s contains a single point:

$$\cap_{n=1}^{+\infty} I^{(n)} = \{x_0\}$$

for some x_0. From (i) we see that $x_0 \in I_n$ for all n and thus $x_0 \in C$. Therefore, $x_0 = x_n$ for some $n \in \mathbf{N}$. But then $x_0 \in I^{(n)}$ and, by (iv), *the same* point x_n does not belong to $I^{(n)}$.

We get a contradiction and, hence, C is uncountable.

4.5 A non-Lebesgue set in R.

We consider the following equivalence relation in the set $[0, 1)$. For any $x, y \in [0, 1)$ we write $x \sim y$ if and only if $x - y \in \mathbf{Q}$. That $\sim$ is an equivalence relation

is easy to see:
(a) $x \sim x$, because $x - x = 0 \in \mathbf{Q}$.
(b) If $x \sim y$, then $x - y \in \mathbf{Q}$, then $y - x = -(x - y) \in \mathbf{Q}$, then $y \sim x$.
(c) If $x \sim y$ and $y \sim z$, then $x - y \in \mathbf{Q}$ and $y - z \in \mathbf{Q}$, then $x - z = (x - y) + (y - z) \in \mathbf{Q}$, then $x \sim z$.

Using the Axiom of Choice, we form a set N containing exactly one element from each equivalence class of $\sim$. This means that:

(i) for every $x \in [0, 1)$ there is exactly one $\overline{x} \in N$ so that $x - \overline{x} \in \mathbf{Q}$.

Indeed, if we consider the equivalence class of x and the element $\overline{x}$ of N from this equivalence class, then $x \sim \overline{x}$ and hence $x - \overline{x} \in \mathbf{Q}$. Moreover, if there are two $\overline{x}, \overline{\overline{x}} \in N$ so that $x - \overline{x} \in \mathbf{Q}$ and $x - \overline{\overline{x}} \in \mathbf{Q}$, then $x \sim \overline{x}$ and $x \sim \overline{\overline{x}}$, implying that N contains two different elements from the equivalence class of x.

Our aim is to prove that N is not a Lebesgue set.

We form the set

$$A = \cup_{r \in \mathbf{Q} \cap [0,1)} (N + r).$$

Diferent $(N + r)'$s are disjoint:

(ii) if $r_1, r_2 \in \mathbf{Q} \cap [0, 1)$ 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 $x \sim x - r_2$, implying that N contains two different (since $r_1 \neq r_2$) elements from the equivalence class of x.

(iii) $A \subseteq [0, 2)$.

This is clear, since $N \subseteq [0, 1)$ implies $N + r \subseteq [0, 2)$ for all $r \in \mathbf{Q} \cap [0, 1)$.

Take an arbitrary $x \in [0, 1)$ and, by (i), the unique $\overline{x} \in N$ with $x - \overline{x} \in \mathbf{Q}$. Since $-1 < x - \overline{x} < 1$ we consider cases: if $r = x - \overline{x} \in [0, 1)$, then $x = \overline{x} + r \in N + r \subseteq A$, while if $r = x - \overline{x} \in (-1, 0)$, then $x + 1 = \overline{x} + (r + 1) \in N + (r + 1) \subseteq A$. Therefore, for every $x \in [0, 1)$ either $x \in A$ or $x + 1 \in A$. It is easy to see that exactly one of these two cases is true. Because if $x \in A$ and $x + 1 \in A$, then $x \in N + r_1$ and $x + 1 \in N + r_2$ for some $r_1, r_2 \in \mathbf{Q} \cap [0, 1)$. Hence, $x - r_1, x + 1 - r_2 \in N$ and N contains two different (since $r_2 - r_1 \neq 1$) elements of the equivalence class of x. Thus, if we define the sets

$$E_1 = \{x \in [0, 1) \,|\, x \in A\}, \quad E_2 = \{x \in [0, 1) \,|\, x + 1 \in A\}$$

then we have proved that

(iv) $E_1 \cup E_2 = [0, 1), \quad E_1 \cap E_2 = \emptyset$.

From (iv) we shall need only that $[0, 1) \subseteq E_1 \cup E_2$.

We can also prove that

(v) $E_1 \cup (E_2 + 1) = A, \quad E_1 \cap (E_2 + 1) = \emptyset$.

In fact, the second is easy because $E_1, E_2 \subseteq [0, 1)$ and hence $E_2 + 1 \subseteq [1, 2)$. The first is also easy. If $x \in E_1$ then $x \in A$. If $x \in E_2 + 1$ then $x - 1 \in E_2$ and then $x = (x - 1) + 1 \in A$. Thus $E_1 \cup (E_2 + 1) \subseteq A$. On the other hand, if $x \in A \subseteq [0, 2)$, then, either $x \in A \cap [0, 1)$ implying $x \in E_1$, or $x \in A \cap [1, 2)$

implying $x - 1 \in E_2$ i.e. $x \in E_2 + 1$. Thus $A \subseteq E_1 \cup (E_2 + 1)$.
From (v) we shall need only that $E_1, E_2 + 1 \subseteq A$.

Suppose N is a Lebesgue set. By (ii) and by the invariance of m_1 under translations, we get that $m_1(A) = \sum_{r \in \mathbf{Q} \cap [0,1)} m_1(N + r) = \sum_{r \in \mathbf{Q} \cap [0,1)} m_1(N)$. If $m_1(N) > 0$, then $m_1(A) = +\infty$, contradicting (iii). If $m_1(N) = 0$, then $m_1(A) = 0$, implying by (v) that $m_1(E_1) = m_1(E_2 + 1) = 0$, hence $m_1(E_1) = m_1(E_2) = 0$, and finally from (iv), $1 = m_1([0,1)) \leq m_1(E_1) + m_1(E_2) = 0$.

We arrive at a contradiction and N is not a Lebesgue set.

4.6 Exercises.

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 $m_n(A) < +\infty$.

2. *The invariance of Lebesgue measure under isometries.*

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

   Prove that, for every $E \in \mathcal{L}_n$, we have $m_n(T(E)) = m_n(E)$.

3. A parallelepiped in $\mathbf{R}^n$ is called **degenerate** if it is included in a hyperplane of $\mathbf{R}^n$, i.e. in a set of the form $b + V$, where $b \in \mathbf{R}^n$ and V is a linear subspace of $\mathbf{R}^n$ with $\dim(V) = n - 1$.

   Prove that a parallelepiped M is degenerate if and only if $m_n(M) = 0$.

4. State in a formal way and prove the rule

   $$\text{volume} \ = \ \text{base area} \ \times \ \text{height}$$

   for parallelepipeds in $\mathbf{R}^n$.

5. *Regularity of Lebesgue measure.*

   Suppose that $A \in \mathcal{L}_n$.
   (i) Prove that there is a decreasing sequence (U_j) of open sets in $\mathbf{R}^n$ so that $A \subseteq U_j$ for all j and $m_n(U_j) \to m_n(A)$. Conclude that

   $$m_n(A) = \inf\{m_n(U) \,|\, U \text{ open } \supseteq A\}.$$

   (ii) Prove that there is an increasing sequence (K_j) of compact sets in $\mathbf{R}^n$ so that $K_j \subseteq A$ for all j and $m_n(K_j) \to m_n(A)$. Conclude that

   $$m_n(A) = \sup\{m_n(K) \,|\, K \text{ compact } \subseteq A\}.$$

   The validity of (i) and (ii) for $(\mathbf{R}^n, \mathcal{L}_n, m_n)$ is called **regularity**. We shall study this notion in chapter 5.

6. *An example of an m_1-null uncountable set, dense in an interval.*

Let $\mathbf{Q} \cap [0,1] = \{x_1, x_2, \ldots\}$. For every $\epsilon > 0$ we define

$$U(\epsilon) = \cup_{j=1}^{+\infty}\Big(x_j - \frac{\epsilon}{2^j}, x_j + \frac{\epsilon}{2^j}\Big), \qquad A = \cap_{n=1}^{+\infty} U\Big(\frac{1}{n}\Big).$$

(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 $\mathbf{Q} \cap [0,1] \subseteq A$ and that A is uncountable.

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

8. Prove that the Cantor set is perfect: it is closed and has no isolated point.

9. *The Cantor set and ternary expansions of numbers.*

(i) Prove that for every sequence (a_n) in $\{0,1,2\}$ the series $\sum_{n=1}^{+\infty} \frac{a_n}{3^n}$ converges to a number in $[0,1]$.
(ii) Conversely, prove that for every number x in $[0,1]$ there is a sequence (a_n) in $\{0,1,2\}$ so that $x = \sum_{n=1}^{+\infty} \frac{a_n}{3^n}$. 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.
(iii) If $x \in [0,1]$ is a rational $\frac{m}{3^N}$, where $m \equiv 1 (\mathrm{mod}\, 3)$ and $N \in \mathbf{N}$, then x has exactly two ternary expansions: one is of the form $0.a_1 \ldots a_{N-1}1000\ldots$ and the other is of the form $0.a_1 \ldots a_{N-1}0222\ldots$.

If $x \in [0,1]$ is either irrational or rational $\frac{m}{3^N}$, where $m \equiv 0$ or $2 (mod\, 3)$ and $N \in \mathbf{N}$, then it has exactly one ternary expansion which is not of either one of the above forms.
(iv) 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.

10. *The Cantor function.*

Let $I_0 = [0,1], I_1, I_2, \ldots$ be the sets used in the construction of the Cantor set C. For each $n \in \mathbf{N}$ define $f_n : [0,1] \to [0,1]$ as follows. If, going from left to right, $J_1^{(n)}, \ldots, J_{2^n-1}^{(n)}$ are the $2^n - 1$ subintervals of $[0,1] \setminus I_n$, then define $f_n(0) = 0, f_n(1) = 1$, define f_n to be constant $\frac{k}{2^n}$ in $J_k^{(n)}$ for all $k = 1, \ldots, 2^n - 1$ and to be linear in each of the subintervals of I_n in such a way that f_n is continuous in $[0,1]$.
(i) Prove that $|f_n(x) - f_{n-1}(x)| \leq \frac{1}{3 \cdot 2^n}$ for all $n \geq 2$ and all $x \in [0,1]$. This implies that for every $x \in [0,1]$ the series $f_1(x) + \sum_{k=2}^{+\infty}(f_k(x) - f_{k-1}(x))$ converges to a real number.
(ii) Define $f(x)$ to be the sum of the series appearing in (i) and prove that $|f(x) - f_n(x)| \leq \frac{1}{3 \cdot 2^n}$ for all $x \in [0,1]$. Therefore, f_n converges to f uniformly in $[0,1]$.
(iii) Prove that $f(0) = 0, f(1) = 1$ and that f is continuous and increasing

in $[0, 1]$.
(iv) Prove that for every n: f is constant $\frac{k}{2^n}$ in $J_k^{(n)}$ for all $k = 1, \dots, 2^n - 1$.
(v) Prove that, if $x, y \in C$ and $x < y$ and x, y are not end-points of the same complementary interval of C, then $f(x) < f(y)$.

This function f is called **the Cantor function**.

11. *The difference set of a set.*

(i) Let $E \subseteq \mathbf{R}$ with $m_1^*(E) > 0$ and $0 \le \alpha < 1$. Prove that there is a non-empty open interval (a, b) so that $m_1^*(E \cap (a, b)) \ge \alpha \cdot (b - a)$.
(ii) Let $E \subseteq \mathbf{R}$ be a Lebesgue set with $m_1(E) > 0$. Taking $\alpha = \frac{3}{4}$ in (i), prove that $E \cap (E + z) \cap (a, b) \ne \emptyset$ for all z with $|z| < \frac{1}{4}(b - a)$.
(iii) Let $E \subseteq \mathbf{R}$ be a Lebesgue set with $m_1(E) > 0$. Prove that the set $D(E) = \{x - y \,|\, x, y \in E\}$, called **the difference set of** E, includes some open interval of the form $(-\epsilon, \epsilon)$.

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

(i) For any $x, y \in \mathbf{R}$ define $x \sim y$ if $x - y \in \mathbf{Q}$. Prove that $\sim$ is an equivalence relation in $\mathbf{R}$.
(ii) Let L be a set containing exactly one element from each of the equivalence classes of $\sim$. Prove that $\mathbf{R} = \cup_{r \in \mathbf{Q}}(L + r)$ and that the sets $L + r,\ r \in \mathbf{Q}$, are pairwise disjoint.
(iii) Prove that the difference set of L (see exercise 4.6.11) contains no rational number $\ne 0$.
(iv) Using the result of exercise 4.6.11, prove that L is not a Lebesgue set.

13. *Non-Lebesgue sets are everywhere, I.*
We shall prove that every $E \subseteq \mathbf{R}$ with $m_1^*(E) > 0$ includes at least one non-Lebesgue set.
(i) Consider the non-Lebesgue set $N \subseteq [0, 1]$ which was constructed in section 4.5 and prove that, if $B \subseteq N$ is a Lebesgue set, then $m_1(B) = 0$. In other words, if $M \subseteq N$ has $m_1^*(M) > 0$, then M is a non-Lebesgue set.
(ii) Consider an arbitrary $E \subseteq \mathbf{R}$ with $m_1^*(E) > 0$. If $\alpha = 1 - m_1^*(N)$, then $0 \le \alpha < 1$, and consider an interval (a, b) so that $m_1^*(E \cap (a, b)) \ge \alpha(b - a)$ (see exercise 4.6.11). Then the set $N' = (b - a)N + a$ is included in $[a, b]$, has $m_1^*(N') = (1 - \alpha) \cdot (b - a)$ and, if $M' \subseteq N'$ has $m_1^*(M') > 0$, then M' is not a Lebesgue set.
(iii) Prove that $E \cap N'$ is not a Lebesgue set.

14. *No-Lebesgue sets are everywhere, II.*

(i) Consider the set L from exercise 4.6.12. Then $E = \cup_{r \in \mathbf{Q}}(E \cap (L + r))$ and prove that the difference set (exercise 4.6.11) of each $E \cap (L + r)$ contains no rational number $\ne 0$.
(ii) Prove that, for at least one $r \in \mathbf{Q}$, the set $E \cap (L + r)$ is not a Lebesgue set (using exercise 4.6.11).

15. *Not all Lebesgue sets in* $\mathbf{R}$ *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 constructed in exercise 4.6.10. Define $g : [0,1] \to [0,2]$ by the formula

    $$g(x) = f(x) + x, \qquad x \in [0,1].$$

    (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 $g([0,1] \setminus C)$, where C is the Cantor set, is an open set with Lebesgue measure equal to 1. Therefore the set $E = g(C)$ has Lebesgue measure equal to 1.
    (iii) Exercises 4.6.13 and 4.6.14 give non-Lebesgue sets $M \subseteq E$. Consider the set $K = g^{-1}(M) \subseteq C$. Prove that K is a Lebesgue set.
    (iv) Using exercise 1.6.8, prove that K is not a Borel set in $\mathbf{R}$.
    (v) g maps K onto M.

16. *More Cantor sets.*

    Take an arbitrary sequence (ϵ_n) so that $0 < \epsilon_n < \frac{1}{2}$ for all n. We split $I_0 = [0,1]$ into the three intervals $[0, \frac{1}{2} - \epsilon_1], (\frac{1}{2} - \epsilon_1, \frac{1}{2} + \epsilon_1), [\frac{1}{2} + \epsilon_1, 1]$ and form I_1 as the union of the two closed intervals. Inductively, if we have already constructed I_{n-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 $\frac{1}{2} - \epsilon_n$. The union of the new intervals is the I_n.

    We set $K = \cap_{n=1}^{+\infty} I_n$.
    (i) Prove that K is compact, has no isolated points and includes no open interval.
    (ii) Prove that K is uncountable.
    (iii) Prove that $m_1(I_n) = (1 - 2\epsilon_1) \cdots (1 - 2\epsilon_n)$ for all n.
    (iv) Prove that $m_1(K) = \lim_{n \to +\infty} (1 - 2\epsilon_1) \cdots (1 - 2\epsilon_n)$.
    (v) Taking $\epsilon_n = \frac{\epsilon}{3^n}$ for all n, prove that $m_1(K) > 1 - \epsilon$.
    (Use that $(1 - a_1) \cdots (1 - a_n) > 1 - (a_1 + \cdots + a_n)$ for all n and all $a_1, \ldots, a_n \in [0,1]$).
    (vi) Prove that $m_1(K) > 0$ if and only if $\sum_{n=1}^{+\infty} \epsilon_n < +\infty$.
    (Use the inequality you used for (v) and also that $1 - a \leq e^{-a}$ for all a.)

17. *Uniqueness of Lebesgue measure.*

    Prove that m_n is the only measure μ on $(\mathbf{R}^n, \mathcal{B}_{\mathbf{R}^n})$ which is invariant under translations (i.e. $\mu(E + x) = \mu(E)$ for all Borel sets E and all x) and which satisfies $\mu(Q_0) = 1$, where $Q_0 = [-1,1] \times \cdots \times [-1,1]$.

18. Let $E \subseteq \mathbf{R}$ be a Lebesgue set and A be a dense subset of $\mathbf{R}$. If $m_1(E \triangle (E + x)) = 0$ for all $x \in A$, prove that $m_1(E) = 0$ or $m_1(E^c) = 0$.

19. Let $E \subseteq \mathbf{R}$ be a Lebesgue set and $\delta > 0$. If $m_1(E \cap (a, b)) \geq \delta(b - a)$ for all intervals (a, b), prove that $m_1(E^c) = 0$.

Chapter 5

Borel measures

5.1 Lebesgue-Stieltjes measures in R.

Lemma 5.1 *If* $-\infty \le a < b \le +\infty$ *and* $F : (a, b) \to \mathbf{R}$ *is increasing, then*
(i) for all $x \in [a, b)$ *we have* $F(x+) = \inf\{F(y) \,|\, x < y\}$*,*
(ii) for all $x \in (a, b]$ *we have* $F(x-) = \sup\{F(y) \,|\, y < x\}$*,*
(iii) if $a < x < y < z < b$*, then* $F(x-) \le F(x) \le F(x+) \le F(y) \le F(z-) \le F(z) \le F(z+)$*,*
(iv) for all $x \in [a, b)$ *we have* $F(x+) = \lim_{y \to x+} F(y\pm)$*,*
(v) for all $x \in (a, b]$ *we have* $F(x-) = \lim_{y \to x-} F(y\pm)$*.*

Proof: (i) Let $M = \inf\{F(y) \,|\, x < y\}$. Then for every $\gamma > M$ there is some $t > x$ so that $F(t) < \gamma$. Hence for all $y \in (x, t)$ we have $M \le F(y) < \gamma$. This says that $F(x+) = M$.
(ii) Similarly, let $m = \sup\{F(y) \,|\, y < x\}$. Then for every $\gamma < m$ there is some $t < x$ so that $\gamma < F(t)$. Hence for all $y \in (t, x)$ we have $\gamma < F(y) \le m$. This says that $F(x-) = m$.
(iii) $F(x)$ is an upper bound of the set $\{F(y) \,|\, y < x\}$ and a lower bound of $\{F(y) \,|\, x < y\}$. This, by (i) and (ii), implies that $F(x-) \le F(x) \le F(x+)$ and, of course, $F(z-) \le F(z) \le F(z+)$. Also, if $x < y < z$, then $F(y)$ is an element of both sets $\{F(y) \,|\, x < y\}$ and $\{F(y) \,|\, y < z\}$. Therefore $F(y)$ is between the infimum of the first, $F(x+)$, and the supremum of the second set, $F(z-)$.
(iv) By the result of (i), for every $\gamma > F(x+)$ there is some $t > x$ so that $F(x+) \le F(t) < \gamma$. This, combined with (iii), implies that $F(x+) \le F(y\pm) < \gamma$ for all $y \in (x, t)$. Thus, $F(x+) = \lim_{y \to x+} F(y\pm)$.
(v) By (ii), for every $\gamma < F(x-)$ there is some $t < x$ so that $\gamma < F(t) \le F(x-)$. This, combined with (iii), implies $\gamma < F(y\pm) \le F(x-)$ for all $y \in (t, x)$. Thus, $F(x-) = \lim_{y \to x-} F(y\pm)$.

Consider now a_0, b_0 with $-\infty \le a_0 < b_0 \le +\infty$ and an increasing function $F : (a_0, b_0) \to \mathbf{R}$ and define a *non-negative* function τ acting on 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 approach it from the outside while, if the end-point is not included in the interval, then approach it from the inside of the interval.

We use the collection of all open subintervals of (a_0, b_0) and the function τ to define, as an application of Theorem 3.2, the following outer measure on (a_0, b_0):

$$\mu_F^*(E) = \inf\left\{ \sum_{j=1}^{+\infty} \tau((a_j, b_j)) \,|\, E \subseteq \cup_{j=1}^{+\infty}(a_j, b_j), (a_j, b_j) \subseteq (a_0, b_0) \text{ for all } j \right\},$$

for every $E \subseteq (a_0, b_0)$.

Theorem 3.1 implies that the collection of μ_F^*-measurable sets is a σ-algebra of subsets of (a_0, b_0), which we denote by Σ_F, and the restriction, denoted μ_F, of μ_F^* on Σ_F is a complete measure.

Definition 5.1 *The measure μ_F is called* ***the Lebesgue-Stieltjes measure induced by the (increasing)*** $F : (a_0, b_0) \to \mathbf{R}$.

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

Following exactly the same procedure as with Lebesgue measure, we shall study the relation between the σ-algebra Σ_F and the Borel sets in (a_0, b_0).

Lemma 5.2 *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 $\tau(P) = \tau(P_1) + \cdots + \tau(P_m)$.*

Proof: A telescoping sum: $\tau(P_1)+\cdots+\tau(P_m) = \sum_{i=1}^{m}(F(c^{(i)}+)-F(c^{(i-1)}+)) = F(b+) - F(a+) = \tau((a, b])$.

Lemma 5.3 *If $P, P_1, \ldots, P_l$ are open-closed subintervals of (a_0, b_0), $P_1, \ldots, P_l$ are pairwise disjoint and $P = P_1 \cup \cdots \cup P_l$, then $\tau(P) = \tau(P_1) + \cdots + \tau(P_l)$.*

Proof: Exactly one of $P_1, \ldots, P_l$ has the same right end-point as P. We rename and call it P_l. Then exactly one of $P_1, \ldots, P_{l-1}$ 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 the same as the result of Lemma 5.2.

Lemma 5.4 *If $P, P_1, \ldots, P_l$ are open-closed subintervals of (a_0, b_0), $P_1, \ldots, P_l$ are pairwise disjoint and $P_1 \cup \cdots \cup P_l \subseteq P$, then $\tau(P_1) + \cdots + \tau(P_l) \leq \tau(P)$.*

Proof: 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 $P'_1, \ldots, P'_k$. By Lemma 5.3 we get $\tau(P) = \tau(P_1) + \cdots + \tau(P_l) + \tau(P'_1) + \cdots + \tau(P'_k) \geq \tau(P_1) + \cdots + \tau(P_l)$.

Lemma 5.5 *Suppose that $P, P_1, \ldots, P_l$ are open-closed subintervals of (a_0, b_0) and $P \subseteq P_1 \cup \cdots \cup P_l$. Then $\tau(P) \leq \tau(P_1) + \cdots + \tau(P_l)$.*

Proof: 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 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: $P'_1 = P'_1$ and

$$P'_j \setminus (P'_1 \cup \cdots \cup P'_{j-1}) = P_1^{(j)} \cup \cdots \cup P_{m_j}^{(j)}$$

for $2 \leq j \leq l$.

Lemma 5.3 (for the equality) and Lemma 5.4 (for the two inequalities) imply

$$\begin{aligned} \tau(P) &= \tau(P'_1) + \sum_{j=2}^{l} \Big(\sum_{m=1}^{m_j} \tau(P_m^{(j)}) \Big) \\ &\leq \tau(P'_1) + \sum_{j=2}^{l} \tau(P'_j) \leq \sum_{j=1}^{l} \tau(P_j). \end{aligned}$$

Lemma 5.6 *Let Q be a closed interval and $R_1, \ldots, R_l$ be open subintervals of (a_0, b_0). If $Q \subseteq R_1 \cup \cdots \cup R_l$, then $\tau(Q) \leq \tau(R_1) + \cdots + \tau(R_l)$.*

Proof: Let $Q = [a, b]$ and $R_j = (a_j, b_j)$ for $j = 1, \ldots, l$. We define for $\epsilon > 0$

$$P_\epsilon = (a - \epsilon, b], \qquad P_{j,\epsilon} = (a_j, b_j - \epsilon].$$

We shall first prove that there is some $\epsilon_0 > 0$ so that for all $\epsilon < \epsilon_0$

$$P_\epsilon \subseteq P_{1,\epsilon} \cup \cdots \cup P_{l,\epsilon}.$$

Suppose that, for all n, the above inclusion is not true for $\epsilon = \frac{1}{n}$. Hence, for all n there is $x_n \in (a - \frac{1}{n}, b]$ so that $x_n \notin \cup_{j=1}^{l} (a_j, b_j - \frac{1}{n}]$. By the Bolzano-Weierstrass theorem, there is a subsequence (x_{n_k}) converging to some $\overline{x}$. Looking carefully at the various inequalities, we get $\overline{x} \in [a, b]$ and $\overline{x} \notin \cup_{j=1}^{l} (a_j, b_j)$. This is a contradiction and the inclusion we want to prove is true for some $\epsilon_0 = \frac{1}{n_0}$. If $\epsilon < \epsilon_0$, then the inclusion is still true because the left side becomes smaller while the right side becomes larger.

Now Lemma 5.5 gives for $\epsilon < \epsilon_0$ that

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

and, using Lemma 5.1 for the limit as $\epsilon \to 0+$,

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

Theorem 5.1 *Let $F : (a_0, b_0) \to \mathbf{R}$ be increasing. Then every 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 $\epsilon > 0$ and, thus, $\mu_F^*(Q) \le F(b+) - F(a-) = \tau(Q)$.

For every covering $Q \subseteq \cup_{j=1}^{+\infty} R_j$ by open subintervals of (a_0, b_0), there is (by compactness) l so that $Q \subseteq \cup_{j=1}^{l} R_j$. Lemma 5.6 implies $\tau(Q) \le \sum_{j=1}^{l} \tau(R_j) \le \sum_{j=1}^{+\infty} \tau(R_j)$. Hence $\tau(Q) \le \mu_F^*(Q)$ and we conclude that

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

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 $\epsilon > 0$. Hence $\mu_F^*(P) \le F(b+) - F(a+) = \tau(P)$.

If $R = (a, b) \subseteq (a_0, b_0)$, then $\mu_F^*(R) \le \tau((a, b)) = \tau(R)$.

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

We write $\mu_F^*(R \cap P) = \mu_F^*\big((P_R \cap P) \cup ((d - \epsilon, d) \cap P)\big) \le \mu_F^*(P_R \cap P) + \mu_F^*((d - \epsilon, d)) \le \tau(P_R \cap P) + F(d-) - F((d - \epsilon)+)$ by the previous results. The same inequalities, with P^c instead of P, give $\mu_F^*(R \cap P^c) \le \mu_F^*(P_R \cap P^c) + F(d-) - F((d - \epsilon)+)$. Taking the sum, 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 write $P_R \cap P^c = P_1 \cup \cdots \cup P_l$ for pairwise disjoint open-closed intervals and get $\tau(P_R \cap P) + \mu_F^*(P_R \cap P^c) \le \tau(P_R \cap P) + \sum_{j=1}^{l} \mu_F^*(P_j) \le \tau(P_R \cap P) + \sum_{j=1}^{l} \tau(P_j) = \tau(P_R)$ by the first results and Lemma 5.3.

Therefore $\mu_F^*(R \cap P) + \mu_F^*(R \cap P^c) \le \tau(P_R) + 2[F(d-) - F((d - \epsilon)+)] = F((d - \epsilon)+) - F(c+) + 2[F(d-) - F((d - \epsilon)+)]$ and, taking limit, $\mu_F^*(R \cap P) + \mu_F^*(R \cap P^c) \le F(d-) - F(c+) = \tau(R)$.

We proved that

$$\mu_F^*(R \cap P) + \mu_F^*(R \cap P^c) \le \tau(R)$$

for all open intervals R and open-closed intervals P which are $\subseteq (a_0, b_0)$.

Now consider arbitrary $E \subseteq (a_0, b_0)$ with $\mu_F^*(E) < +\infty$. Take a covering $E \subseteq \cup_{j=1}^{+\infty} R_j$ by open subintervals of (a_0, b_0) so that $\sum_{j=1}^{+\infty} \tau(R_j) < \mu_F^*(E) + \epsilon$. By σ-subadditivity and the last result we find $\mu_F^*(E \cap P) + \mu_F^*(E \cap P^c) \le \sum_{j=1}^{+\infty} \big(\mu_F^*(R_j \cap P) + \mu_F^*(R_j \cap P^c)\big) \le \sum_{j=1}^{+\infty} \tau(R_j) < \mu_F^*(E) + \epsilon$.

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

$$\mu_F^*(E \cap P) + \mu_F^*(E \cap P^c) \le \mu_F^*(E),$$

concluding that $P \in \Sigma_F$.

If $Q = [a, b] \subseteq (a_0, b_0)$, we take any (a_k) in (a_0, b_0) so that $a_k \uparrow a$ and, then, $Q = \cap_{k=1}^{+\infty} (a_k, b] \in \Sigma_F$. Moreover, by the first results,

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

If $P = (a, b] \subseteq (a_0, b_0)$, we take any (a_k) in $(a, b]$ so that $a_k \downarrow a$ 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 (b_k) in $[a, b)$ so that $b_k \uparrow b$ and we get that $T = \cup_{k=1}^{+\infty}[a, b_k] \in \Sigma_F$. 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 (a_k) and (b_k) in (a, b) so that $a_k \downarrow a$, $b_k \uparrow b$ and $a_1 \leq b_1$. Then $R = \cup_{k=1}^{+\infty}[a_k, b_k] \in \Sigma_F$. Moreover, $\mu_F(R) = \lim_{k\to+\infty} \mu_F([a_k, b_k]) = \lim_{k\to+\infty}(F(b_k+)-F(a_k-)) = F(b-)-F(a+) = \tau(R)$.

Theorem 5.2 *Let $F : (a_0, b_0) \to \mathbf{R}$ be increasing. Then μ_F is σ-finite and it is finite if and only if F is bounded. Also, $\mu_F\big((a_0, b_0)\big) = F(b_0-) - F(a_0+)$.*

Proof: We consider any two sequences (a_k) and (b_k) in (a_0, b_0) so that $a_k \downarrow a_0$, $b_k \uparrow b_0$ and $a_1 \leq b_1$. Then $(a_0, b_0) = \cup_{k=1}^{+\infty}[a_k, b_k]$ and $\mu_F([a_k, b_k]) = F(b_k+) - F(a_k-) < +\infty$ for all k. Hence, μ_F is σ-finite.

Since $\mu_F((a_0, b_0)) = F(b_0-) - F(a_0+)$, if μ_F is finite, then $-\infty < F(a_0+)$ and $F(b_0-) < +\infty$. This implies that all values of F lie in the bounded interval $[F(a_0+), F(b_0-)]$ and F is bounded. Conversely, if F is bounded, then the limits $F(a_0+), F(b_0-)$ are finite and $\mu_F((a_0, b_0)) < +\infty$.

It is easy to prove that *the collection of all subintervals of* (a_0, b_0) *generates the σ-algebra of all Borel sets in* (a_0, b_0). Indeed, let $\mathcal{E}$ be the collection of all intervals in $\mathbf{R}$ and $\mathcal{F}$ be the collection of all subintervals of (a_0, b_0). It is clear that $\mathcal{F} = \mathcal{E}\rceil(a_0, b_0)$ and Theorems 1.2 and 1.3 imply that

$$\mathcal{B}_{(a_0,b_0)} = \mathcal{B}_{\mathbf{R}}\rceil(a_0, b_0) = \Sigma(\mathcal{E})\rceil(a_0, b_0) = \Sigma(\mathcal{F}).$$

Theorem 5.3 *Let $F : (a_0, b_0) \to \mathbf{R}$ be increasing. Then all Borel sets in (a_0, b_0) belong to Σ_F.*

Proof: Theorem 5.1 implies that the collection $\mathcal{F}$ of all subintervals of (a_0, b_0) is included in Σ_F. By the discussion of the previous paragraph, we conclude that $\mathcal{B}_{(a_0,b_0)} = \Sigma(\mathcal{F}) \subseteq \Sigma_F$.

Theorem 5.4 *Let $F : (a_0, b_0) \to \mathbf{R}$ be increasing. Then for every $E \subseteq (a_0, b_0)$ we have*
(i) $E \in \Sigma_F$ if and only if there is $A \subseteq (a_0, b_0)$, a countable intersection of open sets, so that $E \subseteq A$ and $\mu_F^*(A \setminus E) = 0$.
(ii) $E \in \Sigma_F$ if and only if there B, a countable union of compact sets, so that $B \subseteq E$ and $\mu_F^*(E \setminus B) = 0$.

Proof: The proof is exactly the same as the proof of the similar Theorem 4.4. Only the obvious changes have to be made: m_n changes to μ_F and m_n^* to μ_F^*, $\mathbf{R}^n$ changes to (a_0, b_0), vol_n changes to τ and $\mathcal{L}_n$ changes to Σ_F.

Therefore, *every set in Σ_F is, except from a μ_F-null set, equal to a Borel set.*

Theorem 5.5 *Let $F : (a_0, b_0) \to \mathbf{R}$ be increasing. Then*
(i) μ_F is the only measure on $\big((a_0, b_0), \mathcal{B}_{(a_0,b_0)}\big)$ with $\mu_F((a, b]) = F(b+)-F(a+)$ for all intervals $(a, b] \subseteq (a_0, b_0)$.
(ii) $\big((a_0, b_0), \Sigma_F, \mu_F\big)$ 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 Theorem 4.5. Only the obvious notational modifications are needed.

It should be observed that the measure of a set $\{x\}$ consisting of a single point $x \in (a_0, b_0)$ is equal to $\mu_F(\{x\}) = F(x+) - F(x-)$, 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 in 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 5.1 shows that this is not a serious restriction.

Proposition 5.1 *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 $F : (a_0, b_0) \to \mathbf{R}$ we define $F_0 : (a_0, b_0) \to \mathbf{R}$ by the formula

$$F_0(x) = F(x+) , \qquad x \in (a_0, b_0)$$

and it is immediate from Lemma 5.1 that F_0 is increasing, continuous from the right, i.e. $F_0(x+) = F_0(x)$ for all x, and $F_0(x+) = F(x+), F_0(x-) = F(x-)$ for all x. Now, it is obvious that 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 construction of the measures μ_{F_0}, μ_F starts) assign the same values to every interval $S \subseteq (a_0, b_0)$.

The functions F_0 and F of Proposition 5.1 have the same jump at every x and, in particular, they have the same continuity points.

5.2 Borel measures.

Definition 5.2 *Let X be a topological space and (X, Σ, μ) be a measure space. The measure μ is called **a Borel measure on** X if $\mathcal{B}_X \subseteq \Sigma$, i.e. if all Borel sets in X are in Σ.*

*The Borel measure μ is called **locally finite** if for every $x \in X$ there is some open neighborhood U_x of x (i.e. an open set containing x) such that $\mu(U_x) < +\infty$.*

Observe that, for μ to be a Borel measure, it is enough to have that all open sets or all closed sets are in Σ. This is because $\mathcal{B}_X$ is generated by the collections of all open or all closed sets and because Σ is a σ-algebra.

Examples
The Lebesgue measure on $\mathbf{R}^n$ and, more generally, the Lebesgue-Stieltjes measure on any generalized interval (a_0, b_0) (induced by any increasing function) are locally finite Borel measures. In fact, the content of the following theorem is that the only locally finite Borel measures on (a_0, b_0) are exactly the Lebesgue-Stieltjes measures.

Lemma 5.7 *Let X be a topological space and μ a Borel measure on X. If μ is locally finite, then $\mu(K) < +\infty$ for every compact $K \subseteq X$.*

If μ is a locally finite Borel measure on $\mathbf{R}^n$, then $\mu(M) < +\infty$ for every bounded $M \subseteq \mathbf{R}^n$.

Proof: We take for each $x \in K$ an open neighborhood U_x of x so that $\mu(U_x) < +\infty$. Since $K \subseteq \cup_{x \in K} U_x$ and K is compact, there are $x_1, \dots, x_n$ so that $K \subseteq \cup_{k=1}^n U_{x_k}$. Hence, $\mu(K) \leq \sum_{k=1}^n \mu(U_{x_k}) < +\infty$.

If $M \subseteq \mathbf{R}^n$ is bounded, then $\overline{M}$ is compact and $\mu(M) \leq \mu(\overline{M}) < +\infty$.

Theorem 5.6 *Let $-\infty \leq a_0 < b_0 \leq +\infty$ and $c_0 \in (a_0, b_0)$. For every locally finite Borel measure μ on (a_0, b_0) there is a unique increasing and continuous from the right $F : (a_0, b_0) \to \mathbf{R}$ so that $\mu = \mu_F$ on $\mathcal{B}_{(a_0,b_0)}$ and $F(c_0) = 0$. For any other increasing and continuous from the right $G : (a_0, b_0) \to \mathbf{R}$, it is true that $\mu = \mu_G$ if and only if G differs from F by a constant.*

Proof: Define the function

$$F(x) = \begin{cases} \mu((c_0, x]), & \text{if } c_0 \leq x < b_0, \\ -\mu((x, c_0]), & \text{if } a_0 < x < c_0. \end{cases}$$

By Lemma 5.7, F is real valued and it is clear, by the monotonicity of μ, that F is increasing. Now take any decreasing sequence (x_n) so that $x_n \downarrow x$. If $c_0 \leq x$, by continuity of μ from above, $\lim_{n\to+\infty} F(x_n) = \lim_{n\to+\infty} \mu((c_0, x_n]) = \mu((c_0, x]) = F(x)$. Also, if $x < c_0$, then $x_n < c_0$ for large n, and, by continuity of μ from below, $\lim_{n\to+\infty} F(x_n) = -\lim_{n\to+\infty} \mu((x_n, c_0]) = -\mu((x, c_0]) = F(x)$. Therefore, F is continuous from the right at every x.

If we compare μ and the induced μ_F at the intervals $(a, b]$, we get $\mu_F((a, b]) = F(b) - F(a) = \mu((a, b])$, where the second equality becomes trivial by considering cases: $a < b < c_0$, $a < c_0 \leq b$ and $c_0 \leq a < b$. Theorem 5.5 implies that $\mu_F = \mu$ on $\mathcal{B}_{(a_0,b_0)}$.

If G is increasing, continuous from the right with $\mu_G = \mu (= \mu_F)$ on $\mathcal{B}_{(a_0,b_0)}$, then $G(x) - G(c_0) = \mu_G((c_0, x]) = \mu_F((c_0, x]) = F(x) - F(c_0)$ for all $x \geq c_0$ and, similarly, $G(c_0) - G(x) = \mu_G((x, c_0]) = \mu_F((x, c_0]) = F(c_0) - F(x)$ for all $x < c_0$. Therefore F, G differ by a constant: $G - F = G(c_0) - F(c_0)$ on (a_0, b_0). Hence, if $F(c_0) = 0 = G(c_0)$, then F, G are equal on (a_0, b_0).

If the locally finite Borel measure μ on (a_0, b_0) satisfies the $\mu((a_0, c_0]) < +\infty$, then we may make a different choice for F than the one in Theorem 5.6. We add the constant $\mu((a_0, c_0])$ to the function of the theorem and get the function

$$F(x) = \mu((a_0, x]), \qquad x \in (a_0, b_0).$$

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 replace the general Borel set (a somewhat obscure object) by open or closed sets.

Let E be a Borel subset in a topological space X and μ a Borel measure on X. It is clear that $\mu(K) \leq \mu(E) \leq \mu(U)$ for all K compact and U open with $K \subseteq E \subseteq U$. Hence

$$\sup\{\mu(K) \,|\, K \text{ compact } \subseteq E\} \leq \mu(E) \leq \inf\{\mu(U) \,|\, U \text{ open } \supseteq E\}.$$

Definition 5.3 *Let X be a topological space and μ a Borel measure on X. Then μ is called* ***regular*** *if the following are true for every Borel set E in X:*
(i) $\mu(E) = \inf\{\mu(U) \,|\, U \text{ open } \supseteq E\}$,
(ii) $\mu(E) = \sup\{\mu(K) \,|\, K \text{ compact } \subseteq E\}$.

Therefore, 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.

Proposition 5.2 *Let O be any open set in $\mathbf{R}^n$. There is an increasing sequence (K_m) of compact subsets of O so that int$(K_m) \uparrow O$ and, hence, $K_m \uparrow O$ also.*

Proof: Define the sets

$$K_m = \Big\{x \in O \,|\, |x| \leq m \text{ and } |y - x| \geq \frac{1}{m} \text{ for all } y \notin O\Big\},$$

where $|x|^2 = x_1^2 + \cdots + x_n^2$ for all $x = (x_1, \ldots, x_n)$.

The set K_m is bounded, since $|x| \leq m$ for all $x \in K_m$.

If (x_j) is a sequence in K_m converging to some x, then, from $|x_j| \leq m$ for all j, we get $|x| \leq m$, and, from $|y - x_j| \geq \frac{1}{m}$ for all j and for all $y \notin O$, we get $|y - x| \geq \frac{1}{m}$ for all $y \notin O$. Thus, $x \in K_m$ and K_m is closed.

Therefore, K_m is a compact subset of O and, clearly, $K_m \subseteq K_{m+1} \subseteq O$ for all m. Hence, $\text{int}(K_m) \subseteq \text{int}(K_{m+1})$ for every m.

Now take any $x \in O$ and an $\epsilon > 0$ such that $B(x; 2\epsilon) \subseteq O$. Consider, also, $M \geq \max(|x| + \epsilon, \frac{1}{\epsilon})$. It is trivial to see that $B(x; \epsilon) \subseteq K_M$ and thus $x \in \text{int}(K_M)$. Therefore, $\text{int}(K_m) \uparrow O$. Since $\text{int}(K_m) \subseteq K_m \subseteq O$, we conclude that $K_m \uparrow O$.

Theorem 5.7 *Let X be a topological space with the property that for every open set O in X there is an increasing sequence of compact subsets of O whose interiors cover O.*

Suppose that μ is a locally finite Borel measure on X. Then:
(i) For every Borel set E and every $\epsilon > 0$ there is an open U and a closed F so that $F \subseteq E \subseteq U$ and $\mu(U \setminus E), \mu(E \setminus F) < \epsilon$. If also $\mu(E) < +\infty$, then F can be taken compact.
(ii) For every Borel set E in X there is A, a countable intersection of open sets, and B, a countable union of compact sets, so that $B \subseteq E \subseteq A$ and $\mu(A \setminus E) = \mu(E \setminus B) = 0$.
(iii) μ is regular.

Proof: (a) Suppose that $\mu(X) < +\infty$.

Consider the collection $\mathcal{S}$ of all Borel sets E in X with the property expressed in (i), namely, that for every $\epsilon > 0$ there is an open U and a closed F so that $F \subseteq E \subseteq U$ and $\mu(U \setminus E), \mu(E \setminus F) < \epsilon$.

Take any open set $O \subseteq X$ and arbitrary $\epsilon > 0$. If we consider $U = O$, then $\mu(U \setminus O) = 0 < \epsilon$. By assumption there is a sequence (K_m) of compact sets so that $K_m \uparrow O$. Therefore, $O \setminus K_m \downarrow \emptyset$ and, since $\mu(O \setminus K_1) \leq \mu(X) < +\infty$, continuity from above implies that $\lim_{m\to+\infty} \mu(O \setminus K_m) = 0$. Therefore there is some m so that $\mu(O \setminus F) < \epsilon$, if $F = K_m$.

Thus, all open sets belong to $\mathcal{S}$.

If $E \in \mathcal{S}$ and $\epsilon > 0$ is arbitrary, we find an open U and a closed F so that $F \subseteq E \subseteq U$ and $\mu(U \setminus E), \mu(E \setminus F) < \epsilon$. Then F^c is open, U^c is closed, $U^c \subseteq E^c \subseteq F^c$ and $\mu(F^c \setminus E^c) = \mu(E \setminus F) < \epsilon$ and $\mu(E^c \setminus U^c) = \mu(U \setminus E) < \epsilon$. This implies that $E^c \in \mathcal{S}$.

Now, take $E_1, E_2, \ldots \in \mathcal{S}$ and $E = \cup_{j=1}^{+\infty} E_j$. For $\epsilon > 0$ and each E_j take open U_j and closed F_j so that $F_j \subseteq E_j \subseteq U_j$ and $\mu(U_j \setminus E_j), \mu(E_j \setminus F_j) < \frac{\epsilon}{2^j}$. Define $B = \cup_{j=1}^{+\infty} F_j$ and the open $U = \cup_{j=1}^{+\infty} U_j$ so that $B \subseteq E \subseteq U$. Then $U \setminus E \subseteq \cup_{j=1}^{+\infty}(U_j \setminus E_j)$ and $E \setminus B \subseteq \cup_{j=1}^{+\infty}(E_j \setminus F_j)$. This implies $\mu(U \setminus E) \leq \sum_{j=1}^{+\infty} \mu(U_j \setminus E_j) < \sum_{j=1}^{+\infty} \frac{\epsilon}{2^j} = \epsilon$ and, similarly, $\mu(E \setminus B) < \epsilon$. The problem now is that B is not necessarily closed. Consider the closed sets $F_j' = F_1 \cup \cdots \cup F_j$, so that $F_j' \uparrow B$. Then $E \setminus F_j' \downarrow E \setminus B$ and, since $\mu(E \setminus F_1') \leq \mu(X) < +\infty$, continuity from below implies $\mu(E \setminus F_j') \downarrow \mu(E \setminus B)$. Therefore there is some j so that $\mu(E \setminus F_j') < \epsilon$. The inclusion $F_j' \subseteq E$ is clearly true.

We conclude that $E = \cup_{j=1}^{+\infty} E_j \in \mathcal{S}$ and $\mathcal{S}$ is a σ-algebra.

Since $\mathcal{S}$ contains all open sets, we have that $\mathcal{B}_X \subseteq \mathcal{S}$ and finish the proof of the first statement of (i) in the special case $\mu(X) < +\infty$.

(b) Now, consider the general case, and take any Borel set E in X which is included in some compact set $K \subseteq X$. For each $x \in K$ we take an open neighborhood U_x of x with $\mu(U_x) < +\infty$. By the compactness of K, there exist $x_1, \ldots, x_n \in K$ so that $K \subseteq \cup_{k=1}^{n} U_{x_k}$. We form the open set $G = \cup_{k=1}^{n} U_{x_k}$ and have that

$$E \subseteq G, \qquad \mu(G) < +\infty.$$

We next consider the restriction μ_G of μ on G, which is defined by the formula

$$\mu_G(A) = \mu(A \cap G)$$

for all Borel sets A in X. It is clear that μ_G is a Borel measure on X which is finite, since $\mu_G(X) = \mu(G) < +\infty$.

By (a), for every $\epsilon > 0$ there is an open U and a closed F so that $F \subseteq E \subseteq U$ and $\mu_G(U \setminus E), \mu_G(E \setminus F) < \epsilon$. Since $E \subseteq G$, we get $\mu((G \cap U) \setminus E) = \mu(G \cap (U \setminus E)) = \mu_G(U \setminus E) < \epsilon$ and $\mu(E \setminus F) = \mu(G \cap (E \setminus F)) = \mu_G(E \setminus F) < \epsilon$.

Therefore, if we consider the open set $U' = G \cap U$, we get $F \subseteq E \subseteq U'$ and $\mu(U' \setminus E), \mu(E \setminus F) < \epsilon$ and the first statement of (i) is now proved with no restriction on $\mu(X)$ but only for Borel sets in X which are included in compact subsets of X.

(c) We take an increasing sequence (K_m) of compact sets so that $\mathrm{int}(K_m) \uparrow X$. For any Borel set E in X we consider the sets $E_1 = E \cap K_1$ and $E_m = E \cap (K_m \setminus K_{m-1})$ for all $m \geq 2$ and we have that $E = \cup_{m=1}^{+\infty} E_m$. Since $E_m \subseteq K_m$, (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 $\mu(U_m \setminus E_m), \mu(E_m \setminus F_m) < \frac{\epsilon}{2^m}$. Now define the open $U = \cup_{m=1}^{+\infty} U_m$ and the closed (why?) $F = \cup_{m=1}^{+\infty} F_m$, so that $F \subseteq E \subseteq U$. As in the proof of (a), we easily get $\mu(U \setminus E), \mu(E \setminus F) < \epsilon$.

This concludes the proof of the first statement of (i).

(d) Let $\mu(E) < +\infty$. Take a closed F so that $F \subseteq E$ and $\mu(E \setminus F) < \epsilon$, and consider the compact sets K_m of part (c). Then the sets $F_m = F \cap K_m$ are compact and $F_m \uparrow F$. Therefore, $E \setminus F_m \downarrow E \setminus F$ and, by continuity of μ from above, $\mu(E \setminus F_m) \to \mu(E \setminus F)$. Thus there is a large enough m so that $\mu(E \setminus F_m) < \epsilon$. This proves the second statement of (i).

(e) Take open U_j and closed F_j so that $F_j \subseteq E \subseteq U_j$ and $\mu(U_j \setminus E), \mu(E \setminus F_j) < \frac{1}{j}$. Define $A = \cap_{j=1}^{+\infty} U_j$ and $B = \cup_{j=1}^{+\infty} F_j$ so that $B \subseteq E \subseteq A$. Now, for all j we have $\mu(A \setminus E) \leq \mu(U_j \setminus E) < \frac{1}{j}$ and $\mu(E \setminus B) \leq \mu(E \setminus F_j) < \frac{1}{j}$. Therefore, $\mu(A \setminus E) = \mu(E \setminus B) = 0$. We define the compact sets $K_{j,m} = F_j \cap K_m$ and observe that $B = \cup_{(j,m) \in \mathbf{N} \times \mathbf{N}} K_{j,m}$. This is the proof of (ii).

(f) If $\mu(E) = +\infty$, it is clear that $\mu(E) = \inf\{\mu(U) \,|\, U \text{ open and } E \subseteq U\}$. Also, from (ii), there is some $B = \cup_{m=1}^{+\infty} K'_m$, where all K'_m are compact, so that $B \subseteq E$ and $\mu(B) = \mu(E) = +\infty$. Consider the compact sets $K_m = K'_1 \cup \cdots \cup K'_m$ which satisfy $K_m \uparrow B$. Then $\mu(K_m) \to \mu(B) = \mu(E)$ and thus $\sup\{\mu(K) \,|\, K \text{ compact and } K \subseteq E\} = \mu(E)$.

If $\mu(E) < +\infty$, then, from (a), for every $\epsilon > 0$ there is a compact K and an open U so that $K \subseteq E \subseteq U$ and $\mu(U \setminus E), \mu(E \setminus K) < \epsilon$. This implies $\mu(E) - \epsilon < \mu(K)$ and $\mu(U) < \mu(E) + \epsilon$ and, thus, the proof of (iii) is complete.

Lemma 5.8 *Let X be a topological space which satisfies the assumptions of Theorem 5.7. Let Y be an open or a closed subset of X with its subspace topology. Then Y also satisfies the assumptions of Theorem 5.7.*

Proof Let Y be open in X. If O is an open subset of Y, then it is also an open subset of X. Therefore, there is an increasing sequence (K_m) of compact subsets of O so that $\mathrm{int}_X(K_m) \uparrow O$, where $\mathrm{int}_X(K_m)$ is the interior of K_m with respect to X. Since $K_m \subseteq Y$ and Y is open in X, it is clear that $\mathrm{int}_Y(K_m) = \mathrm{int}_X(K_m)$ and, thus, $\mathrm{int}_Y(K_m) \uparrow O$.

Let Y be closed in X and take any $O \subseteq Y$ which is open in Y. Then

$O = O' \cap Y$ for some $O' \subseteq X$ which is open in X and, hence, there is an increasing sequence (K'_m) of compact subsets of O' so that $\text{int}_X(K'_m) \uparrow O'$. We set $K_m = K'_m \cap Y$ and have that each K_m is a compact subset of O. Moreover, $\text{int}_X(K'_m) \cap Y \subseteq \text{int}_Y(K_m)$ for every m and, thus, $\text{int}_Y(K_m) \uparrow O$.

Examples

1. Proposition 5.2 implies that the euclidean space $\mathbf{R}^n$ satisfies the assumptions of Theorem 5.7. Therefore, *every locally finite Borel measure on* $\mathbf{R}^n$ *is regular.*

A special case of this is the Lebesgue measure in $\mathbf{R}^n$ (see Theorem 4.4 and Exercice 4.6.5).

2. If Y is an *open* or a *closed* subset of $\mathbf{R}^n$ with the subspace topology, then Lemma 5.8 together with Theorem 5.7 imply that *every locally finite Borel measure on* Y *is regular.*

As a special case, if $Y = (a_0, b_0)$ is a generalized interval in $\mathbf{R}$, then every locally finite Borel measure on Y is regular. Since Theorem 5.6 says that any such measure is a Lebesgue-Stieltjes measure, this result is, also, easily implied by Theorem 5.4.

5.3 Metric outer measures.

Let (X, d) be a metric space. We recall that, if E, F are non-empty subsets of X, the quantity

$$d(E, F) = \inf\{d(x, y) \mid x \in E, y \in F\}$$

is *the distance between* E *and* F.

Definition 5.4 *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 $E, F \subseteq X$ *with* $d(E, F) > 0$*.*

Theorem 5.8 *Let* (X, d) *be a metric space and* μ^* *an outer measure on* X*. Then, the measure* μ *which is induced by* μ^* *on* (X, Σ_{μ^*}) *is a Borel measure (i.e. all Borel sets in* X *are* μ^**-measurable) if and only if* μ^* *is a metric outer measure.*

Proof: Suppose that all Borel sets in X are μ^*-measurable and take arbitrary non-empty $E, F \subseteq X$ with $d(E, F) > 0$. We consider $r = d(E, F)$ and the open set $U = \cup_{x \in E} B(x; r)$. 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 and consider an open $U \subseteq X$.

If A is a non-empty subset of U, we define

$$A_n = \Big\{x \in A \,|\, d(x, y) \geq \frac{1}{n} \text{ for every } y \notin U\Big\}.$$

It is obvious that $A_n \subseteq A_{n+1}$ for all n. If $x \in A \subseteq U$, there is $r > 0$ so that $B(x; r) \subseteq U$ and, if we take $n \in \mathbf{N}$ so that $\frac{1}{n} \leq r$, then $x \in A_n$. Therefore,

$$A_n \uparrow A.$$

We define, now, $B_1 = A_1$ and $B_n = A_n \setminus A_{n-1}$ for all $n \geq 2$ and have that the sets $B_1, B_2, \dots$ are pairwise disjoint and that $A = \cup_{n=1}^{+\infty} B_n$. If $x \in A_n$ and $z \in B_{n+2}$, then $z \notin A_{n+1}$ and there is some $y \notin U$ so that $d(y, z) < \frac{1}{n+1}$. Then $d(x, z) \geq d(x, y) - d(y, z) > \frac{1}{n} - \frac{1}{n+1} = \frac{1}{n(n+1)}$. 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

$$\mu^*(B_1) + \mu^*(B_3) + \cdots + \mu^*(B_{2n+1}) \leq \mu^*(A_{2n+1})$$

and

$$\mu^*(B_2) + \mu^*(B_4) + \cdots + \mu^*(B_{2n}) \leq \mu^*(A_{2n})$$

for all n. If at least one of the series $\mu^*(B_1)+\mu^*(B_3)+\cdots$ and $\mu^*(B_2)+\mu^*(B_4)+\cdots$ diverges to $+\infty$, then either $\mu^*(A_{2n+1}) \to +\infty$ or $\mu^*(A_{2n}) \to +\infty$. Since the sequence $(\mu^*(A_n))$ is increasing, we get that in both cases it diverges to $+\infty$. Since, also $\mu^*(A_n) \leq \mu^*(A)$ for all n, we get that $\mu^*(A_n) \uparrow \mu^*(A)$. If both series $\mu^*(B_1)+\mu^*(B_3)+\cdots$ and $\mu^*(B_2)+\mu^*(B_4)+\cdots$ converge, for every $\epsilon > 0$ there is n so that $\sum_{k=n+1}^{+\infty} \mu^*(B_k) < \epsilon$. Now, $\mu^*(A) \leq \mu^*(A_n) + \sum_{k=n+1}^{+\infty} \mu^*(B_k) \leq \mu^*(A_n) + \epsilon$. This implies that $\mu^*(A_n) \uparrow \mu^*(A)$. Therefore, in any case,

$$\mu^*(A_n) \uparrow \mu^*(A).$$

We consider an arbitrary $E \subseteq X$ and we take $A = E \cap U$. Since $E \cup U^c \subseteq U^c$, we have that $d(A_n, E \cap U^c) > 0$ for all n and, hence, $\mu^*(E) \geq \mu^*\big(A_n \cup (E \cap U^c)\big) = \mu^*(A_n) + \mu^*(E \cap U^c)$ for all n. Taking the limit as $n \to +\infty$, we find

$$\mu^*(E) \geq \mu^*(E \cap U) + \mu^*(E \cap U^c).$$

We conclude that every U open in X is μ^*-measurable and, hence, every Borel set in X is μ^*-measurable.

5.4 Hausdorff measure.

Let (X, d) be a metric space. The *diameter* of a non-empty set $E \subseteq X$ is defined as $\text{diam}(E) = \sup\{d(x, y) \,|\, x, y \in E\}$ and the diameter of the $\emptyset$ is defined as $\text{diam}(\emptyset) = 0$.

We take an arbitrary $\delta > 0$ and consider the collection $\mathcal{C}_\delta$ of all subsets of X of diameter not larger than δ. We, then, fix some α with $0 < \alpha < +\infty$ and consider the function $\tau_{\alpha,\delta}:\mathcal{C}_\delta\to[0,+\infty]$ defined by $\tau_{\alpha,\delta}(E)=\big(\mathrm{diam}(E)\big)^\alpha$ for every $E\in\mathcal{C}_\delta$. We are, now, ready to apply Theorem 3.2 and define

$$h^*_{\alpha,\delta}(E)=\inf\Big\{\sum_{j=1}^{+\infty}\big(\mathrm{diam}(E_j)\big)^\alpha \mid E\subseteq\cup_{j=1}^{+\infty}E_j\,,\mathrm{diam}(E_j)\leq\delta \text{ for all } j\Big\}.$$

We have that $h^*_{\alpha,\delta}$ is an outer measure on X and we further define

$$h^*_\alpha(E)=\sup_{\delta>0}h^*_{\alpha,\delta}(E)\,,\qquad E\subseteq X.$$

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, $h^*_{\alpha,\delta_2}(E)\leq h^*_{\alpha,\delta_1}(E)$ and, hence,

$$h^*_\alpha(E)=\lim_{\delta\to0+}h^*_{\alpha,\delta}(E)\,,\qquad E\subseteq X.$$

Theorem 5.9 *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 $h^*_{\alpha,\delta}(E)\leq h^*_{\alpha,\delta}(F)\leq h^*_\alpha(F)$. Taking the supremum of the left side, we find $h^*_\alpha(E)\leq h^*_\alpha(F)$.

If $E=\cup_{j=1}^{+\infty}E_j$, then for every $\delta>0$ we have $h^*_{\alpha,\delta}(E)\leq\sum_{j=1}^{+\infty}h^*_{\alpha,\delta}(E_j)\leq\sum_{j=1}^{+\infty}h^*_\alpha(E_j)$ and, taking the supremum of the left side, we find $h^*_\alpha(E)\leq\sum_{j=1}^{+\infty}h^*_\alpha(E_j)$.

Therefore, h^*_α is an outer measure on X.

Now, take any $E,F\subseteq X$ with $d(E,F)>0$. If $h^*_\alpha(E\cup F)=+\infty$, then the equality $h^*_\alpha(E\cup F)=h^*_\alpha(E)+h^*_\alpha(F)$ is clearly true. We suppose that $h^*_\alpha(E\cup F)<+\infty$ and, hence, $h^*_{\alpha,\delta}(E\cup F)<+\infty$ for every $\delta>0$. We take arbitrary $\delta<d(E,F)$ and an arbitrary covering $E\cup F\subseteq\cup_{j=1}^{+\infty}A_j$ with $\mathrm{diam}(A_j)\leq\delta$ for every j. It is obvious that each A_j intersects at most one of the E and F. We set $B_j=A_j$ when A_j intersects E and $B_j=\emptyset$ otherwise and, similarly, $C_j=A_j$ when A_j intersects F and $C_j=\emptyset$ otherwise. Then, $E\subseteq\cup_{j=1}^{+\infty}B_j$ and $F\subseteq\cup_{j=1}^{+\infty}C_j$ and, hence, $h^*_{\alpha,\delta}(E)\leq\sum_{j=1}^{+\infty}(\mathrm{diam}(B_j))^\alpha$ and $h^*_{\alpha,\delta}(F)\leq\sum_{j=1}^{+\infty}(\mathrm{diam}(C_j))^\alpha$. Adding, we find $h^*_{\alpha,\delta}(E)+h^*_{\alpha,\delta}(F)\leq\sum_{j=1}^{+\infty}(\mathrm{diam}(A_j))^\alpha$ and, taking the infimum of the right side, $h^*_{\alpha,\delta}(E)+h^*_{\alpha,\delta}(F)\leq h^*_{\alpha,\delta}(E\cup F)$. Taking the limit as $\delta\to0+$ we find $h^*_\alpha(E)+h^*_\alpha(F)\leq h^*_\alpha(E\cup F)$ and, since the opposite inequality is obvious, we conclude that

$$h^*_\alpha(E)+h^*_\alpha(F)=h^*_\alpha(E\cup F).$$

Definition 5.5 *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 $(X,\Sigma_{h^*_\alpha})$ is called* ***the α-dimensional Hausdorff measure on*** *X.*

Theorem 5.10 *If (X, d) is a metric space and $0 < \alpha < +\infty$, then h_α is a Borel measure on X. Namely, $\mathcal{B}_X \subseteq \Sigma_{h^*_\alpha}$.*

Proof: Immediate, by Theorems 5.8 and 5.9.

Proposition 5.3 *Let (X, d) be a metric space, E a Borel set in X and let $0 < \alpha_1 < \alpha_2 < +\infty$. 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 consider a covering $E \subseteq \cup_{j=1}^{+\infty} A_j$ by subsets of X with $\text{diam}(A_j) \leq \delta$ for all j so that $\sum_{j=1}^{+\infty} \big(\text{diam}(A_j)\big)^{\alpha_1} < h^*_{\alpha_1,\delta}(E) + 1 \leq h^*_{\alpha_1}(E) + 1$.

Therefore, $h^*_{\alpha_2,\delta}(E) \leq \sum_{j=1}^{+\infty} \big(\text{diam}(A_j)\big)^{\alpha_2} \leq \delta^{\alpha_2-\alpha_1} \sum_{j=1}^{+\infty} \big(\text{diam}(A_j)\big)^{\alpha_1} \leq (h^*_{\alpha_1}(E) + 1)\delta^{\alpha_2-\alpha_1}$ and, taking the limit as $\delta \to 0+$, we find $h^*_{\alpha_2}(E) = 0$.

Hence, $h_{\alpha_2}(E) = 0$.

Proposition 5.4 *If E is any Borel set in a metric space (X, d), there is an $\alpha_0 \in [0, +\infty]$ with the property 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.
1. $h_\alpha(E) = 0$ for every $\alpha > 0$. In this case we set $\alpha_0 = 0$.
2. $h_\alpha(E) = +\infty$ for every $\alpha > 0$. We, now, set $\alpha_0 = +\infty$.
3. There are α_1 and α_2 in $(0, +\infty)$ so that $0 < h_{\alpha_1}(E)$ and $h_{\alpha_2}(E) < +\infty$.

Proposition 5.3 implies that $\alpha_1 \leq \alpha_2$ and that $h_\alpha(E) = +\infty$ for every $\alpha \in (0, \alpha_1)$ and $h_\alpha(E) = 0$ for every $\alpha \in (\alpha_2, +\infty)$. We consider the set $\{\alpha \in (0, +\infty) \,|\, h_\alpha(E) = +\infty\}$ and its supremum $\alpha_0 \in [\alpha_1, \alpha_2]$. The same Proposition 5.3 implies that $h_\alpha(E) = +\infty$ for every $\alpha \in (0, \alpha_0)$ and $h_\alpha(E) = 0$ for every $\alpha \in (\alpha_0, +\infty)$.

Definition 5.6 *If E is any Borel set in a metric space (X, d), the a_0 of Proposition 5.4 is called* ***the Hausdorff dimension of*** *E and it is denoted*

$$\dim_h(E)\,.$$

5.5 Exercises.

1. If $-\infty < x_1 < x_2 < \cdots < x_N < +\infty$ and $0 < \lambda_1, \ldots, \lambda_N < +\infty$, then find (and draw) the cumulative distribution function of $\mu = \sum_{k=1}^{N} \lambda_k \delta_{x_k}$.

2. *The Cantor measure.*

   Consider the Cantor function f (exercise 4.6.10) extended to $\mathbf{R}$ by $f(x) = 0$ for all $x < 0$ and $f(x) = 1$ for all $x > 1$. Then $f : \mathbf{R} \to [0, 1]$ is increasing, continuous and bounded.
   (i) f is the cumulative distribution function of μ_f.
   (ii) Prove that $\mu_f(C) = \mu_f(\mathbf{R}) = 1$.
   (iii) Each one of the 2^n subintervals of I_n (look at the construction of C) has measure equal to $\frac{1}{2^n}$.

3. Let μ be a locally finite Borel measure on $\mathbf{R}$ such that $\mu((-\infty, 0]) < +\infty$. Prove that there is a unique $f : \mathbf{R} \to \mathbf{R}$ increasing and continuous from the right so that $\mu = \mu_f$ and $f(-\infty) = 0$. Which is this function?

4. *Linear combinations of regular Borel measures.*

   If μ, μ_1, μ_2 are regular Borel measures on the topological space X and $\lambda \in [0, +\infty)$, prove that $\lambda\mu$ and $\mu_1 + \mu_2$ (exercise 2.6.2) are regular Borel measures on X.

5. Prove that every locally finite Borel measure on $\mathbf{R}^n$ is σ-finite.

6. *The support of a regular Borel measure.*

   Let μ be a regular 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

   $$\text{supp}(\mu) = \{x \in X \mid x \text{ is a support point for } \mu\}$$

   is called **the support of** μ.
   (i) Prove that $\text{supp}(\mu)$ is a closed set in X.
   (ii) Prove that $\mu(K) = 0$ for all compact sets $K \subseteq (\text{supp}(\mu))^c$.
   (iii) Using the regularity of μ, prove that $\mu\big((\text{supp}(\mu))^c\big) = 0$.
   (iv) Prove that $(\text{supp}(\mu))^c$ is the largest open set in X which is μ-null.

7. If f is the Cantor function (exercise 5.5.2), prove that the support (exercise 5.5.6) of μ_f is the Cantor set C.

8. *Supports of Lebesgue-Stieltjes measures.*

   Let $F : \mathbf{R} \to \mathbf{R}$ be any increasing function. Prove that the complement of the support (exercise 5.5.6) of the measure μ_F is the union of all *open* intervals on each of which F is constant.

9. Let $a : \mathbf{R} \to [0, +\infty]$ induce the point-mass distribution μ on $(\mathbf{R}, \mathcal{P}(\mathbf{R}))$. Then μ is a Borel measure on $\mathbf{R}$.
   (i) Prove that μ is locally finite if and only if $\sum_{-R \leq x \leq R} a_x < +\infty$ for all $R > 0$.
   (ii) In particular, prove that, if μ is locally finite, then $A = \{x \in \mathbf{R} \mid a_x > 0\}$ is countable.
   (iii) In case μ is locally finite, find an increasing, continuous from the right $F : \mathbf{R} \to \mathbf{R}$ (in terms of the function a) so that $\mu = \mu_F$ on $\mathcal{B}_{\mathbf{R}}$. Describe the sets E such that $\mu_F^*(E) = 0$ and find the σ-algebra Σ_F of all μ_F^*-measurable sets. Is $\Sigma_F = \mathcal{P}(\mathbf{R})$?

10. *Restrictions of regular Borel measures.*

    Let μ be a σ-finite regular Borel measure on the topological space X and Y be a Borel subset of X. Prove that the restriction μ_Y is a regular Borel measure on X.

11. *Continuous regular Borel measures.*

    Let μ be a regular Borel measure on the 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 in X with $0 < \mu(A) < +\infty$ and every $t \in (0, \mu(A))$ there is some Borel set B in X so that $B \subseteq A$ and $\mu(B) = t$.

12. Let X be a separable, complete metric space and μ be a Borel measure on X so that $\mu(X) = 1$. Prove that there is some B, a countable union of compact subsets of X, so that $\mu(B) = 1$.

Chapter 6

Measurable functions

6.1 Measurability.

Definition 6.1 *Let (X, Σ) and (Y, Σ') be measurable spaces and $f : X \to Y$. We say that f is (Σ, Σ')-**measurable** if $f^{-1}(E) \in \Sigma$ for all $E \in \Sigma'$.*

Example

A constant function is measurable. In fact, let (X, Σ) and (Y, Σ') be measurable spaces and $f(x) = y_0 \in Y$ for all $x \in X$. Take arbitrary $E \in \Sigma'$. If $y_0 \in E$, then $f^{-1}(E) = X \in \Sigma$. If $y_0 \notin E$, then $f^{-1}(E) = \emptyset \in \Sigma$.

Proposition 6.1 *Let (X, Σ) and (Y, Σ') measurable spaces and $f : X \to Y$. Suppose that $\mathcal{E}$ is a collection of subsets of Y so that $\Sigma(\mathcal{E}) = \Sigma'$. If $f^{-1}(E) \in \Sigma$ for all $E \in \mathcal{E}$, then f is (Σ, Σ')-measurable.*

Proof: We consider the collection $\mathcal{S} = \{E \subseteq Y \mid f^{-1}(E) \in \Sigma\}$.

Since $f^{-1}(\emptyset) = \emptyset \in \Sigma$, it is clear that $\emptyset \in \mathcal{S}$.

If $E \in \mathcal{S}$, then $f^{-1}(E^c) = (f^{-1}(E))^c \in \Sigma$ and thus $E^c \in \mathcal{S}$.

If $E_1, E_2, \ldots \in \mathcal{S}$, then $f^{-1}(\cup_{j=1}^{+\infty} E_j) = \cup_{j=1}^{+\infty} f^{-1}(E_j) \in \Sigma$, implying that $\cup_{j=1}^{+\infty} E_j \in \mathcal{S}$.

Therefore $\mathcal{S}$ is a σ-algebra of subsets of Y. $\mathcal{E}$ is, by hypothesis, included in $\mathcal{S}$ and, thus, $\Sigma' = \Sigma(\mathcal{E}) \subseteq \mathcal{S}$. This concludes the proof.

Proposition 6.2 *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 $\mathcal{E}$ be the collection of all open subsets of Y. Then, by continuity, $f^{-1}(E)$ is an open and, hence, Borel subset of X for all $E \in \mathcal{E}$. Since $\Sigma(\mathcal{E}) = \mathcal{B}_Y$, Proposition 6.1 implies that f is $(\mathcal{B}_X, \mathcal{B}_Y)$-measurable.

6.2 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 $(f\rceil A)(x) = f(x)$ for all $x \in A$, is the usual **restriction of f on A**.

Recall that, if Σ is a σ-algebra of subsets of X and $A \in \Sigma$ is non-empty, then, by Lemma 2.1, $\Sigma\rceil A = \{E \subseteq A \mid E \in \Sigma\}$ is a σ-algebra of subsets of A. We call $\Sigma\rceil A$ **the restriction of** Σ **on** A.

Proposition 6.3 *Let (X, Σ), (Y, Σ') be measurable spaces and $f : X \to Y$. Let the non-empty $A_1, \dots, A_n \in \Sigma$ be pairwise disjoint and $A_1 \cup \dots \cup A_n = X$.*

Then f is (Σ, Σ')-measurable if and only if $f\rceil A_j$ is $(\Sigma\rceil A_j, \Sigma')$-measurable for all $j = 1, \dots, n$.

Proof: Let f be (Σ, Σ')-measurable. For all $E \in \Sigma'$ we have $(f\rceil A_j)^{-1}(E) = f^{-1}(E) \cap A_j \in \Sigma\rceil A_j$ because the set $f^{-1}(E) \cap A_j$ belongs to Σ and is included in A_j. Hence $f\rceil A_j$ is $(\Sigma\rceil A_j, \Sigma')$-measurable for all j.

Now, let $f\rceil A_j$ be $(\Sigma\rceil A_j, \Sigma')$-measurable for all j. For every $E \in \Sigma'$ we have that $f^{-1}(E) \cap A_j = (f\rceil A_j)^{-1}(E) \in \Sigma\rceil A_j$ and, hence, $f^{-1}(E) \cap A_j \in \Sigma$ for all j. Therefore $f^{-1}(E) = (f^{-1}(E) \cap A_1) \cup \dots \cup (f^{-1}(E) \cap A_n) \in \Sigma$, implying that f is (Σ, Σ')-measurable.

In a free language: *measurability of a function separately on complementary (measurable) pieces of the space is equivalent to measurability on the whole space.*

There are two operations on measurable functions that are taken care of by Proposition 6.3. 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 finitely many A_j's are non-empty, pairwise disjoint and cover X. The rules are: *restriction of measurable functions on measurable sets are measurable* and *gluing of measurable functions defined on measurable subsets results to a measurable function.*

6.3 Functions with arithmetical values.

Definition 6.2 *Let (X, Σ) be measurable space and $f : X \to \mathbf{R}$ or $\overline{\mathbf{R}}$ or $\mathbf{C}$ or $\overline{\mathbf{C}}$. We say f is Σ-**measurable** if it is $(\Sigma, \mathcal{B}_{\mathbf{R}}$ or $\mathcal{B}_{\overline{\mathbf{R}}}$ or $\mathcal{B}_{\mathbf{C}}$ or $\mathcal{B}_{\overline{\mathbf{C}}})$-measurable, respectively.*

*In the particular case when (X, Σ) is $(\mathbf{R}^n, \mathcal{B}_{\mathbf{R}^n})$ or $(\mathbf{R}^n, \mathcal{L}_n)$, then we use the term **Borel measurable** or, respectively, **Lebesgue measurable** for f.*

If $f : X \to \mathbf{R}$, then it is also true that $f : X \to \overline{\mathbf{R}}$. Thus, according to the definition we have given, there might be a conflict between the two meanings of Σ-measurability of f. But, actually, there is no such conflict. Suppose, for example, that f is assumed $(\Sigma, \mathcal{B}_{\mathbf{R}})$-measurable. If $E \in \mathcal{B}_{\overline{\mathbf{R}}}$, then $E \cap \mathbf{R} \in \mathcal{B}_{\mathbf{R}}$ and, thus, $f^{-1}(E) = f^{-1}(E \cap \mathbf{R}) \in \Sigma$. Hence f is $(\Sigma, \mathcal{B}_{\overline{\mathbf{R}}})$-measurable. Let, conversely, f be $(\Sigma, \mathcal{B}_{\overline{\mathbf{R}}})$-measurable. If $E \in \mathcal{B}_{\mathbf{R}}$, then $E \in \mathcal{B}_{\overline{\mathbf{R}}}$ and, thus, $f^{-1}(E) \in \Sigma$. Hence f is $(\Sigma, \mathcal{B}_{\mathbf{R}})$-measurable.

The same question arises when $f : X \to \mathbf{C}$, because it is then also true that $f : X \to \overline{\mathbf{C}}$. Exactly as before, we may prove that f is $(\Sigma, \mathcal{B}_{\mathbf{C}})$-measurable if and only if it is $(\Sigma, \mathcal{B}_{\overline{\mathbf{C}}})$-measurable and there is no conflict in the definition.

Proposition 6.4 *Let (X,Σ) be measurable space and $f : X \to \mathbf{R}^n$. Let, for each $j = 1,\dots,n$, $f_j : X \to \mathbf{R}$ denote the j-th component function of f. Namely, $f(x) = (f_1(x),\dots,f_n(x))$ for all $x \in X$.*

Then f is $(\Sigma,\mathcal{B}_{\mathbf{R}^n})$-measurable if and only if every f_j is Σ-measurable.

Proof: Let f be $(\Sigma,\mathcal{B}_{\mathbf{R}^n})$-measurable. For all intervals $(a,b]$ we have

$$f_j^{-1}((a,b]) = f^{-1}(\mathbf{R}\times\cdots\times\mathbf{R}\times(a,b]\times\mathbf{R}\times\cdots\times\mathbf{R})$$

which belongs to Σ. Since the collection of all $(a,b]$ generates $\mathcal{B}_{\mathbf{R}}$, Proposition 6.1 implies that f_j is Σ-measurable.

Now let every f_j be Σ-measurable. Then

$$f^{-1}((a_1,b_1]\times\cdots\times(a_n,b_n]) = f_1^{-1}((a_1,b_1])\cap\cdots\cap f_n^{-1}((a_n,b_n])$$

which is an element of Σ. The collection of all open-closed intervals generates $\mathcal{B}_{\mathbf{R}^n}$ and Proposition 6.1, again, implies that f is $(\Sigma,\mathcal{B}_{\mathbf{R}^n})$- measurable.

In a free language: *measurability of a vector function is equivalent to measurability of all component functions.*

The next two results give simple criteria for measurability of real or complex valued functions.

Proposition 6.5 *Let (X,Σ) be measurable space and $f : X \to \mathbf{R}$. Then f is Σ-measurable if and only if $f^{-1}((a,+\infty)) \in \Sigma$ for all $a \in \mathbf{R}$.*

Proof: Since $(a,+\infty) \in \mathcal{B}_{\mathbf{R}}$, one direction is trivial.

If $f^{-1}((a,+\infty)) \in \Sigma$ for all $a \in \mathbf{R}$, then $f^{-1}((a,b]) = f^{-1}((a,+\infty)) \setminus f^{-1}((b,+\infty)) \in \Sigma$ for all $(a,b]$. Now the collection of all intervals $(a,b]$ generates $\mathcal{B}_{\mathbf{R}}$ and Proposition 6.1 implies that f is Σ-measurable.

Of course, in the statement of Proposition 6.5 one may replace the intervals $(a,+\infty)$ by the intervals $[a,+\infty)$ or $(-\infty,b)$ or $(-\infty,b]$.

If $f : X \to \mathbf{C}$, then the functions $\Re(f),\Im(f) : X \to \mathbf{R}$ are defined by $\Re(f)(x) = \Re(f(x))$ and $\Im(f)(x) = \Im(f(x))$ for all $x \in X$ and they are called **the real part** and **the imaginary part of** f, respectively.

Proposition 6.6 *Let (X,Σ) be measurable space and $f : X \to \mathbf{C}$. Then f is Σ-measurable if and only if both $\Re(f)$ and $\Im(f)$ are Σ-measurable.*

Proof: An immediate application of Proposition 6.4.

The next two results investigate extended-real or extended-complex valued functions.

Proposition 6.7 *Let (X,Σ) be measurable space and $f : X \to \overline{\mathbf{R}}$. The following are equivalent.*
(i) f is Σ-measurable.
(ii) $f^{-1}(\{+\infty\}), f^{-1}(\mathbf{R}) \in \Sigma$ and, if $A = f^{-1}(\mathbf{R})$ is non-empty, the function $f\rceil A : A \to \mathbf{R}$ is $\Sigma\rceil A$-measurable.
(iii) $f^{-1}((a,+\infty]) \in \Sigma$ for all $a \in \mathbf{R}$.

Proof: It is trivial that (i) implies (iii), since $(a, +\infty] \in \mathcal{B}_{\overline{\mathbf{R}}}$ for all $a \in \mathbf{R}$.

Assume (ii) and consider $B = f^{-1}(\{+\infty\}) \in \Sigma$ and $C = f^{-1}(\{-\infty\}) = (A \cup B)^c \in \Sigma$. The restrictions $f\rceil B = +\infty$ and $f\rceil C = -\infty$ are constants and hence are, respectively, $\Sigma\rceil B$-measurable and $\Sigma\rceil C$-measurable. Proposition 6.3 implies that f is Σ-measurable and thus (ii) implies (i).

Now assume (iii). Then $f^{-1}(\{+\infty\}) = \cap_{n=1}^{+\infty} f^{-1}((n, +\infty]) \in \Sigma$ and then $f^{-1}((a, +\infty)) = f^{-1}((a, +\infty]) \setminus f^{-1}(\{+\infty\}) \in \Sigma$ for all $a \in \mathbf{R}$. Moreover, $f^{-1}(\mathbf{R}) = \cup_{n=1}^{+\infty} f^{-1}((-n, +\infty)) \in \Sigma$. For all $a \in \mathbf{R}$ we get $(f\rceil A)^{-1}((a, +\infty)) = f^{-1}((a, +\infty)) \in \Sigma\rceil A$, because the last set belongs to Σ and is included in A. Proposition 6.5 implies that $f\rceil A$ is $\Sigma\rceil A$-measurable and (ii) is now proved.

Proposition 6.8 *Let (X, Σ) be measurable space and $f : X \to \overline{\mathbf{C}}$. The following are equivalent.*
(i) f is Σ-measurable.
(ii) $f^{-1}(\mathbf{C}) \in \Sigma$ and, if $A = f^{-1}(\mathbf{C})$ is non-empty, the $f\rceil A : A \to \mathbf{C}$ is $\Sigma\rceil A$-measurable.

Proof: Assume (ii) and consider $B = f^{-1}(\{\infty\}) = (f^{-1}(\mathbf{C}))^c \in \Sigma$. The restriction $f\rceil B$ is constant ∞ and hence $\Sigma\rceil B$-measurable. Proposition 6.3 implies that f is Σ-measurable. Thus (ii) implies (i).

Now assume (i). Then $A = f^{-1}(\mathbf{C}) \in \Sigma$ since $\mathbf{C} \in \mathcal{B}_{\overline{\mathbf{C}}}$. Proposition 6.3 implies that $f\rceil A$ is $\Sigma\rceil A$-measurable and (i) implies (ii).

6.4 Composition.

Proposition 6.9 *Let (X, Σ), (Y, Σ'), (Z, Σ'') be measurable spaces and let $f : X \to Y$, $g : Y \to Z$. If f is (Σ, Σ')-measurable and g is (Σ', Σ'')-measurable, then $g \circ f : X \to Z$ is (Σ, Σ'')-measurable.*

Proof: For all $E \in \Sigma''$ we have $(g \circ f)^{-1}(E) = f^{-1}\big(g^{-1}(E)\big) \in \Sigma$, because $g^{-1}(E) \in \Sigma'$.

Hence: *composition of measurable functions is measurable.*

6.5 Sums and products.

The next result is: *sums and products of real or complex valued measurable functions are measurable functions.*

Proposition 6.10 *Let (X, Σ) be a measurable space and $f, g : X \to \mathbf{R}$ or $\mathbf{C}$ be Σ-measurable. Then $f + g, fg$ are Σ-measurable.*

Proof: (a) We consider $H : X \to \mathbf{R}^2$ by the formula $H(x) = (f(x), g(x))$ for all $x \in X$. Proposition 6.4 implies that H is $(\Sigma, \mathcal{B}_{\mathbf{R}^2})$-measurable. Now consider $\phi, \psi : \mathbf{R}^2 \to \mathbf{R}$ by the formulas $\phi(y, z) = y + z$ and $\psi(y, z) = yz$. These functions are continuous and Proposition 6.2 implies that they are $(\mathcal{B}_{\mathbf{R}^2}, \mathcal{B}_{\mathbf{R}})$-measurable.

Therefore the compositions $\phi \circ H, \psi \circ H : X \to \mathbf{R}$ are Σ-measurable. But $(\phi \circ H)(x) = f(x) + g(x) = (f+g)(x)$ and $(\psi \circ H)(x) = f(x)g(x) = (fg)(x)$ for all $x \in X$ and we conclude that $f+g = \phi \circ H$ and $fg = \psi \circ H$ are Σ-measurable.
(b) In the case $f, g : X \to \mathbf{C}$ we consider $\Re(f), \Im(f), \Re(g), \Im(g) : X \to \mathbf{R}$, which, by Proposition 6.6, are all Σ-measurable. Then, part (a) implies that $\Re(f+g) = \Re(f) + \Re(g), \Im(f+g) = \Im(f) + \Im(g), \Re(fg) = \Re(f)\Re(g) - \Im(f)\Im(g), \Im(fg) = \Re(f)\Im(g) + \Im(f)\Re(g)$ are all Σ-measurable. Proposition 6.6 again, gives that $f+g, fg$ are Σ-measurable.

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 $\overline{\mathbf{R}}$ and neither is $\infty + \infty$ defined in $\overline{\mathbf{C}}$. Hence, when we add $f, g : X \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$, we must agree on how to treat the summation on, respectively, the set $B = \{x \in X \,|\, f(x) = +\infty, g(x) = -\infty \text{ or } f(x) = -\infty, g(x) = +\infty\}$ or the set $B = \{x \in X \,|\, f(x) = \infty, g(x) = \infty\}$. There are two standard ways to do this. One is to ignore the bad set and consider $f+g$ defined on $B^c \subseteq X$, 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 6.11 *Let (X, Σ) be a measurable space and $f, g : X \to \overline{\mathbf{R}}$ be Σ-measurable. Then the set*

$$B = \{x \in X \,|\, f(x) = +\infty, g(x) = -\infty \text{ or } f(x) = -\infty, g(x) = +\infty\}$$

belongs to Σ.
(i) The function $f+g : B^c \to \overline{\mathbf{R}}$ is $\Sigma\rceil B^c$-measurable.
(ii) If $h : B \to \overline{\mathbf{R}}$ is $\Sigma\rceil B$-measurable and we define

$$(f+g)(x) = \begin{cases} f(x) + g(x), & \text{if } x \in B^c, \\ h(x), & \text{if } x \in B, \end{cases}$$

then $f+g : X \to \overline{\mathbf{R}}$ is Σ-measurable.

Similar results hold if $f, g : X \to \overline{\mathbf{C}}$ and $B = \{x \in X \,|\, f(x) = \infty, g(x) = \infty\}$.

Proof: We have

$$B = \big(f^{-1}(\{+\infty\}) \cap g^{-1}(\{-\infty\})\big) \cup \big(f^{-1}(\{-\infty\}) \cap g^{-1}(\{+\infty\})\big) \in \Sigma.$$

(i) Consider the sets $A = \{x \in X \,|\, f(x), g(x) \in \mathbf{R}\}$, $C_1 = \{x \in X \,|\, f(x) = +\infty, g(x) \neq -\infty \text{ or } f(x) \neq -\infty, g(x) = +\infty\}$ and $C_2 = \{x \in X \,|\, f(x) = -\infty, g(x) \neq +\infty \text{ or } f(x) \neq +\infty, g(x) = -\infty\}$. It is clear that $A, C_1, C_2 \in \Sigma$, that $B^c = A \cup C_1 \cup C_2$ and that the three sets are pairwise disjoint.

The restriction of $f+g$ on A is the sum of the real valued $f\rceil A, g\rceil A$. By Proposition 6.3, both $f\rceil A, g\rceil A$ are $\Sigma\rceil A$-measurable and, by Proposition 6.10, $(f+g)\rceil A = f\rceil A + g\rceil A$ is $\Sigma\rceil A$-measurable. The restriction $(f+g)\rceil C_1$ is constant $+\infty$, and is thus $\Sigma\rceil C_1$-measurable. Also the restriction $(f+g)\rceil C_2 = -\infty$ is

$\Sigma\rceil C_2$-measurable. Proposition 6.3 implies that $f+g : B^c \to \overline{\mathbf{R}}$ is $\Sigma\rceil B^c$-measurable.
(ii) This is immediate after the result of (i) and Proposition 6.3.
The case $f, g : X \to \overline{\mathbf{C}}$ is similar, if not simpler.

For multiplication we make the following

Convention: $(\pm\infty)\cdot 0 = 0\cdot(\pm\infty) = 0$ *in* $\overline{\mathbf{R}}$ *and* $\infty\cdot 0 = 0\cdot\infty = 0$ *in* $\overline{\mathbf{C}}$.

Thus, multiplication is always defined and we may state that: *the product of measurable functions is measurable.*

Proposition 6.12 *Let* (X,Σ) *be a measurable space and* $f, g : X \to \overline{\mathbf{R}}$ *or* $\overline{\mathbf{C}}$ *be* Σ*-measurable. Then the function* fg *is* Σ*-measurable.*

Proof: Let $f, g : X \to \overline{\mathbf{R}}$.
Consider the sets $A = \{x \in X \,|\, f(x), g(x) \in \mathbf{R}\}$, $C_1 = \{x \in X \,|\, f(x) = +\infty, g(x) > 0$ or $f(x) = -\infty, g(x) < 0$ or $f(x) > 0, g(x) = +\infty$ or $f(x) < 0, g(x) = -\infty\}$, $C_2 = \{x \in X \,|\, f(x) = -\infty, g(x) > 0$ or $f(x) = +\infty, g(x) < 0$ or $f(x) > 0, g(x) = -\infty$ or $f(x) < 0, g(x) = +\infty\}$ and $D = \{x \in X \,|\, f(x) = \pm\infty, g(x) = 0$ or $f(x) = 0, g(x) = \pm\infty\}$. These four sets are pairwise disjoint, their union is X and they all belong to Σ.
The restriction of fg on A is equal to the product of the real valued $f\rceil A, g\rceil A$, which, by Propositions 6.3 and 6.10, is $\Sigma\rceil A$-measurable. The restriction $(fg)\rceil C_1$ is constant $+\infty$ and, hence, $\Sigma\rceil C_1$-measurable. Similarly, $(fg)\rceil C_2 = -\infty$ is $\Sigma\rceil C_2$-measurable. Finally, $(fg)\rceil D = 0$ is $\Sigma\rceil D$-measurable.
Proposition 6.3 implies now that fg is Σ-measurable.
If $f, g : X \to \overline{\mathbf{C}}$, the proof is similar and slightly simpler.

6.6 Absolute value and signum.

The action of the absolute value on infinities is: $|+\infty| = |-\infty| = +\infty$ and $|\infty| = +\infty$.

Proposition 6.13 *Let* (X,Σ) *be a measurable space and* $f : X \to \overline{\mathbf{R}}$ *or* $\overline{\mathbf{C}}$ *be* Σ*-measurable. Then the function* $|f| : X \to [0,+\infty]$ *is* Σ*-measurable.*

Proof: Let $f : X \to \overline{\mathbf{R}}$. The function $|\cdot| : \overline{\mathbf{R}} \to [0,+\infty]$ is continuous and, hence, $(\mathcal{B}_{\overline{\mathbf{R}}}, \mathcal{B}_{\overline{\mathbf{R}}})$-measurable. Therefore, $|f|$, the composition of $|\cdot|$ and f, is Σ-measurable.
The same proof applies in the case $f : X \to \overline{\mathbf{C}}$.

Definition 6.3 *For every* $z \in \overline{\mathbf{C}}$ *we define*

$$sign(z) = \begin{cases} \frac{z}{|z|}, & \text{if } z \neq 0, \\ 0, & \text{if } z = 0, \\ \infty, & \text{if } z = \infty. \end{cases}$$

If we denote $\mathbf{C}^* = \overline{\mathbf{C}} \setminus \{0, \infty\}$, then the restriction $\text{sign}\rceil\mathbf{C}^* : \mathbf{C}^* \to \overline{\mathbf{C}}$ is continuous. This implies that, for every Borel set E in $\overline{\mathbf{C}}$, the set $(\text{sign}\rceil\mathbf{C}^*)^{-1}(E)$ is a Borel set contained in $\mathbf{C}^*$. The restriction $\text{sign}\rceil\{0\}$ is constant 0 and the restriction $\text{sign}\rceil\{\infty\}$ is constant ∞. Therefore, for every Borel set E in $\overline{\mathbf{C}}$, the sets $(\text{sign}\rceil\{0\})^{-1}(E), (\text{sign}\rceil\{\infty\})^{-1}(E)$ are Borel sets. Altogether, $\text{sign}^{-1}(E) = (\text{sign}\rceil\mathbf{C}^*)^{-1}(E) \cup (\text{sign}\rceil\{0\})^{-1}(E) \cup (\text{sign}\rceil\{\infty\})^{-1}(E)$ is a Borel set in $\overline{\mathbf{C}}$. This means that $\text{sign} : \overline{\mathbf{C}} \to \overline{\mathbf{C}}$ is $(\mathcal{B}_{\overline{\mathbf{C}}}, \mathcal{B}_{\overline{\mathbf{C}}})$-measurable.

All this applies in the same way to the function $\text{sign} : \overline{\mathbf{R}} \to \overline{\mathbf{R}}$ with the simple formula

$$\text{sign}(x) = \begin{cases} 1, & \text{if } 0 < x \leq +\infty, \\ -1, & \text{if } -\infty \leq x < 0, \\ 0, & \text{if } x = 0. \end{cases}$$

Hence $\text{sign} : \overline{\mathbf{R}} \to \overline{\mathbf{R}}$ is $(\mathcal{B}_{\overline{\mathbf{R}}}, \mathcal{B}_{\overline{\mathbf{R}}})$-measurable.

For all $z \in \overline{\mathbf{C}}$ we may write

$$z = \text{sign}(z) \cdot |z|$$

and this is called **the polar decomposition of** z.

Proposition 6.14 *Let (X, Σ) be a measurable space and $f : X \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$ be Σ-measurable. Then the function sign(f) is Σ-measurable.*

Proof: If $f : X \to \overline{\mathbf{R}}$, then $\text{sign}(f)$ is the composition of $\text{sign} : \overline{\mathbf{R}} \to \overline{\mathbf{R}}$ and f and the result is clear by Proposition 6.9. The same applies if $f : X \to \overline{\mathbf{C}}$.

6.7 Maximum and minimum.

Proposition 6.15 *Let (X, Σ) be measurable space and $f_1, \ldots, f_n : X \to \overline{\mathbf{R}}$ be Σ-measurable. Then the functions $\max\{f_1, \ldots, f_n\}, \min\{f_1, \ldots, f_n\} : X \to \overline{\mathbf{R}}$ are Σ-measurable.*

Proof: If $h = \max\{f_1, \ldots, f_n\}$, then for all $a \in \mathbf{R}$ we have $h^{-1}((a, +\infty]) = \cup_{j=1}^n f_j^{-1}((a, +\infty]) \in \Sigma$. Proposition 6.7 implies that h is Σ-measurable and from $\min\{f_1, \ldots, f_n\} = -\max\{-f_1, \ldots, -f_n\}$ we see that $\min\{f_1, \ldots, f_n\}$ is also Σ-measurable.

The next result is about *comparison of measurable functions.*

Proposition 6.16 *Let (X, Σ) be a measurable space and $f, g : X \to \overline{\mathbf{R}}$ be Σ-measurable. Then $\{x \in X \,|\, f(x) = g(x)\}, \{x \in X \,|\, f(x) < g(x)\} \in \Sigma$.*

If $f, g : X \to \overline{\mathbf{C}}$ is Σ-measurable, then $\{x \in X \,|\, f(x) = g(x)\} \in \Sigma$.

Proof: Consider the set $A = \{x \in X \,|\, f(x), g(x) \in \mathbf{R}\} \in \Sigma$. Then the functions $f\rceil A, g\rceil A$ are $\Sigma\rceil A$-measurable and thus $f\rceil A - g\rceil A$ is $\Sigma\rceil A$-measurable. Hence the sets $\{x \in A \,|\, f(x) = g(x)\} = (f\rceil A - g\rceil A)^{-1}(\{0\})$ and $\{x \in A \,|\, f(x) < g(x)\} = (f\rceil A - g\rceil A)^{-1}((-\infty, 0))$ belong to $\Sigma\rceil A$. This, of course, means that these sets belong to Σ (and that they are subsets of A).

We can obviously write $\{x \in X \,|\, f(x) = g(x)\} = \{x \in A \,|\, f(x) = g(x)\} \cup \big(f^{-1}(\{-\infty\}) \cap g^{-1}(\{-\infty\})\big) \cup \big(f^{-1}(\{+\infty\}) \cap g^{-1}(\{+\infty\})\big) \in \Sigma$. In a similar manner, $\{x \in X \,|\, f(x) < g(x)\} = \{x \in A \,|\, f(x) < g(x)\} \cup \big(f^{-1}(\{-\infty\}) \cap g^{-1}((-\infty, +\infty])\big) \cup \big(f^{-1}([-\infty, +\infty)) \cap g^{-1}(\{+\infty\})\big) \in \Sigma$.

The case of $f, g : X \to \overline{\mathbf{C}}$ and of $\{x \in X \,|\, f(x) = g(x)\}$ is even simpler.

6.8 Truncation.

There are many possible truncations of a function.

Definition 6.4 *Let $f : X \to \overline{\mathbf{R}}$ and consider $\alpha, \beta \in \overline{\mathbf{R}}$ with $\alpha \leq \beta$. We define*

$$f_{(\alpha)}^{(\beta)}(x) = \begin{cases} f(x), & \text{if } \alpha \leq f(x) \leq \beta, \\ \alpha, & \text{if } f(x) < \alpha, \\ \beta, & \text{if } \beta < f(x). \end{cases}$$

We write $f^{(\beta)}$ instead of $f_{(-\infty)}^{(\beta)}$ and $f_{(\alpha)}$ instead of $f_{(\alpha)}^{(+\infty)}$.

*The functions $f_{(\alpha)}^{(\beta)}, f^{(\beta)}, f_{(\alpha)}$ are called **truncations of** f.*

Proposition 6.17 *Let (X, Σ) be a measurable space and $f : X \to \overline{\mathbf{R}}$ be a Σ-measurable function. Then all truncations $f_{(\alpha)}^{(\beta)}$ are Σ-measurable.*

Proof: The proof is obvious after the formula $f_{(\alpha)}^{(\beta)} = \min\big\{\max\{f, \alpha\}, \beta\big\}$.

An important role is played by the following special truncations.

Definition 6.5 *Let $f : X \to \overline{\mathbf{R}}$. The $f^+ : X \to [0, +\infty]$ and $f^- : X \to [0, +\infty]$ defined by the formulas*

$$f^+(x) = \begin{cases} f(x), & \text{if } 0 \leq f(x), \\ 0, & \text{if } f(x) < 0, \end{cases} \qquad f^-(x) = \begin{cases} 0, & \text{if } 0 \leq f(x), \\ -f(x), & \text{if } f(x) < 0, \end{cases}$$

*are called, respectively, **the positive part** and **the negative part of** f.*

It is clear that $f^+ = f_{(0)}$ and $f^- = -f^{(0)}$. Hence if Σ is a σ-algebra of subsets of X and f is Σ-measurable, then both f^+ and f^- are Σ-measurable. It is also trivial to see that at every $x \in X$ either $f^+(x) = 0$ or $f^-(x) = 0$ and that

$$f^+ + f^- = |f|, \qquad f^+ - f^- = f.$$

There is another type of truncations used mainly for extended-complex valued functions.

Definition 6.6 *Let $f : X \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$ and consider $r \in [0, +\infty]$. We define*

$$^{(r)}f(x) = \begin{cases} f(x), & \text{if } \ |f(x)| \leq r, \\ r \cdot \mathit{sign}(f(x)), & \text{if } \ r < |f(x)|. \end{cases}$$

*The functions $^{(r)}f$ are called **truncations of** f.*

Observe that, if $f : X \to \overline{\mathbf{R}}$, then ${}^{(r)}f = f^{(r)}_{(-r)}$.

Proposition 6.18 *Let (X, Σ) be a measurable space and $f : X \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$ a Σ-measurable function. Then all truncations ${}^{(r)}f$ are Σ-measurable.*

Proof: Observe that the function $\phi_r : \overline{\mathbf{R}} \to \overline{\mathbf{R}}$ with formula

$$\phi_r(x) = \begin{cases} x, & \text{if } |x| \leq r, \\ r \cdot \text{sign}(x), & \text{if } r < |x|, \end{cases}$$

is continuous on $\overline{\mathbf{R}}$ and hence $(\mathcal{B}_{\overline{\mathbf{R}}}, \mathcal{B}_{\overline{\mathbf{R}}})$-measurable. Now ${}^{(r)}f = \phi_r \circ f$ is Σ-measurable.

The proof in the case $f : X \to \overline{\mathbf{C}}$ is similar.

6.9 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 6.19 *Let (X, Σ) be a measurable space and (f_j) a sequence of Σ-measurable functions $f_j : X \to \overline{\mathbf{R}}$. Then all the functions $\sup_{j \in \mathbf{N}} f_j$, $\inf_{j \in \mathbf{N}} f_j$, $\limsup_{j \to +\infty} f_j$ and $\liminf_{j \to +\infty} f_j$ are Σ-measurable.*

Proof: Let $h = \sup_{j \in \mathbf{N}} f_j : X \to \overline{\mathbf{R}}$. For every $a \in \mathbf{R}$ we have $h^{-1}((a, +\infty]) = \cup_{j=1}^{+\infty} f_j^{-1}((a, +\infty]) \in \Sigma$. Proposition 6.7 implies that h is Σ-measurable.

Now $\inf_{j \in \mathbf{N}} f_j = -\sup_{j \in \mathbf{N}}(-f_j)$ is also Σ-measurable.

And, finally, $\limsup_{j \to +\infty} f_j = \inf_{j \in \mathbf{N}} \big(\sup_{k \geq j} f_k \big)$ and $\liminf_{j \to +\infty} f_j = \sup_{j \in \mathbf{N}} \big(\inf_{k \geq j} f_k \big)$ are Σ-measurable.

Proposition 6.20 *Let (X, Σ) be a measurable space and (f_j) a sequence of Σ-measurable functions $f_j : X \to \overline{\mathbf{R}}$. Then the set*

$$A = \{x \in X \mid \lim_{j \to +\infty} f_j(x) \text{ exists in } \overline{\mathbf{R}}\}$$

belongs to Σ.
(i) The function $\lim_{j \to +\infty} f_j : A \to \overline{\mathbf{R}}$ is $\Sigma \rceil A$-measurable.
(ii) If $h : A^c \to \overline{\mathbf{R}}$ is $\Sigma \rceil A^c$-measurable and we define

$$(\lim_{j \to +\infty} f_j)(x) = \begin{cases} \lim_{j \to +\infty} f_j(x), & \text{if } x \in A, \\ h(x), & \text{if } x \in A^c, \end{cases}$$

then $\lim_{j \to +\infty} f_j : X \to \overline{\mathbf{R}}$ is Σ-measurable.

Similar results hold if $f_j : X \to \overline{\mathbf{C}}$ for all j and we consider the set $A = \{x \in X \mid \lim_{j \to +\infty} f_j(x) \text{ exists in } \overline{\mathbf{C}}\}$.

Proof: (a) Suppose that $f_j : X \to \overline{\mathbf{R}}$ for all j.

Proposition 6.19 implies that $\limsup_{j\to+\infty} f_j$ and $\liminf_{j\to+\infty} f_j$ are both Σ-measurable. Since $\lim_{j\to+\infty} f_j(x)$ exists if and only if $\limsup_{j\to+\infty} f_j(x) = \liminf_{j\to+\infty} f_j(x)$, we have that

$$A = \{x \in X \mid \limsup_{j\to+\infty} f_j(x) = \liminf_{j\to+\infty} f_j(x)\}$$

and Proposition 6.16 implies that $A \in \Sigma$.

(i) It is clear that the function $\lim_{j\to+\infty} f_j : A \to \overline{\mathbf{R}}$ is just the restriction of $\limsup_{j\to+\infty} f_j$ (or of $\liminf_{j\to+\infty} f_j$) to A and hence it is $\Sigma\rceil A$-measurable.

(ii) The proof of (ii) is a direct consequence of (i) and Proposition 6.3.

(b) Let now $f_j : X \to \mathbf{C}$ for all j.

Consider the set $B = \{x \in X \mid \lim_{j\to+\infty} f_j(x) \text{ exists in } \mathbf{C}\}$ and the set $C = \{x \in X \mid \lim_{j\to+\infty} f_j(x) = \infty\}$. Clearly, $B \cup C = A$.

Now, $C = \{x \in X \mid \lim_{j\to+\infty} |f_j|(x) = +\infty\}$. Since $|f_j| : X \to \mathbf{R}$ for all j, part (a) implies that the function $\lim_{j\to+\infty} |f_j|$ is measurable on the set on which it exists. Therefore, $C \in \Sigma$.

B is the intersection of $B_1 = \{x \in X \mid \lim_{j\to+\infty} \Re(f_j)(x) \text{ exists in } \mathbf{R}\}$ and $B_2 = \{x \in X \mid \lim_{j\to+\infty} \Im(f_j)(x) \text{ exists in } \mathbf{R}\}$. By part (a) applied to the sequences $(\Re(f_j)), (\Im(f_j))$ of real valued functions, we see that the two functions $\lim_{j\to+\infty} \Re(f_j), \lim_{j\to+\infty} \Im(f_j)$ are both measurable on the set on which each of them exists. Hence, both B_1, B_2 (the inverse images of $\mathbf{R}$ under these functions) belong to Σ and thus $B = B_1 \cap B_2 \in \Sigma$.

Therefore $A = B \cup C \in \Sigma$.

We have just seen that the functions $\lim_{j\to+\infty} \Re(f_j), \lim_{j\to+\infty} \Im(f_j)$ are measurable on the set where each of them exists and hence their restrictions to B are both $\Sigma\rceil B$-measurable. These functions are, respectively, the real and the imaginary part of the restriction to B of $\lim_{j\to+\infty} f_j$ and Proposition 6.6 says that $\lim_{j\to+\infty} f_j$ is $\Sigma\rceil B$-measurable. Finally, the restriction to C of this limit is constant ∞ and thus it is $\Sigma\rceil C$-measurable. By Proposition 6.3, $\lim_{j\to+\infty} f_j$ is $\Sigma\rceil A$-measurable.

This is the proof of (i) in the case of complex valued functions and the proof of (ii) is immediate after Proposition 6.3.

(c) Finally, let $f_j : X \to \overline{\mathbf{C}}$ for all j.

For each j we consider the function

$$g_j(x) = \begin{cases} f_j(x), & \text{if } f_j(x) \neq \infty, \\ j, & \text{if } f_j(x) = \infty. \end{cases}$$

If we set $A_j = f_j^{-1}(\mathbf{C}) \in \Sigma$, then $g_j\rceil A_j = f_j\rceil A_j$ is $\Sigma\rceil A_j$-measurable. Also $g_j\rceil A_j^c$ is constant j and hence $\Sigma\rceil A_j^c$-measurable. Therefore $g_j : X \to \mathbf{C}$ is Σ-measurable.

It is easy to show that the two limits $\lim_{j\to+\infty} g_j(x)$ and $\lim_{j\to+\infty} f_j(x)$ either both exist or both do not exist and, if they do exist, they are equal. In fact, let $\lim_{j\to+\infty} f_j(x) = p \in \overline{\mathbf{C}}$. If $p \in \mathbf{C}$, then for large enough j we shall have that $f_j(x) \neq \infty$, implying $g_j(x) = f_j(x)$ and thus $\lim_{j\to+\infty} g_j(x) = p$. If $p = \infty$, then $|f_j(x)| \to +\infty$. Therefore $|g_j(x)| \geq \min\{|f_j(x)|, j\} \to +\infty$ and hence $\lim_{j\to+\infty} g_j(x) = \infty = p$ in this case also. The converse is similarly proved. If $\lim_{j\to+\infty} g_j(x) = p \in \mathbf{C}$, then, for large enough j, $g_j(x) \neq j$ and thus $f_j(x) = g_j(x)$ implying $\lim_{j\to+\infty} f_j(x) = \lim_{j\to+\infty} g_j(x) = p$. If $\lim_{j\to+\infty} g_j(x) = \infty$, then $\lim_{j\to+\infty} |g_j(x)| = +\infty$. Since $|f_j(x)| \geq |g_j(x)|$ we get $\lim_{j\to+\infty} |f_j(x)| = +\infty$ and thus $\lim_{j\to+\infty} f_j(x) = \infty$.

Therefore $A = \{x \in X \mid \lim_{j\to+\infty} g_j(x) \text{ exists in } \overline{\mathbf{C}}\}$ and, applying the result of (b) to the functions $g_j : X \to \mathbf{C}$, we get that $A \in \Sigma$. For the same reason, the function $\lim_{j\to+\infty} f_j$, which on A is equal to $\lim_{j\to+\infty} g_j$, is $\Sigma\rceil A$-measurable.

6.10 Simple functions.

Definition 6.7 *Let $E \subseteq X$. The function $\chi_E : X \to \mathbf{R}$ defined by*

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

is called ***the characteristic function of*** E.

Observe that, not only E determines its χ_E, but also χ_E determines the set E by $E = \{x \in X \mid \chi_E(x) = 1\} = \chi_E^{-1}(\{1\})$.

The following are trivial:

$$\lambda\chi_E + \kappa\chi_F = \lambda\chi_{E\setminus F} + (\lambda+\kappa)\chi_{E\cap F} + \kappa\chi_{F\setminus E} \qquad \chi_E\chi_F = \chi_{E\cap F} \qquad \chi_{E^c} = 1 - \chi_E$$

for all $E, F \subseteq X$ and all $\lambda, \kappa \in \mathbf{C}$.

Proposition 6.21 *Let (X, Σ) be a measurable space and $E \subseteq X$. Then χ_E is Σ-measurable if and only if $E \in \Sigma$.*

Proof: If χ_E is Σ-measurable, then $E = \chi_E^{-1}(\{1\}) \in \Sigma$.

Conversely, let $E \in \Sigma$. Then for an arbitrary $F \in \mathcal{B}_{\mathbf{R}}$ or $\mathcal{B}_{\mathbf{C}}$ we have $\chi_E^{-1}(F) = \emptyset$ if $0, 1 \notin F$, $\chi_E^{-1}(F) = E$ if $1 \in F$ and $0 \notin F$, $\chi_E^{-1}(F) = E^c$ if $1 \notin F$ and $0 \in F$ and $\chi_E^{-1}(F) = X$ if $0, 1 \in F$. In any case, $\chi_E^{-1}(F) \in \Sigma$ and χ_E is Σ-measurable.

Definition 6.8 *A function defined on a non-empty set X is called **a simple function on** X if its range is a finite subset of $\mathbf{C}$.*

The following proposition completely describes the structure of simple functions.

Proposition 6.22 *(i) A function $\phi : X \to \mathbf{C}$ 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 $m \in \mathbf{N}$, different $\kappa_1, \ldots, \kappa_m \in \mathbf{C}$ and non-empty pairwise disjoint $E_1, \ldots, E_m \subseteq X$ with $\cup_{j=1}^m E_j = X$ 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 Σ is a σ-algebra of subsets of X, then ϕ is Σ-measurable if and only if all E_k's in the representation of ϕ described in (ii) belong to Σ.

Proof: Let $\phi = \sum_{j=1}^{n} \lambda_j \chi_{F_j}$, where $\lambda_j \in \mathbf{C}$ and $F_j \subseteq X$ for all $j = 1, \dots, n$. Taking an arbitrary $x \in X$, 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 $\phi(x) = \lambda_{j_1} + \cdots + \lambda_{j_k}$. Therefore the range of ϕ contains at most all the possible sums $\lambda_{j_1} + \cdots + \lambda_{j_k}$ together with 0 and hence it is finite. Thus ϕ is simple on X.

Conversely, suppose ϕ is simple on X and let its range consist of the different $\kappa_1, \dots, \kappa_m \in \mathbf{C}$. We consider $E_j = \{x \in X \mid \phi(x) = \kappa_j\} = \phi^{-1}(\{\kappa_j\})$. Then every $x \in X$ belongs to exactly one of these sets, so that they are pairwise disjoint and $X = E_1 \cup \cdots \cup E_m$. Now it is clear that $\phi = \sum_{j=1}^{m} \kappa_j \chi_{E_j}$, because both sides take the same value at every x.

If $\phi = \sum_{i=1}^{m'} \kappa'_i \chi_{E'_i}$ is another representation of ϕ with different κ'_i's and non-empty pairwise disjoint E'_i's 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 $j = 1, \dots, m$. We conclude that the representation is unique.

Now if all E_j's belong to the σ-algebra Σ, then, by Proposition 6.21, all χ_{E_j}'s are Σ-measurable and hence ϕ is also Σ-measurable. Conversely, if ϕ is Σ-measurable, then all $E_j = \phi^{-1}(\{\kappa_j\})$ belong to Σ.

Definition 6.9 *The unique representation of the simple function ϕ, which is described in part (ii) of Proposition 6.22, 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 6.23 *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 simple functions.*

Proof: Let ϕ, ψ be simple functions on X and $p, q \in \mathbf{C}$. Assume that $\lambda_1, \dots, \lambda_n$ are the values of ϕ and $\kappa_1, \dots, \kappa_m$ 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 $n+m$ numbers λ_i, κ_j.

Theorem 6.1 *(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, it converges to f uniformly on every subset on which f is bounded.*

(ii) Given $f : X \to \overline{\mathbf{C}}$, there is a sequence (ϕ_n) of 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 Σ is a σ-algebra of subsets of X and f is Σ-measurable, then the ϕ_n in (i) and (ii) can be taken to be Σ-measurable.

Proof: (i) For every $n, k \in \mathbf{N}$ with $1 \leq k \leq 2^{2n}$, we define the sets

$$E_n^{(k)} = f^{-1}\Big(\Big(\frac{k-1}{2^n}, \frac{k}{2^n}\Big]\Big), \qquad F_n = f^{-1}((2^n, +\infty])$$

and the simple function

$$\phi_n = \sum_{k=1}^{2^{2n}} \frac{k-1}{2^n} \chi_{E_n^{(k)}} + 2^n \chi_{F_n}.$$

For each n the sets $E_n^{(1)}, \ldots, E_n^{(2^{2n})}, F_n$ are pairwise disjoint and their union is the set $f^{-1}((0, +\infty])$, while their complementary set is $G = f^{-1}(\{0\})$. Observe that if f is Σ-measurable then all $E_n^{(k)}$ and F_n belong to Σ and hence ϕ_n is Σ-measurable.

In G we have $0 = \phi_n = f$, in each $E_n^{(k)}$ we have $\phi_n = \frac{k-1}{2^n} < f \leq \frac{k}{2^n} = \phi_n + \frac{1}{2^n}$ and in F_n we have $\phi_n = 2^n < f$.

Now, if $f(x) = +\infty$, then $x \in F_n$ for every n and hence $\phi_n(x) = 2^n \to +\infty = f(x)$. If $0 \leq f(x) < +\infty$, then for all large n we have $0 \leq f(x) \leq 2^n$ and hence $0 \leq f(x) - \phi_n(x) \leq \frac{1}{2^n}$, which implies that $\phi_n(x) \to f(x)$. Therefore, $\phi_n \to f$ pointwise on X.

If $K \subseteq X$ and f is bounded on K, then there is an n_0 so that $f(x) \leq 2^{n_0}$ for all $x \in K$. Hence for all $n \geq n_0$ we have $0 \leq f(x) - \phi_n(x) \leq \frac{1}{2^n}$ for all $x \in K$. This says that $\phi_n \to f$ uniformly on K.

It remains to prove that (ϕ_n) is increasing. If $x \in G$, then $\phi_n(x) = \phi_{n+1}(x) = f(x) = 0$. Now observe the relations

$$E_{n+1}^{(2k-1)} \cup E_{n+1}^{(2k)} = E_n^{(k)}, \qquad 1 \leq k \leq 2^{2n},$$

and

$$(\cup_{l=2^{2n+1}+1}^{2^{2(n+1)}} E_{n+1}^{(l)}) \cup F_{n+1} = F_n.$$

The first relation implies that, if $x \in E_n^{(k)}$ then $\phi_n(x) = \frac{k-1}{2^n}$ and $\phi_{n+1}(x) = \frac{(2k-1)-1}{2^{n+1}}$ or $\frac{2k-1}{2^{n+1}}$. Therefore, if $x \in E_n^{(k)}$, then $\phi_n(x) \leq \phi_{n+1}(x)$.
The second relation implies that, if $x \in F_n$, then $\phi_n(x) = 2^n$ and $\phi_{n+1}(x) = \frac{(2^{2n+1}+1)-1}{2^{n+1}}$ or $\ldots$ or $\frac{2^{2(n+1)}-1}{2^{n+1}}$ or 2^{n+1}. Hence, if $x \in F_n$, then $\phi_n(x) \leq \phi_{n+1}(x)$.
(ii) Let $A = f^{-1}(\mathbf{C})$, whence $f = \infty$ on A^c. Consider the restriction $f \rceil A : A \to \mathbf{C}$ and the functions

$$(\Re(f\rceil A))^+, (\Re(f\rceil A))^-, (\Im(f\rceil A))^+, (\Im(f\rceil A))^- : A \to [0, +\infty).$$

If f is Σ-measurable, then $A \in \Sigma$ and these four functions are $\Sigma\rceil A$-measurable.

By the result of part (i) there are increasing sequences (p_n), (q_n), (r_n) and (s_n) of non-negative (real valued) simple functions on A so that each converges to, respectively, $(\Re(f\rceil A))^+$, $(\Re(f\rceil A))^-$, $(\Im(f\rceil A))^+$ and $(\Im(f\rceil A))^-$ pointwise on A and uniformly on every subset of A on which $f\rceil A$ is bounded (because on such a subset all four functions are also bounded). Now it is obvious that, if we set $\phi_n = (p_n - q_n) + i(r_n - s_n)$, then ϕ_n is a simple function on A which is $\Sigma\rceil A$-measurable if f is Σ-measurable. It is clear that $\phi_n \to f\rceil A$ pointwise on A and uniformly on every subset of A on which $f\rceil A$ is bounded.

Also $|\phi_n| = \sqrt{(p_n - q_n)^2 + (r_n - s_n)^2} = \sqrt{p_n^2 + q_n^2 + r_n^2 + s_n^2}$ and thus the sequence $(|\phi_n|)$ is increasing on A.

If we define ϕ_n as the constant n on A^c, then the proof is complete.

6.11 The role of null sets.

Definition 6.10 *Let (X, Σ, μ) 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 $\{x \in X \,|\, P(x) \text{ is not true}\}$ is included in a (μ-)null set.*

We also use the short expressions: $P(x)$ *holds (μ-)a.e. on* X and $P(x)$ *holds for (μ-)a.e.* $x \in X$.

It is obvious that if $P(x)$ holds for a.e. $x \in X$ and μ is *complete* then the set $\{x \in X \,|\, P(x) \text{ is not true}\}$ is contained in Σ and hence its complement $\{x \in X \,|\, P(x) \text{ is true}\}$ is also in Σ.

Proposition 6.24 *Let (X, Σ, μ) be a measure space and $(X, \overline{\Sigma}, \overline{\mu})$ be its completion. Let (Y, Σ') be a measurable space and $f : X \to Y$ be (Σ, Σ')-measurable. If $g : X \to Y$ is equal to f a.e on X, then g is $(\overline{\Sigma}, \Sigma')$-measurable.*

Proof: There exists $N \in \Sigma$ so that $\{x \in X \,|\, f(x) \neq g(x)\} \subseteq N$ and $\mu(N) = 0$.

Take an arbitrary $E \in \Sigma'$ and write $g^{-1}(E) = \{x \in X \,|\, g(x) \in E\} = \{x \in N^c \,|\, g(x) \in E\} \cup \{x \in N \,|\, g(x) \in E\} = \{x \in N^c \,|\, f(x) \in E\} \cup \{x \in N \,|\, g(x) \in E\}$.

The first set is $= N^c \cap f^{-1}(E)$ and belongs to Σ and the second set is $\subseteq N$. By the definiton of the completion we get that $g^{-1}(E) \in \overline{\Sigma}$ and hence g is $(\overline{\Sigma}, \Sigma')$-measurable.

In the particular case of a *complete* measure space (X, Σ, μ) we have the rule: *if f is measurable on X and g is equal to f a.e. on X, then g is also measurable on X.*

Proposition 6.25 *Let (X, Σ, μ) be a measure space and $(X, \overline{\Sigma}, \overline{\mu})$ be its completion. Let (f_j) be a sequence of Σ-measurable functions $f_j : X \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$. If $g : X \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$ is such that $g(x) = \lim_{j \to +\infty} f_j(x)$ for a.e. $x \in X$, then g is $\overline{\Sigma}$-measurable.*

Proof: $\{x \in X \,|\, \lim_{j \to +\infty} f_j(x) \text{ does not exist or is } \neq g(x)\} \subseteq N$ for some $N \in \Sigma$ with $\mu(N) = 0$.

N^c belongs to Σ and the restrictions $f_j\rceil N^c$ are all $\Sigma\rceil N^c$-measurable. By Proposition 6.20, the restriction $g\rceil N^c = \lim_{j\to+\infty} f_j\rceil N^c$ is $\Sigma\rceil N^c$-measurable. This, of course, means that for every $E \in \Sigma'$ we have $\{x \in N^c \,|\, g(x) \in E\} \in \Sigma$.

Now we write $g^{-1}(E) = \{x \in N^c \,|\, g(x) \in E\} \cup \{x \in N \,|\, g(x) \in E\}$. The first set belongs to Σ and the second is $\subseteq N$. Therefore $g^{-1}(E) \in \overline{\Sigma}$ and g is $\overline{\Sigma}$-measurable.

Again, in the particular case of a *complete* measure space (X, Σ, μ) the rule is: *if (f_j) is a sequence of measurable functions on X and its limit is equal to g a.e. on X, then g is also measurable on X.*

Proposition 6.26 *Let (X, Σ, μ) be a measure space and $(X, \overline{\Sigma}, \overline{\mu})$ be its completion. Let (Y, Σ') be a measurable space and $f : A \to Y$ be $(\Sigma\rceil A, \Sigma')$-measurable, where $A \in \Sigma$ with $\mu(A^c) = 0$. If we extend f to X in an arbitrary manner, then the extended function is $(\overline{\Sigma}, \Sigma')$-measurable.*

Proof: Let $h : A^c \to Y$ be an arbitrary function and let

$$F(x) = \begin{cases} f(x), & \text{if } x \in A, \\ h(x), & \text{if } x \in A^c. \end{cases}$$

Take an arbitrary $E \in \Sigma'$ and 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 $\Sigma\rceil A$ and hence to Σ, while the second set is $\subseteq A^c$. Therefore $F^{-1}(E) \in \overline{\Sigma}$ and F is $(\overline{\Sigma}, \Sigma')$-measurable.

If (X, Σ, μ) is a *complete* measure space, the rule is: *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.*

6.12 Exercises.

1. Let (X, Σ) be a measurable space and $f : X \to \overline{\mathbf{R}}$. Prove that f is measurable if $f^{-1}((a, +\infty]) \in \Sigma$ for all *rational* $a \in \mathbf{R}$.

2. Let $f : X \to \overline{\mathbf{R}}$. If $g, h : X \to \overline{\mathbf{R}}$ are such that $g, h \geq 0$ and $f = g - h$ on X, prove that $f^+ \leq g$ and $f^- \leq h$ on X.

3. Let (X, Σ) be a measurable space and $f : X \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$ be measurable. We agree that $0^p = +\infty, (+\infty)^p = 0$ if $p < 0$ and $0^0 = (+\infty)^0 = 1$. Prove that, for all $p \in \mathbf{R}$, the function $|f|^p$ is measurable.

4. Prove that every monotone $f : \mathbf{R} \to \mathbf{R}$ is Borel measurable.

5. *Translates and dilates of functions.*

   Let $f : \mathbf{R}^n \to Y$ and take arbitrary $y \in \mathbf{R}^n$ and $\lambda \in (0, +\infty)$. We define $g, h : \mathbf{R}^n \to Y$ by

   $$g(x) = f(x - y), \qquad h(x) = f\Big(\frac{x}{\lambda}\Big)$$

for all $x \in \mathbf{R}^n$. g is called **the translate of f by y** and h is called **the dilate of f by λ**.

Let (Y, Σ') be a measurable space. Prove that, if f is $(\mathcal{L}_n, \Sigma')$-measurable, then the same is true for g and h.

6. *Functions with prescribed level sets.*

   Let (X, Σ) be a measurable space and assume that the collection $\{E_\lambda\}_{\lambda\in\mathbf{R}}$ of subsets of X, which belong to Σ, has the properties:
   (i) $E_\lambda \subseteq E_\kappa$ for all λ, κ with $\lambda \leq \kappa$,
   (ii) $\cup_{\lambda\in\mathbf{R}} E_\lambda = X$, $\cap_{\lambda\in\mathbf{R}} E_\lambda = \emptyset$,
   (iii) $\cap_{\kappa,\kappa>\lambda} E_\kappa = E_\lambda$ for all $\lambda \in \mathbf{R}$.

   Consider the function $f : X \to \mathbf{R}$ defined by $f(x) = \inf\{\lambda \in \mathbf{R} \,|\, x \in E_\lambda\}$. Prove that f is measurable and that $E_\lambda = \{x \in X \,|\, f(x) \leq \lambda\}$ for every $\lambda \in \mathbf{R}$.

   How will the result change if we drop any of the assumptions in (ii) and (iii)?

7. *Not all functions are Lebesgue measurable and not all Lebesgue measurable functions are Borel measurable.*

   (i) Prove that a Borel measurable $g : \mathbf{R} \to \mathbf{R}$ is also Lebesgue measurable.
   (ii) Find a function $f : \mathbf{R} \to \mathbf{R}$ which is not Lebesgue measurable.
   (iii) Using exercise 4.6.15, find a function $g : \mathbf{R} \to \mathbf{R}$ which is Lebesgue measurable but not Borel measurable.

8. Give an example of a non-Lebesgue measurable $f : \mathbf{R} \to \mathbf{R}$ so that $|f|$ is Lebesgue measurable.

9. Starting with an appropriate non-Lebesgue measurable function, give an example of an uncountable collection $\{f_i\}_{i\in I}$ of Lebesgue measurable functions $f_i : \mathbf{R} \to \mathbf{R}$ so that $\sup_{i\in I} f_i$ is non-Lebesgue measurable.

10. (i) Prove that, if $G : \mathbf{R} \to \mathbf{R}$ is continuous and $H : \mathbf{R} \to \mathbf{R}$ is Borel measurable, then $H \circ G : \mathbf{R} \to \mathbf{R}$ is Borel measurable.
    (ii) Using exercise 4.6.15, construct a continuous $G : \mathbf{R} \to \mathbf{R}$ and a Lebesgue measurable $H : \mathbf{R} \to \mathbf{R}$ so that $H \circ G : \mathbf{R} \to \mathbf{R}$ is not Lebesgue measurable.

11. Let (X, Σ, μ) be a measure space and $f : X \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$ be measurable. Assume that $\mu(\{x \in X \,|\, |f(x)| = +\infty\}) = 0$ and that there is $M < +\infty$ so that $\mu(\{x \in X \,|\, |f(x)| > M\}) < +\infty$.

    Prove that for every $\epsilon > 0$ there is a *bounded* measurable $g : X \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$ so that $\mu(\{x \in X \,|\, g(x) \neq f(x)\}) < \epsilon$. You may try a suitable truncation of f.

12. We say that $\phi : X \to \mathbf{C}$ 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 elementary functions on X so that $\phi_n \to f$ uniformly on X. If Σ is a σ-algebra of subsets of X and f is measurable, prove that the ϕ_n's can be taken measurable.

13. *We can add, multiply and take limits of equalities holding almost everywhere.*

    Let (X, Σ, μ) 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, then $f = h$ a.e. on X.
    (ii) Let $f_1, f_2, g_1, g_2 : X \to \mathbf{R}$. If $f_1 = f_2$ a.e. on X and $g_1 = g_2$ a.e. on X, then $f_1 + g_1 = f_2 + g_2$ and $f_1 g_1 = f_2 g_2$ a.e. on X.
    (iii) Let $f_j, g_j : X \to \overline{\mathbf{R}}$ so that $f_j = g_j$ a.e. on X for all $j \in \mathbf{N}$. Then $\sup_{j \in \mathbf{N}} f_j = \sup_{j \in \mathbf{N}} g_j$ a.e. on X. Similar results hold for $\inf, \limsup$ and $\liminf$.
    (iv) Let $f_j, g_j : X \to \overline{\mathbf{R}}$ so that $f_j = g_j$ a.e. on X for all $j \in \mathbf{N}$. If $A = \{x \in X \mid \lim_{j \to +\infty} f_j(x) \text{ exists}\}$ and $B = \{x \in X \mid \lim_{j \to +\infty} g_j(x) \text{ exists}\}$, then $A \triangle B \subseteq N$ for some $N \in \Sigma$ with $\mu(N) = 0$ and $\lim_{j \to +\infty} f_j = \lim_{j \to +\infty} g_j$ 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 $(A \cap B)^c$, then $\lim_{j \to +\infty} f_j = \lim_{j \to +\infty} g_j$ a.e. on X.

14. Let (X, Σ, μ) be a measure space and $(X, \overline{\Sigma}, \overline{\mu})$ be its completion.
    (i) If $E \in \overline{\Sigma}$, then there is $A \in \Sigma$ so that $\chi_E = \chi_A$ a.e. on X.
    (ii) If $\phi : X \to \mathbf{C}$ is a $\overline{\Sigma}$-measurable simple function, then there is a Σ-measurable simple function $\psi : X \to \mathbf{C}$ so that $\phi = \psi$ a.e. on X.
    (iii) Use Theorem 6.1 to prove that, if $g : X \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$ is $\overline{\Sigma}$-measurable, then there is a Σ-measurable $f : X \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$ so that $g = f$ a.e. on X.

15. Let X, Y be topological spaces of which Y is Hausdorff. This means that, if $y_1, y_2 \in Y$ and $y_1 \neq y_2$, then there are *disjoint* open neighborhoods V_{y_1}, V_{y_2} of y_1, y_2, respectively. Assume that μ is 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.

16. *The support of a function.*

    (a) Let X be a topological space and a continuous $f : X \to \mathbf{C}$. The set $\text{supp}(f) = \overline{f^{-1}(\mathbf{C} \setminus \{0\})}$ is called **the support of** f. Prove that $\text{supp}(f)$ is the smallest closed subset of X outside of which $f = 0$.
    (b) Let μ be a regular Borel measure on the topological space X and $f : X \to \mathbf{C}$ be a Borel measurable function. 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 in X.

(ii) Prove that $\mu(\{x \in K \,|\, f(x) \neq 0\}) = 0$ for all compact sets $K \subseteq (\text{supp}(f))^c$.
(iii) Using the regularity of μ, prove that $f = 0$ a.e on $(\text{supp}(f))^c$.
(iv) Prove that $(\text{supp}(f))^c$ is the largest open set in X on which $f = 0$ a.e.
(c) Assume that the μ appearing in (b) has the additional property that $\mu(U) > 0$ for every open $U \subseteq X$. Use exercise 6.12.15 to prove that for any continuous $f : X \to \mathbf{C}$ the two definitions of $\text{supp}(f)$ (the one in (a) and the one in (b)) coincide.

17. *The Theorem of Lusin.*

We shall prove that every Lebesgue measurable function which is finite a.e. on $\mathbf{R}^n$ is equal to a continuous function except on a set of arbitrarily small Lebesgue measure.
(i) For each $a < a+\delta < b-\delta < b$ we consider the function $\tau_{a,b,\delta} : \mathbf{R} \to \mathbf{R}$ which: is 0 outside (a, b), is 1 on $[a+\delta, b-\delta]$ and is linear on $[a, a+\delta]$ and on $[b-\delta, b]$ so that it is continuous on $\mathbf{R}$. Now, let $R = (a_1, b_1) \times \cdots \times (a_n, b_n)$ and, for small enough $\delta > 0$, we consider the function $\tau_{R,\delta} : \mathbf{R}^n \to \mathbf{R}$ by the formula

$$\tau_{R,\delta}(x_1, \ldots, x_n) = \tau_{a_1,b_1,\delta}(x_1) \cdots \tau_{a_n,b_n,\delta}(x_n).$$

If $R_\delta = (a_1 + \delta, b_1 - \delta) \times \cdots \times (a_n + \delta, b_n - \delta)$, prove that $\tau_{R,\delta} = 1$ on $\overline{R_\delta}$, $\tau_{R,\delta} = 0$ outside R, $0 \leq \tau_{R,\delta} \leq 1$ on $\mathbf{R}^n$ and $\tau_{R,\delta}$ is continuous on $\mathbf{R}^n$. Therefore, prove that for every $\epsilon > 0$ there is $\delta > 0$ so that $m_n(\{x \in \mathbf{R}^n \,|\, \tau_{R,\delta}(x) \neq \chi_R(x)\}) < \epsilon$.
(ii) Let $E \in \mathcal{L}_n$ with $m_n(E) < +\infty$. Use Theorem 4.6 to prove that for every $\epsilon > 0$ there is a continuous $\tau : \mathbf{R}^n \to \mathbf{R}$ so that $0 \leq \tau \leq 1$ on $\mathbf{R}^n$ and $m_n(\{x \in \mathbf{R}^n \,|\, \tau(x) \neq \chi_E(x)\}) < \epsilon$.
(iii) Let ϕ be a non-negative Lebesgue measurable simple function on $\mathbf{R}^n$ which is 0 outside some set of finite Lebesgue measure. Prove that for all $\epsilon > 0$ there is a continuous $\tau : \mathbf{R}^n \to \mathbf{R}$ so that $0 \leq \tau \leq \max_{\mathbf{R}^n} \phi$ on $\mathbf{R}^n$ and $m_n(\{x \in \mathbf{R}^n \,|\, \tau(x) \neq \phi(x)\}) < \epsilon$.
(iv) Let $f : \mathbf{R}^n \to [0, 1]$ be a Lebesgue measurable function which is 0 outside some set of finite Lebesgue measure. Use Theorem 6.1 to prove that $f = \sum_{k=1}^{+\infty} \psi_k$ uniformly on $\mathbf{R}^n$, where all ψ_k are Lebesgue measurable simple functions with $0 \leq \psi_k \leq \frac{1}{2^k}$ on $\mathbf{R}^n$ for all k. Now apply the result of (iii) to each ψ_k and prove that for all $\epsilon > 0$ there is a continuous $g : \mathbf{R}^n \to [0, 1]$ so that $m_n(\{x \in \mathbf{R}^n \,|\, g(x) \neq f(x)\}) < \epsilon$.
(v) Let $f : \mathbf{R}^n \to [0, +\infty]$ be a Lebesgue measurable function which is 0 outside some set of finite Lebesgue measure and finite a.e. on $\mathbf{R}^n$. By taking an appropriate truncation of f prove that for all $\epsilon > 0$ there is a bounded Lebesgue measurable function $h : \mathbf{R}^n \to [0, +\infty]$ which is 0 outside some set of finite Lebesgue measure so that $m_n(\{x \in \mathbf{R}^n \,|\, h(x) \neq f(x)\}) < \epsilon$. Now apply the result of (iv) to find a continuous $g : \mathbf{R}^n \to \mathbf{R}$ so that $m_n(\{x \in \mathbf{R}^n \,|\, g(x) \neq f(x)\}) < \epsilon$.

(vi) Find pairwise disjoint open-closed qubes $P^{(k)}$ so that $\mathbf{R}^n = \cup_{k=1}^{+\infty} P^{(k)}$ and let $R^{(k)}$ be the open qube with the same edges as $P^{(k)}$. Consider for each k a small enough $\delta_k > 0$ so that $m_n(\{x \in \mathbf{R}^n \,|\, \tau_{R^{(k)},\delta_k}(x) \neq \chi_{R^{(k)}}(x)\}) < \frac{\epsilon}{2^{k+1}}$.
(vii) Let $f : \mathbf{R}^n \to [0, +\infty]$ be Lebesgue measurable and finite a.e. on $\mathbf{R}^n$. If $R^{(k)}$ are the qubes from (vi), then each $f\chi_{R^{(k)}} : \mathbf{R}^n \to [0, +\infty]$ is Lebesgue measurable, finite a.e. on $\mathbf{R}^n$ and 0 outside $R^{(k)}$. Apply (v) to find continuous $g_k : \mathbf{R}^n \to \mathbf{R}$ so that $m_n(\{x \in \mathbf{R}^n \,|\, g_k(x) \neq f(x)\chi_{R^{(k)}}(x)\}) < \frac{\epsilon}{2^{k+1}}$.

Prove that $m_n(\{x \in \mathbf{R}^n \,|\, \tau_{R^{(k)},\delta_k}(x) g_k(x) \neq f(x)\chi_{R^{(k)}}(x)\}) < \frac{\epsilon}{2^k}$.

Define $g = \sum_{k=1}^{+\infty} \tau_{R^{(k)},\delta_k} g_k$ and prove that g is continuous on $\mathbf{R}^n$ and that $m_n(\{x \in \mathbf{R}^n \,|\, g(x) \neq f(x)\}) < \epsilon$.
(viii) Extend the result of (vii) to all $f : \mathbf{R}^n \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$ which are Lebesgue measurable and finite a.e. on $\mathbf{R}^n$.

18. Let $f : \mathbf{R}^n \to \mathbf{R}$ be continuous at m_n-a.e. $x \in \mathbf{R}^n$. Prove that f is Lebesgue measurable on $\mathbf{R}^n$.

Chapter 7

Integrals

7.1 Integrals of non-negative simple functions.

In this whole section (X, Σ, μ) will be a fixed measure space.

Definition 7.1 *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*

$$\int_X \phi \, d\mu = \sum_{k=1}^{m} \kappa_k \mu(E_k)$$

and call it ***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 then we write $\int_X \phi(x) \, d\mu(x)$.

From now on, if it is obvious which measure space (X, Σ, μ) we are talking about, we shall simply say *integral*, instead of μ-integral.

We can make the following observations.
(i) 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. In other words, *the set where $\phi = 0$ does not matter for the calculation of the integral of ϕ.*
(ii) We also see that $\int_X \phi \, d\mu < +\infty$ if and only if $\mu(E_k) < +\infty$ for all k for which $\kappa_k > 0$. Taking the union of all these E_k's we see that $\int_X \phi \, d\mu < +\infty$ if and only if $\mu(\{x \in X \mid \phi(x) > 0\}) < +\infty$. In other words, *$\phi$ has a finite integral if and only if $\phi = 0$ outside a set of finite measure.*
(iii) Moreover, $\int_X \phi \, d\mu = 0$ if and only if $\mu(E_k) = 0$ for all k for which $\kappa_k > 0$. Taking, as before, the union of these E_k's we see that $\int_X \phi \, d\mu = 0$ if and only if $\mu(\{x \in X \mid \phi(x) > 0\}) = 0$. In other words, *$\phi$ has vanishing integral if and only if $\phi = 0$ outside a null set.*

Lemma 7.1 *Let $\phi = \sum_{j=1}^n \lambda_j \chi_{F_j}$, where $0 \leq \lambda_j < +\infty$ for all j and the sets $F_j \in \Sigma$ are pairwise disjoint. Then $\int_X \phi \, d\mu = \sum_{j=1}^n \lambda_j \mu(F_j)$.*

The representation $\phi = \sum_{j=1}^n \lambda_j \chi_{F_j}$ in the statement may not be the standard representation of ϕ. In fact, the λ_j's are not assumed different and it is not assumed either that the F_j's are non-empty or that they cover X.
Proof: (a) In case all F_j's are empty, then their characteristic functions are 0 on X and we get $\phi = 0 = 0 \cdot \chi_X$ as the standard representation of ϕ. Therefore $\int_X \phi \, d\mu = 0 \cdot \mu(X) = 0 = \sum_{j=1}^n \lambda_j \mu(F_j)$, since all measures are 0. In this particular case the result of the lemma is proved.
(b) In case some, but not all, of the F_j's are empty, we rearrange so that $F_1, \dots, F_l \neq \emptyset$ and $F_{l+1}, \dots, F_n = \emptyset$. (We include the case $l = n$.) Then we have $\phi = \sum_{j=1}^l \lambda_j \chi_{F_j}$, where all F_j's are non-empty, and the equality to be proved becomes $\int_X \phi \, d\mu = \sum_{j=1}^l \lambda_j \mu(F_j)$.

In case the F_j's 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 $\phi = \sum_{j=1}^{l+1} \lambda_j \chi_{F_j}$ for the assumed equality and $\int_X \phi \, d\mu = \sum_{j=1}^{l+1} \lambda_j \mu(F_j)$ for the one to be proved.

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 $F_j \in \Sigma$ 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 $\lambda_1, \dots, \lambda_k$ are *all* the values of ϕ on X, perhaps with repetitions. We rearrange in groups, so that

$$\begin{aligned}
&\lambda_1 = \cdots = \lambda_{k_1} = \kappa_1, \\
&\lambda_{k_1+1} = \cdots = \lambda_{k_1+k_2} = \kappa_2, \\
&\dots \\
&\lambda_{k_1+\cdots+k_{m-1}+1} = \cdots = \lambda_{k_1+\cdots+k_m} = \kappa_m
\end{aligned}$$

are the *different* values of ϕ on X (and, of course, $k_1 + \cdots + k_m = k$). For every $i = 1, \dots, m$ we define $E_i = \bigcup_{j=k_1+\cdots+k_{i-1}+1}^{k_1+\cdots+k_i} F_j = \{x \in X \,|\, \phi(x) = \kappa_i\}$, and then

$$\phi = \sum_{i=1}^m \kappa_i \chi_{E_i}$$

is the standard representation of ϕ. By definition

$$\begin{aligned}
\int_X \phi \, d\mu &= \sum_{i=1}^m \kappa_i \mu(E_i) = \sum_{i=1}^m \kappa_i \sum_{j=k_1+\cdots+k_{i-1}+1}^{k_1+\cdots+k_i} \mu(F_j) \\
&= \sum_{i=1}^m \sum_{j=k_1+\cdots+k_{i-1}+1}^{k_1+\cdots+k_i} \lambda_j \mu(F_j) = \sum_{j=1}^k \lambda_j \mu(F_j).
\end{aligned}$$

Lemma 7.2 *If ϕ, ψ are non-negative 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 $\int_X \lambda\phi \, d\mu = \lambda \int_X \phi \, d\mu$.*

Proof: (a) If $\lambda = 0$, then $\lambda\phi = 0 = 0 \cdot \chi_X$ is the standard representation of $\lambda\phi$ and hence $\int_X \lambda\phi \, d\mu = 0 \cdot \mu(X) = 0 = \lambda \int_X \phi \, d\mu$.

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 $\int_X \lambda\phi \, d\mu = \sum_{j=1}^m \lambda\kappa_j \mu(E_j) = \lambda \sum_{j=1}^m \kappa_j \mu(E_j) = \lambda \int_X \phi \, d\mu$.
(b) 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 $X = \cup_{1\leq j\leq m, 1\leq i\leq n}(E_j \cap F_i)$ and that the sets $E_j \cap F_i \in \Sigma$ are pairwise disjoint. It is also clear that $\phi + \psi$ is constant $\kappa_j + \lambda_i$ on each $E_j \cap F_i$ and thus

$$\phi + \psi = \sum_{1\leq j\leq m, 1\leq i\leq n} (\kappa_j + \lambda_i)\chi_{E_j \cap F_i}.$$

Lemma 7.1 implies that

$$\begin{aligned}
\int_X (\phi + \psi) \, d\mu &= \sum_{1\leq j\leq m, 1\leq i\leq n} (\kappa_j + \lambda_i)\mu(E_j \cap F_i) \\
&= \sum_{1\leq j\leq m, 1\leq i\leq n} \kappa_j \mu(E_j \cap F_i) + \sum_{1\leq j\leq m, 1\leq i\leq n} \lambda_i \mu(E_j \cap F_i) \\
&= \sum_{j=1}^m \kappa_j \sum_{i=1}^n \mu(E_j \cap F_i) + \sum_{i=1}^n \lambda_i \sum_{j=1}^m \mu(E_j \cap F_i) \\
&= \sum_{j=1}^m \kappa_j \mu(E_j) + \sum_{i=1}^n \lambda_i \mu(F_i) = \int_X \phi \, d\mu + \int_X \psi \, d\mu.
\end{aligned}$$

Lemma 7.3 *If ϕ, ψ are non-negative measurable simple functions so that $\phi \leq \psi$ on X, then $\int_X \phi \, d\mu \leq \int_X \psi \, d\mu$.*

Proof: 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 $E_j \cap F_i \neq \emptyset$, we take any $x \in E_j \cap F_i$ and find $\kappa_j = \phi(x) \leq \psi(x) = \lambda_i$. Therefore, since in the calculation below only the non-empty intersections really matter,

$$\begin{aligned}
\int_X \phi \, d\mu &= \sum_{j=1}^m \kappa_j \mu(E_j) = \sum_{1\leq j\leq m, 1\leq i\leq n} \kappa_j \mu(E_j \cap F_i) \\
&\leq \sum_{1\leq j\leq m, 1\leq i\leq n} \lambda_i \mu(E_j \cap F_i) = \sum_{i=1}^n \lambda_i \mu(F_i) = \int_X \psi \, d\mu.
\end{aligned}$$

Lemma 7.4 *Let ϕ be non-negative measurable simple function and (A_n) an increasing sequence in Σ with $\cup_{n=1}^{+\infty} A_n = X$. Then $\int_X \phi\chi_{A_n} \, d\mu \to \int_X \phi \, d\mu$.*

Proof: Let $\phi = \sum_{j=1}^m \kappa_j \chi_{E_j}$ be the standard representation of ϕ. Then $\phi\chi_{A_n} = \sum_{j=1}^m \kappa_j \chi_{E_j}\chi_{A_n} = \sum_{j=1}^m \kappa_j \chi_{E_j \cap A_n}$. Lemma 7.1 implies that $\int_X \phi\chi_{A_n} \, d\mu = \sum_{j=1}^m \kappa_j \mu(E_j \cap A_n)$.

For each j we see that $\mu(E_j \cap A_n) \to \mu(E_j)$ by the continuity of μ from below. Therefore $\int_X \phi\chi_{A_n} \, d\mu \to \sum_{j=1}^m \kappa_j \mu(E_j) = \int_X \phi \, d\mu$.

Lemma 7.5 *Let $\phi, \phi_1, \phi_2, \dots$ be non-negative measurable simple functions so that $\phi_n \le \phi_{n+1}$ on X for all n.*
(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: Lemma 7.3 implies that $\int_X \phi_n \, d\mu \le \int_X \phi_{n+1} \, d\mu$ for all n and hence the limit $\lim_{n\to+\infty} \int_X \phi_n \, d\mu$ exists in $[0, +\infty]$.
(i) Since, by Lemma 7.3, $\int_X \phi_n \, d\mu \le \int_X \phi \, d\mu$, we get $\lim_{n\to+\infty} \int_X \phi_n \, d\mu \le \int_X \phi \, d\mu$.
(ii) Consider arbitrary $\alpha \in [0, 1)$ and define $A_n = \{x \in X \mid \alpha\phi(x) \le \phi_n(x)\} \in \Sigma$. It is easy to see that (A_n) is increasing and that $\cup_{n=1}^{+\infty} A_n = X$. Indeed, if there is any $x \notin \cup_{n=1}^{+\infty} A_n$, then $\phi_n(x) < \alpha\phi(x)$ for all n, implying that $0 < \phi(x) \le \alpha\phi(x)$ which cannot be true.

Now we have that $\alpha\phi\chi_{A_n} \le \phi_n$ on X. Lemmas 7.2, 7.3 and 7.4 imply that

$$\begin{aligned} \alpha \int_X \phi \, d\mu &= \int_X \alpha\phi \, d\mu \\ &= \lim_{n\to+\infty} \int_X \alpha\phi\chi_{A_n} \, d\mu \le \lim_{n\to+\infty} \int_X \phi_n \, d\mu. \end{aligned}$$

We now take the limit as $\alpha \to 1-$ and get $\int_X \phi \, d\mu \le \lim_{n\to+\infty} \int_X \phi_n \, d\mu$.

Lemma 7.6 *If (ϕ_n) and (ψ_n) are two increasing sequences of non-negative measurable simple functions and if $\lim_{n\to+\infty} \phi_n = \lim_{n\to+\infty} \psi_n$ holds 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 every k we have that $\psi_k \le \lim_{n\to+\infty} \phi_n$ on X. Lemma 7.5 implies that $\int_X \psi_k \, d\mu \le \lim_{n\to+\infty} \int_X \phi_n \, d\mu$. Taking the limit in k, we find that $\lim_{n\to+\infty} \int_X \psi_n \, d\mu \le \lim_{n\to+\infty} \int_X \phi_n \, d\mu$.

The opposite inequality is proved symmetrically.

7.2 Integrals of non-negative functions.

Again in this section, (X, Σ, μ) will be a fixed measure space.

Definition 7.2 *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 may use the symbol $\int_X f(x) \, d\mu(x)$ if we want to see the independent variable in the integral.

Lemma 7.6 guarantees that $\int_X f \, d\mu$ is well defined and Theorem 6.1 implies the existence of at least one (ϕ_n) as in the definition.

Proposition 7.1 *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 $\int_X \lambda f\, d\mu = \lambda \int_X f\, d\mu$.*

Proof: We take increasing sequences (ϕ_n), (ψ_n) of non-negative measurable simple functions on X with $\lim_{n\to+\infty} \phi_n = f$, $\lim_{n\to+\infty} \psi_n = g$ on X. Now $(\phi_n + \psi_n)$ is an increasing sequence of non-negative measurable simple functions with $\lim_{n\to+\infty}(\phi_n + \psi_n) = f + g$ on X. By Lemma 7.2, $\int_X (f+g)\, d\mu = \lim_{n\to+\infty} \int_X (\phi_n + \psi_n)\, d\mu = \lim_{n\to+\infty} \int_X \phi_n\, d\mu + \lim_{n\to+\infty} \int_X \psi_n\, d\mu = \int_X f\, d\mu + \int_X g\, d\mu$.

Also, $(\lambda\phi_n)$ is an increasing sequence of non-negative measurable simple functions on X such that $\lim_{n\to+\infty} \lambda\phi_n = \lambda f$ on X. Lemma 7.2 implies again that $\int_X \lambda f\, d\mu = \lim_{n\to+\infty} \int_X \lambda\phi_n\, d\mu = \lambda \lim_{n\to+\infty} \int_X \phi_n\, d\mu = \lambda \int_X f\, d\mu$.

Proposition 7.2 *Let $f, g : X \to [0, +\infty]$ be measurable functions such that $f \le g$ on X. Then $\int_X f\, d\mu \le \int_X g\, d\mu$.*

Proof: Consider arbitrary increasing sequences (ϕ_n) and (ψ_n) of non-negative measurable simple functions with $\lim_{n\to+\infty} \phi_n = f$, $\lim_{n\to+\infty} \psi_n = g$ on X. Then for every k we have that $\phi_k \le f \le g = \lim_{n\to+\infty} \psi_n$ on X. Lemma 7.5 implies that $\int_X \phi_k\, d\mu \le \lim_{n\to+\infty} \int_X \psi_n\, d\mu = \int_X g\, d\mu$. Taking the limit in k we conclude that $\int_X f\, d\mu \le \int_X g\, d\mu$.

Proposition 7.3 *Let $f, g : X \to [0, +\infty]$ be measurable functions on X.*
(i) $\int_X f\, d\mu = 0$ if and only if $f = 0$ a.e. on X.
(ii) If $f = g$ a.e. on X, then $\int_X f\, d\mu = \int_X g\, d\mu$.

Proof: (i) Suppose that $\int_X f\, d\mu = 0$. Define $A_n = \{x \in X \mid \frac{1}{n} \le f(x)\} = f^{-1}([\frac{1}{n}, +\infty])$ for every $n \in \mathbf{N}$. Then $\frac{1}{n}\chi_{A_n} \le f$ on X and Proposition 7.2 says that $\frac{1}{n}\mu(A_n) = \int_X \frac{1}{n}\chi_{A_n}\, d\mu \le \int_X f\, d\mu = 0$. Thus $\mu(A_n) = 0$ for all n and, since $\{x \in X \mid f(x) \ne 0\} = \cup_{n=1}^{+\infty} A_n$, we find that $\mu(\{x \in X \mid f(x) \ne 0\}) = 0$.

Conversely, let $f = 0$ a.e. on X. Consider an arbitrary increasing sequence (ϕ_n) of non-negative measurable simple functions with $\lim_{n\to+\infty} \phi_n = f$ on X. Clearly, $\phi_n = 0$ a.e. on X for all n. Observation (iii) after Definition 7.1 says that $\int_X \phi_n\, d\mu = 0$ for all n. Hence $\int_X f\, d\mu = \lim_{n\to+\infty} \int_X \phi_n\, d\mu = 0$.
(ii) Consider $A = \{x \in X \mid f(x) = g(x)\} \in \Sigma$. Then there is some $B \in \Sigma$ so that $A^c \subseteq B$ and $\mu(B) = 0$. Define $D = B^c \subseteq A$. Then $f\chi_D, g\chi_D$ are measurable and $f\chi_D = g\chi_D$ on X. Also, $f\chi_B = 0$ a.e. on X and $g\chi_B = 0$ a.e. on X.

By part (i), we have that $\int_X f\chi_B\, d\mu = \int_X g\chi_B\, d\mu = 0$ and then Proposition 7.1 implies $\int_X f\, d\mu = \int_X (f\chi_D + f\chi_B)\, d\mu = \int_X f\chi_D\, d\mu = \int_X g\chi_D\, d\mu = \int_X (g\chi_D + g\chi_B)\, d\mu = \int_X g\, d\mu$.

The next three theorems, together with Theorems 7.10 and 7.11 in the next section, are the most important results of integration theory.

Theorem 7.1 ***(The Monotone Convergence Theorem)*** *(Lebesgue, Levi) Let $f, f_n : X \to [0, +\infty]$ ($n \in \mathbf{N}$) be measurable functions on X so that $f_n \le f_{n+1}$ a.e. on X 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 $f_n \le f_{n+1}$ on X and $\lim_{n\to+\infty} f_n = f$ on X.

Proposition 7.2 implies that $\int_X f_n\,d\mu \le \int_X f_{n+1}\,d\mu \le \int_X f\,d\mu$ for all n and hence the $\lim_{n\to+\infty}\int_X f_n\,d\mu$ exists and it is $\le \int_X f\,d\mu$.
(i) Take an arbitrary increasing sequence (ϕ_n) of non-negative measurable simple functions so that $\lim_{n\to+\infty}\phi_n = f$ on X. Then for every k we have $\phi_k \le f = \lim_{n\to+\infty} f_n$. We now take an arbitrary $\alpha \in [0,1)$ and define the set $A_n = \{x \in X \,|\, \alpha\phi_k(x) \le f_n(x)\} \in \Sigma$. It is clear that (A_n) is increasing and $X = \cup_{n=1}^{+\infty} A_n$. It is also true that $\alpha\phi_k\chi_{A_n} \le f_n$ on X and, using Lemma 7.5, $\alpha\int_X \phi_k\,d\mu = \int_X \alpha\phi_k\,d\mu = \lim_{n\to+\infty}\int_X \alpha\phi_k\chi_{A_n}\,d\mu \le \lim_{n\to+\infty}\int_X f_n\,d\mu$. Taking limit as $\alpha \to 1-$, we find $\int_X \phi_k\,d\mu \le \lim_{n\to+\infty}\int_X f_n\,d\mu$. Finally, taking limit in k, we conclude that $\int_X f\,d\mu \le \lim_{n\to+\infty}\int_X f_n\,d\mu$ and the proof has finished.
(ii) If we want to avoid the use of Lemma 7.5, here is an alternative proof of the inequality $\int_X f\,d\mu \le \lim_{n\to+\infty}\int_X f_n\,d\mu$.

Take an increasing sequence $(\psi_n^{(k)})$ of non-negative measurable simple functions so that $\lim_{n\to+\infty}\psi_n^{(k)} = f_k$ on X. Next, define the non-negative measurable simple functions $\phi_n = \max\{\psi_n^{(1)},\ldots,\psi_n^{(n)}\}$.

It easy to see that (ϕ_n) is increasing, that $\phi_n \le f_n \le f$ on X and that $\phi_n \to f$ on X. For the last one, take any $x \in X$ and any $t < f(x)$. Find k so that $t < f_k(x)$ and, then, a large $n \ge k$ so that $t < \psi_n^{(k)}(x)$. Then $t < \phi_n(x) \le f(x)$ and this means that $\phi_n(x) \to f(x)$.

Thus $\int_X f\,d\mu = \lim_{n\to+\infty}\int_X \phi_n\,d\mu \le \lim_{n\to+\infty}\int_X f_n\,d\mu$.
(b) In the general case, Theorem 2.2 implies that there is some $A \in \Sigma$ 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. These imply that $f_n\chi_A \le f_{n+1}\chi_A$ on X for all n and $\lim_{n\to+\infty} f_n\chi_A = f\chi_A$ on X. From part (a) we have that $\lim_{n\to+\infty}\int_X f_n\chi_A\,d\mu = \int_X f\chi_A\,d\mu$.

Since $f = f\chi_A$ a.e. on X and $f_n = f_n\chi_A$ a.e. on X, Proposition 7.3 implies that $\int_X f\,d\mu = \int_X f\chi_A\,d\mu$ and $\int_X f_n\,d\mu = \int_X f_n\chi_A\,d\mu$ for all n. Hence, $\lim_{n\to+\infty}\int_X f_n\,d\mu = \lim_{n\to+\infty}\int_X f_n\chi_A\,d\mu = \int_X f\chi_A\,d\mu = \int_X f\,d\mu$.

Theorem 7.2 *Let $f, f_n : X \to [0,+\infty]$ $(n \in \mathbf{N})$ be measurable on X so that $\sum_{n=1}^{+\infty} f_n = f$ a.e. on X. Then*

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

Proof: We write $g_n = f_1 + \cdots + f_n$ for each n. (g_n) is an increasing sequence of non-negative measurable functions with $g_n \to f$ a.e. on X. Proposition 7.1 and Theorem 7.1 imply that $\sum_{k=1}^{n}\int_X f_k\,d\mu = \int_X g_n\,d\mu \to \int_X f\,d\mu$.

Theorem 7.3 ***(The Lemma of Fatou)*** *Let $f, f_n : X \to [0,+\infty]$ $(n \in \mathbf{N})$ be measurable. If $f = \liminf_{n\to+\infty} f_n$ a.e. on X, then*

$$\int_X f\,d\mu \le \liminf_{n\to+\infty}\int_X f_n\,d\mu.$$

Proof: We define $g_n = \inf_{k\geq n} f_k$. Then each $g_n : X \to [0,+\infty]$ is measurable, the sequence (g_n) is increasing and $g_n \leq f_n$ on X for all n. By hypothesis, $f = \lim_{n\to+\infty} g_n$ a.e. on X. Proposition 7.2 and Theorem 7.1 imply that $\int_X f\,d\mu = \lim_{n\to+\infty}\int_X g_n\,d\mu \leq \liminf_{n\to+\infty}\int_X f_n\,d\mu$.

7.3 Integrals of complex valued functions.

Let (X,Σ,μ) be a fixed measure space.

Definition 7.3 *Let $f : X \to \overline{\mathbf{R}}$ be a measurable function and consider its positive and negative parts $f^+, f^- : X \to [0,+\infty]$. If at least one of $\int_X f^+\,d\mu$ and $\int_X f^-\,d\mu$ is $< +\infty$, we define*

$$\int_X f\,d\mu = \int_X f^+\,d\mu - \int_X f^-\,d\mu$$

and call it ***the integral of f over X (with respect to μ)*** *or, simply,* ***the (μ-)integral of f****.*

We say that f is ***integrable on X (with respect to μ)*** *or* ***(μ-)integrable*** *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.

Lemma 7.7 *Let $f : X \to \overline{\mathbf{R}}$ be a measurable function. Then the following are equivalent:*
(i) f is integrable
(ii) $\int_X f^+\,d\mu < +\infty$ and $\int_X f^-\,d\mu < +\infty$
(iii) $\int_X |f|\,d\mu < +\infty$.

Proof: The equivalence of (i) and (ii) is clear from the definition.

We know that $|f| = f^+ + f^-$ and, hence, $f^+, f^- \leq |f|$ on X. Therefore, $\int_X |f|\,d\mu = \int_X f^+\,d\mu + \int_X f^-\,d\mu$ and $\int_X f^+\,d\mu, \int_X f^-\,d\mu \leq \int_X |f|\,d\mu$. The equivalence of (ii) and (iii) is now obvious.

Proposition 7.4 *Let $f : X \to \overline{\mathbf{R}}$ be a measurable function. If f is integrable, then*
(i) $f(x) \in \mathbf{R}$ for a.e. $x \in X$ and
(ii) the set $\{x \in X \mid f(x) \neq 0\}$ is of σ-finite measure.

Proof: (i) Let f be integrable. Lemma 7.7 implies $\int_X |f|\,d\mu < +\infty$. Consider the set $B = \{x \in X \mid |f(x)| = +\infty\} \in \Sigma$. For every $r \in (0,+\infty)$ we have that $r\chi_B \leq |f|$ on X and hence $r\mu(B) = \int_X r\chi_B\,d\mu \leq \int_X |f|\,d\mu < +\infty$. This implies that $\mu(B) \leq \frac{1}{r}\int_X |f|\,d\mu$ and, taking the limit as $r \to +\infty$, we find $\mu(B) = 0$.
(ii) Consider the sets $A = \{x \in X \mid f(x) \neq 0\}$ and $A_n = \{x \in X \mid |f(x)| \geq \frac{1}{n}\}$.

From $\frac{1}{n}\chi_{A_n} \leq |f|$ on X, we get $\frac{1}{n}\mu(A_n) = \int_X \frac{1}{n}\chi_{A_n}\,d\mu \leq \int_X |f|\,d\mu < +\infty$. Thus $\mu(A_n) < +\infty$ for all n and, since $A = \cup_{n=1}^{+\infty} A_n$, we conclude that A is of σ-finite measure.

Definition 7.4 *Let $f : X \to \overline{\mathbf{C}}$ be measurable. Then $|f| : X \to [0, +\infty]$ is measurable and we say that f is* ***integrable on*** X ***(with respect to*** μ***)*** *or, simply,* ***(μ-)integrable****, if $\int_X |f| \, d\mu < +\infty$.*

Proposition 7.5 *Let $f : X \to \overline{\mathbf{C}}$ be measurable. If f is integrable, then*
(i) $f(x) \in \mathbf{C}$ for a.e. $x \in X$ and
(ii) the set $\{x \in X \mid f(x) \neq 0\}$ is of σ-finite measure.

Proof: Immediate application of Proposition 7.4 to $|f|$.

Assume now that $f : X \to \overline{\mathbf{C}}$ is a measurable *integrable* function. By Proposition 7.5, the set $D_f = \{x \in X \mid f(x) \in \mathbf{C}\} = f^{-1}(\mathbf{C}) \in \Sigma$ has a null complement. The function

$$f\chi_{D_f} = \begin{cases} f, & \text{on } D_f \\ 0, & \text{on } D_f^c \end{cases} : X \to \mathbf{C}$$

is measurable and $f\chi_{D_f} = f$ a.e. on X. The advantage of $f\chi_{D_f}$ over f is that $f\chi_{D_f}$ is *complex valued* and, hence, the $\Re(f\chi_{D_f}), \Im(f\chi_{D_f}) : X \to \mathbf{R}$ are defined on X. We also have that $|\Re(f\chi_{D_f})| \leq |f\chi_{D_f}| \leq |f|$ on X and similarly $|\Im(f\chi_{D_f})| \leq |f|$ on X. Therefore $\int_X |\Re(f\chi_{D_f})| \, d\mu \leq \int_X |f| \, d\mu < +\infty$, implying that $\Re(f\chi_{D_f})$ is an integrable real valued function. The same is true for $\Im(f\chi_{D_f})$ and thus the integrals $\int_X \Re(f\chi_{D_f}) \, d\mu$ and $\int_X \Im(f\chi_{D_f}) \, d\mu$ are defined and they are (finite) real numbers.

Definition 7.5 *Let $f : X \to \overline{\mathbf{C}}$ be a measurable integrable function and let $D_f = \{x \in X \mid f(x) \in \mathbf{C}\}$. We define*

$$\int_X f \, d\mu = \int_X \Re(f\chi_{D_f}) \, d\mu + i \int_X \Im(f\chi_{D_f}) \, d\mu$$

and call it ***the integral of*** f ***over*** X ***(with respect to*** μ***)*** *or* ***the (μ-)integral of*** f.

We shall make a few observations 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 its integral is a (finite) complex number. Observe that the integral of an extended-real valued function is defined if the function is integrable (and the value of its integral is a finite real number) and *also* in certain other cases when the value of its integral can be either $+\infty$ or $-\infty$.
(ii) We used the function $f\chi_{D_f}$, which changes the value ∞ of f to the value 0, simply because we need complex values in order to be able to consider their real and imaginary parts. We may allow more freedom and see what happens if we use a function

$$F = \begin{cases} f, & \text{on } D_f \\ h, & \text{on } D_f^c \end{cases} : X \to \mathbf{C},$$

where h is an arbitrary $\Sigma\rceil D_f^c$-measurable *complex* valued function on D_f^c. It is clear that $F = f\chi_{D_f}$ a.e. on X and hence $\Re(F) = \Re(f\chi_{D_f})$ a.e. on X.

Of course, this implies that $\Re(F)^+ = \Re(f\chi_{D_f})^+$ and $\Re(F)^- = \Re(f\chi_{D_f})^-$ a.e. on X. From Proposition 7.3, $\int_X \Re(F)\,d\mu = \int_X \Re(F)^+\,d\mu - \int_X \Re(F)^-\,d\mu = \int_X \Re(f\chi_{D_f})^+\,d\mu - \int_X \Re(f\chi_{D_f})^-\,d\mu = \int_X \Re(f\chi_{D_f})\,d\mu$. Similarly, $\int_X \Im(F)\,d\mu = \int_X \Im(f\chi_{D_f})\,d\mu$. Therefore there is no difference between the possible definition $\int_X f\,d\mu = \int_X \Re(F)\,d\mu + i\int_X \Im(F)\,d\mu$ and the one we have given. Of course, the function 0 on D_f^c is the simplest of all choices for h.
(iii) If $f : X \to \mathbf{C}$ is complex valued on X, then $D_f = X$ and the definition takes the simpler form

$$\int_X f\,d\mu = \int_X \Re(f)\,d\mu + i\int_X \Im(f)\,d\mu.$$

We also have

$$\Re\Big(\int_X f\,d\mu\Big) = \int_X \Re(f)\,d\mu, \qquad \Im\Big(\int_X f\,d\mu\Big) = \int_X \Im(f)\,d\mu.$$

The next is helpful and we shall make use of it very often.

Lemma 7.8 *If $f : X \to \overline{\mathbf{C}}$ is integrable, there is $F : X \to \mathbf{C}$ so that $F = f$ a.e. on X and $\int_X F\,d\mu = \int_X f\,d\mu$.*

Proof: We take $F = f\chi_{D_f}$, where $D_f = f^{-1}(\mathbf{C})$.

Theorem 7.4 *Let $f, g : X \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$ be measurable so that $f = g$ a.e. on X and $\int_X f\,d\mu$ is defined. Then $\int_X g\,d\mu$ is also defined and $\int_X g\,d\mu = \int_X f\,d\mu$.*

Proof: (a) Let $f, g : X \to \overline{\mathbf{R}}$. If $f = g$ a.e. on X, then $f^+ = g^+$ a.e. on X and $f^- = g^-$ a.e. on X. Proposition 7.3 implies that $\int_X f^+\,d\mu = \int_X g^+\,d\mu$ and $\int_X f^-\,d\mu = \int_X g^-\,d\mu$. Now if $\int_X f^+\,d\mu$ or $\int_X f^-\,d\mu$ is finite, then, respectively, $\int_X g^+\,d\mu$ or $\int_X g^-\,d\mu$ is also finite. Therefore $\int_X g\,d\mu$ is defined and $\int_X f\,d\mu = \int_X g\,d\mu$.
(b) Let $f, g : X \to \overline{\mathbf{C}}$ and $f = g$ a.e. on X.

If f is integrable, from $|f| = |g|$ a.e. on X and from Proposition 7.3, we find $\int_X |g|\,d\mu = \int_X |f|\,d\mu < +\infty$ and, hence, g is also integrable.

Now, Lemma 7.8 says that there are $F, G : X \to \mathbf{C}$ so that $F = f$ and $G = g$ a.e. on X and also $\int_X F\,d\mu = \int_X f\,d\mu$ and $\int_X G\,d\mu = \int_X g\,d\mu$. From $f = g$ a.e. on X we see that $F = G$ a.e. on X. This implies that $\Re(F) = \Re(G)$ a.e. on X and, from (a), $\int_X \Re(F)\,d\mu = \int_X \Re(G)\,d\mu$. Similarly, $\int_X \Im(F)\,d\mu = \int_X \Im(G)\,d\mu$.

Therefore, $\int_X f\,d\mu = \int_X F\,d\mu = \int_X \Re(F)\,d\mu + i\int_X \Im(F)\,d\mu = \int_X \Re(G)\,d\mu + i\int_X \Im(G)\,d\mu = \int_X G\,d\mu = \int_X g\,d\mu$.

Theorem 7.5 *Let $f : X \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$ be measurable. Then the following are equivalent:*
(i) $f = 0$ a.e. on X
(ii) $\int_X |f|\,d\mu = 0$
(iii) $\int_X f\chi_A\,d\mu = 0$ for every $A \in \Sigma$.

Proof: If $\int_X |f|\,d\mu = 0$, Proposition 7.3 implies that $|f| = 0$ and, hence, $f = 0$ a.e. on X.

If $f = 0$ a.e. on X, then $f\chi_A = 0$ a.e. on X for all $A \in \Sigma$. Theorem 7.4 implies that $\int_X f\chi_A\,d\mu = 0$.

Finally, let $\int_X f\chi_A\,d\mu = 0$ for every $A \in \Sigma$.

(a) If $f : X \to \overline{\mathbf{R}}$ we take $A = f^{-1}([0,+\infty])$ and find $\int_X f^+\,d\mu = \int_X f\chi_A\,d\mu = 0$. Similarly, $\int_X f^-\,d\mu = 0$ and thus $\int_X |f|\,d\mu = \int_X f^+\,d\mu + \int_X f^-\,d\mu = 0$.

(b) If $f : X \to \overline{\mathbf{C}}$, we first take $A = X$ and find $\int_X f\,d\mu = 0$. This says, in particular, that f is integrable. We take some $F : X \to \mathbf{C}$ so that $F = f$ a.e. on X.

For every $A \in \Sigma$ we have $F\chi_A = f\chi_A$ a.e. on X and, from Theorem 7.4, $\int_X F\chi_A\,d\mu = \int_X f\chi_A\,d\mu = 0$. This implies $\int_X \Re(F)\chi_A\,d\mu = \int_X \Re(F\chi_A)\,d\mu = \Re(\int_X F\chi_A\,d\mu) = 0$ and, from part (a), $\Re(F) = 0$ a.e. on X. Similarly, $\Im(F) = 0$ a.e. on X and thus $F = 0$ a.e. on X. We conclude that $f = 0$ a.e. on X.

Theorem 7.6 *Let $f : X \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$ be measurable and $\lambda \in \mathbf{R}$ or $\mathbf{C}$.*
(i) If $f : X \to \overline{\mathbf{R}}$, $\lambda \in \mathbf{R}$ and $\int_X f\,d\mu$ is defined, then $\int_X \lambda f\,d\mu$ is also defined and

$$\int_X \lambda f\,d\mu = \lambda \int_X f\,d\mu.$$

(ii) If f is integrable, then λf is also integrable and the previous equality is again true.

Proof: (i) Let $f : X \to \overline{\mathbf{R}}$ and $\int_X f\,d\mu$ be defined and, hence, either $\int_X f^+\,d\mu < +\infty$ or $\int_X f^-\,d\mu < +\infty$.

If $0 < \lambda < +\infty$, then $(\lambda f)^+ = \lambda f^+$ and $(\lambda f)^- = \lambda f^-$. Therefore, at least one of $\int_X (\lambda f)^+\,d\mu = \lambda \int_X f^+\,d\mu$ and $\int_X (\lambda f)^-\,d\mu = \lambda \int_X f^-\,d\mu$ is finite. This means that $\int_X \lambda f\,d\mu$ 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 $-\infty < \lambda < 0$, then $(\lambda f)^+ = -\lambda f^-$ and $(\lambda f)^- = -\lambda f^+$ and the previous argument can be repeated with no essential change.

If $\lambda = 0$, the result is trivial.

(ii) If $f : X \to \overline{\mathbf{R}}$ is integrable and $\lambda \in \mathbf{R}$, then $\int_X |\lambda f|\,d\mu = |\lambda| \int_X |f|\,d\mu < +\infty$, which means that λf is also integrable. The equality $\int_X \lambda f\,d\mu = \lambda \int_X f\,d\mu$ has been proved in (i).

If $f : X \to \overline{\mathbf{C}}$ is integrable and $\lambda \in \mathbf{C}$, the same argument gives that λf is also integrable.

We, now, take $F : X \to \mathbf{C}$ so that $F = f$ a.e. on X. Then, also $\lambda F = \lambda f$ a.e. on X and Theorem 7.4 implies that $\int_X \lambda F\,d\mu = \int_X \lambda f\,d\mu$ and $\int_X F\,d\mu = \int_X f\,d\mu$. Hence, it is enough to prove that $\int_X \lambda F\,d\mu = \lambda \int_X F\,d\mu$.

From $\Re(\lambda F) = \Re(\lambda)\Re(F) - \Im(\lambda)\Im(F)$ and from the real valued case we get that

$$\int_X \Re(\lambda F)\,d\mu = \Re(\lambda) \int_X \Re(F)\,d\mu - \Im(\lambda) \int_X \Im(F)\,d\mu.$$

Similarly,

$$\int_X \Im(\lambda F)\,d\mu = \Re(\lambda)\int_X \Im(F)\,d\mu + \Im(\lambda)\int_X \Re(F)\,d\mu.$$

From these two equalities

$$\int_X \lambda F\,d\mu = \lambda\int_X \Re(F)\,d\mu + i\lambda\int_X \Im(F)\,d\mu = \lambda\int_X F\,d\mu.$$

Theorem 7.7 *Let $f, g : X \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$ be measurable and consider any measurable definition of $f+g$.*
(i) If $f, g : X \to \overline{\mathbf{R}}$ and $\int_X f\,d\mu, \int_X g\,d\mu$ are both defined and they are not opposite infinities, then $\int_X (f+g)\,d\mu$ is also defined and

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

(ii) If $f, g : X \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$ are integrable, then $f+g$ is also integrable and the previous equality is again true.

Proof: (i) Considering the integrals $\int_X f^+\,d\mu, \int_X f^-\,d\mu, \int_X g^+\,d\mu, \int_X g^-\,d\mu$, the assumptions imply that at most the $\int_X f^+\,d\mu, \int_X g^+\,d\mu$ are $+\infty$ or at most the $\int_X f^-\,d\mu, \int_X g^-\,d\mu$ are $+\infty$.

Let $\int_X f^-\,d\mu < +\infty$ and $\int_X g^-\,d\mu < +\infty$.

Proposition 7.4 implies that, if $B = \{x \in X \mid f(x) \neq -\infty, g(x) \neq -\infty\}$, then $\mu(B^c) = 0$. We define the functions $F = f\chi_B$ and $G = g\chi_B$. Then $F, G : X \to (-\infty, +\infty]$ are measurable and $F = f$ 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* $x \in X$.

Observe that for all measurable definitions of $f+g$, we have $F+G = f+g$ a.e. on X. Because of Theorem 7.4, it is enough to prove that the $\int_X (F+G)\,d\mu$ is defined and that $\int_X (F+G)\,d\mu = \int_X F\,d\mu + \int_X G\,d\mu$.

From $F = F^+ - F^- \leq F^+$ and $G = G^+ - G^- \leq G^+$ on X we get $F+G \leq F^+ + G^+$ on X. Hence $(F+G)^+ \leq F^+ + G^+$ on X and similarly $(F+G)^- \leq F^- + G^-$ on X.

From $(F+G)^- \leq F^- + G^-$ on X we find $\int_X (F+G)^-\,d\mu \leq \int_X F^-\,d\mu + \int_X G^-\,d\mu < +\infty$. Therefore, $\int_X (F+G)\,d\mu$ is defined.

We now have $(F+G)^+ - (F+G)^- = F+G = (F^+ + G^+) - (F^- + G^-)$ or, equivalently, $(F+G)^+ + F^- + G^- = (F+G)^- + F^+ + G^+$.

Proposition 7.1 implies that

$$\int_X (F+G)^+\,d\mu + \int_X F^-\,d\mu + \int_X G^-\,d\mu = \int_X (F+G)^-\,d\mu + \int_X F^+\,d\mu + \int_X G^+\,d\mu.$$

Because of the finiteness of $\int_X (F+G)^-\,d\mu, \int_X F^-\,d\mu, \int_X G^-\,d\mu$, we get

$$\int_X (F+G)\,d\mu = \int_X (F+G)^+\,d\mu - \int_X (F+G)^-\,d\mu$$

$$= \int_X F^+ \, d\mu + \int_X G^+ \, d\mu - \int_X F^- \, d\mu - \int_X G^- \, d\mu$$
$$= \int_X F \, d\mu + \int_X G \, d\mu.$$

The proof in the case when $\int_X f^+ \, d\mu < +\infty$ and $\int_X g^+ \, d\mu < +\infty$ is similar.
(ii) By Lemma 7.8, there are $F, G : X \to \mathbf{C}$ so that $F = f$ 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. Now, by Theorem 7.4, it is enough to prove that $F+G$ is integrable and $\int_X (F+G) \, d\mu = \int_X F \, d\mu + \int_X G \, d\mu$.

Now $\int_X |F+G| \, d\mu \le \int_X |F| \, d\mu + \int_X |G| \, d\mu < +\infty$ and, hence, $F+G$ is integrable.

By part (i) we have $\int_X \Re(F+G) \, d\mu = \int_X \Re(F) \, d\mu + \int_X \Re(G) \, d\mu$ and the same equality with the imaginary parts. Combining, we get $\int_X (F+G) \, d\mu = \int_X F \, d\mu + \int_X G \, d\mu$.

Theorem 7.8 *Let $f, g : X \to \overline{\mathbf{R}}$ be measurable. If $\int_X f \, d\mu$ and $\int_X g \, d\mu$ are both defined and $f \le g$ on X, then*

$$\int_X f \, d\mu \le \int_X g \, d\mu.$$

Proof: From $f \le g = g^+ - g^- \le g^+$ we get $f^+ \le g^+$. Similarly, $g^- \le f^-$. Therefore, if $\int_X g^+ \, d\mu < +\infty$, then $\int_X f^+ \, d\mu < +\infty$ and, if $\int_X f^- \, d\mu < +\infty$, then $\int_X g^- \, d\mu < +\infty$.

Hence we can subtract the two inequalities

$$\int_X f^+ \, d\mu \le \int_X g^+ \, d\mu, \qquad \int_X g^- \, d\mu \le \int_X f^- \, d\mu$$

and find that $\int_X f \, d\mu \le \int_X g \, d\mu$.

Theorem 7.9 *Let $f : X \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$ be measurable.*
(i) If $f : X \to \overline{\mathbf{R}}$ and $\int_X f \, d\mu$ is defined, then

$$\Big| \int_X f \, d\mu \Big| \le \int_X |f| \, d\mu.$$

(ii) If $f : X \to \overline{\mathbf{C}}$ is integrable, then the inequality in (i) is again true.

Proof: (i) We write $| \int_X f \, d\mu| = | \int_X f^+ \, d\mu - \int_X f^- \, d\mu| \le \int_X f^+ \, d\mu + \int_X f^- \, d\mu = \int_X |f| \, d\mu$.
(ii) Consider $F : X \to \mathbf{C}$ so that $F = f$ a.e. on X. By Theorem 7.4, it is enough to prove $| \int_X F \, d\mu| \le \int_X |F| \, d\mu$.

If $\int_X F \, d\mu = 0$, then the inequality is trivially true. Let $0 \ne \int_X F \, d\mu \in \mathbf{C}$ and take $\lambda = \overline{\mathrm{sign}(\int_X F \, d\mu)} \ne 0$. Then

$$\Big| \int_X F \, d\mu \Big| = \lambda \int_X F \, d\mu = \int_X \lambda F \, d\mu = \Re\Big(\int_X \lambda F \, d\mu \Big) = \int_X \Re(\lambda F) \, d\mu$$
$$\le \int_X |\Re(\lambda F)| \, d\mu \le \int_X |\lambda F| \, d\mu = \int_X |F| \, d\mu.$$

Theorem 7.10 ***(The Dominated Convergence Theorem)*** *(Lebesgue) Consider the measurable $f, f_n : X \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$ ($n \in \mathbf{N}$) and $g : X \to [0, +\infty]$. Assume that $f = \lim_{n\to+\infty} f_n$ a.e. on X, that, for all n, $|f_n| \le g$ a.e. on X and that $\int_X g\,d\mu < +\infty$. Then all f_n and f are integrable and*

$$\int_X f_n\,d\mu \to \int_X f\,d\mu.$$

Proof: From the $|f_n| \le g$ a.e. on X we find $\int_X |f_n|\,d\mu \le \int_X g\,d\mu < +\infty$ and hence f_n is integrable. Also, from $|f_n| \le g$ a.e. on X and $f = \lim_{n\to+\infty} f_n$ a.e. on X, we get that $|f| \le g$ a.e. on X and, for the same reason, f is also integrable.

We may now take $F, F_n : X \to \mathbf{R}$ or $\mathbf{C}$ so that $F = f$ and $F_n = f_n$ a.e. on X for all n. We, then, have $|F_n| \le g$ a.e. on X and $F = \lim_{n\to+\infty} F_n$ a.e. on X and it is enough to prove $\int_X F_n\,d\mu \to \int_X F\,d\mu$.

(i) Let $F, F_n : X \to \mathbf{R}$. Since $0 \le g + F_n, g - F_n$ on X, the Lemma of Fatou implies that

$$\int_X g\,d\mu \pm \int_X F\,d\mu \le \liminf_{n\to+\infty} \int_X (g \pm F_n)\,d\mu$$

and hence

$$\int_X g\,d\mu \pm \int_X F\,d\mu \le \int_X g\,d\mu + \liminf_{n\to+\infty} \pm \int_X F_n\,d\mu.$$

Since $\int_X g\,d\mu$ is finite, we get that $\pm\int_X F\,d\mu \le \liminf_{n\to+\infty} \pm \int_X F_n\,d\mu$ and hence

$$\limsup_{n\to+\infty} \int_X F_n\,d\mu \le \int_X F\,d\mu \le \liminf_{n\to+\infty} \int_X F_n\,d\mu.$$

This implies $\int_X F_n\,d\mu \to \int_X F\,d\mu$.

(ii) Let $F, F_n : X \to \mathbf{C}$. From $|\Re(F_n)| \le |F_n| \le g$ a.e. on X and from $\Re(F_n) \to \Re(F)$ a.e. on X, part (i) implies $\int_X \Re(F_n)\,d\mu \to \int_X \Re(F)\,d\mu$. Similarly, $\int_X \Im(F_n)\,d\mu \to \int_X \Im(F)\,d\mu$ and, from these two, $\int_X F_n\,d\mu \to \int_X F\,d\mu$.

Theorem 7.11 ***(The Series Theorem)*** *Consider the measurable $f, f_n : X \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$ ($n \in \mathbf{N}$). If $\sum_{n=1}^{+\infty} \int_X |f_n|\,d\mu < +\infty$, then*

(i) $\sum_{n=1}^{+\infty} f_n(x)$ exists for a.e. $x \in X$,

(ii) if $f = \sum_{n=1}^{+\infty} f_n$ a.e. on X, then

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

Proof: (i) Define $g = \sum_{n=1}^{+\infty} |f_n| : X \to [0, +\infty]$ on X. From Theorem 7.2 we have $\int_X g\,d\mu = \sum_{n=1}^{+\infty} \int_X |f_n|\,d\mu < +\infty$. This implies that $g < +\infty$ a.e. on X, which means that the series $\sum_{n=1}^{+\infty} f_n(x)$ converges absolutely, and hence converges, for a.e. $x \in X$.

(ii) Consider $s_n = \sum_{k=1}^{n} f_k$ for all n. Then $\lim_{n\to+\infty} s_n = f$ a.e. on X and $|s_n| \le g$ on X. Theorem 7.10 implies that $\sum_{k=1}^{n} \int_X f_k\,d\mu = \int_X s_n\,d\mu \to \int_X f\,d\mu$.

Theorem 7.12 ***(Approximation)*** *Let $f : X \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$ be integrable. Then for every $\epsilon > 0$ there is some measurable simple function $\phi : X \to \mathbf{R}$ or $\mathbf{C}$ so that $\int_X |f - \phi|\, d\mu < \epsilon$.*

Proof: If $f : X \to [0, +\infty]$ is measurable with $\int_X f\, d\mu < +\infty$, there is an increasing sequence (ϕ_n) of non-negative measurable simple functions so that $\phi_n \uparrow f$ on X and $\int_X \phi_n\, d\mu \uparrow \int_X f\, d\mu$. Therefore, for some n we have $\int_X f\, d\mu - \epsilon < \int_X \phi_n\, d\mu \leq \int_X f\, d\mu$. Thus $\int_X |f - \phi_n|\, d\mu = \int_X (f - \phi_n)\, d\mu < \epsilon$.

Now if $f : X \to \overline{\mathbf{R}}$ is integrable, then $\int_X f^+\, d\mu < +\infty$ and $\int_X f^-\, d\mu < +\infty$. From the first case considered, there are non-negative measurable simple functions ϕ_1, ϕ_2 so that $\int_X |f^+ - \phi_1|\, d\mu < \frac{\epsilon}{2}$ and $\int_X |f^- - \phi_2|\, d\mu < \frac{\epsilon}{2}$. We define the simple function $\phi = \phi_1 - \phi_2 : X \to \mathbf{R}$ and get $\int_X |f - \phi|\, d\mu \leq \int_X |f^+ - \phi_1|\, d\mu + \int_X |f^- - \phi_2|\, d\mu < \epsilon$.

Finally, let $f : X \to \overline{\mathbf{C}}$ be integrable. Then there is $F : X \to \mathbf{C}$ so that $F = f$ a.e. on X. The functions $\Re(F), \Im(F) : X \to \mathbf{R}$ are both integrable, and hence we can find real valued measurable simple functions ϕ_1, ϕ_2 so that $\int_X |\Re(F) - \phi_1|\, d\mu < \frac{\epsilon}{2}$ and $\int_X |\Im(F) - \phi_2|\, d\mu < \frac{\epsilon}{2}$. We define $\phi = \phi_1 + i\phi_2$ and get $\int_X |f - \phi|\, d\mu = \int_X |F - \phi|\, d\mu < \epsilon$.

7.4 Integrals over subsets.

Let (X, Σ, μ) be a measure space.

Let $A \in \Sigma$ and $f : X \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$ be measurable. In order to define *an integral of f over A* we have two *natural* choices. One way is to take $f\chi_A$, which is f in A and 0 outside A, and consider $\int_X f\chi_A\, d\mu$. Another way 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 $\mu\rceil A$ on $(A, \Sigma\rceil A)$. The following lemma says that the two procedures are equivalent and give the same results.

Lemma 7.9 *Let $A \in \Sigma$ and $f : X \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$ be measurable.*
(i) If $f : X \to \overline{\mathbf{R}}$ and either $\int_X f\chi_A\, d\mu$ or $\int_A (f\rceil A)\, d(\mu\rceil A)$ exists, then the other also exists and they are equal.
(ii) If $f : X \to \overline{\mathbf{C}}$ 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 the integrals $\int_X f\chi_A\, d\mu$ and $\int_A (f\rceil A)\, d(\mu\rceil A)$ are equal.

Proof: (a) Take a non-negative measurable simple function $\phi = \sum_{j=1}^m \kappa_j \chi_{E_j}$ with its standard representation. Now $\phi\chi_A = \sum_{j=1}^m \kappa_j \chi_{E_j \cap A} : X \to [0, +\infty)$ has $\int_X \phi\chi_A\, d\mu = \sum_{j=1}^m \kappa_j \mu(E_j \cap A)$. On the other hand, $\phi\rceil A = \sum_{j=1}^m \kappa_j \chi_{E_j \cap A} : A \to [0, +\infty)$ (where we omit the terms for which $E_j \cap A = \emptyset$) has exactly the same integral $\int_A (\phi\rceil A)\, d(\mu\rceil A) = \sum_{j=1}^m \kappa_j (\mu\rceil A)(E_j \cap A) = \sum_{j=1}^m \kappa_j \mu(E_j \cap A)$.
(b) Now let $f : X \to [0, +\infty]$ be measurable. Take an increasing sequence (ϕ_n) of non-negative measurable simple $\phi_n : X \to [0, +\infty)$ with $\phi_n \to f$. Then $(\phi_n\chi_A)$ is increasing and $\phi_n\chi_A \to f\chi_A$. Also, $(\phi_n\rceil A)$ is increasing and $\phi_n\rceil A \to f\rceil A$. Hence, by (a) we get, $\int_X f\chi_A = \lim_{n\to+\infty} \int_X \phi_n\chi_A\, d\mu = \lim_{n\to+\infty} \int_A (\phi_n\rceil A)\, d(\mu\rceil A) = \int_A (f\rceil A)\, d(\mu\rceil A)$.
(c) If $f : X \to \overline{\mathbf{R}}$ is measurable, then $f^+\chi_A = (f\chi_A)^+$ and $f^-\chi_A = (f\chi_A)^-$ and

also $(f\rceil A)^+ = f^+\rceil A$ and $(f\rceil A)^- = f^-\rceil A$. Hence, by (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 also $\int_X (f\chi_A)^-\, d\mu = \int_A (f\rceil A)^-\, d(\mu\rceil A)$. These show (i).
(d) Finally, let $f : X \to \overline{\mathbf{C}}$ be measurable. Then $|f\chi_A| = |f|\chi_A$ and $|f\rceil A| = |f|\rceil A$. By (b) we have $\int_X |f\chi_A|\, d\mu = \int_X |f|\chi_A\, d\mu = \int_A (|f|\rceil A)\, d(\mu\rceil A) = \int_A |f\rceil A|\, d(\mu\rceil A)$, implying that $f\chi_A$ and $f\rceil A$ are simultaneously integrable or non-integrable.

Assuming integrability, there is an $F : X \to \mathbf{C}$ 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, also, $F\rceil A = f\rceil A$ a.e. on A. Therefore, it is enough to prove that $\int_X F\chi_A\, d\mu = \int_A (F\rceil A)\, d(\mu\rceil A)$.

Part (c) implies $\int_X \Re(F\chi_A)\, d\mu = \int_X \Re(F)\chi_A\, d\mu = \int_A (\Re(F)\rceil A)\, d(\mu\rceil A) = \int_A \Re(F\rceil A)\, d(\mu\rceil A)$. Similarly, $\int_X \Im(F\chi_A)\, d\mu = \int_A \Im(F\rceil A)\, d(\mu\rceil A)$ and we conclude that $\int_X F\chi_A\, d\mu = \int_A (F\rceil A)\, d(\mu\rceil A)$.

Definition 7.6 *Let $f : X \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$ be measurable and $A \in \Sigma$.*
*(i) If $f : X \to \overline{\mathbf{R}}$ and $\int_X f\chi_A\, d\mu$ or, equivalently, $\int_A (f\rceil A)\, d(\mu\rceil A)$ is defined, we say that **the** $\int_A f\, d\mu$ **is defined** and define*

$$\int_A f\, d\mu = \int_X f\chi_A\, d\mu = \int_A (f\rceil A)\, d(\mu\rceil A).$$

*(ii) If $f : X \to \overline{\mathbf{C}}$ and $f\chi_A$ is integrable on X or, equivalently, $f\rceil A$ is integrable on A, we say that f **is integrable on** A and define $\int_A f\, d\mu$ exactly as in (i).*

Lemma 7.10 *Let $f : X \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$ be measurable.*
(i) If $f : X \to \overline{\mathbf{R}}$ and $\int_X f\, d\mu$ is defined, then $\int_A f\, d\mu$ is defined for every $A \in \Sigma$.
(ii) If $f : X \to \overline{\mathbf{C}}$ is integrable then f is integrable on every $A \in \Sigma$.

Proof: (i) We have $(f\chi_A)^+ = f^+\chi_A \leq f^+$ and $(f\chi_A)^- = f^-\chi_A \leq f^-$ on X. Therefore, either $\int_X (f\chi_A)^+\, d\mu \leq \int_X f^+\, d\mu < +\infty$ or $\int_X (f\chi_A)^-\, d\mu \leq \int_X f^-\, d\mu < +\infty$. This says that $\int_X f\chi_A\, d\mu$ is defined and, hence, $\int_A f\, d\mu$ is also defined.
(ii) If $f : X \to \overline{\mathbf{C}}$ is integrable, then $\int_X |f\chi_A|\, d\mu \leq \int_X |f|\, d\mu < +\infty$ and $f\chi_A$ is also integrable.

Proposition 7.6 *Let $f : X \to \overline{\mathbf{R}}$ be measurable and $\int_X f\, d\mu$ be defined. Then either $\int_A f\, d\mu \in (-\infty, +\infty]$ for all $A \in \Sigma$ or $\int_A f\, d\mu \in [-\infty, +\infty)$ for all $A \in \Sigma$.*

Proof: Let $\int_X f^-\, d\mu < +\infty$. Then $\int_X (f\chi_A)^-\, d\mu \leq \int_X f^-\, d\mu < +\infty$ and hence $\int_A f\, d\mu = \int_X f\chi_A\, d\mu > -\infty$ for all $A \in \Sigma$.

Similarly, if $\int_X f^+\, d\mu < +\infty$, then $\int_A f\, d\mu < +\infty$ for all $A \in \Sigma$.

Theorem 7.13 *If $f : X \to \overline{\mathbf{R}}$ and $\int_X f\, d\mu$ is defined or $f : X \to \overline{\mathbf{C}}$ and f is integrable, then*
(i) $\int_A f\, d\mu = 0$ for all $A \in \Sigma$ with $\mu(A) = 0$,
(ii) $\int_A f\, d\mu = \sum_{n=1}^{+\infty} \int_{A_n} f\, d\mu$ for all pairwise disjoint $A_1, A_2, \ldots \in \Sigma$ with $A = \cup_{n=1}^{+\infty} A_n$,

(iii) $\int_{A_n} f\,d\mu \to \int_A f\,d\mu$ *for all increasing* (A_n) *in* Σ *with* $A = \cup_{n=1}^{+\infty} A_n$*,*
(iv) $\int_{A_n} f\,d\mu \to \int_A f\,d\mu$ *for all decreasing* (A_n) *in* Σ *with* $A = \cap_{n=1}^{+\infty} A_n$ *and* $|\int_{A_1} f\,d\mu| < +\infty$.

Proof: (i) This is easy because $f\chi_A = 0$ a.e. on X.
(ii) Let $A_1, A_2, \ldots \in \Sigma$ be pairwise disjoint and $A = \cup_{n=1}^{+\infty} A_n$.

If $f : X \to [0, +\infty]$ is measurable, since $f\chi_A = \sum_{n=1}^{+\infty} f\chi_{A_n}$ on X, Theorem 7.2 implies $\int_A f\,d\mu = \int_X f\chi_A\,d\mu = \sum_{n=1}^{+\infty} \int_X f\chi_{A_n}\,d\mu = \sum_{n=1}^{+\infty} \int_{A_n} f\,d\mu$.

If $f : X \to \overline{\mathbf{C}}$ and f is integrable, we have by the previous case that $\sum_{n=1}^{+\infty} \int_X |f\chi_{A_n}|\,d\mu = \sum_{n=1}^{+\infty} \int_{A_n} |f|\,d\mu = \int_A |f|\,d\mu < +\infty$. Because of $f\chi_A = \sum_{n=1}^{+\infty} f\chi_{A_n}$ on X, Theorem 7.11 implies that $\int_A f\,d\mu = \sum_{n=1}^{+\infty} \int_{A_n} f\,d\mu$.

If $f : X \to \overline{\mathbf{R}}$ and $\int_X f^-\,d\mu < +\infty$, we apply the first case and get $\sum_{n=1}^{+\infty} \int_{A_n} f^+\,d\mu = \int_A f^+\,d\mu$ and $\sum_{n=1}^{+\infty} \int_{A_n} f^-\,d\mu = \int_A f^-\,d\mu < +\infty$. Subtracting, we find $\sum_{n=1}^{+\infty} \int_{A_n} f\,d\mu = \int_A f\,d\mu$.

If $\int_X f^+\,d\mu < +\infty$, the proof is similar.
(iii) Write $A = A_1 \cup \cup_{k=2}^{+\infty}(A_k \setminus A_{k-1})$, where the sets in the union are pairwise disjoint. Apply (ii) to get $\int_A f\,d\mu = \int_{A_1} f\,d\mu + \sum_{k=2}^{+\infty} \int_{A_k \setminus A_{k-1}} f\,d\mu = \int_{A_1} f\,d\mu + \lim_{n\to+\infty} \sum_{k=2}^{n} \int_{A_k \setminus A_{k-1}} f\,d\mu = \lim_{n\to+\infty} \int_{A_n} f\,d\mu$.
(iv) Write $A_1 \setminus A = \cup_{n=1}^{+\infty}(A_1 \setminus A_n)$, where $(A_1 \setminus A_n)$ is increasing. Apply (iii) to get $\int_{A_1 \setminus A_n} f\,d\mu \to \int_{A_1 \setminus A} f\,d\mu$.

From $\int_{A_1 \setminus A} f\,d\mu + \int_A f\,d\mu = \int_{A_1} f\,d\mu$ and from $|\int_{A_1} f\,d\mu| < +\infty$ we immediately get that also $|\int_A f\,d\mu| < +\infty$. From the same equality we then get $\int_{A_1 \setminus A} f\,d\mu = \int_{A_1} f\,d\mu - \int_A f\,d\mu$. Similarly, $\int_{A_1 \setminus A_n} f\,d\mu = \int_{A_1} f\,d\mu - \int_{A_n} f\,d\mu$ and hence $\int_{A_1} f\,d\mu - \int_{A_n} f\,d\mu \to \int_{A_1} f\,d\mu - \int_A f\,d\mu$. Because of $|\int_{A_1} f\,d\mu| < +\infty$ again, we finally have $\int_{A_n} f\,d\mu \to \int_A f\,d\mu$.

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 $A \in \Sigma$. 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.

Assume that $f, f_n : X \to \overline{\mathbf{R}}$ *or* $\overline{\mathbf{C}}$ *are measurable, that* $g : X \to [0, +\infty]$ *has* $\int_A g\,d\mu < +\infty$*, that* $|f_n| \leq g$ *a.e. on* A *and* $f_n \to f$ *a.e. on* A*. The result is that* $\int_A f_n\,d\mu \to \int_A f\,d\mu$.

Indeed, we have then that $\int_X g\chi_A\,d\mu < +\infty$, that $|f_n\chi_A| \leq g\chi_A$ a.e. on X and $f_n\chi_A \to f\chi_A$ a.e. on X. The usual form of the dominated convergence theorem (for X) implies that $\int_A f_n\,d\mu = \int_X f_n\chi_A\,d\mu \to \int_X f\chi_A\,d\mu = \int_A f\,d\mu$.

Alternatively, we observe that $\int_A (g\rceil A)\,d(\mu\rceil A) < +\infty$, that $|f_n\rceil A| \leq g\rceil A$ a.e. on A and $f_n\rceil A \to f\rceil A$ a.e. on A. The dominated convergence theorem (for A) implies that $\int_A f_n\,d\mu = \int_A (f_n\rceil A)\,d(\mu\rceil A) \to \int_A (f\rceil A)\,d(\mu\rceil A) = \int_A f\,d\mu$.

7.5 Point-mass distributions.

Consider the point-mass distribution μ induced by a function $a : X \to [0, +\infty]$ through the formula

$$\mu(E) = \sum_{x \in E} a_x$$

for all $E \subseteq X$.

We observe that *all* functions $f : X \to Y$, no matter what the (Y, Σ') is, are (Σ, Σ')-measurable.

If $\phi = \sum_{j=1}^n \kappa_j \chi_{E_j}$ is any non-negative simple function on X with its standard representation, then $\int_X \phi \, d\mu = \sum_{j=1}^n \kappa_j \mu(E_j) = \sum_{j=1}^n \kappa_j \big(\sum_{x \in E_j} a_x \big) = \sum_{j=1}^n \big(\sum_{x \in E_j} \kappa_j a_x \big) = \sum_{j=1}^n \big(\sum_{x \in E_j} \phi(x) a_x \big)$. We apply Proposition 2.6 to get

$$\int_X \phi \, d\mu = \sum_{x \in X} \phi(x) \, a_x .$$

Proposition 7.7 *If* $f : X \to [0, +\infty]$ *then*

$$\int_X f \, d\mu = \sum_{x \in X} f(x) \, a_x .$$

Proof: Consider an increasing sequence (ϕ_n) of non-negative simple functions so that $\phi_n \uparrow f$ on X and $\int_X \phi_n \, d\mu \uparrow \int_X f \, d\mu$.

Then $\int_X \phi_n \, d\mu = \sum_{x \in X} \phi_n(x) \, a_x \leq \sum_{x \in X} f(x) \, a_x$ and, taking limit in n, we find $\int_X f \, d\mu \leq \sum_{x \in X} f(x) \, a_x$.

If F is a finite subset of X, then $\sum_{x \in F} \phi_n(x) \, a_x \leq \sum_{x \in X} \phi_n(x) \, a_x = \int_X \phi_n \, d\mu$. Using the obvious $\lim_{n \to +\infty} \sum_{x \in F} \phi_n(x) \, a_x = \sum_{x \in F} f(x) \, a_x$, we find $\sum_{x \in F} f(x) \, a_x \leq \int_X f \, d\mu$. Taking supremum over F, $\sum_{x \in X} f(x) \, a_x \leq \int_X f \, d\mu$ and, combining with the opposite inequality, the proof is finished.

We would like to extend the validity of this Proposition 7.7 to real valued or complex valued functions, but we do not have a definition for sums of 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 identify $b_i = b(i)$). This is a simple application of Proposition 7.7.

Using properties of integrals we may prove corresponding properties of sums. For example, it is true that

$$\sum_{i \in I} (b_i + c_i) = \sum_{i \in I} b_i + \sum_{i \in I} c_i, \qquad \sum_{i \in I} \lambda b_i = \lambda \sum_{i \in I} b_i$$

for every non-negative b_i, c_i and λ. The proof consists in rewriting $\int_I (b+c)\, d\sharp = \int_I b\, d\sharp + \int_I c\, d\sharp$ and $\int_I \lambda b\, d\sharp = \lambda \int_I b\, d\sharp$ in terms of sums.

For every $b \in \overline{\mathbf{R}}$ we write $b^+ = \max\{b, 0\}$ and $b^- = -\min\{b, 0\}$ and, clearly, $b = b^+ - b^-$ and $|b| = b^+ + b^-$.

Definition 7.7 *If I is any index set and $b : I \to \overline{\mathbf{R}}$, 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 counting measure $\sharp$ or, equivalently, if and only if $\sum_{i\in I} |b_i| = \int_I |b|\, d\sharp < +\infty$.* Also, *the $\sum_{i\in I} b_i$ is defined if and only if the $\int_I b\, d\sharp$ is defined and they are equal.*

Further exploiting the analogy between sums and integrals we have

Definition 7.8 *If I is any index set and $b : I \to \overline{\mathbf{C}}$, 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 $b : I \to \overline{\mathbf{R}}$.

Proposition 7.8 *Let $b : I \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$. Then $\{b_i\}_{i\in I}$ is summable over I if and only if the set $\{i \in I \mid b_i \neq 0\}$ is countable and, taking an arbitrary enumeration $\{i_1, i_2, \ldots\}$ of it, $\sum_{k=1}^{+\infty} |b_{i_k}| < +\infty$.*

Proof: An application of Propositions 2.3 and 2.4.

In particular, if $\{b_i\}_{i\in I}$ is summable over I then b_i is *finite* for all $i \in I$. This allows us to give the

Definition 7.9 *Let $b : I \to \overline{\mathbf{C}}$ be summable over I. We define* ***the sum of $\{b_i\}_{i\in I}$ over I*** *as*

$$\sum_{i\in I} b_i = \sum_{i\in I} \Re(b_i) + i \sum_{i\in I} \Im(b_i).$$

Therefore, the sum of complex valued terms is defined only when the sum is summable and, hence, this sum always has a finite value. Again, we can say that if $b : I \to \overline{\mathbf{C}}$ is summable over I (which is equivalent to b being integrable over I with respect to 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 of 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.* Proofs are not needed.

Theorem 7.14 *(i) (**The Monotone Convergence Theorem**) Let* $b, b^{(k)} : I \to [0, +\infty]$ *(*$k \in \mathbf{N}$*). If* $b_i^{(k)} \uparrow b_i$ *for all* i*, then* $\sum_{i\in I} b_i^{(k)} \uparrow \sum_{i \in I} b_i$*.*
(ii) Let $b^{(k)} : I \to [0, +\infty]$ *(*$k \in \mathbf{N}$*). Then* $\sum_{i\in I}(\sum_{k=1}^{+\infty} b_i^{(k)}) = \sum_{k=1}^{+\infty}(\sum_{i\in I} b_i^{(k)})$*.*
*(iii) (**The Lemma of Fatou**) Let* $b, b^{(k)} : I \to [0, +\infty]$ *(*$k \in \mathbf{N}$*). If* $b_i = \liminf_{k\to+\infty} b_i^{(k)}$ *for all* $i \in I$*, then* $\sum_{i\in I} b_i \leq \liminf_{k\to+\infty} \sum_{i\in I} b_i^{(k)}$*.*
*(iv) (**The Dominated Convergence Theorem**) Let* $b, b^{(k)} : I \to \overline{\mathbf{R}}$ *or* $\overline{\mathbf{C}}$ *(*$k \in \mathbf{N}$*) and* $c : I \to [0, +\infty]$*. If* $|b_i^{(k)}| \leq c_i$ *for all* i *and* k*, if* $\sum_{i\in I} c_i < +\infty$ *and if* $b_i^{(k)} \to b_i$ *for all* i*, then* $\sum_{i\in I} b_i^{(k)} \to \sum_{i\in I} b_i$*.*
*(v) (**The Series Theorem**) Let* $b^{(k)} : I \to \overline{\mathbf{R}}$ *or* $\overline{\mathbf{C}}$ *(*$k \in \mathbf{N}$*). Assuming that* $\sum_{k=1}^{+\infty}(\sum_{i\in I} |b_i^{(k)}|) < +\infty$*, then* $\sum_{k=1}^{+\infty} b_i^{(k)}$ *converges for every* i*. Moreover,* $\sum_{i\in I}(\sum_{k=1}^{+\infty} b_i^{(k)}) = \sum_{k=1}^{+\infty}(\sum_{i\in I} b_i^{(k)})$*.*

Observe that the only $\sharp$-null set is the $\emptyset$. Therefore, saying that a property holds $\sharp$-a.e. on I is equivalent to saying that it holds at every point of I.

Going back to the general case, if μ is the point-mass distribution induced by the function $a : X \to [0, +\infty]$, and $f : X \to \overline{\mathbf{R}}$, then $\int_X f \, d\mu$ is defined if and only if either $\sum_{x\in X} f^+(x)a_x = \int_X f^+ \, d\mu < +\infty$ or $\sum_{x\in X} f^-(x)a_x = \int_X f^- \, d\mu < +\infty$, 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 $\sum_{x\in X} |f(x)|a_x = \int_X |f| \, d\mu < +\infty$. This is also true when $f : X \to \overline{\mathbf{C}}$, and in this case we have

$$\int_X f \, d\mu = \sum_{x\in X} \Re(f(x)\chi_{D_f}(x))a_x + i \sum_{x\in X} \Im(f(x)\chi_{D_f}(x))a_x,$$

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 $f(x)a_x \in \mathbf{C}$ 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

$$\int_X f\,d\mu = \sum_{x\in X} \Re(f(x)a_x) + i\sum_{x\in X} \Im(f(x)a_x) = \sum_{x\in X} f(x)a_x.$$

Now we have arrived at the complete interpretation of sums as integrals.

Theorem 7.15 *Let μ be a point-mass distribution induced by $a : X \to [0,+\infty]$. If $f : X \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$, then the $\int_X f\,d\mu$ exists if and only if the $\sum_{x\in X} f(x)a_x$ exists and, in this case,*

$$\int_X f\,d\mu = \sum_{x\in X} f(x)a_x.$$

A simple particular case of a point-mass distribution is the Dirac mass δ_{x_0} at $x_0 \in X$. We remember that this is induced by $a_x = 1$ if $x = x_0$ and $a_x = 0$ if $x \neq x_0$. In this case the integrals become very simple:

$$\int_X f\,d\delta_{x_0} = f(x_0)$$

for every f. It is clear that f is integrable if and only if $f(x_0) \in \mathbf{C}$. Thus, *integration with respect to the Dirac mass at x_0 coincides with the so-called* ***point evaluation at*** x_0.

7.6 Lebesgue integral.

A function $f : \mathbf{R}^n \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$ is called Lebesgue integrable if it is integrable with respect to m_n.

It is easy to see that every continuous $f : \mathbf{R}^n \to \mathbf{R}$ or $\mathbf{C}$ which is 0 outside some bounded set is Lebesgue integrable. Indeed, f is then Borel measurable and if Q is any closed interval in $\mathbf{R}^n$ outside of which f is 0, then $|f| \leq K\chi_Q$, where $K = \max\{|f(x)|\,|\,x \in Q\} < +\infty$. Therefore, $\int_{\mathbf{R}^n} |f|\,dm_n \leq K m_n(Q) < +\infty$ and f is Lebesgue integrable.

Theorem 7.16 ***(Approximation)*** *Let $f : \mathbf{R}^n \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$ be Lebesgue integrable. Then for every $\epsilon > 0$ there is some continuous function $g : \mathbf{R}^n \to \mathbf{R}$ or $\mathbf{C}$ which is 0 outside some bounded set so that $\int_{\mathbf{R}^n} |g - f|\,dm_n < \epsilon$.*

Proof: (a) Let $-\infty < a < b < +\infty$ and for each $\delta \in (0, \frac{b-a}{2})$ consider the *continuous* function $\tau_{a,b,\delta} : \mathbf{R} \to [0,1]$ which is 1 on $(a+\delta, b-\delta)$, is 0 outside (a,b) and is linear in each of $[a, a+\delta]$ and $[b-\delta, b]$.

Let $R = (a_1,b_1) \times \cdots \times (a_n,b_n)$ be an open interval in $\mathbf{R}^n$. Consider, for small $\delta > 0$, the open interval $R_\delta = (a_1+\delta, b_1-\delta) \times \cdots \times (a_n+\delta, b_n-\delta) \subseteq R$. Then it is clear that, by choosing δ small enough, we can have $m_n(R \setminus R_\delta) < \epsilon$. Define the function $\tau_{R,\delta} : \mathbf{R}^n \to [0,1]$ by the formula

$$\tau_{R,\delta}(x_1,\ldots,x_n) = \tau_{a_1,b_1,\delta}(x_1)\cdots\tau_{a_n,b_n,\delta}(x_n).$$

Then, $\tau_{R,\delta}$ is continuous on $\mathbf{R}^n$, it is 1 on R_δ and it is 0 outside R. Therefore, $\int_{\mathbf{R}^n} |\tau_{R,\delta} - \chi_R|\, m_n \leq m_n(R \setminus R_\delta) < \epsilon$.
(b) Let $E \in \mathcal{L}_n$ have $m_n(E) < +\infty$. Theorem 4.6 implies that there are pairwise disjoint open intervals $R_1, \ldots, R_l$ so that $m_n(E \triangle (R_1 \cup \cdots \cup R_l)) < \frac{\epsilon}{2}$. The functions χ_E and $\chi_{R_1} + \cdots + \chi_{R_l}$ differ (by at most 1) only in the set $E \triangle (R_1 \cup \cdots \cup R_l)$. Hence, $\int_{\mathbf{R}^n} |\sum_{i=1}^{l} \chi_{R_i} - \chi_E|\, dm_n < \frac{\epsilon}{2}$.

By (a), we can take small enough $\delta > 0$ so that, for each R_i, we have $\int_{\mathbf{R}^n} |\tau_{R_i,\delta} - \chi_{R_i}|\, m_n < \frac{\epsilon}{2l}$. This implies $\int_{\mathbf{R}^n} |\sum_{i=1}^{l} \tau_{R_i,\delta} - \sum_{i=1}^{l} \chi_{R_i}|\, dm_n < \sum_{i=1}^{l} \frac{\epsilon}{2l} = \frac{\epsilon}{2}$.

Denoting $\psi = \sum_{i=1}^{l} \tau_{R_i,\delta} : \mathbf{R}^n \to \mathbf{R}$, we have $\int_{\mathbf{R}^n} |\psi - \chi_E|\, dm_n < \epsilon$. Observe that ψ is a continuous function which is 0 outside the bounded set $\cup_{i=1}^{l} R_i$.
(c) Let now $f : \mathbf{R}^n \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$ be Lebesgue integrable. From Theorem 7.12 we know that there is some Lebesgue measurable simple $\psi : \mathbf{R}^n \to \mathbf{R}$ or $\mathbf{C}$ so that $\int_{\mathbf{R}^n} |\psi - f|\, dm_n < \frac{\epsilon}{2}$. Let $\psi = \sum_{j=1}^{m} \kappa_j \chi_{E_j}$ be the standard representation of ψ, where we omit the possible value $\kappa = 0$. From $\sum_{j=1}^{m} |\kappa_j| m_n(E_j) = \int_{\mathbf{R}^n} |\psi|\, dm_n \leq \int_{\mathbf{R}^n} |f|\, dm_n + \int_{\mathbf{R}^n} |f - \psi|\, dm_n < +\infty$, we get that $m_n(E_j) < +\infty$ for all j. By part (b), for each E_j we can find a continuous $\psi_j : \mathbf{R}^n \to \mathbf{R}$ so that $\int_{\mathbf{R}^n} |\psi_j - \chi_{E_j}|\, dm_n < \frac{\epsilon}{2m|\kappa_j|}$.

If we set $g = \sum_{j=1}^{m} \kappa_j \psi_j$, then this is continuous on $\mathbf{R}^n$ and

$$\begin{aligned}
\int_{\mathbf{R}^n} |g - f|\, dm_n &\leq \int_{\mathbf{R}^n} |g - \psi|\, dm_n + \int_{\mathbf{R}^n} |\psi - f|\, dm_n \\
&< \sum_{j=1}^{m} \int_{\mathbf{R}^n} |\kappa_j \psi_j - \kappa_j \chi_{E_j}|\, dm_n + \frac{\epsilon}{2} \\
&< \sum_{j=1}^{m} |\kappa_j| \frac{\epsilon}{2m|\kappa_j|} + \frac{\epsilon}{2} = \epsilon.
\end{aligned}$$

Since each ψ_j is 0 outside a bounded set, g is also 0 outside a bounded set.

We shall now investigate the relation between the Lebesgue integral and the Riemann integral. We recall the definition of the latter.

Assume that $Q = [a_1, b_1] \times \cdots \times [a_n, b_n]$ is a closed interval in $\mathbf{R}^n$ and consider a *bounded* function $f : Q \to \mathbf{R}$.

If $l \in \mathbf{N}$ 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, in 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 **upper Darboux sum** and the **lower Darboux sum of f with respect to Δ** as, respectively,

$$\overline{\Sigma}(f;\Delta) = \sum_{j=1}^{l} M_j \operatorname{vol}_n(Q_j),$$

$$\underline{\Sigma}(f;\Delta) = \sum_{j=1}^{l} m_j \operatorname{vol}_n(Q_j).$$

If $m = \inf\{f(x) \,|\, x \in Q\}$ and $M = \sup\{f(x) \,|\, x \in Q\}$, we have that $m \le m_j \le M_j \le M$ for every j and, using Lemma 4.2, we see that

$$m \operatorname{vol}_n(Q) \le \underline{\Sigma}(f;\Delta) \le \overline{\Sigma}(f;\Delta) \le M \operatorname{vol}_n(Q).$$

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 $Q_i^{(2)}$'s of Δ_2 which are included in $Q_j^{(1)}$ cover it and hence form a partition of it. Therefore, from Lemma 4.2 again,

$$\begin{aligned} m_j^{(1)} \operatorname{vol}_n(Q_j^{(1)}) &\le \sum_{Q_i^{(2)} \subseteq Q_j^{(1)}} m_i^{(2)} \operatorname{vol}_n(Q_i^{(2)}) \\ &\le \sum_{Q_i^{(2)} \subseteq Q_j^{(1)}} M_i^{(2)} \operatorname{vol}_n(Q_i^{(2)}) \le M_j^{(1)} \operatorname{vol}_n(Q_j^{(1)}). \end{aligned}$$

Summing over all $j = 1, \ldots, l_1$ we find

$$\underline{\Sigma}(f;\Delta_1) \le \underline{\Sigma}(f;\Delta_2) \le \overline{\Sigma}(f;\Delta_2) \le \overline{\Sigma}(f;\Delta_1).$$

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 $\Delta = \{Q_j^{(1)} \cap Q_i^{(2)} \,|\, 1 \le j \le l_1, 1 \le i \le l_2\}$. Then, $\underline{\Sigma}(f;\Delta_1) \le \underline{\Sigma}(f;\Delta) \le \overline{\Sigma}(f;\Delta) \le \overline{\Sigma}(f;\Delta_2)$ and we conclude that

$$m \operatorname{vol}_n(Q) \le \underline{\Sigma}(f;\Delta_1) \le \overline{\Sigma}(f;\Delta_2) \le M \operatorname{vol}_n(Q)$$

for all partitions Δ_1, Δ_2 of Q. We now define

$$(\mathcal{R}_n)\underline{\int}_Q f = \sup\{\underline{\Sigma}(f;\Delta) \,|\, \Delta \text{ partition of } Q\}$$

$$(\mathcal{R}_n)\overline{\int}_Q f = \inf\{\overline{\Sigma}(f;\Delta) \,|\, \Delta \text{ partition of } Q\}$$

and call them, respectively, **the lower Riemann integral** and **the upper Riemann integral of f over Q**. It is then clear that

$$m \operatorname{vol}_n(Q) \le (\mathcal{R}_n)\underline{\int}_Q f \le (\mathcal{R}_n)\overline{\int}_Q f \le M \operatorname{vol}_n(Q).$$

We say that f is **Riemann integrable over** Q if $(\mathcal{R}_n)\underline{\int}_Q f = (\mathcal{R}_n)\overline{\int}_Q f$ and in this case we define

$$(\mathcal{R}_n)\int_Q f = (\mathcal{R}_n)\underline{\int}_Q f = (\mathcal{R}_n)\overline{\int}_Q f$$

to be the **Riemann integral of** f **over** Q.

Lemma 7.11 *Let Q be a closed interval in $\mathbf{R}^n$ and $f : Q \to \mathbf{R}$ be bounded. Then f is Riemann integrable over Q if and only if for every $\epsilon > 0$ there is some partition Δ of Q so that $\overline{\Sigma}(f;\Delta) - \underline{\Sigma}(f;\Delta) < \epsilon$.*

Proof: To prove the sufficiency, take arbitrary $\epsilon > 0$ and the corresponding Δ. Then $0 \le (\mathcal{R}_n)\overline{\int}_Q f - (\mathcal{R}_n)\underline{\int}_Q f \le \overline{\Sigma}(f;\Delta) - \underline{\Sigma}(f;\Delta) < \epsilon$. Taking the limit as $\epsilon \to 0+$, we prove the equality of the upper Riemann integral and the lower Riemann integral of f over Q.

For the necessity, assume $(\mathcal{R}_n)\underline{\int}_Q f = (\mathcal{R}_n)\overline{\int}_Q f$ and for each $\epsilon > 0$ take partitions Δ_1, Δ_2 of Q so that $(\mathcal{R}_n)\int_Q f - \frac{\epsilon}{2} < \underline{\Sigma}(f;\Delta_1)$ and $\overline{\Sigma}(f;\Delta_2) < (\mathcal{R}_n)\int_Q f + \frac{\epsilon}{2}$. Therefore, if Δ is the common refinement of Δ_1 and Δ_2, then $\overline{\Sigma}(f;\Delta) - \underline{\Sigma}(f;\Delta) \le \overline{\Sigma}(f;\Delta_2) - \underline{\Sigma}(f;\Delta_1) < \epsilon$.

Proposition 7.9 *Let Q be a closed interval in $\mathbf{R}^n$ and $f : Q \to \mathbf{R}$ be continuous on Q. Then f is Riemann integrable over Q.*

Proof: Since f is uniformly continuous on Q, given 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 in the same Q_j. This implies that for every Q_j we have $M_j - m_j = \max\{f(x) \mid x \in Q_j\} - \min\{f(y) \mid y \in Q_j\} < \frac{\epsilon}{\mathrm{vol}_n(Q)}$.

Hence

$$\overline{\Sigma}(f;\Delta) - \underline{\Sigma}(f;\Delta) = \sum_{j=1}^{l}(M_j - m_j)\,\mathrm{vol}_n(Q_j) < \frac{\epsilon}{\mathrm{vol}_n(Q)} \sum_{j=1}^{l} \mathrm{vol}_n(Q_j) = \epsilon$$

and Lemma 7.11 implies that f is Riemann integrable over Q.

Theorem 7.17 *Let Q be a closed interval in $\mathbf{R}^n$ and $f : Q \to \mathbf{R}$ be Riemann integrable over Q. If we extend f as 0 outside Q, then f is Lebesgue integrable and*

$$\int_{\mathbf{R}^n} f\,dm_n = \int_Q f\,dm_n = (\mathcal{R}_n)\int_Q f.$$

Proof: Lemma 7.11 implies that, for all $k \in \mathbf{N}$, there is a partition $\Delta_k = \{Q_1^{(k)}, \dots, Q_{l_k}^{(k)}\}$ of Q so that $\overline{\Sigma}(f;\Delta_k) - \underline{\Sigma}(f;\Delta_k) < \frac{1}{k}$. We consider the simple functions

$$\psi_k = \sum_{j=1}^{l_k} m_j^{(k)}\,\chi_{P_j^{(k)}}, \qquad \phi_k = \sum_{j=1}^{l_k} M_j^{(k)}\,\chi_{P_j^{(k)}},$$

where $P_j^{(k)}$ is the open-closed interval with the same sides as $Q_j^{(k)}$ and $m_j^{(k)} = \inf\{f(x) \,|\, x \in Q_j^{(k)}\}$, $M_j^{(k)} = \sup\{f(x) \,|\, x \in Q_j^{(k)}\}$.

From $\underline{\Sigma}(f;\Delta_k) \leq (\mathcal{R}_n)\int_Q f \leq \overline{\Sigma}(f;\Delta_k)$ we get that

$$\overline{\Sigma}(f;\Delta_k), \underline{\Sigma}(f;\Delta_k) \to (\mathcal{R}_n)\int_Q f.$$

It is clear that $\psi_k \leq f\chi_P \leq \phi_k$ for all k, where P is the open-closed interval with the same sides as Q. It is also clear that

$$\int_{\mathbf{R}^n} \psi_k \, dm_n = \sum_{j=1}^{l_k} m_j^{(k)} \operatorname{vol}_n(P_j^{(k)}) = \underline{\Sigma}(f;\Delta_k)$$

$$\int_{\mathbf{R}^n} \phi_k \, dm_n = \sum_{j=1}^{l_k} M_j^{(k)} \operatorname{vol}_n(P_j^{(k)}) = \overline{\Sigma}(f;\Delta_k).$$

Hence $\int_{\mathbf{R}^n}(\phi_k - \psi_k)\, dm_n < \frac{1}{k}$ for all k.

We define $g = \limsup_{k\to+\infty} \psi_k$ and $h = \liminf_{k\to+\infty} \phi_k$ and then, of course, $g \leq f\chi_P \leq h$. The Lemma of Fatou implies that

$$0 \leq \int_{\mathbf{R}^n}(h-g)\, dm_n \leq \liminf_{k\to+\infty}\int_{\mathbf{R}^n}(\phi_k - \psi_k)\, dm_n = 0.$$

By Proposition 7.3, $g = h$ a.e. on $\mathbf{R}^n$ and, thus, $f = g = h$ a.e. on $\mathbf{R}^n$. Since g, h are Borel measurable, Proposition 6.24 implies that f is Lebesgue measurable. f is also bounded and is 0 outside Q. Hence $|f| \leq K\chi_Q$, where $K = \sup\{|f(x)| \,|\, x \in Q\}$. Thus, $\int_{\mathbf{R}^n}|f|\, dm_n \leq K m_n(Q) < +\infty$ and f is Lebesgue integrable.

Another application of the Lemma of Fatou gives

$$\begin{aligned}\int_{\mathbf{R}^n}(h - f\chi_P)\, dm_n &\leq \liminf_{k\to+\infty}\int_{\mathbf{R}^n}(\phi_k - f\chi_P)\, dm_n \\ &= \liminf_{k\to+\infty}\overline{\Sigma}(f;\Delta_k) - \int_{\mathbf{R}^n} f\chi_P\, dm_n \\ &= (\mathcal{R}_n)\int_Q f - \int_{\mathbf{R}^n} f\chi_P\, dm_n.\end{aligned}$$

Hence $\int_{\mathbf{R}^n} h\, dm_n \leq (\mathcal{R}_n)\int_Q f$ and, similarly, $(\mathcal{R}_n)\int_Q f \leq \int_{\mathbf{R}^n} g\, dm_n$. Since $f = g = h$ a.e. on $\mathbf{R}^n$, we conclude that

$$(\mathcal{R}_n)\int_Q f = \int_{\mathbf{R}^n} f\, dm_n.$$

The converse of Theorem 7.17 does not hold. There are examples of bounded functions $f : Q \to \mathbf{R}$ which are Lebesgue integrable but not Riemann integrable over Q.

Example:
Define $f(x) = 1$, if $x \in Q$ has all its coordinates rational, and $f(x) = 0$, if $x \in Q$ has at least one of its coordinates irrational. If $\Delta = \{Q_1, \dots, Q_k\}$ is any partition of Q, then all Q_j's with non-empty interior (the rest do not matter because they have zero volume) contain at least one x with $f(x) = 1$ and at least one x with $f(x) = 0$. Hence, for all such Q_j we have $M_j = 1$ and $m_j = 0$. Hence, $\overline{\Sigma}(f; \Delta) = \mathrm{vol}_n(Q)$ and $\underline{\Sigma}(f; \Delta) = 0$ for every Δ. This says that $(\mathcal{R}_1)\overline{\int}_Q f = \mathrm{vol}_n(Q)$ and $(\mathcal{R}_n)\underline{\int}_Q f = 0$ and f is not Riemann integrable over Q.

On the other hand f extended as 0 outside Q is 0 a.e on $\mathbf{R}^n$ and hence it is Lebesgue integrable on $\mathbf{R}^n$ with $\int_{\mathbf{R}^n} f\,dm_n = \int_Q f\,dm_n = 0$.

Theorem 7.17 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 $\mathbf{R}$. The standard symbol of calculus for the Riemann integral $(\mathcal{R}_1) \int_{[a,b]} f$ is the familiar

$$\int_a^b f \qquad \text{or} \qquad \int_a^b f(x)\,dx.$$

We may now use the same symbol for the Lebesgue integral $\int_{[a,b]} f\,dm_1$ 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. Bear also in mind that 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, since the one-point sets $\{a\}$, $\{b\}$ have zero Lebesgue measure. 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 intgral does not apply. For example, we may use the symbol

$$\int_{-\infty}^{+\infty} f(x)\,dx$$

for the Lebesgue integral $\int_{\mathbf{R}} f\,dm_1$ 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 7.17 provides also with a tool to *calculate* Lebesgue integrals, at least in the case of $\mathbf{R}$. If a function is Riemann integrable over a closed interval $[a, b] \subseteq \mathbf{R}$, we have many techniques (integration by parts, change of variable, antiderivatives etc) to calculate its $(\mathcal{R}_1) \int_{[a,b]} f$ which is the same as $\int_{[a,b]} f\,dm_1$. In case the given f is Riemann integrable over intervals $[a_k, b_k]$ with $a_k \downarrow -\infty$ and $b_k \uparrow +\infty$ and we can calculate the integrals $(\mathcal{R}_1) \int_{[a_k,b_k]} f = \int_{[a_k,b_k]} f\,dm_1$, then it is a matter of being able to pass to the limit $\int_{[a_k,b_k]} f\,dm_1 \to \int_{\mathbf{R}} f\,dm_1$ to calculate the Lebesgue integral over $\mathbf{R}$. To do this we may try to use the Monotone Convergence Theorem or the Dominated Convergence Theorem.

Another topic is the change of Lebesgue integral under linear transformations of the space.

Proposition 7.10 *Let $T : \mathbf{R}^n \to \mathbf{R}^n$ be a linear transformation with $\det(T) \neq 0$. If (Y, Σ') is a measurable space and $f : \mathbf{R}^n \to Y$ is $(\mathcal{L}_n, \Sigma')$-measurable, then $f \circ T^{-1} : \mathbf{R}^n \to Y$ is also $(\mathcal{L}_n, \Sigma')$-measurable.*

Proof: For every $E \in \Sigma'$ we have $(f \circ T^{-1})^{-1}(E) = T(f^{-1}(E)) \in \mathcal{L}_n$, because of Theorem 4.8.

Theorem 7.18 *Let $T : \mathbf{R}^n \to \mathbf{R}^n$ be a linear transformation with $\det(T) \neq 0$ and $f : \mathbf{R}^n \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$ be Lebesgue measurable.*
(i) If $f : \mathbf{R}^n \to \overline{\mathbf{R}}$ and the $\int_{\mathbf{R}^n} f\,dm_n$ exists, then the $\int_{\mathbf{R}^n} f \circ T^{-1}\,dm_n$ also exists and

$$\int_{\mathbf{R}^n} f \circ T^{-1}\,dm_n = |\det(T)| \int_{\mathbf{R}^n} f\,dm_n.$$

(ii) If $f : \mathbf{R}^n \to \overline{\mathbf{C}}$ is integrable, then $f \circ T^{-1}$ is also integrable and the equality of (i) is again true.

Proof: (a) Let $\phi : \mathbf{R}^n \to [0, +\infty)$ be a non-negative Lebesgue measurable simple function and $\phi = \sum_{j=1}^m \kappa_j \chi_{E_j}$ be its standard representation. Then $\int_{\mathbf{R}^n} \phi\,dm_n = \sum_{j=1}^m \kappa_j m_n(E_j)$.

It is clear that $\phi \circ T^{-1} = \sum_{j=1}^m \kappa_j \chi_{E_j} \circ T^{-1} = \sum_{j=1}^m \kappa_j \chi_{T(E_j)}$ from which we get $\int_{\mathbf{R}^n} \phi \circ T^{-1}\,dm_n = \sum_{j=1}^m \kappa_j m_n(T(E_j)) = |\det(T)| \sum_{j=1}^m \kappa_j m_n(E_j) = |\det(T)| \int_{\mathbf{R}^n} \phi\,dm_n$.
(b) Let $f : \mathbf{R}^n \to [0, +\infty]$ be Lebesgue measurable. Take any increasing sequence (ϕ_k) of non-negative Lebesgue measurable simple functions so that $\phi_k \to f$ on $\mathbf{R}^n$. Then $(\phi_k \circ T^{-1})$ is increasing and $\phi_k \circ T^{-1} \to f \circ T^{-1}$ on $\mathbf{R}^n$. From part (a), $\int_{\mathbf{R}^n} f \circ T^{-1}\,dm_n = \lim_{k\to+\infty} \int_{\mathbf{R}^n} \phi_k \circ T^{-1}\,dm_n = |\det(T)| \lim_{k\to+\infty} \int_{\mathbf{R}^n} \phi_k\,dm_n = |\det(T)| \int_{\mathbf{R}^n} f\,dm_n$.
(c) Let $f : \mathbf{R}^n \to \overline{\mathbf{R}}$ and the $\int_{\mathbf{R}^n} f\,dm_n$ exist. Then $(f \circ T^{-1})^+ = f^+ \circ T^{-1}$ and $(f \circ T^{-1})^- = f^- \circ T^{-1}$, and from (b) we get $\int_{\mathbf{R}^n} (f \circ T^{-1})^+\,dm_n = |\det(T)| \int_{\mathbf{R}^n} f^+\,dm_n$ and $\int_{\mathbf{R}^n} (f \circ T^{-1})^-\,dm_n = |\det(T)| \int_{\mathbf{R}^n} f^-\,dm_n$. Now (i) is obvious.
(d) Let $f : \mathbf{R}^n \to \overline{\mathbf{C}}$ be integrable. From $|f \circ T^{-1}| = |f| \circ T^{-1}$ and from (b) we have that $\int_{\mathbf{R}^n} |f \circ T^{-1}|\,dm_n = |\det(T)| \int_{\mathbf{R}^n} |f|\,dm_n < +\infty$. Hence $f \circ T^{-1}$ is also integrable.

We take an $F : \mathbf{R}^n \to \mathbf{C}$ so that $F = f$ a.e. on $\mathbf{R}^n$.

If $A = \{x \in \mathbf{R}^n \,|\, F(x) \neq f(x)\}$ and $B = \{x \in \mathbf{R}^n \,|\, F \circ T^{-1}(x) \neq f \circ T^{-1}(x)\}$, then $B = T(A)$. Hence, $m_n(B) = |\det(T)| m_n(A) = 0$ and, thus, $F \circ T^{-1} = f \circ T^{-1}$ a.e. on $\mathbf{R}^n$. Therefore, to prove (ii) it is enough to prove $\int_{\mathbf{R}^n} F \circ T^{-1}\,dm_n = |\det(T)| \int_{\mathbf{R}^n} F\,dm_n$.

We have $\Re(F \circ T^{-1}) = \Re(F) \circ T^{-1}$ and, from part (c), $\int_{\mathbf{R}^n} \Re(F \circ T^{-1})\, dm_n = |\det(T)| \int_{\mathbf{R}^n} \Re(F)\, dm_n$. We, similarly, prove the same equality with the imaginary parts and, combining, we get the desired equality.

The equality of the two integrals in Theorem 7.18 is nothing but the **(linear) change of variable formula**. If we write $y = T^{-1}(x)$ or, equivalently, $x = T(y)$, then the equality reads

$$\int_{\mathbf{R}^n} f(T^{-1}(x))\, dm_n(x) = |\det(T)| \int_{\mathbf{R}^n} f(y)\, dm_n(y).$$

Thus, the *informal rule* for the change of differentials is

$$dm_n(x) = |\det(T)| dm_n(y).$$

7.7 Lebesgue-Stieltjes integral.

Let $-\infty \le a_0 < b_0 \le +\infty$. Every monotone $f : (a_0, b_0) \to \overline{\mathbf{R}}$ is Borel measurable. This is seen by observing that $f^{-1}((a, b])$ is an interval, and hence a Borel set, for every $(a, b]$. If, now, $F : (a_0, b_0) \to \mathbf{R}$ is another increasing function and μ_F is the induced Borel measure, then f satisfies the necessary measurability condition and the $\int_{(a_0 b_0)} f\, d\mu_F$ exists provided, as usual, that either $\int_{(a_0,b_0)} f^+\, d\mu_F < +\infty$ or $\int_{(a_0,b_0)} f^-\, d\mu_F < +\infty$.

The same can, of course, be said when f is continuous.

In particular, *if f is continuous or monotone in a (bounded) interval $S \subseteq (a_0, b_0)$ and it is bounded on S, then it is integrable over S with respect to μ_F.*

We shall prove three classical results about Lebesgue-Stieltjes integrals.

Observe that the four integrals which we get from $\int_S f\, d\mu_F$, by taking $S = [a, b], [a, b), (a, b]$ and (a, b), may be different. This is because the $\int_{\{a\}} f\, d\mu_F = f(a)\mu_F(\{a\}) = f(a)(F(a+) - F(a-))$ and $\int_{\{b\}} f\, d\mu_F = f(b)(F(b+) - F(b-))$ may not be zero.

Proposition 7.11 *(Integration by parts) Let $F, G : (a_0, b_0) \to \mathbf{R}$ 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 $a \le b$. 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 in the right side of the above equality.

Proof: We introduce a sequence of partitions $\Delta_k = \{c_0^{(k)}, \dots, c_{l_k}^{(k)}\}$ of $[a, b]$ so that $a = c_0^{(k)} < c_1^{(k)} < \cdots < c_{l_k}^{(k)} = b$ for each k and so that

$$\lim_{k \to +\infty} \max\{c_j^{(k)} - c_{j-1}^{(k)} \,|\, 1 \le j \le l_k\} = 0.$$

We also introduce the simple functions

$$g_k = \sum_{j=1}^{l_k} G(c_j^{(k)}+)\chi_{(c_{j-1}^{(k)},c_j^{(k)}]}, \qquad f_k = \sum_{j=1}^{l_k} F(c_{j-1}^{(k)}+)\chi_{(c_{j-1}^{(k)},c_j^{(k)}]}.$$

It is clear that $G(a+) \leq g_k \leq G(b+)$ and $F(a+) \leq f_k \leq F(b-)$ for all k.

If, for an arbitrary $x \in (a,b]$ we take the interval $(c_{j-1}^{(k)}, c_j^{(k)}]$ containing x (observe that $j = j(k,x)$), then $g_k(x) = G(c_j^{(k)}+)$ and $f_k(x) = F(c_{j-1}^{(k)}+)$. Since $\lim_{k\to+\infty}(c_j^{(k)} - c_{j-1}^{(k)}) \to 0$, we have that $c_j^{(k)} \to x$ and $c_{j-1}^{(k)} \to x$. Therefore,

$$g_k(x) \to G(x+), \qquad f_k(x) \to F(x-)$$

as $k \to +\infty$.

We apply the Dominated Convergence Theorem to find

$$\sum_{j=1}^{l_k} G(c_j^{(k)}+)(F(c_j^{(k)}+) - F(c_{j-1}^{(k)}+)) = \int_{(a,b]} g_k(x)\, d\mu_F(x) \to \int_{(a,b]} G(x+)\, d\mu_F(x)$$

$$\sum_{j=1}^{l_k} F(c_{j-1}^{(k)}+)(G(c_j^{(k)}+) - G(c_{j-1}^{(k)}+)) = \int_{(a,b]} f_k(x)\, d\mu_G(x) \to \int_{(a,b]} F(x-)\, d\mu_G(x)$$

as $k \to +\infty$.

Adding the two last relations we find

$$G(b+)F(b+) - G(a+)F(a+) = \int_{(a,b]} G(x+)\, d\mu_F(x) + \int_{(a,b]} F(x-)\, d\mu_G(x).$$

If we want the integrals over (a,b), we have to *subtract* from the right side of the equality the quantity $\int_{\{b\}} G(x+)\, d\mu_F(x) + \int_{\{b\}} F(x-)\, d\mu_G(x)$ which is equal to $G(b+)(F(b+) - F(b-)) + F(b-)(G(b+) - G(b-)) = G(b+)F(b+) - G(b-)F(b-)$. Then, subtracting the same quantity from the left side of the equality, this becomes $F(b-)G(b-) - F(a+)G(a+)$. We work in the same way for all other types of intervals.

The next two results concern the reduction of Lebesgue-Stieltjes integrals to Lebesgue integrals. This makes calculation of the former more accessible in many situations.

Proposition 7.12 *Assume that* $F : (a_0, b_0) \to \mathbf{R}$ *is increasing and has a continuous derivative on* (a_0, b_0). *Then*

$$\mu_F(E) = \int_E F'(x)\, dm_1(x)$$

for every Borel set $E \subseteq (a_0, b_0)$ *and*

$$\int_{(a_0,b_0)} f(x)\, d\mu_F(x) = \int_{(a_0,b_0)} f(x)F'(x)\, dm_1(x)$$

for every Borel measurable $f : (a_0, b_0) \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$ for which either of the two integrals exists.

Proof: (i) The assumptions on F imply that it is continuous and that $F' \geq 0$ on (a_0, b_0). For every $[a, b] \subseteq (a_0, b_0)$ we have, by the Fundamental Theorem of Calculus for Riemann integrals, that $\int_{(a,b]} F'(x)\, dm_1(x) = \int_{[a,b]} F'(x)\, dm_1(x) = F(b) - F(a) = F(b+) - F(a+) = \mu_F((a, b])$. If we apply this to a strictly monotone sequence $a_n \downarrow a$, we get, by the Monotone Convergence Theorem, that $\int_{(a,b]} F'(x)\, dm_1(x) = \mu_F((a, b])$ for every $(a, b] \subseteq (a_0, b_0)$.

We now introduce the Borel measure μ on (a_0, b_0) by the formula

$$\mu(E) = \int_E F'(x)\, dm_1(x)$$

for every Borel set $E \subseteq (a_0, b_0)$. Clearly $\mu(\emptyset) = 0$ and $\mu(E) \geq 0$ for all Borel $E \subseteq (a_0, b_0)$. The σ-additivity of μ is an immediate consequence of Theorem 7.13.

Now, from the first paragraph, we have $\mu((a, b]) = \mu_F((a, b])$ for every $(a, b] \subseteq (a_0, b_0)$. Theorem 5.5 implies that $\mu = \mu_F$ and hence

$$\mu_F(E) = \int_E F'(x)\, dm_1(x)$$

for every Borel set $E \subseteq (a_0, b_0)$.

Taking arbitrary linear combinations of characteristic functions, we find that $\int_{(a_0,b_0)} \phi(x)\, d\mu_F(x) = \int_{(a_0,b_0)} \phi(x) F'(x)\, dm_1(x)$ for all Borel measurable simple functions $\phi : (a_0, b_0) \to [0, +\infty)$. Now, applying the Monotone Convergence Theorem to an appropriate increasing sequence of simple functions, we get

$$\int_{(a_0,b_0)} f(x)\, d\mu_F(x) = \int_{(a_0,b_0)} f(x) F'(x)\, dm_1(x)$$

for every Borel measurable $f : (a_0, b_0) \to [0, +\infty]$. The proof is easily concluded for any $f : (a_0, b_0) \to \overline{\mathbf{R}}$, by taking its positive and negative parts, and then for any $f : (a_0, b_0) \to \overline{\mathbf{C}}$, by taking its real and imaginary parts (and paying attention to the set where $f = \infty$).

Proposition 7.13 *Assume that $F : (a_0, b_0) \to \mathbf{R}$ is increasing and $G : (a, b) \to \mathbf{R}$ has a bounded, continuous derivative on (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) By the assumptions on G we have that it is continuous on (a, b) and that the limits $G(b-)$ and $G(a+)$ exist and they are numbers. We then extend G as $G(b-)$ on $[b, b_0)$ and as $G(a+)$ on $(a_0, a]$ and G becomes continuous on

(a_0, b_0). We use the same partitions Δ_k as in the proof of Proposition 7.11 and the same simple functions

$$g_k = \sum_{j=1}^{l_k} G(c_j^k+)\chi_{(c_{j-1}^k, c_j^k]} = \sum_{j=1}^{l_k} G(c_j^k)\chi_{(c_{j-1}^k, c_j^k]}.$$

We have again that $|g_k| \leq M$ where $M = \sup\{|G(x)| \,|\, x \in [a,b]\}$ and that $g_k(x) \to G(x+) = G(x)$ for every $x \in (a,b]$. By the Dominated Convergence Theorem,

$$\sum_{j=1}^{l_k} G(c_j^k)(F(c_j^k+) - F(c_{j-1}^k+)) = \int_{(a,b]} g_k(x)\, d\mu_F(x) \to \int_{(a,b]} G(x)\, d\mu_F(x)$$

as $k \to +\infty$.

By the mean value theorem, for every j with $j = 1, \ldots, l_k$, we have

$$G(c_j^k) - G(c_{j-1}^k) = G'(\xi_j^k)(c_j^k - c_{j-1}^k)$$

for some $\xi_j^k \in (c_{j-1}^k, c_j^k)$. Hence

$$\begin{aligned}\sum_{j=1}^{l_k} F(c_{j-1}^k+)(G(c_j^k) - G(c_{j-1}^k)) &= \sum_{j=1}^{l_k} F(c_{j-1}^k+)G'(\xi_j^k)(c_j^k - c_{j-1}^k) \\ &= \int_{(a,b)} \phi_k(x)\, dm_1(x),\end{aligned}$$

where we set $\phi_k = \sum_{j=1}^{l_k} F(c_{j-1}^k+)G'(\xi_j^k)\chi_{(c_{j-1}^k, c_j^k]}$.

We have that $\phi_k(x) \to F(x-)G'(x)$ for every $x \in (a,b)$ and that $|\phi_k| \leq K$ on (a,b) for some K which does not depend on k. By the Dominated Convergence Theorem, $\int_{(a,b)} \phi_k(x)\, dm_1(x) \to \int_{(a,b)} F(x-)G'(x)\, dm_1(x)$. We combine to get

$$G(b)F(b+) - G(a)F(a+) = \int_{(a,b]} G(x)\, d\mu_F(x) + \int_{(a,b)} F(x-)G'(x)\, dm_1(x).$$

From both sides we subtract $\int_{\{b\}} G(x)\, d\mu_F(x) = G(b)(F(b+) - F(b-))$ to find

$$G(b)F(b-) - G(a)F(a+) = \int_{(a,b)} G(x)\, d\mu_F(x) + \int_{(a,b)} F(x-)G'(x)\, dm_1(x),$$

which is the first equality in the statement of the proposition. The second equality is proved in a similar way.

(B) There is a second proof making no use of partitions.

Assume first that G is also *increasing* in (a,b). Then its extension as $G(a+)$ on $(a_0, a]$ and as $G(b-)$ on $[b, b_0)$ is increasing in (a_0, b_0). We apply Proposition 7.11 to get

$$\int_{(a,b)} G(x)\, d\mu_F(x) = G(b-)F(b-) - G(a+)F(a+) - \int_{(a,b)} F(x-)\, d\mu_G(x),$$

which, by Proposition 7.12, becomes the desired

$$\int_{(a,b)} G(x)\,d\mu_F(x) = G(b-)F(b-) - G(a+)F(a+) - \int_{(a,b)} F(x-)G'(x)\,dm_1(x).$$

If G is not increasing, we take an arbitrary $x_0 \in (a,b)$ and write $G(x) = G(x_0) + \int_{(x_0,x)} G'(t)\,dm_1(t)$ for every $x \in (a,b)$. Now, $(G')^+$ and $(G')^-$ are non-negative, continuous, bounded functions on (a,b) and we can write $G = G_1 - G_2$ on (a,b), where

$$G_1(x) = G(x_0) + \int_{(x_0,x)} (G')^+(t)\,dm_1(t), \qquad G_2(x) = \int_{(x_0,x)} (G')^-(t)\,dm_1(t)$$

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 bounded, continuous derivative on (a,b) and from the previous paragraph 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 the two equalities and prove the desired equality.

It is worth keeping in mind the fact, which is included in the second proof of Proposition 7.13, that an arbitrary G with a continuous, bounded derivative on an interval (a,b) can be decomposed as a difference, $G = G_1 - G_2$, of two *increasing* functions with a continuous, bounded derivative on (a,b). We shall generalise it later in the context of functions of bounded variation.

7.8 Reduction to integrals over R.

Let (X, Σ, μ) be a measure space.

Definition 7.10 *Let $f : X \to [0,+\infty]$ be measurable. Then the function $\lambda_f : [0,+\infty) \to [0,+\infty]$, defined by*

$$\lambda_f(t) = \mu(\{x \in X \mid t < f(x)\}),$$

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 $[0,+\infty)$. Since $\{x \in X \mid t_n < f(x)\} \uparrow \{x \in X \mid t < f(x)\}$ for every $t_n \downarrow t$, we see 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 in the interval $(t_0,+\infty)$ (which may be empty).

Proposition 7.14 *Let $f : X \to [0,+\infty]$ be measurable and $G : \mathbf{R} \to \mathbf{R}$ 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 $\phi = \sum_{j=1}^m \kappa_j \chi_{E_j}$ be a non-negative measurable simple function on X with its standard representation, where we omit the value 0. Rearrange so that $0 < \kappa_1 < \cdots < \kappa_m$ and then

$$\lambda_\phi(t) = \begin{cases} \mu(E_1) + \mu(E_2) + \cdots + \mu(E_m), & \text{if } 0 \le t < \kappa_1 \\ \mu(E_2) + \cdots + \mu(E_m), & \text{if } \kappa_1 \le t < \kappa_2 \\ \cdots & \\ \mu(E_m), & \text{if } \kappa_{m-1} \le t < \kappa_m \\ 0, & \text{if } \kappa_m \le t \end{cases}$$

Then

$$\begin{aligned} \int_{[0,+\infty)} \lambda_\phi(t)\, d\mu_G(t) &= \big(\mu(E_1) + \mu(E_2) + \cdots + \mu(E_m)\big)\big(G(\kappa_1-) - G(0-)\big) \\ &\quad + \big(\mu(E_2) + \cdots + \mu(E_m)\big)\big(G(\kappa_2-) - G(\kappa_1-)\big) \\ &\quad \cdots \\ &\quad + \mu(E_m)\big(G(\kappa_m-) - G(\kappa_{m-1}-)\big) \\ &= G(\kappa_1-)\mu(E_1) + G(\kappa_2-)\mu(E_2) + \cdots + G(\kappa_m-)\mu(E_m) \\ &= \int_X G(\phi(x)-)\, d\mu(x). \end{aligned}$$

because $G(\phi(x)-)$ 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) Take arbitrary measurable $f : X \to [0,+\infty]$ and any increasing sequence (ϕ_n) of non-negative measurable simple $\phi_n : X \to [0,+\infty)$ so that $\phi_n \uparrow f$ on X. Then $0 \le G(\phi_n(x)-) \uparrow G(f(x)-)$ for every $x \in X$ and, by the Monotone Convergence Theorem,

$$\int_X G(\phi_n(x)-)\, d\mu(x) \to \int_X G(f(x)-)\, d\mu(x).$$

Since $\{x \in X \,|\, t < \phi_n(x)\} \uparrow \{x \in X \,|\, t < f(x)\}$, we have that $\lambda_{\phi_n}(t) \uparrow \lambda_f(t)$ for every $t \in [0,+\infty)$. Again by the Monotone Convergence Theorem,

$$\int_{[0,+\infty)} \lambda_{\phi_n}(t)\, d\mu_G(t) \to \int_{[0,+\infty)} \lambda_f(t)\, d\mu_G(t).$$

By the result of (a), we get $\int_X G(f(x)-)\,d\mu(x) = \int_{[0,+\infty)} \lambda_f(t)\,d\mu_G(t)$.

Proposition 7.12 implies the second equality of the statement and the special case $G(t) = t$ implies the last equality.

Proposition 7.15 *Let $\mu(X) < +\infty$ and $f : X \to [0,+\infty]$ be measurable. We define $F : \mathbf{R} \to \mathbf{R}$ by*

$$F_f(t) = \mu(\{x \in X \,|\, f(x) \le t\}) = \begin{cases} \mu(X) - \lambda_f(t), & \textit{if } 0 \le t < +\infty \\ 0, & \textit{if } -\infty < t < 0 \end{cases}$$

Then F_f is increasing and continuous from the right and, for every increasing $G : \mathbf{R} \to \mathbf{R}$ with $G(0-) = 0$, we have

$$\int_X G(f(x)-)\,d\mu(x) = \int_{[0,+\infty)} G(t-)\,d\mu_{F_f}(t) + G(+\infty)\mu(f^{-1}(+\infty)).$$

Proof: It is obvious that F_f is increasing. If $t_n \downarrow t$, then $\{x \in X \,|\, f(x) \le t_n\} \downarrow \{x \in X \,|\, f(x) \le t\}$. By the continuity of μ from above, we get $F_f(t_n) \downarrow F_f(t)$ and F_f is continuous from the right.

We take any $n \in \mathbf{N}$ and apply Proposition 7.11 to find

$$\begin{aligned} \int_{[0,n]} G(t-)\,d\mu_{F_f}(t) &= G(n+)F_f(n) - \int_{[0,n]} F_f(t)\,d\mu_G(t) \\ &= \int_{[0,n]} (F_f(n) - F_f(t))\,d\mu_G(t). \end{aligned}$$

The left side is $= \int_{[0,+\infty)} G(t-)\chi_{[0,n]}(t)\,d\mu_{F_f}(t) \uparrow \int_{[0,+\infty)} G(t-)\,d\mu_{F_f}(t)$, by the Monotone Convergence Theorem.

The right side is $= \int_{[0,+\infty)} \mu(\{x \in X \,|\, t < f(x) \le n\})\chi_{[0,n]}(t)\,d\mu_G(t) \uparrow \int_{[0,+\infty)} \mu(\{x \in X \,|\, t < f(x) < +\infty\})\,d\mu_G(t)$, again by the Monotone Convergence Theorem.

Thus, $\int_{[0,+\infty)} G(t-)\,d\mu_{F_f}(t) = \int_{[0,+\infty)} \mu(\{x \in X \,|\, t < f(x) < +\infty\})\,d\mu_G(t)$ and, adding to both sides the quantity $G(+\infty)\mu(\{x \in X \,|\, f(x) = +\infty\})$ we find

$$\int_{[0,+\infty)} G(t-)\,d\mu_{F_f}(t) + G(+\infty)\mu(\{x \in X \,|\, f(x) = +\infty\}) = \int_{[0,+\infty)} \lambda_f(t)\,d\mu_G(t)$$

and the equality of the statement is an implication of Proposition 7.14.

7.9 Exercises.

1. *The graph and the area under the graph of a function.*

   Let $f : \mathbf{R}^n \to [0,+\infty]$ be Lebesgue measurable. If

$$\begin{aligned} A_f &= \{(x_1,\ldots,x_n,x_{n+1}) \,|\, 0 \le x_{n+1} < f(x_1,\ldots,x_n)\} \subseteq \mathbf{R}^{n+1}, \\ G_f &= \{(x_1,\ldots,x_n,x_{n+1}) \,|\, x_{n+1} = f(x_1,\ldots,x_n)\} \subseteq \mathbf{R}^{n+1}, \end{aligned}$$

prove that $A_f, G_f \in \mathcal{L}_{n+1}$ and

$$m_{n+1}(A_f) = \int_{\mathbf{R}^n} f\, dm_n, \qquad m_{n+1}(G_f) = 0.$$

2. *An equivalent definition of the integral.*

   Let (X, Σ, μ) be a measure space and $f : X \to [0, +\infty]$ be measurable. Let $\Delta = \{E_1, \dots, E_l\}$, where $l \in \mathbf{N}$ and the non-empty sets $E_1, \dots, E_l \in \Sigma$ are pairwise disjoint and cover X. Such Δ are called **Σ-partitions of** X. Define $\underline{\Sigma}(f, \Delta) = \sum_{j=1}^{l} m_j \mu(E_j)$, where $m_j = \inf\{f(x) \,|\, x \in E_j\}$. Prove that
   $$\int_X f\, d\mu = \sup\{\underline{\Sigma}(f, \Delta) \,|\, \Delta \text{ is a } \Sigma\text{-partition of } X\}.$$

3. If (X, Σ, μ) is a measure space, $f, g, h : X \to \overline{\mathbf{R}}$ are measurable, g, h are integrable and $g \leq f \leq h$ a.e. on X, prove that f is also integrable.

4. *The Uniform Convergence Theorem.*

   Let (X, Σ, μ) be a measure space, all $f_n : X \to \mathbf{R}$ or $\mathbf{C}$ be integrable and let $f_n \to f$ uniformly on X. If $\mu(X) < +\infty$, prove that f is integrable and that $\int_X f_n\, d\mu \to \int_X f\, d\mu$.

5. *The Bounded Convergence Theorem.*

   Let (X, Σ, μ) be a measure space and $f, f_n : X \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$ be measurable. If $\mu(X) < +\infty$ and there is $M < +\infty$ so that $|f_n| \leq M$ a.e. on X and $f_n \to f$ a.e. on X, prove that $\int_X f_n\, d\mu \to \int_X f\, d\mu$.

6. Let (X, Σ, μ) be a measure space, $f, f_n : X \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$ be measurable and $g : X \to [0, +\infty]$ be integrable. If $|f_n| \leq g$ a.e. on X for every n and $f_n \to f$ a.e. on X, prove that $\int_X |f_n - f|\, d\mu \to 0$.

7. Let (X, Σ, μ) be a measure space, $f, f_n : X \to [0, +\infty]$ be measurable with $f_n \leq f$ a.e. on X and $f_n \to f$ a.e. on X. Prove that $\int_X f_n\, d\mu \to \int_X f\, d\mu$.

8. Let (X, Σ, μ) be a measure space, $f, f_n : X \to [0, +\infty]$ be measurable and $f_n \to f$ a.e. on X. If there is $M < +\infty$ so that $\int_X f_n\, d\mu < M$ for infinitely many n's, prove that f is integrable.

9. *Generalisation of the Lemma of Fatou.*

   Assume (X, Σ, μ) is a measure space, $f, g, f_n : X \to \overline{\mathbf{R}}$ are measurable and $\int_X g^-\, d\mu < +\infty$. If $g \leq f_n$ a.e. on X and $f = \liminf_{n \to +\infty} f_n$ a.e. on X, prove that $\int_X f\, d\mu \leq \liminf_{n \to +\infty} \int_X f_n\, d\mu$.

10. Let (X, Σ, μ) be a measure space, $f, f_n : X \to [0, +\infty]$ be measurable with $f_n \downarrow f$ a.e. on X and $\int_X f_1\, d\mu < +\infty$. Prove that $\int_X f_n\, d\mu \downarrow \int_X f\, d\mu$.

11. Use either the Lemma of Fatou or the Series Theorem 7.2 to prove the Monotone Convergence Theorem.

12. *Generalisation of the Dominated Convergence Theorem.*

Let (X, Σ, μ) be a measure space and $f, f_n : X \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$, $g, g_n : X \to [0, +\infty]$ be measurable. If $|f_n| \leq g_n$ a.e. on X, if $\int_X g_n \, d\mu \to \int_X g \, d\mu < +\infty$ and if $f_n \to f$ a.e. on X and $g_n \to g$ a.e. on X, prove that $\int_X f_n \, d\mu \to \int_X f \, d\mu$.

13. Assume (X, Σ, μ) is a measure space, all $f, f_n : X \to [0, +\infty]$ are measurable, $f_n \to f$ a.e. on X and $\int_X f_n \, d\mu \to \int_X f \, d\mu < +\infty$. Prove that $\int_A f_n \, d\mu \to \int_A f \, d\mu$ for every $A \in \Sigma$.

14. Let (X, Σ, μ) be a measure space, $f, f_n : X \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$ be integrable and $f_n \to f$ a.e. on X. Prove that $\int_X |f_n - f| \, d\mu \to 0$ if and only if $\int_X |f_n| \, d\mu \to \int_X |f| \, d\mu$.

15. *Improper Integrals.*

Let $f : [a, b) \to \mathbf{R}$, where $-\infty < a < b \leq +\infty$. 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 $\overline{\mathbf{R}}$, 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 a 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 $f : [a, b) \to \mathbf{R}$ 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 7.9.17.

(iv) If $\int_a^{\to b} |f(x)| \, dx < +\infty$ (we say that the improper integral is **absolutely convergent**), prove that the $\int_a^{\to b} f(x) \, dx$ exists and is a real number (we say that the improper integral is **convergent**.)

16. Using improper integrals (see exercise 7.9.15), find the Lebesgue integral $\int_{-\infty}^{+\infty} f(x) \, dx$, where $f(x)$ is any of the following:

$$\frac{1}{1 + x^2}, \quad e^{-|x|}, \quad \frac{1}{x^2} \chi_{[0,+\infty)}(x), \quad \frac{1}{x}, \quad \frac{1}{|x|}, \quad \frac{1}{\sqrt{|x|}} \chi_{[-1,1]}(x).$$

17. Using improper integrals (see exercise 7.9.15), find the Lebesgue integral $\int_{-\infty}^{+\infty} f(x) \, dx$, where $f(x)$ is any of the following:

$$\sum_{n=1}^{+\infty} \frac{1}{2^n} \chi_{(n,n+1]}(x), \qquad \sum_{n=1}^{+\infty} \frac{(-1)^{n+1}}{2^n} \chi_{(n,n+1]}(x),$$

$$\sum_{n=1}^{+\infty} \frac{1}{n}\chi_{(n,n+1]}(x), \qquad \sum_{n=1}^{+\infty} \frac{(-1)^{n+1}}{n}\chi_{(n,n+1]}(x).$$

18. Apply the Lemma of Fatou for Lebesgue measure on $\mathbf{R}$ and the sequences:

$$\chi_{(n,n+1)}(x), \quad \chi_{(n,+\infty)}(x), \quad n\chi_{0,\frac{1}{n})}(x), \quad 1+\mathrm{sign}\Big(\sin\Big(2^n\frac{x}{2\pi}\Big)\Big).$$

19. If f is Lebesgue integrable on $[-1,1]$, prove $\lim_{n\to+\infty}\int_{-1}^{1} x^n f(x)\,dx = 0$.

20. *The discontinuous factor.*

Prove that

$$\lim_{t\to+\infty}\frac{1}{\pi}\int_a^{+\infty}\frac{t}{1+t^2x^2}\,dx = \begin{cases} 0, & \text{if } 0<a<+\infty,\\ \frac{1}{2}, & \text{if } a=0,\\ 1, & \text{if } -\infty<a<0.\end{cases}$$

21. Prove that

$$\lim_{n\to+\infty}\int_0^n\Big(1+\frac{x}{n}\Big)^n e^{-\alpha x}\,dx = \begin{cases} \frac{1}{\alpha-1}, & \text{if } 1<\alpha,\\ +\infty, & \text{if } \alpha\le 1.\end{cases}$$

22. Let (X,Σ,μ) be a measure space and $f : X \to [0,+\infty]$ be measurable with $0 < c = \int_X f\,d\mu < +\infty$. Prove that

$$\lim_{n\to+\infty} n\int_X \log\Big(1+\Big(\frac{f}{n}\Big)^\alpha\Big)\,d\mu = \begin{cases} +\infty, & \text{if } 0<\alpha<1,\\ c, & \text{if } \alpha=1,\\ 0, & \text{if } 1<\alpha<+\infty.\end{cases}$$

23. Consider $\mathbf{Q}\cap[0,1] = \{r_1, r_2, \ldots\}$ and a sequence (a_n) of real numbers so that $\sum_{n=1}^{+\infty}|a_n| < +\infty$. Prove that the series

$$\sum_{n=1}^{+\infty}\frac{a_n}{\sqrt{|x-r_n|}}$$

converges absolutely for m_1-a.e. $x\in[0,1]$.

24. *The measure induced by a function.*

Let (X,Σ,μ) be a measure space and $f : X \to [0,+\infty]$ be measurable. Define $\nu : \Sigma \to [0,+\infty]$ by

$$\nu(E) = \int_E f\,d\mu$$

for all $E\in\Sigma$. Prove that ν is a measure on (X,Σ) which is called **the measure induced by** f. Prove that:
(i) $\int_X g\,d\nu = \int_X gf\,d\mu$ for every measurable $g : X\to[0,+\infty]$,
(ii) if $g : X\to\overline{\mathbf{R}}$ is measurable, then $\int_X g\,d\nu$ exists if and only if $\int_X gf\,d\mu$ exists and in such a case the equality of (i) is true,
(iii) if $g : X\to\overline{\mathbf{C}}$ is measurable, then g is integrable with respect to ν if and only if gf is integrable with respect to μ and in such a case the equality of (i) is true.

25. Let (X, Σ, μ) be a measure space and $f : X \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$ be integrable. Prove that for every $\epsilon > 0$ there is an $E \in \Sigma$ with $\mu(E) < +\infty$ and $\int_{E^c} |f| \, d\mu < \epsilon$.

26. *Absolute continuity of the integral of f.*

    Let (X, Σ, μ) be a measure space and $f : X \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$ be integrable. Prove that for every $\epsilon > 0$ there is $\delta > 0$ so that: $|\int_E f \, d\mu| < \epsilon$ for all $E \in \Sigma$ with $\mu(E) < \delta$.
    (Hint: One may prove it first for simple functions and then use the Approximation Theorem 7.12.)

27. Let $f : \mathbf{R} \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$ be Lebesgue integrable. Prove $F(x) = \int_{-\infty}^{x} f(t) \, dt$ is a continuous function of x on $\mathbf{R}$.

28. *Continuity of translations.*

    Assume that $f : \mathbf{R}^n \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$ is Lebesgue integrable. Prove that

    $$\lim_{\mathbf{R}^n \ni h \to 0} \int_{\mathbf{R}^n} |f(x-h) - f(x)| \, dx = 0.$$

    (Hint: Prove it first for continuous functions which are 0 outside a bounded set and then use the Approximation Theorem 7.16.)

29. *The Riemann-Lebesgue Lemma.*

    Assume that $f : \mathbf{R} \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$ is Lebesgue integrable. Prove that

    $$\lim_{x \to +\infty} \int_{-\infty}^{+\infty} f(t) \cos(xt) \, dt = \lim_{x \to +\infty} \int_{-\infty}^{+\infty} f(t) \sin(xt) \, dt = 0$$

    in the two following ways:

    Prove the limits when f is the characteristic function of any interval and then use an approximation theorem.

    Prove that $|\int_{-\infty}^{+\infty} f(t) \cos(xt) \, dt| = \frac{1}{2} |\int_{-\infty}^{+\infty} (f(t - \frac{\pi}{x}) - f(t)) \cos(xt) \, dt| \leq \frac{1}{2} \int_{-\infty}^{+\infty} |f(t - \frac{\pi}{x}) - f(t)| \, dt$ and then use the result of exercise 7.9.28.

30. Let $Q \subseteq \mathbf{R}^n$ be a closed interval and $x_0 \in Q$. If $f : Q \to \mathbf{R}$ is Riemann integrable over Q and $g : Q \to \mathbf{R}$ coincides with f on $Q \setminus \{x_0\}$, prove that g is also Riemann integrable over Q and that $(\mathcal{R}_n) \int_Q g = (\mathcal{R}_n) \int_Q f$.

31. Let $Q \subseteq \mathbf{R}^n$ be a closed interval, $\lambda \in \mathbf{R}$ and $f, g : Q \to \mathbf{R}$ be Riemann integrable over Q. Prove that $f + g, \lambda f$ and fg are all Riemann integrable over Q and

    $$(\mathcal{R}_n) \int_Q (f+g) = (\mathcal{R}_n) \int_Q f + (\mathcal{R}_n) \int_Q g, \qquad (\mathcal{R}_n) \int_Q \lambda f = \lambda (\mathcal{R}_n) \int_Q f.$$

32. Let $Q \subseteq \mathbf{R}^n$ be a closed interval.
(i) If the bounded functions $f, f_k : Q \to \mathbf{R}$ are all Riemann integrable over Q and $0 \leq f_k \uparrow f$ on Q, prove that $(\mathcal{R}_n) \int_Q f_k \to (\mathcal{R}_n) \int_Q f$.
(ii) Find bounded functions $f, f_k : Q \to \mathbf{R}$ so that $0 \leq f_k \uparrow f$ on Q and so that all f_k are Riemann integrable over Q, but f is not Riemann integrable over Q.

33. *Continuity of an integral as a function of a parameter.*

Let (X, Σ, μ) be a measure space and $f : X \times (a, b) \to \mathbf{R}$ 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 $t \in (a, b)$, $|f(x, t)| \leq g(x)$ a.e. $x \in X$.

Prove that $F(t) = \int_X f(x, t)\, d\mu(x)$ is continuous as a function of t on (a, b).

34. *Differentiability of an integral as a function of a parameter.*

Let (X, Σ, μ) be a measure space and $f : X \times (a, b) \to \mathbf{R}$ 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 $|\frac{\partial f}{\partial t}(x, t)| \leq g(x)$ for every $t \in (a, b)$. Thus, $\frac{\partial f}{\partial t} : A \times (a, b) \to \mathbf{R}$ for some $A \in \Sigma$ with $\mu(X \setminus A) = 0$.

Prove that $F(t) = \int_X f(x, t)\, d\mu(x)$ is differentiable as a function of t on (a, b) and that

$$\frac{dF}{dt}(t) = \int_X \frac{\partial f}{\partial t}(x, t)\, d\mu(x), \qquad a < t < b.$$

35. *The integral of Gauss.*

Consider the functions $f, h : [0, +\infty) \to \mathbf{R}$ defined by

$$f(x) = \frac{1}{2}\Big(\int_0^x e^{-\frac{1}{2}t^2}\, dt\Big)^2, \qquad h(x) = \int_0^1 \frac{e^{-\frac{1}{2}x^2(t^2+1)}}{t^2+1}\, dt.$$

(i) Using Exercise 7.9.34, prove that $f'(x) + h'(x) = 0$ for every $x \in (0, +\infty)$ and, hence, that $f(x) + h(x) = \frac{\pi}{4}$ for every $x \in [0, +\infty)$.
(ii) Prove that

$$\int_{-\infty}^{+\infty} e^{-\frac{1}{2}t^2}\, dt = \sqrt{2\pi}.$$

36. *The distribution (or measure) of Gauss.*

Consider the function $g : \mathbf{R} \to \mathbf{R}$ 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 $g(+\infty) = 1$ and with continuous derivative $g'(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$, $x \in \mathbf{R}$.
(ii) The Lebesgue-Stieltjes measure μ_g induced by g is called **the distribution** or **the measure of Gauss**. Prove that $\mu_g(\mathbf{R}) = 1$, that

$$\mu_g(E) = \frac{1}{\sqrt{2\pi}} \int_E e^{-\frac{1}{2}x^2}\, dx$$

for every Borel set in $\mathbf{R}$ and that

$$\int_{\mathbf{R}} f(x)\, d\mu_g(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(x) e^{-\frac{1}{2}x^2}\, dx$$

for every Borel measurable $f : \mathbf{R} \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$ for which either of the two integrals exists.

37. (i) Using Exercise 7.9.34, prove that the function $F : (0, +\infty) \to \mathbf{R}$ defined by

$$F(t) = \int_0^{+\infty} e^{-tx} \frac{\sin x}{x}\, dx$$

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 $\int_0^{\to+\infty} \frac{\sin x}{x}\, dx$ 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

$$\lim_{t\to+\infty} \frac{1}{\pi} \int_a^{\to+\infty} \frac{\sin(tx)}{x}\, dx = \begin{cases} 0, & \text{if } 0 < a < +\infty, \\ \frac{1}{2}, & \text{if } a = 0, \\ 1, & \text{if } -\infty < a < 0. \end{cases}$$

38. *The gamma-function.*
Let $H_+ = \{z = x + iy \in \mathbf{C} \mid x > 0\}$ and consider the function $\Gamma : H_+ \to \mathbf{C}$ 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 $z \in H_+$.
(ii) Using Exercise 7.9.34, prove that

$$\frac{\partial \Gamma}{\partial x}(z) = -i \frac{\partial \Gamma}{\partial y}(z)$$

for every $z \in H_+$. This means that Γ is holomorphic in H_+.
(iii) Prove that $\Gamma(n) = (n-1)!$ for every $n \in \mathbf{N}$.

39. *The invariance of Lebesgue integral and of Lebesgue measure under isometries.*

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

$$\int_{\mathbf{R}^n} f \circ T^{-1} \, dm_n = \int_{\mathbf{R}^n} f \, dm_n$$

for every Lebesgue measurable $f : \mathbf{R}^n \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$, provided that at least one of the two integrals exists. (See also exercise 4.6.2.)

40. (i) Consider the Cantor set C and the $I_0 = [0,1], I_1, I_2, \dots$ which were used for its construction. Prove that the 2^{n-1} subintervals of $I_{n-1} \setminus I_n$, $n \in \mathbf{N}$, can be described as

$$\Big(\frac{a_1}{3} + \cdots + \frac{a_{n-1}}{3^{n-1}} + \frac{1}{3^n}, \frac{a_1}{3} + \cdots + \frac{a_{n-1}}{3^{n-1}} + \frac{2}{3^n}\Big),$$

where each of $a_1, \dots, a_{n-1}$ takes the values 0 and 2.
(ii) Let f be the Cantor function, which was introduced in exercise 4.6.10, extended as 0 in $(-\infty, 0)$ and as 1 in $(1, +\infty)$. Prove that f is constant

$$f(x) = \frac{a_1}{2^2} + \cdots + \frac{a_{n-1}}{2^n} + \frac{1}{2^n}$$

in the above subinterval $(\frac{a_1}{3} + \cdots + \frac{a_{n-1}}{3^{n-1}} + \frac{1}{3^n}, \frac{a_1}{3} + \cdots + \frac{a_{n-1}}{3^{n-1}} + \frac{2}{3^n})$.
(iii) If $G : (0,1) \to \mathbf{R}$ is another function with bounded derivative in $(0,1)$, prove that

$$\int_{(0,1)} G(x) \, d\mu_f(x) = G(1-) - \sum_{n=1}^{+\infty} \sum_{a_1, \dots, a_{n-1} = 0,\, 2} \Big(\frac{a_1}{2^2} + \cdots + \frac{a_{n-1}}{2^n} + \frac{1}{2^n}\Big) \cdot$$

$$\cdot \Big(G\Big(\frac{a_1}{3} + \cdots + \frac{a_{n-1}}{3^{n-1}} + \frac{2}{3^n}\Big) - G\Big(\frac{a_1}{3} + \cdots + \frac{a_{n-1}}{3^{n-1}} + \frac{1}{3^n}\Big)\Big).$$

(iv) In particular, $\int_{(0,1)} x \, d\mu_f(x) = \frac{1}{2}$.
(v) Prove that

$$\int_{(0,1)} e^{itx} \, d\mu_f(x) = e^{\frac{1}{2}it} \lim_{n \to +\infty} \prod_{k=1}^{n} \cos\Big(\frac{t}{3^k}\Big)$$

for every $t \in \mathbf{R}$.

41. Let $F, G : \mathbf{R} \to \mathbf{R}$ be increasing and assume that FG is also increasing. 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 $E \subseteq \mathbf{R}$ and

$$\int_{\mathbf{R}} f(x)\, d\mu_{FG}(x) = \int_{\mathbf{R}} f(x)G(x+)\, d\mu_F(x) + \int_{\mathbf{R}} f(x)F(x-)\, d\mu_G(x)$$

for every Borel measurable $f : \mathbf{R} \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$ for which at least two of the three integrals exist.

42. If $F : \mathbf{R} \to \mathbf{R}$ is increasing and continuous and $f : \mathbf{R} \to [0, +\infty]$ is Borel measurable, prove that $\int_{\mathbf{R}} f(F(x))\, d\mu_F(x) = \int_{F(-\infty)}^{F(+\infty)} f(t)\, dt$.

    Show, by example, that this may not be true if F is not continuous.

43. *Riemann's criterion for convergence of a series.*

    Assume $F : \mathbf{R} \to [0, +\infty)$ 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 $\sum_{n=1}^{+\infty} a_n < +\infty$.

44. *Mean values.*

    Let (X, Σ, μ) be a measure space, $f : X \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$ be integrable and F be a closed subset of $\mathbf{R}$ or $\mathbf{C}$. If $\frac{1}{\mu(E)} \int_E f\, d\mu \in F$ for every $E \in \Sigma$ with $0 < \mu(E)$, prove that $f(x) \in F$ for a.e. $x \in X$.

45. Let (X, Σ, μ) be a measure space and $E \in \Sigma$ have σ-finite measure. Prove that there is an $f : X \to [0, +\infty]$ with $\int_X f\, d\mu < +\infty$ and $f(x) > 0$ for every $x \in E$.

46. Let (X, Σ, μ) be a measure space and $f : X \to [0, +\infty]$ be measurable. Prove that

$$\frac{1}{2} \sum_{n \in \mathbf{Z}} 2^n \lambda_f(2^n) \leq \int_X f(x)\, d\mu(x) \leq \sum_{n \in \mathbf{Z}} 2^n \lambda_f(2^n)$$

    and, hence, that f is integrable if and only if the $\sum_{n \in \mathbf{Z}} 2^n \lambda_f(2^n)$ is finite.

47. *Equidistributed functions.*

    Let (X, Σ, μ) 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 $t \in [0, +\infty)$.

    Prove that, if f, g are equidistributed, then

$$\int_X f^p(x)\, d\mu(x) = \int_X g^p(x)\, d\mu(x)$$

    for every $p \in (0, +\infty)$.

48. Let (X, Σ, μ) 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 (exercise 7.9.47) if and only if $m = n$, $\kappa_1 = \lambda_1, \dots, \kappa_m = \lambda_m$ and $\mu(E_1) = \mu(F_1), \dots, \mu(E_m) = \mu(F_m)$.

49. *The inequality of Chebychev.*

    Let (X, Σ, μ) be a measure space and $f : X \to [0, +\infty]$ be measurable. Prove that

$$\mu(\{x \in X \mid t < f(x)\}) = \lambda_f(t) \leq \frac{1}{t} \int_X f(x) \, d\mu(x)$$

    for every $t \in (0, +\infty)$. Prove also that, if f is integrable, then

$$\lim_{t \to +\infty} t\lambda_f(t) = 0.$$

50. Let (X, Σ, μ) be a measure space, $f : X \to [0, +\infty]$ be measurable and $0 < p < +\infty$. Prove that

$$\int_X f^p(x) \, d\mu(x) = p \int_0^{+\infty} t^{p-1} \lambda_f(t) \, dt.$$

    If, also, $f < +\infty$ a.e. on X, prove that

$$\int_X f^p(x) \, d\mu(x) = \int_{[0,+\infty)} t^p \, d\mu_{F_f}(t),$$

    where F_f is defined in Proposition 7.15.

51. *The Jordan content of sets in* $\mathbf{R}^n$.

    If $E \subseteq R^n$ is bounded we define its **inner Jordan content**

$$c_n^{(i)}(E) = \sup \Big\{ \sum_{j=1}^m vol_n(R_j) \,|\, m \in \mathbf{N}, R_1, \dots, R_m \text{ pairwise disjoint open intervals with } \cup_{j=1}^m R_j \subseteq E \Big\}$$

    and its **outer Jordan content**

$$c_n^{(o)}(E) = \inf \Big\{ \sum_{j=1}^m vol_n(R_j) \,|\, m \in \mathbf{N}, R_1, \dots, R_m \text{ open intervals with } \cup_{j=1}^m R_j \supseteq E \Big\}.$$

    (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 $E \subseteq \mathbf{R}^n$.

The bounded 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 E is bounded and $c_n^{(o)}(E) = 0$, prove that E is a Jordan set.
(iv) Prove that all intervals S are Jordan sets and $c_n(S) = vol_n(S)$.
(v) If E is bounded, prove that it is a Jordan set if and only if for every $\epsilon > 0$ there exist pairwise disjoint open intervals $R_1, \ldots, R_m$ and open intervals $R'_1, \ldots, R'_k$ so that $\cup_{j=1}^m R_j \subseteq E \subseteq \cup_{i=1}^k R'_i$ and

$$\sum_{i=1}^k vol_n(R'_i) - \sum_{j=1}^m vol_n(R_j) < \epsilon.$$

(vi) If E is bounded, prove that E is a Jordan set if and only if $c_n^{(o)}(\partial E) = 0$.
(vii) Prove that the collection of bounded Jordan sets is closed under finite unions and set-theoretic differences. Moreover, if $E_1, \ldots, E_l$ are pairwise disjoint Jordan sets, prove that

$$c_n(E) = \sum_{j=1}^l c_n(E_j).$$

(viii) Prove that if the bounded set E is closed, then $m_n(E) = 0$ implies $c_n(E) = 0$. If E is not closed, then this result may not be true. For example, if $E = \mathbf{Q} \cap [0,1] \subseteq \mathbf{R}$, then $m_1(E) = 0$, but $c_1^{(i)}(E) = 0 < 1 = c_1^{(o)}(E)$ and, hence, E is not a Jordan set. (See exercise 4.6.6.)
(ix) If the bounded set E is a Jordan set, prove that it is a Lebesgue set and

$$m_n(E) = c_n(E).$$

(x) Let E be bounded and take any closed interval Q so that $E \subseteq Q$. 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 closed interval, $f, g : Q \to \mathbf{R}$ 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

$$(\mathcal{R}_n) \int_Q f = (\mathcal{R}_n) \int_Q g.$$

52. *Lebesgue's characterisation of Riemann integrable functions.*

Let $Q \subseteq \mathbf{R}^n$ be a closed interval and $f : Q \to \mathbf{R}$ be bounded. Prove that f is Riemann integrable if and only if $\{x \in Q \,|\, f \text{ is discontinuous at } x\}$ is a null set.

Chapter 8

Product measures

8.1 Product σ-algebra.

If I is a general set of indices, the elements of the cartesian product $\prod_{i\in I} X_i$ are all functions $x : I \to \cup_{i\in I} X_i$ 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, $I = \{1, \dots, n\}$, we use the traditional notation $x = (x_1, \dots, x_n)$ for the element $x = (x_i)_{i\in I}$ and we use the notation $\prod_{i=1}^n X_i$ or $X_1 \times \cdots \times X_n$ for $\prod_{i\in I} X_i$. And if I is countable, say $I = \mathbf{N} = \{1, 2, \dots\}$, we write $x = (x_1, x_2, \dots)$ for the element $x = (x_i)_{i\in I}$ and we write $\prod_{i=1}^{+\infty} X_i$ or $X_1 \times X_2 \times \cdots$ for $\prod_{i\in I} X_i$.

Definition 8.1 *If I is a set of indices, then, for every $j \in I$, 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.***

In case $I = \{1, \dots, n\}$ or $I = \mathbf{N}$, the formula of the j-th projection is

$$\pi_j(x) = x_j$$

for all $x = (x_1, \dots, x_n) \in X_1 \times \cdots \times X_n = \prod_{i=1}^n X_i$ or, respectively, $x = (x_1, x_2, \dots) \in X_1 \times X_2 \times \cdots = \prod_{i=1}^{+\infty} X_i$.

Clearly, the inverse image $\pi_j^{-1}(A_j) = \{x = (x_i)_{i\in I} \in \prod_{i\in I} X_i \,|\, x_j \in A_j\}$ of an arbitrary $A_j \subseteq X_j$ is the cartesian product

$$\pi_j^{-1}(A_j) = \prod_{i\in I} Y_i, \qquad \text{where} \quad Y_i = \begin{cases} X_i, & \text{if } i \neq j \\ A_j, & \text{if } i = j \end{cases}$$

In particular, if $I = \{1, \dots, n\}$, then

$$\pi_j^{-1}(A_j) = X_1 \times \cdots \times X_{j-1} \times A_j \times X_{j+1} \times \cdots \times X_n$$

and, if $I = \mathbf{N}$, then

$$\pi_j^{-1}(A_j) = X_1 \times \cdots \times X_{j-1} \times A_j \times X_{j+1} \times \cdots.$$

Definition 8.2 *If (X_i, Σ_i) is a measurable space for every $i \in I$, we consider the σ-algebra of subsets of the cartesian product $\prod_{i\in I} X_i$*

$$\otimes_{i\in I}\Sigma_i = \Sigma\big(\{\pi_j^{-1}(A_j) \,|\, j \in I, A_j \in \Sigma_j\}\big),$$

*called **the product σ-algebra** of Σ_i $(i \in I)$.*

In particular, $\otimes_{i=1}^n \Sigma_i$ is generated by the collection of all sets of the form $X_1 \times \cdots \times X_{j-1} \times A_j \times X_{j+1} \times \cdots \times X_n$, where $1 \leq j \leq n$ and $A_j \in \Sigma_j$.

Similarly, $\otimes_{i=1}^{+\infty} \Sigma_i$ is generated by the collection of all sets of the form $X_1 \times \cdots \times X_{j-1} \times A_j \times X_{j+1} \times \cdots$, where $j \in \mathbf{N}$ and $A_j \in \Sigma_j$.

Proposition 8.1 *Let (X_i, Σ_i) be a measurable space for each $i \in I$. Then $\otimes_{i\in I}\Sigma_i$ is the smallest σ-algebra Σ of subsets of $\prod_{i\in I} X_i$ for which all projections $\pi_j : \prod_{i\in I} X_i \to X_j$ are (Σ, Σ_j)-measurable.*

Proof: For every j and every $A_j \in \Sigma_j$ we have that $\pi_j^{-1}(A_j) \in \otimes_{i\in I}\Sigma_i$ and, hence, every π_j is $(\otimes_{i\in I}\Sigma_i, \Sigma_j)$-measurable.

Now, let Σ be a σ-algebra of subsets of $\prod_{i\in I} X_i$ for which all projections $\pi_j : \prod_{i\in I} X_i \to X_j$ are (Σ, Σ_j)-measurable. Then for every j and every $A_j \in \Sigma_j$ we have that $\pi_j^{-1}(A_j) \in \Sigma$. This implies that $\{\pi_j^{-1}(A_j) \,|\, j \in I, A_j \in \Sigma_j\} \subseteq \Sigma$ and, hence, $\otimes_{i\in I}\Sigma_i \subseteq \Sigma$.

Proposition 8.2 *Let (X_i, Σ_i) be a measurable space for each $i \in I$. If $\mathcal{E}_i$ is a collection of subsets of X_i with $\Sigma_i = \Sigma(\mathcal{E}_i)$ for all $i \in I$, then $\otimes_{i\in I}\Sigma_i = \Sigma(\mathcal{E})$, where*

$$\mathcal{E} = \{\pi_j^{-1}(E_j) \,|\, j \in I, E_j \in \mathcal{E}_j\}.$$

Proof: Since $\mathcal{E} \subseteq \{\pi_j^{-1}(A_j) \,|\, j \in I, A_j \in \Sigma_j\} \subseteq \otimes_{i\in I}\Sigma_i$, it is immediate that $\Sigma(\mathcal{E}) \subseteq \otimes_{i\in I}\Sigma_i$.

We, now, fix $j \in I$ and consider the $\pi_j : \prod_{i\in I} X_i \to X_j$. We have that $\pi_j^{-1}(E_j) \in \mathcal{E} \subseteq \Sigma(\mathcal{E})$ for every $E_j \in \mathcal{E}_j$. Proposition 6.1 implies that π_j is $(\Sigma(\mathcal{E}), \Sigma_j)$-measurable and, since j is arbitrary, Proposition 8.1 implies that $\otimes_{i\in I}\Sigma_i \subseteq \Sigma(\mathcal{E})$.

Proposition 8.3 *Let (X_i, Σ_i) be measurable spaces. If $\mathcal{E}_i$ is a collection of subsets of X_i so that $\Sigma_i = \Sigma(\mathcal{E}_i)$ for every $i \in I$, then $\otimes_{i\in I}\Sigma_i = \Sigma(\tilde{\mathcal{E}})$, where*

$$\tilde{\mathcal{E}} = \{\prod_{i\in I} E_i \,|\, E_i \neq X_i \textit{ for at most countably many } i \in I \textit{ and } E_i \in \mathcal{E}_i \textit{ if } E_i \neq X_i\}.$$

Proof: We observe that $\pi_j^{-1}(E_j) \in \tilde{\mathcal{E}}$ for every $j \in I$ and every $E_j \in \mathcal{E}_j$ and, hence, $\mathcal{E} \subseteq \tilde{\mathcal{E}} \subseteq \Sigma(\tilde{\mathcal{E}})$. This implies $\Sigma(\mathcal{E}) \subseteq \Sigma(\tilde{\mathcal{E}})$.

Now take any $\prod_{i\in I} E_i \in \tilde{\mathcal{E}}$. We set $\{i_1, i_2, \ldots\} = \{i \in I \,|\, E_i \neq X_i\}$ and observe that

$$\prod_{i\in I} E_i = \cap_{n=1}^{+\infty} \pi_{i_n}^{-1}(E_{i_n}) \in \Sigma(\mathcal{E}).$$

Thus, $\tilde{\mathcal{E}} \subseteq \Sigma(\mathcal{E})$ and, hence, $\Sigma(\tilde{\mathcal{E}}) \subseteq \Sigma(\mathcal{E})$. Proposition 8.2 finishes the proof.

In particular, $\otimes_{i=1}^n \Sigma_i$ is generated by the collection of all cartesian products of the form $E_1 \times \cdots \times E_n$, where $E_j \in \mathcal{E}_j$ for all $j = 1, \ldots, n$.

Also, $\otimes_{i=1}^{+\infty} \Sigma_i$ is generated by the collection of all cartesian products of the form $E_1 \times E_2 \times \cdots$, where $E_j \in \mathcal{E}_j$ for all $j \in \mathbf{N}$.

Example
If we consider $\mathbf{R}^n = \prod_{i=1}^n \mathbf{R}$ and, for each copy of $\mathbf{R}$, we take the collection of all open-closed 1-dimensional intervals as a generator of $\mathcal{B}_{\mathbf{R}}$, then Proposition 8.3 implies that the collection of all open-closed n-dimensional intervals is a generator of $\otimes_{i=1}^n \mathcal{B}_{\mathbf{R}}$. But we already know that the same collection is a generator of $\mathcal{B}_{\mathbf{R}^n}$. Therefore,

$$\mathcal{B}_{\mathbf{R}^n} = \otimes_{i=1}^n \mathcal{B}_{\mathbf{R}}.$$

This can be generalised. If $n_1 + \cdots + n_k = n$, we formally identify the typical element $(x_1, \ldots, x_n) \in \mathbf{R}^n$ with

$$\big((x_1, \ldots, x_{n_1}), \ldots, (x_{n_1+\cdots+n_{k-1}+1}, \ldots, x_{n_1+\cdots+n_k})\big),$$

i.e. with the typical element of $\prod_{j=1}^k \mathbf{R}^{n_j}$. We thus identify

$$\mathbf{R}^n = \prod_{j=1}^k \mathbf{R}^{n_j}.$$

Now, $\otimes_{j=1}^k \mathcal{B}_{\mathbf{R}^{n_j}}$ is generated by the collection of all products $\prod_{j=1}^k A_j$, where each A_j is an n_j-dimensional open-closed interval. By the above identification, $\prod_{j=1}^k A_j$ is the typical n-dimensional open-closed interval and, hence, $\otimes_{j=1}^k \mathcal{B}_{\mathbf{R}^{n_j}}$ is generated by the collection of all open-closed intervals in $\mathbf{R}^n$. Therefore,

$$\mathcal{B}_{\mathbf{R}^n} = \otimes_{j=1}^k \mathcal{B}_{\mathbf{R}^{n_j}}.$$

8.2 Product measure.

In this section we shall limit ourselves to cartesian products of finitely many spaces. We fix the measure spaces $(X_1, \Sigma_1, \mu_1), \ldots, (X_n, \Sigma_n, \mu_n)$ and the measurable space $(\prod_{j=1}^n X_j, \otimes_{j=1}^n \Sigma_j)$.

From Proposition 8.3 and the paragraph after it, we know that $\otimes_{j=1}^n \Sigma_j$ is generated by the collection $\tilde{\mathcal{E}}$ of all sets of the form $\prod_{j=1}^n A_j$, where $A_j \in \Sigma_j$ for all j. Observe that $\prod_{j=1}^n X_j$ belongs to $\tilde{\mathcal{E}}$ and also $\emptyset = \prod_{j=1}^n \emptyset$ belongs to $\tilde{\mathcal{E}}$.

The elements of $\tilde{\mathcal{E}}$ play the same role that open-closed intervals play for the introduction of Lebesgue measure on $\mathbf{R}^n$. We agree to call these sets **measurable intervals in** $\prod_{j=1}^n X_j$, a term which will be justified by Theorem 8.3, and denote them by

$$\tilde{R} = \prod_{j=1}^n A_j.$$

Proposition 8.4 *Let (X_j, Σ_j) be a measurable space for every $j = 1, \ldots, n$. The collection*

$$\mathcal{A} = \{\tilde{R}_1 \cup \cdots \cup \tilde{R}_m \,|\, m \in \mathbf{N}, \tilde{R}_1, \ldots, \tilde{R}_m \textit{ pairwise disjoint elements of } \tilde{\mathcal{E}}\}$$

is an algebra of subsets of $\prod_{j=1}^n X_j$.

Proof: If $\tilde{R} = \prod_{j=1}^n A_j$ and $\tilde{R}' = \prod_{j=1}^n B_j$ are elements of $\tilde{\mathcal{E}}$, then $\tilde{R} \cap \tilde{R}' = \prod_{j=1}^n (A_j \cap B_j)$ is an element of $\tilde{\mathcal{E}}$.

Moreover, if $\tilde{R} = \prod_{j=1}^n A_j$ is an element of $\tilde{\mathcal{E}}$, then

$$\begin{aligned}
\tilde{R}^c = & \, (A_1^c \times A_2 \times \cdots \times A_n) \cup \\
& \cdots\cdots \\
& \cup (X_1 \times X_2 \times \cdots \times X_{j-1} \times A_j^c \times A_{j+1} \times \cdots \times A_n) \cup \\
& \cdots\cdots \\
& \cup (X_1 \times X_2 \times \cdots \times X_{n-1} \times A_n^c)
\end{aligned}$$

is a disjoint union of elements of $\tilde{\mathcal{E}}$, i.e. an element of $\mathcal{A}$.

Now, if $\tilde{R}_1 \cup \cdots \cup \tilde{R}_m$ and $\tilde{R}_1' \cup \cdots \cup \tilde{R}_k'$ are any two elements of $\mathcal{A}$, then $(\tilde{R}_1 \cup \cdots \cup \tilde{R}_m) \cap (\tilde{R}_1' \cup \cdots \cup \tilde{R}_k') = \bigcup_{1 \leq j \leq m, 1 \leq i \leq k} (\tilde{R}_j \cap \tilde{R}_i')$, is, by the result of the first paragraph, also an element of $\mathcal{A}$. Hence, $\mathcal{A}$ is closed under finite intersections. Also, if $\tilde{R}_1 \cup \cdots \cup \tilde{R}_m$ is an element of $\mathcal{A}$, then $(\tilde{R}_1 \cup \cdots \cup \tilde{R}_m)^c = \tilde{R}_1^c \cap \cdots \cap \tilde{R}_m^c$ is, by the result of the second paragraph, a finite intersection of elements of $\mathcal{A}$ and, hence, an element of $\mathcal{A}$.

Therefore, $\mathcal{A}$ is closed under finite intersections and under complements. This implies that it is an algebra of subsets of $\prod_{j=1}^n X_j$.

For each $\tilde{R} = \prod_{j=1}^n A_j \in \tilde{\mathcal{E}}$, we define the quantity

$$\tau(\tilde{R}) = \prod_{j=1}^n \mu_j(A_j),$$

which plays the role of *volume* of the measurable interval $\tilde{R}$.

Definition 8.3 *Let (X_j, Σ_j, μ_j) be a measure space for every $j = 1, \ldots, n$. For every $E \subseteq \prod_{j=1}^n X_j$ we define*

$$\mu^*(E) = \inf\Big\{ \sum_{i=1}^{+\infty} \tau(\tilde{R}_i) \,|\, \tilde{R}_i \in \tilde{\mathcal{E}} \textit{ for all } i \textit{ and } E \subseteq \cup_{i=1}^{+\infty} \tilde{R}_i \Big\}$$

Theorem 3.2 implies that the function $\mu^* : \mathcal{P}\big(\prod_{j=1}^n X_j\big) \to [0, +\infty]$ is an outer measure on $\prod_{j=1}^n X_j$.

Proposition 8.5 *Let (X_j, Σ_j, μ_j) be a measure space for every $j = 1, \dots, n$ and $\tilde{R}, \tilde{R}_i$ be measurable intervals for every $i \in \mathbf{N}$.*
(i) If $\tilde{R} \subseteq \cup_{i=1}^{+\infty} \tilde{R}_i$, then $\tau(\tilde{R}) \leq \sum_{i=1}^{+\infty} \tau(\tilde{R}_i)$.
(ii) If $\tilde{R} = \cup_{i=1}^{+\infty} \tilde{R}_i$ and all $\tilde{R}_i$ are pairwise disjoint, then $\tau(\tilde{R}) = \sum_{i=1}^{+\infty} \tau(\tilde{R}_i)$.

Proof: (i) Let $\tilde{R} = \prod_{j=1}^n A_j$ and $\tilde{R}_i = \prod_{j=1}^n A_j^{(i)}$, where $A_j, A_j^{(i)} \in \Sigma_j$ for every $i \in \mathbf{N}$ and j with $1 \leq j \leq n$.

From $\prod_{j=1}^n A_j \subseteq \cup_{i=1}^{+\infty} \prod_{j=1}^n A_j^{(i)}$, we get that

$$\begin{aligned}\prod_{j=1}^n \chi_{A_j}(x_j) &= \chi_{\prod_{j=1}^n A_j}(x_1, \dots, x_n) \\ &\leq \sum_{i=1}^{+\infty} \chi_{\prod_{j=1}^n A_j^{(i)}}(x_1, \dots, x_n) = \sum_{i=1}^{+\infty} \prod_{j=1}^n \chi_{A_j^{(i)}}(x_j)\end{aligned}$$

for every $x_1 \in X_1, \dots, x_n \in X_n$. Integrating over X_1 with respect to μ_1, we find

$$\mu_1(A_1) \prod_{j=2}^n \chi_{A_j}(x_j) \leq \sum_{i=1}^{+\infty} \mu_1(A_1^{(i)}) \prod_{j=2}^n \chi_{A_j^{(i)}}(x_j)$$

for every $x_2 \in X_2, \dots, x_n \in X_n$. Integrating over X_2 with respect to μ_2, we get

$$\mu_1(A_1)\mu_2(A_2) \prod_{j=3}^n \chi_{A_j}(x_j) \leq \sum_{i=1}^{+\infty} \mu_1(A_1^{(i)})\mu_2(A_2^{(i)}) \prod_{j=3}^n \chi_{A_j^{(i)}}(x_j)$$

for every $x_3 \in X_3, \dots, x_n \in X_n$. We continue until we have integrated all variables.
(ii) We use equalities everywhere in the above calculations.

The next result justifies the term *measurable interval* for each $\tilde{R} \in \tilde{\mathcal{E}}$.

Theorem 8.1 *Let (X_i, Σ_i, μ_i) be a measure space for every $i = 1, \dots, n$ and μ^* the outer measure of Definition 8.5. Every measurable interval $\tilde{R} = \prod_{j=1}^n A_j$ is μ^*-measurable and*

$$\mu^*(\tilde{R}) = \tau(\tilde{R}) = \prod_{j=1}^n \mu_j(A_j).$$

Also, $\otimes_{j=1}^n \Sigma_j$ is included in the σ-algebra of μ^*-measurable subsets of $\prod_{j=1}^n X_j$.

Proof: (a) If $\tilde{R}$ is a measurable interval, then $\tilde{R} \in \tilde{\mathcal{E}}$ and, from $\tilde{R} \subseteq \tilde{R}$, we obviously get $\mu^*(\tilde{R}) \leq \tau(\tilde{R})$.

Proposition 8.5 implies $\tau(\tilde{R}) \leq \sum_{i=1}^{+\infty} \tau(\tilde{R}_i)$ for every covering $\tilde{R} \subseteq \cup_{i=1}^{+\infty} \tilde{R}_i$ with $\tilde{R}_i \in \tilde{\mathcal{E}}$ for all $i \in \mathbf{N}$. Hence, $\tau(\tilde{R}) \leq \mu^*(\tilde{R})$ and we conclude that

$$\mu^*(\tilde{R}) = \tau(\tilde{R}).$$

(b) We take any two measurable intervals $\tilde{R}, \tilde{R}'$ and Proposition 8.4 implies that there are pairwise disjoint measurable intervals $\tilde{R}_1, \ldots, \tilde{R}_m$ so that $\tilde{R}' \setminus \tilde{R} = \tilde{R}_1 \cup \cdots \cup \tilde{R}_m$. By the subadditivity of μ^*, the result of (a) and Proposition 8.5,

$$\begin{aligned}\mu^*(\tilde{R}' \cap \tilde{R}) + \mu^*(\tilde{R}' \setminus \tilde{R}) &\le \mu^*(\tilde{R}' \cap \tilde{R}) + \mu^*(\tilde{R}_1) + \cdots + \mu^*(\tilde{R}_n) \\ &= \tau(\tilde{R}' \cap \tilde{R}) + \tau(\tilde{R}_1) + \cdots + \tau(\tilde{R}_n) \\ &= \tau(\tilde{R}').\end{aligned}$$

(c) Let $\tilde{R} \in \tilde{\mathcal{E}}$ and consider an arbitrary $E \subseteq \prod_{j=1}^n X_j$ with $\mu^*(E) < +\infty$. For any $\epsilon > 0$ we consider a covering $E \subseteq \cup_{i=1}^{+\infty} \tilde{R}_i$ with $\tilde{R}_i \in \tilde{\mathcal{E}}$ for all $i \in \mathbf{N}$, such that $\sum_{i=1}^{+\infty} \tau(\tilde{R}_i) < \mu^*(E) + \epsilon$. By the result of (b) and the subadditivity of μ^*,

$$\mu^*(E \cap \tilde{R}) + \mu^*(E \setminus \tilde{R}) \le \sum_{i=1}^{+\infty} \big(\mu^*(\tilde{R}_i \cap \tilde{R}) + \mu^*(\tilde{R}_i \setminus \tilde{R})\big) \le \sum_{i=1}^{+\infty} \tau(\tilde{R}_i) < \mu^*(E) + \epsilon.$$

Since ϵ is arbitrary, $\mu^*(E \cap \tilde{R}) + \mu^*(E \setminus \tilde{R}) \le \mu^*(E)$ and we conclude that $\tilde{R}$ is μ^*-measurable.

Since $\bigotimes_{j=1}^n \Sigma_j$ is generated by the collection of all measurable intervals, it is included in the σ-algebra of all μ^*-measurable sets.

Definition 8.4 *Let (X_i, Σ_i, μ_i) be a measure space for each $i = 1, \ldots, n$ and μ^* be the outer measure of Definition 8.5. The measure induced from μ^* by Theorem 3.1 is called* ***the product measure of*** *μ_j, $1 \le j \le n$, and it is denoted*

$$\otimes_{j=1}^n \mu_j.$$

We denote by $\Sigma_{\otimes_{j=1}^n \mu_j}$ the σ-algebra of μ^*-measurable subsets of $\prod_{j=1}^n X_j$. Therefore, $(\prod_{j=1}^n X_j, \Sigma_{\otimes_{j=1}^n \mu_j}, \otimes_{j=1}^n \mu_j)$ is a complete measure space.

Theorem 8.3 implies that

$$\otimes_{j=1}^n \Sigma_j \subseteq \Sigma_{\otimes_{j=1}^n \mu_j}$$

and

$$(\otimes_{j=1}^n \mu_j)\big(\prod_{j=1}^n A_j\big) = \prod_{j=1}^n \mu_j(A_j)$$

for every $A_1 \in \Sigma_1, \ldots, A_n \in \Sigma_n$.

It is very common to consider the restriction, also denoted by $\otimes_{j=1}^n \mu_j$, of $\otimes_{j=1}^n \mu_j$ on $\otimes_{j=1}^n \Sigma_j$.

Theorem 8.2 *Let (X_i, Σ_i, μ_i) be a measure space for each $i = 1, \ldots, n$. If $\mu_1, \ldots, \mu_n$ are σ-finite measures, then*
(i) $\otimes_{j=1}^n \mu_j$ is the unique measure on $(\prod_{j=1}^n X_j, \otimes_{j=1}^n \Sigma_j)$ with the property: $(\otimes_{j=1}^n \mu_j)\big(\prod_{j=1}^n A_j\big) = \prod_{j=1}^n \mu_j(A_j)$ for every $A_1 \in \Sigma_1, \ldots, A_n \in \Sigma_n$ and
(ii) the measure space $(\prod_{j=1}^n X_j, \Sigma_{\otimes_{j=1}^n \mu_j}, \otimes_{j=1}^n \mu_j)$ is the completion of the measure space $(\prod_{j=1}^n X_j, \otimes_{j=1}^n \Sigma_j, \otimes_{j=1}^n \mu_j)$.

Proof: (i) Take the algebra $\mathcal{A}$ of subsets of $\prod_{j=1}^n X_j$ described in Proposition 8.4. If μ is any measure on $(\prod_{j=1}^n X_j, \otimes_{j=1}^n \Sigma_j)$ such that $\mu(\tilde{R}) = (\otimes_{j=1}^n \mu_j)(\tilde{R})$ for every $\tilde{R} \in \tilde{\mathcal{E}}$, then, by additivity of the measures, we have that $\mu(\tilde{R}_1 \cup \cdots \cup \tilde{R}_m) = (\otimes_{j=1}^n \mu_j)(\tilde{R}_1 \cup \cdots \cup \tilde{R}_m)$ for all pairwise disjoint $\tilde{R}_1, \ldots, \tilde{R}_m \in \tilde{\mathcal{E}}$. Therefore, the measures μ and $\otimes_{j=1}^n \mu_j$ are equal on $\mathcal{A}$.

Since all measures μ_j are σ-finite, there exist $A_j^{(i)} \in \Sigma_j$ with $\mu_j(A_j^{(i)}) < +\infty$ for every i, j and $A_j^{(i)} \uparrow X_j$ for every j. This implies that the measurable intervals $\tilde{S}_i = \prod_{j=1}^n A_j^{(i)}$ have the property that $\tilde{S}_i \uparrow \prod_{j=1}^n X_j$ and that $\mu(\tilde{S}_i) = (\otimes_{j=1}^n \mu_j)(\tilde{S}_i) = \prod_{j=1}^n \mu_j(A_j^{(i)}) < +\infty$ for every i.

Since $\otimes_{j=1}^n \Sigma_j = \Sigma(\tilde{\mathcal{E}}) = \Sigma(\mathcal{A})$, Theorem 2.4 implies that μ and $\otimes_{j=1}^n \mu_j$ are equal on $\otimes_{j=1}^n \Sigma_j$.
(ii) We already know that $(\prod_{j=1}^n X_j, \Sigma_{\otimes_{j=1}^n \mu_j}, \otimes_{j=1}^n \mu_j)$ is a complete extension of $(\prod_{j=1}^n X_j, \otimes_{j=1}^n \Sigma_j, \otimes_{j=1}^n \mu_j)$. Therefore, it is also an extension of the completion $(\prod_{j=1}^n X_j, \overline{\otimes_{j=1}^n \Sigma_j}, \overline{\otimes_{j=1}^n \mu_j})$ and it is enough to prove that every $E \in \Sigma_{\otimes_{j=1}^n \mu_j}$ belongs to $\overline{\otimes_{j=1}^n \Sigma_j}$.

Take any $E \in \Sigma_{\otimes_{j=1}^n \mu_j}$ and assume, at first, that $(\otimes_{j=1}^n \mu_j)(E) < +\infty$.

We take arbitrary $k \in \mathbf{N}$ and we find a covering $E \subseteq \cup_{i=1}^{+\infty} \tilde{R}_i^{(k)}$ by pairwise disjoint measurable intervals so that $\sum_{i=1}^{+\infty} \tau(\tilde{R}_i^{(k)}) < (\otimes_{j=1}^n \mu_j)(E) + \frac{1}{k}$. We define $B_k = \cup_{i=1}^{+\infty} \tilde{R}_i^{(k)} \in \otimes_{j=1}^n \Sigma_j$ and have that $E \subseteq B_k$ and $(\otimes_{j=1}^n \mu_j)(E) \leq (\otimes_{j=1}^n \mu_j)(B_k) < (\otimes_{j=1}^n \mu_j)(E) + \frac{1}{k}$. Now, define $A = \cap_{k=1}^{+\infty} B_k \in \otimes_{j=1}^n \Sigma_j$. Then $E \subseteq A$ and $(\otimes_{j=1}^n \mu_j)(E) = (\otimes_{j=1}^n \mu_j)(A)$. Therefore $(\otimes_{j=1}^n \mu_j)(A \setminus E) = 0$.

In case $(\otimes_{j=1}^n \mu_j)(E) = +\infty$, we consider the specific sets $\tilde{S}_i$, which were constructed in the proof of part (i), and take the sets $E_i = E \cap \tilde{S}_i$. These sets have $(\otimes_{j=1}^n \mu_j)(E_i) < +\infty$ and, by the previous paragraph, we can find $A_i \in \otimes_{j=1}^n \Sigma_j$ so that $E_i \subseteq A_i$ and $(\otimes_{j=1}^n \mu_j)(A_i \setminus E_i) = 0$. We define $A = \cup_{i=1}^{+\infty} A_i \in \otimes_{j=1}^n \Sigma_j$ so that $E \subseteq A$ and, since $A \setminus E \subseteq \cup_{i=1}^{+\infty} (A_i \setminus E_i)$, we conclude that $(\otimes_{j=1}^n \mu_j)(A \setminus E) = 0$.

We have proved that for every $E \in \Sigma_{\otimes_{j=1}^n \mu_j}$ there exists $A \in \otimes_{j=1}^n \Sigma_j$ so that $E \subseteq A$ and $(\otimes_{j=1}^n \mu_j)(A \setminus E) = 0$.

Considering $A \setminus E$ instead of E, we find a set $B \in \otimes_{j=1}^n \Sigma_j$ so that $A \setminus E \subseteq B$ and $(\otimes_{j=1}^n \mu_j)\big(B \setminus (A \setminus E)\big) = 0$. Of course, $(\otimes_{j=1}^n \mu_j)(B) = 0$.

Now we observe that $E = (A \setminus B) \cup (E \cap B)$, where $A \setminus B \in \otimes_{j=1}^n \Sigma_j$ and $E \cap B \subseteq B \in \otimes_{j=1}^n \Sigma_j$ with $(\otimes_{j=1}^n \mu_j)(B) = 0$. This says that $E \in \overline{\otimes_{j=1}^n \Sigma_j}$.

We shall examine, now, the influence to the product measure of replacing the measure spaces (X_j, Σ_j, μ_j) by their completions $(X_j, \overline{\Sigma_j}, \overline{\mu_j})$.

Theorem 8.3 *Let (X_j, Σ_j, μ_j) and $(X_j, \overline{\Sigma_j}, \overline{\mu_j})$ be a measure space and its completion for every $j = 1, \ldots, n$.*
(i) The measure spaces (X_j, Σ_j, μ_j) induce the same product measure space as

their completions $(X_j,\overline{\Sigma_j},\overline{\mu_j})$. *Namely,*

$$(\prod_{j=1}^n X_j,\Sigma_{\otimes_{j=1}^n\mu_j},\otimes_{j=1}^n\mu_j)=(\prod_{j=1}^n X_j,\Sigma_{\otimes_{j=1}^n\overline{\mu_j}},\otimes_{j=1}^n\overline{\mu_j}).$$

Moreover, the above product measure space is an extension of both measure spaces $(\prod_{j=1}^n X_j,\otimes_{j=1}^n\Sigma_j,\otimes_{j=1}^n\mu_j)$ *and* $(\prod_{j=1}^n X_j,\otimes_{j=1}^n\overline{\Sigma_j},\otimes_{j=1}^n\overline{\mu_j})$, *of which the second is an extension of the first.*
(ii) If each (X_j,Σ_j,μ_j) *is* σ*-finite, then* $(\prod_{j=1}^n X_j,\Sigma_{\otimes_{j=1}^n\mu_j},\otimes_{j=1}^n\mu_j)$ *is the completion of both* $(\prod_{j=1}^n X_j,\otimes_{j=1}^n\Sigma_j,\otimes_{j=1}^n\mu_j)$ *and* $(\prod_{j=1}^n X_j,\otimes_{j=1}^n\overline{\Sigma_j},\otimes_{j=1}^n\overline{\mu_j})$.

Proof: (i) To construct the product measure space $(\prod_{j=1}^n X_j,\Sigma_{\otimes_{j=1}^n\mu_j},\otimes_{j=1}^n\mu_j)$, we first consider all $\otimes_{j=1}^n\Sigma_j$-measurable intervals of the form $\tilde{R}=\prod_{j=1}^n A_j$ for arbitrary $A_j\in\Sigma_j$ and then define the outer measure

$$\mu_1^*(E)=\inf\big\{\sum_{i=1}^{+\infty}\tau(\tilde{R}_i)\,|\,\tilde{R}_i \text{ are } \otimes_{j=1}^n\Sigma_j\text{-measurable intervals and } E\subseteq\cup_{i=1}^{+\infty}\tilde{R}_i\big\}$$

where $\tau(\tilde{R})=\prod_{j=1}^n\mu_j(A_j)$ for all $\tilde{R}=\prod_{j=1}^n A_j$.

To construct the product measure space $(\prod_{j=1}^n X_j,\Sigma_{\otimes_{j=1}^n\overline{\mu_j}},\otimes_{j=1}^n\overline{\mu_j})$, we now consider all $\otimes_{j=1}^n\overline{\Sigma_j}$-measurable intervals of the form $\tilde{R}=\prod_{j=1}^n A_j$ for arbitrary $A_j\in\overline{\Sigma_j}$ and define the outer measure

$$\mu_2^*(E)=\inf\big\{\sum_{i=1}^{+\infty}\tau(\tilde{R}_i)\,|\,\tilde{R}_i \text{ are } \otimes_{j=1}^n\overline{\Sigma_j}\text{-measurable intervals and } E\subseteq\cup_{i=1}^{+\infty}\tilde{R}_i\big\}$$

where $\tau(\tilde{R})=\prod_{j=1}^n\overline{\mu_j}(A_j)$ for all $\tilde{R}=\prod_{j=1}^n A_j$.

Our first task will be to prove that the two outer measures μ_1^* and μ_2^* are identical.

We observe that all $\otimes_{j=1}^n\Sigma_j$-measurable intervals are at the same time $\otimes_{j=1}^n\overline{\Sigma_j}$-measurable and, hence, $\mu_2^*(E)\leq\mu_1^*(E)$ for every $E\subseteq\mathbf{R}^n$.

Now take any $E\subseteq\mathbf{R}^n$ with $\mu_2^*(E)<+\infty$ and an arbitrary $\epsilon>0$. Then there exists a covering $E\subseteq\cup_{i=1}^{+\infty}\tilde{R}_i$ with $\otimes_{j=1}^n\overline{\Sigma_j}$-measurable intervals $\tilde{R}_i$ so that $\sum_{i=1}^{+\infty}\tau(\tilde{R}_i)<\mu_2^*(E)+\epsilon$. For each i, write $\tilde{R}_i=\prod_{j=1}^n A_j^{(i)}$ with $A_j^{(i)}\in\overline{\Sigma_j}$. It is clear that there exist $B_j^{(i)}\in\Sigma_j$ so that $A_j^{(i)}\subseteq B_j^{(i)}$ and $\overline{\mu_j}(A_j^{(i)})=\mu_j(B_j^{(i)})$. We form the $\otimes_{j=1}^n\Sigma_j$-measurable intervals $\tilde{R}_i'=\prod_{j=1}^n B_j^{(i)}$ and have $\tilde{R}_i\subseteq\tilde{R}_i'$ and $\tau(\tilde{R}_i)=\tau(\tilde{R}_i')$ for all i. We now have a covering $E\subseteq\cup_{i=1}^{+\infty}\tilde{R}_i'$ with $\otimes_{j=1}^n\Sigma_j$-measurable intervals, and this implies $\mu_1^*(E)\leq\sum_{i=1}^{+\infty}\tau(\tilde{R}_i')=\sum_{i=1}^{+\infty}\tau(\tilde{R}_i)<\mu_2^*(E)+\epsilon$. Since ϵ is arbitrary, we find $\mu_1^*(E)\leq\mu_2^*(E)$. In the remaining case $\mu_2^*(E)=+\infty$ the inequality $\mu_1^*(E)\leq\mu_2^*(E)$ is obviously true and we conclude that

$$\mu_1^*(E)=\mu_2^*(E)$$

for every $E\subseteq\mathbf{R}^n$.

The next step in forming the product measure is to apply the process of Caratheodory to the common outer measure $\mu^* = \mu_1^* = \mu_2^*$ and find the common complete product measure space

$$(\prod_{j=1}^{n} X_j, \Sigma_{\otimes_{j=1}^n \mu_j}, \otimes_{j=1}^n \mu_j) = (\prod_{j=1}^{n} X_j, \Sigma_{\otimes_{j=1}^n \overline{\mu_j}}, \otimes_{j=1}^n \overline{\mu_j}).$$

where $\Sigma_{\otimes_{j=1}^n \mu_j} = \Sigma_{\otimes_{j=1}^n \overline{\mu_j}}$ is the symbol we use for Σ_{μ^*}, the σ-algebra of μ^*-measurable sets, and $\otimes_{j=1}^n \mu_j = \otimes_{j=1}^n \overline{\mu_j}$ is the restriction of μ^* on Σ_{μ^*}.

Theorem 8.3 says that $\otimes_{j=1}^n \Sigma_j$ and $\otimes_{j=1}^n \overline{\Sigma_j}$ are included in $\Sigma_{\otimes_{j=1}^n \mu_j}$ and, since every $\otimes_{j=1}^n \Sigma_j$-measurable interval is also a $\otimes_{j=1}^n \overline{\Sigma_j}$-measurable interval, we have that $\otimes_{j=1}^n \Sigma_j$ is included in $\otimes_{j=1}^n \overline{\Sigma_j}$. Thus

$$\otimes_{j=1}^n \Sigma_j \subseteq \otimes_{j=1}^n \overline{\Sigma_j} \subseteq \Sigma_{\otimes_{j=1}^n \mu_j}.$$

(ii) The proof is immediate from Theorem 8.4.

The most basic application of Theorem 8.5 is related to the n-dimensional Lebesgue measure. The next result is no surprise, since the n-dimensional Lebesgue measure of any interval in $\mathbf{R}^n$ is equal to the product of the 1-dimensional Lebesgue measure 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 8.4 *(i) The Lebesgue measure space* $(\mathbf{R}^n, \mathcal{L}_n, m_n)$ *is the product measure space of* n *copies of* $(\mathbf{R}, \mathcal{B}_{\mathbf{R}}, m_1)$ *and, at the same time, the product measure space of* n *copies of* $(\mathbf{R}, \mathcal{L}_1, m_1)$.
(ii) The Lebesgue measure space $(\mathbf{R}^n, \mathcal{L}_n, m_n)$ *is the completion of both measure spaces* $(\mathbf{R}^n, \otimes_{j=1}^n \mathcal{B}_{\mathbf{R}}, m_n) = (\mathbf{R}^n, \mathcal{B}_{\mathbf{R}^n}, m_n)$ *and* $(\mathbf{R}^n, \otimes_{j=1}^n \mathcal{L}_1, m_n)$*, of which the second is an extension of the first.*

Proof: We know that $\otimes_{j=1}^n \mathcal{B}_{\mathbf{R}} = \mathcal{B}_{\mathbf{R}^n}$, that $(\mathbf{R}, \mathcal{L}_1, m_1)$ is the completion of $(\mathbf{R}, \mathcal{B}_{\mathbf{R}}, m_1)$ and that m_1 is a σ-finite measure.

Hence, Theorem 8.5 implies immediately that the n copies of $(\mathbf{R}, \mathcal{B}_{\mathbf{R}}, m_1)$ and, at the same time, the n copies of $(\mathbf{R}, \mathcal{L}_1, m_1)$ induce the same product measure space $(\mathbf{R}^n, \Sigma_{\otimes_{j=1}^n m_1}, \otimes_{j=1}^n m_1)$, which is the completion of both measure spaces $(\mathbf{R}^n, \mathcal{B}_{\mathbf{R}^n}, \otimes_{j=1}^n m_1)$ and $(\mathbf{R}^n, \otimes_{j=1}^n \mathcal{L}_1, \otimes_{j=1}^n m_1)$, of which the second is an extension of the first.

Theorem 8.3 says that, for every Borel measurable interval $\tilde{R} = \prod_{j=1}^n A_j$, we have $(\otimes_{j=1}^n m_1)(\tilde{R}) = \prod_{j=1}^n m_1(A_j)$. In particular, $(\otimes_{j=1}^n m_1)(P) = \mathrm{vol}_n(P)$ for every open-closed interval P in $\mathbf{R}^n$ and Theorem 4.5 implies that $\otimes_{j=1}^n m_1 = m_n$ on $\mathcal{B}_{\mathbf{R}^n}$. Hence

$$(\mathbf{R}^n, \mathcal{B}_{\mathbf{R}^n}, \otimes_{j=1}^n m_1) = (\mathbf{R}^n, \mathcal{B}_{\mathbf{R}^n}, m_n).$$

The proof finishes because $(\mathbf{R}^n, \mathcal{L}_n, m_n)$ is the completion of $(\mathbf{R}^n, \mathcal{B}_{\mathbf{R}^n}, m_n)$.

It is, perhaps, surprising that, although the measure space $(\mathbf{R}, \mathcal{L}_1, m_1)$ is complete, the product $(\mathbf{R}^n, \otimes_{j=1}^n \mathcal{L}_1, m_n)$ is *not* complete (when $n \geq 2$, of course). It is easy to see this. Take any non Lebesgue measurable set $A \subseteq \mathbf{R}$ and form the set $E = A \times \{0\} \times \cdots \times \{0\} \subseteq \mathbf{R}^n$. Consider, also, the Lebesgue measurable interval $\tilde{R} = \mathbf{R} \times \{0\} \times \cdots \times \{0\} \subseteq \mathbf{R}^n$. We have that $E \subseteq \tilde{R}$ and $m_n(\tilde{R}) = m_1(\mathbf{R}) m_1(\{0\}) \cdots m_1(\{0\}) = 0$. If we assume that $(\mathbf{R}^n, \otimes_{j=1}^n \mathcal{L}_1, m_n)$ is complete, then we conclude that $E \in \otimes_{j=1}^n \mathcal{L}_1$. We now take $z = (0, \ldots, 0) \in \mathbf{R}^{n-1}$ and, then, the section $E_z = A$ must belong to $\mathcal{L}_1$. This is not true and we arrive at a contradiction.

8.3 Multiple integrals.

The purpose of this section is to give 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 over sets of lower dimension.

For the sake of simplicity, we further restrict to the case of two measure spaces.

Theorem 8.5 *Let (X_1, Σ_1, μ_1) and (X_2, Σ_2, μ_2) be two measure spaces and $(X_1 \times X_2, \Sigma_{\mu_1 \otimes \mu_2}, \mu_1 \otimes \mu_2)$ be their product measure space.*

If $E \in \Sigma_{\mu_1 \otimes \mu_2}$ has σ-finite $\mu_1 \otimes \mu_2$-measure, then $E_{x_1} \in \overline{\Sigma_2}$ and $E_{x_2} \in \overline{\Sigma_1}$ for μ_1-a.e. $x_1 \in X_1$ and μ_2-a.e. $x_2 \in X_2$ and the a.e. defined functions

$$x_1 \mapsto \overline{\mu_2}(E_{x_1}), \qquad x_2 \mapsto \overline{\mu_1}(E_{x_2})$$

are $\overline{\Sigma_1}$-measurable and, respectively, $\overline{\Sigma_2}$-measurable. Also,

$$(\mu_1 \otimes \mu_2)(E) = \int_{X_1} \overline{\mu_2}(E_{x_1})\, d\overline{\mu_1}(x_1) = \int_{X_2} \overline{\mu_1}(E_{x_2})\, d\overline{\mu_2}(x_2).$$

Proof: As shown by Theorem 8.5, it is true that $\Sigma_{\overline{\mu_1} \otimes \overline{\mu_2}} = \Sigma_{\mu_1 \otimes \mu_2}$ and $\overline{\mu_1} \otimes \overline{\mu_2} = \mu_1 \otimes \mu_2$. It is also immediate that $E_{x_1} \in \overline{\Sigma_2}$ for μ_1-a.e. $x_1 \in X_1$ if and only if $E_{x_1} \in \overline{\Sigma_2}$ for $\overline{\mu_1}$-a.e. $x_1 \in X_1$ and, similarly, $E_{x_2} \in \overline{\Sigma_1}$ for μ_2-a.e. $x_2 \in X_2$ if and only if $E_{x_2} \in \overline{\Sigma_1}$ for $\overline{\mu_2}$-a.e. $x_2 \in X_2$. Hence, the whole statement of the theorem remains the same if we replace at each occurence the measure spaces (X_1, Σ_1, μ_1) and (X_2, Σ_2, μ_2) by their completions $(X_1, \overline{\Sigma_1}, \overline{\mu_1})$ and $(X_2, \overline{\Sigma_2}, \overline{\mu_2})$. Renaming, we restate the theorem as follows:

Let (X_1, Σ_1, μ_1), (X_2, Σ_2, μ_2) and $(X_1 \times X_2, \Sigma_{\mu_1 \otimes \mu_2}, \mu_1 \otimes \mu_2)$ be two complete measure spaces and their product measure space. If $E \in \Sigma_{\mu_1 \otimes \mu_2}$ has σ-finite $\mu_1 \otimes \mu_2$-measure, then $E_{x_1} \in \Sigma_2$ and $E_{x_2} \in \Sigma_1$ for μ_1-a.e. $x_1 \in X_1$ and μ_2-a.e. $x_2 \in X_2$ and the a.e. defined functions

$$x_1 \mapsto \mu_2(E_{x_1}), \qquad x_2 \mapsto \mu_1(E_{x_2})$$

are Σ_1*-measurable and, respectively,* Σ_2*-measurable. Also,*

$$(\mu_1\otimes\mu_2)(E)=\int_{X_1}\mu_2(E_{x_1})\,d\mu_1(x_1)=\int_{X_2}\mu_1(E_{x_2})\,d\mu_2(x_2).$$

We are, now, going to prove the theorem in this equivalent form and we denote $\mathcal{N}$ the collection of all sets $E\in\Sigma_{\mu_1\otimes\mu_2}$ which have all the properties in the conclusion of the theorem.

(a) Every measurable interval $\tilde{R}=A_1\times A_2$ belongs to $\mathcal{N}$.

Indeed, $\tilde{R}_{x_1}=\emptyset$, if $x_1\notin A_1$, and $\tilde{R}_{x_1}=A_2$, if $x\in A_1$. Hence, $\mu_2(\tilde{R}_{x_1})=\mu_2(A_2)\chi_{A_1}(x_1)$ for every $x_1\in X_1$, implying that the function $x_1\mapsto\mu_2(\tilde{R}_{x_1})$ is Σ_1-measurable. Moreover, we have $\int_{X_1}\mu_2(\tilde{R}_{x_1})\,d\mu_1=\mu_2(A_2)\int_{X_1}\chi_{A_1}\,d\mu_1=\mu_2(A_2)\mu_1(A_1)=(\mu_1\otimes\mu_2)(\tilde{R})$. The same arguments hold for x_2-sections.

(b) Assume that the sets $E_1,\dots E_m\in\mathcal{N}$ are pairwise disjoint. Then $E=E_1\cup\dots\cup E_m\in\mathcal{N}$.

Indeed, from $E_{x_1}=(E_1)_{x_1}\cup\dots\cup(E_m)_{x_1}$ for every $x_1\in X_1$, we have that $E_{x_1}\in\Sigma_2$ for μ_1-a.e. $x_1\in X_1$ and $\mu_2(E_{x_1})=\mu_2((E_1)_{x_1})+\dots+\mu_2((E_m)_{x_1})$ for μ_1-a.e. $x_1\in X_1$. By the completeness of μ_1, the function $x_1\mapsto\mu_2(E_{x_1})$ is Σ_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 argument holds for x_2-sections.

(c) Assume that $E_n\in\mathcal{N}$ for every $n\in\mathbf{N}$. If $E_n\uparrow E$, then $E\in\mathcal{N}$.

From $(E_n)_{x_1}\uparrow E_{x_1}$ for every $x_1\in X_1$, we have that $E_{x_1}\in\Sigma_2$ for μ_1-a.e. $x_1\in X_1$. Continuity of μ_2 from below implies that $\mu_2((E_n)_{x_1})\uparrow\mu_2(E_{x_1})$ for μ_1-a.e. $x_1\in X_1$. By the completeness of μ_1, the function $x_1\mapsto\mu_2(E_{x_1})$ is Σ_1- measurable. By continuity of $\mu_1\otimes\mu_2$ from below and from 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) Now, fix any measurable interval $\tilde{R}$ with $(\mu_1\otimes\mu_2)(\tilde{R})<+\infty$ and consider the collection $\mathcal{N}_{\tilde{R}}$ of all sets $E\in\Sigma_{\mu_1\otimes\mu_2}$ for which $E\cap\tilde{R}\in\mathcal{N}$.

If $E_n\in\mathcal{N}_{\tilde{R}}$ for all n and $E_n\downarrow E$, then $E\in\mathcal{N}_{\tilde{R}}$.

Indeed, we have that $E_n\cap\tilde{R}\downarrow E\cap\tilde{R}$ and, hence, $(E_n\cap\tilde{R})_{x_1}\downarrow(E\cap\tilde{R})_{x_1}$ for every $x_1\in X_1$. This implies that $(E\cap\tilde{R})_{x_1}\in\Sigma_2$ for μ_1-a.e. $x_1\in X_1$. From the result of (a), $\int_{X_1}\mu_2(\tilde{R}_{x_1})\,d\mu_1(x_1)=(\mu_1\otimes\mu_2)(\tilde{R})<+\infty$ and, hence, $\mu_2(\tilde{R}_{x_1})<+\infty$ for μ_1-a.e. $x_1\in X_1$. Therefore, $\mu_2\big((E_1\cap\tilde{R})_{x_1}\big)<+\infty$ for μ_1-a.e. $x_1\in X_1$ and, by the continuity of μ_2 from above, we find $\mu_2\big((E_n\cap\tilde{R})_{x_1}\big)\downarrow\mu_2\big((E\cap\tilde{R})_{x_1}\big)$ for μ_1-a.e. $x_1\in X_1$. By the completeness of μ_1, the function $x_1\mapsto\mu_2\big((E\cap\tilde{R})_{x_1}\big)$ is Σ_1-measurable. Another application of continuity from above gives $(\mu_1\otimes\mu_2)(E\cap\tilde{R})=\int_{X_1}\mu_2\big((E\cap\tilde{R})_{x_1}\big)\,d\mu_1(x_1)$ and, since all arguments hold for x_2-sections as well, we conclude that $E\cap\tilde{R}\in\mathcal{N}$ and, hence, $E\in\mathcal{N}_{\tilde{R}}$.

If $E_n\in\mathcal{N}_{\tilde{R}}$ for all n and $E_n\uparrow E$, then $E_n\cap\tilde{R}\uparrow E\cap\tilde{R}$ and, from the result of (c), $E\in\mathcal{N}_{\tilde{R}}$.

We have proved that the collection $\mathcal{N}_{\tilde{R}}$ is a monotone class of subsets of $X_1\times X_2$.

If the $E_1,\dots,E_m\in\mathcal{N}_{\tilde{R}}$ are pairwise disjoint and $E=E_1\cup\dots\cup E_m$, then $E\cap\tilde{R}=(E_1\cap\tilde{R})\cup\dots\cup(E_m\cap\tilde{R})$ and, by the result of (b), $E\in\mathcal{N}_{\tilde{R}}$. From (a), we have that $\mathcal{N}_{\tilde{R}}$ contains all measurable rectangles and, hence, $\mathcal{N}_{\tilde{R}}$ contains all elements of the algebra $\mathcal{A}$ of Proposition 8.4. Therefore, $\mathcal{N}_{\tilde{R}}$ 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 $\Sigma_1 \otimes \Sigma_2$.

This says that $E \cap \tilde{R} \in \mathcal{N}$ for every $E \in \Sigma_1 \otimes \Sigma_2$ and every measurable interval $\tilde{R}$ with $(\mu_1 \otimes \mu_2)(\tilde{R}) < +\infty$.

(e) If $\mathcal{A}$ is, again, the algebra of Proposition 8.4, an application of the results of (b) and (d) implies that $E \cap F \in \mathcal{N}$ for every $E \in \Sigma_1 \otimes \Sigma_2$ and every $F \in \mathcal{A}$ with $(\mu_1 \otimes \mu_2)(F) < +\infty$.

(f) Now, let $E \in \Sigma_1 \otimes \Sigma_2$ with $(\mu_1 \otimes \mu_2)(E) < +\infty$. We find a covering $E \subseteq \cup_{i=1}^{+\infty} \tilde{R}_i$ by measurable intervals so that $\sum_{i=1}^{+\infty} (\mu_1 \otimes \mu_2)(\tilde{R}_i) < (\mu_1 \otimes \mu_2)(E) + 1 < +\infty$. We define $F_n = \cup_{i=1}^{n} \tilde{R}_i \in \mathcal{A}$ and we have that $(\mu_1 \otimes \mu_2)(F_n) < +\infty$ for every n. The result of (e) implies that $E \cap F_n \in \mathcal{N}$ and, since, $E \cap F_n \uparrow E$, we have, by the result of (c), that $E \in \mathcal{N}$.

Hence, $E \in \mathcal{N}$ for every $E \in \Sigma_1 \otimes \Sigma_2$ with $(\mu_1 \otimes \mu_2)(E) < +\infty$.

(g) Now let $E \in \Sigma_{\mu_1 \otimes \mu_2}$ with $(\mu_1 \otimes \mu_2)(E) = 0$. We shall prove that $E \in \mathcal{N}$.

We find, for every $k \in \mathbf{N}$, a covering $E \subseteq \cup_{i=1}^{+\infty} \tilde{R}_i^{(k)}$ by measurable intervals so that $\sum_{i=1}^{+\infty} (\mu_1 \otimes \mu_2)(\tilde{R}_i^{(k)}) < \frac{1}{k}$. We define $A_k = \cup_{i=1}^{+\infty} \tilde{R}_i^{(k)} \in \Sigma_1 \otimes \Sigma_2$ and have that $E \subseteq A_k$ and $(\mu_1 \otimes \mu_2)(A_k) < \frac{1}{k}$. We then write $A = \cap_{k=1}^{+\infty} A_k \in \Sigma_1 \otimes \Sigma_2$ and have that $E \subseteq A$ and $(\mu_1 \otimes \mu_2)(A) = 0$. From the result of (f) we have that $A \in \mathcal{N}$ and, in particular, $0 = \int_{X_1} \mu_2(A_{x_1})\, d\mu_1(x_1) = \int_{X_2} \mu_1(A_{x_2})\, d\mu_2(x_2)$. The first equality implies that $\mu_2(A_{x_1}) = 0$ for μ_1-a.e. $x_1 \in X_1$. From $E_{x_1} \subseteq A_{x_1}$ and from the completeness of μ_2, we see that $E_{x_1} \in \Sigma_2$ and $\mu_2(E_{x_1}) = 0$ for μ_1-a.e. $x_1 \in X_1$. Now, from the completeness of μ_1, we get that the function $x_1 \mapsto \mu_2(E_{x_1})$ is Σ_1-measurable. Moreover, $(\mu_1 \otimes \mu_2)(E) = 0 = \int_{X_1} \mu_2(E_{x_1})\, d\mu_1(x_1)$ and the same arguments hold for x_2-sections. Therefore, $E \in \mathcal{N}$.

(h) If $E \in \Sigma_{\mu_1 \otimes \mu_2}$ has $(\mu_1 \otimes \mu_2)(E) < +\infty$, then $E \in \mathcal{N}$.

Indeed, for every $k \in \mathbf{N}$ we find a covering $E \subseteq \cup_{i=1}^{+\infty} \tilde{R}_i^{(k)}$ by measurable intervals so that $\sum_{i=1}^{+\infty} (\mu_1 \otimes \mu_2)(\tilde{R}_i^{(k)}) < (\mu_1 \otimes \mu_2)(E) + \frac{1}{k}$. We define $A_k = \cup_{i=1}^{+\infty} \tilde{R}_i^{(k)} \in \Sigma_1 \otimes \Sigma_2$ and have that $E \subseteq A_k$ and $(\mu_1 \otimes \mu_2)(A_k) < (\mu_1 \otimes \mu_2)(E) + \frac{1}{k}$. We then write $A = \cap_{k=1}^{+\infty} A_k \in \Sigma_1 \otimes \Sigma_2$ and have that $E \subseteq A$ and $(\mu_1 \otimes \mu_2)(A) = (\mu_1 \otimes \mu_2)(E)$. Hence $A \setminus E \in \Sigma_{\mu_1 \otimes \mu_2}$ has $(\mu_1 \otimes \mu_2)(A \setminus E) = 0$. As in part (g), we can find $A' \in \Sigma_1 \otimes \Sigma_2$ so that $A \setminus E \subseteq A'$ and $(\mu_1 \otimes \mu_2)(A') = 0$. We set $B = A \setminus A' \in \Sigma_1 \otimes \Sigma_2$ and we have $B \subseteq E$ and $(\mu_1 \otimes \mu_2)(E \setminus B) = 0$. By the result of (g), we have $E \setminus B \in \mathcal{N}$ and, by the result of (f), $B \in \mathcal{N}$. By the result of (b), $E = B \cup (E \setminus B) \in \mathcal{N}$.

(i) Finally, if $E \in \Sigma_{\mu_1 \otimes \mu_2}$ has σ-finite $(\mu_1 \otimes \mu_2)$-measure, we find $E_n \in \Sigma_{\mu_1 \otimes \mu_2}$ with $(\mu_1 \otimes \mu_2)(E_n) < +\infty$ for every n and so that $E_n \uparrow E$. Another application of the result of (c) implies that $E \in \mathcal{N}$.

Theorem 8.6 *Let (X_1, Σ_1, μ_1) and (X_2, Σ_2, μ_2) be σ-finite measure spaces and $(X_1 \times X_2, \Sigma_1 \otimes \Sigma_2, \mu_1 \otimes \mu_2)$ be their (restricted) product measure space.*

If $E \in \Sigma_1 \otimes \Sigma_2$, then $E_{x_1} \in \Sigma_2$ and $E_{x_2} \in \Sigma_1$ for every $x_1 \in X_1$ and $x_2 \in X_2$ and the functions

$$x_1 \mapsto \mu_2(E_{x_1}), \qquad x_2 \mapsto \mu_1(E_{x_2})$$

are Σ_1*-measurable and, respectively,* Σ_2*-measurable. Also,*

$$(\mu_1 \otimes \mu_2)(E) = \int_{X_1} \mu_2(E_{x_1})\, d\mu_1(x_1) = \int_{X_2} \mu_1(E_{x_2})\, d\mu_2(x_2).$$

Proof: Exactly as in the proof of Theorem 8.7, we denote $\mathcal{N}$ the collection of all $E \in \Sigma_1 \otimes \Sigma_2$ which satisfy all the properties in the conclusion of this theorem.
(a) If $\tilde{R}$ is any measurable interval, then $\tilde{R} \in \mathcal{N}$.

The proof is identical to the proof of the result of (a) of Theorem 8.7. Observe that, now, all statements hold for every $x_1 \in X_1$ and $x_2 \in X_2$ and there is no need of completeness.
(b) If the sets $E_1, \dots E_m \in \mathcal{N}$ are pairwise disjoint, then $E = E_1 \cup \dots \cup E_m \in \mathcal{N}$.

The proof is identical to the proof of the result of (b) of Theorem 8.7.
(c) If $E_n \in \mathcal{N}$ for every $n \in \mathbf{N}$ and $E_n \uparrow E$, then $E \in \mathcal{N}$.

The proof is identical to the proof of the result of (c) of Theorem 8.7.
(d) We fix any measurable interval $\tilde{R} = A_1 \times A_2$ with $\mu_1(A_1) < +\infty$ and $\mu_2(A_2) < +\infty$ and consider the collection $\mathcal{N}_{\tilde{R}}$ of all sets $E \in \Sigma_1 \otimes \Sigma_2$ for which $E \cap \tilde{R} \in \mathcal{N}$. The rest of the proof of part (d) of Theorem 8.7 continues unchanged and we get that $\mathcal{N}_{\tilde{R}}$ is a monotone class of subsets of $X_1 \times X_2$ which includes the algebra $\mathcal{A}$ of Proposition 8.4. Hence, $\mathcal{N}_{\tilde{R}}$ includes $\Sigma_1 \otimes \Sigma_2$ and this says that $E \cap \tilde{R} \in \mathcal{N}$ for every $E \in \Sigma_1 \otimes \Sigma_2$ and every measurable interval $\tilde{R} = A_1 \times A_2$ with $\mu_1(A_1) < +\infty$ and $\mu_2(A_2) < +\infty$.
(e) Since μ_1 is σ-finite, we can find an increasing sequence $(A_1^{(n)})$ so that $A_1^{(n)} \in \Sigma_1$, $A_1^{(n)} \uparrow X_1$ and $0 < \mu_1(A_1^{(n)}) < +\infty$ for every n. Similarly, we can find an increasing sequence $(A_2^{(n)})$ so that $A_2^{(n)} \in \Sigma_2$, $A_2^{(n)} \uparrow X_2$ and $0 < \mu_2(A_2^{(n)}) < +\infty$ for every n and we form the measurable intervals $\tilde{R}_n = A_1^{(n)} \times A_2^{(n)}$.

We take any $E \in \Sigma_1 \otimes \Sigma_2$ and, from the result of (d), we have that all sets $E_n = E \cap \tilde{R}_n$ belong to $\mathcal{N}$. Since $E_n \uparrow E$, an application of the result of (c) implies that $E \in \mathcal{N}$.

Theorem 8.7 *(Tonelli) Let* (X_1, Σ_1, μ_1) *and* (X_2, Σ_2, μ_2) *be measure spaces and* $(X_1 \times X_2, \Sigma_{\mu_1 \otimes \mu_2}, \mu_1 \otimes \mu_2)$ *be their product measure space.*

If $f : X_1 \times X_2 \to [0, +\infty]$ *is* $\Sigma_{\mu_1 \otimes \mu_2}$*-measurable and if* $f^{-1}((0, +\infty])$ *has* σ*-finite* $\mu_1 \otimes \mu_2$*-measure, then* f_{x_1} *is* $\overline{\Sigma_2}$*-measurable for* μ_1*-a.e.* $x_1 \in X_1$ *and* f_{x_2} *is* $\overline{\Sigma_1}$*-measurable for* μ_2*-a.e.* $x_2 \in X_2$ *and the a.e. defined functions*

$$x_1 \mapsto \int_{X_2} f_{x_1}\, d\overline{\mu_2}, \qquad x_2 \mapsto \int_{X_1} f_{x_2}\, d\overline{\mu_1}$$

are $\overline{\Sigma_1}$*-measurable and, respectively,* $\overline{\Sigma_2}$*-measurable. Also,*

$$\int_{X_1 \times X_2} f\, d(\mu_1 \otimes \mu_2) = \int_{X_1} \Big(\int_{X_2} f_{x_1}\, d\overline{\mu_2} \Big)\, d\overline{\mu_1}(x_1) = \int_{X_2} \Big(\int_{X_1} f_{x_2}\, d\overline{\mu_1} \Big)\, d\overline{\mu_2}(x_2).$$

Proof: (a) A first particular case is when $f = \chi_E$ is the characteristic function of an $E \in \Sigma_{\mu_1 \otimes \mu_2}$ with σ-finite $\mu_1 \otimes \mu_2$-measure.

Theorem 8.7 implies that $(\chi_E)_{x_1} = \chi_{E_{x_1}}$ is $\overline{\Sigma_2}-$measurable for μ_1-a.e. $x_1 \in X_1$ and the function $x_1 \mapsto \int_{X_2} (\chi_E)_{x_1} \, d\overline{\mu_2} = \overline{\mu_2}(E_{x_1})$ is $\overline{\Sigma_1}$-measurable. Finally, we have $\int_{X_1 \times X_2} \chi_E \, d(\mu_1 \otimes \mu_2) = (\mu_1 \otimes \mu_2)(E) = \int_{X_1} \overline{\mu_2}(E_{x_1}) \, d\overline{\mu_1}(x_1) = \int_{X_1} \Big(\int_{X_2} (\chi_E)_{x_1} \, d\overline{\mu_2} \Big) \, d\overline{\mu_1}(x_1)$. The argument for x_2-sections is the same.
(b) Next, we take $\phi = \sum_{j=1}^m \kappa_j \chi_{E_j}$ to be the standard representation of a simple $\phi : X_1 \times X_2 \to [0, +\infty)$, where we omit the possible value $\kappa = 0$, and which is $\Sigma_{\mu_1 \otimes \mu_2}$-measurable and so that $\cup_{j=1}^m E_j = \phi^{-1}((0, +\infty])$ has σ-finite $\mu_1 \otimes \mu_2$-measure. Then, $\phi_{x_1} = \sum_{j=1}^m \kappa_j (\chi_{E_j})_{x_1}$ and $\phi_{x_2} = \sum_{j=1}^m \kappa_j (\chi_{E_j})_{x_2}$ for every $x_1 \in X_1$ and $x_2 \in X_2$. Therefore, this case reduces, by linearity, to (a).
(c) Finally, we take any $\Sigma_{\mu_1 \otimes \mu_2}$-measurable $f : X_1 \times X_2 \to [0, +\infty]$ with $f^{-1}((0, +\infty])$ having σ-finite $\mu_1 \otimes \mu_2$-measure. We take an increasing sequence (ϕ_n) of $\Sigma_{\mu_1 \otimes \mu_2}$-measurable simple functions $\phi_n : X_1 \times X_2 \to [0, +\infty]$ so that $\phi_n \uparrow f$ on $X_1 \times X_2$. From $\phi_n \leq f$, it is clear that $\phi_n^{-1}((0, +\infty])$ has σ-finite $\mu_1 \otimes \mu_2$-measure for every n. Part (b) says that every ϕ_n satisfies the conclusion of the theorem and, since $(\phi_n)_{x_1} \uparrow f_{x_1}$ and $(\phi_n)_{x_2} \uparrow f_{x_2}$ for every $x_1 \in X_1$ and $x_2 \in X_2$, an application of the Monotone Convergence Theorem implies that f also satisfies the conclusion of the theorem.

Theorem 8.8 *(Fubini) Let (X_1, Σ_1, μ_1) and (X_2, Σ_2, μ_2) two measure spaces and $(X_1 \times X_2, \Sigma_{\mu_1 \otimes \mu_2}, \mu_1 \otimes \mu_2)$ their product measure space.*

If $f : X_1 \times X_2 \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$ is integrable with respect to $\mu_1 \otimes \mu_2$, then f_{x_1} is integrable with respect to $\overline{\mu_2}$ for μ_1-a.e. $x_1 \in X_1$ and f_{x_2} is integrable with respect to $\overline{\mu_1}$ for μ_2-a.e. $x_2 \in X_2$ and the a.e. defined functions

$$x_1 \mapsto \int_{X_2} f_{x_1} \, d\overline{\mu_2}, \qquad x_2 \mapsto \int_{X_1} f_{x_2} \, d\overline{\mu_1}$$

are integrable with respect to $\overline{\mu_1}$ and, respectively, integrable with respect to $\overline{\mu_2}$. Also,

$$\int_{X_1 \times X_2} f \, d(\mu_1 \otimes \mu_2) = \int_{X_1} \Big(\int_{X_2} f_{x_1} \, d\overline{\mu_2} \Big) \, d\overline{\mu_1}(x_1) = \int_{X_2} \Big(\int_{X_1} f_{x_2} \, d\overline{\mu_1} \Big) \, d\overline{\mu_2}(x_2).$$

Proof: (a) If $f : X_1 \times X_2 \to [0, +\infty]$ is integrable with respect to $\mu_1 \otimes \mu_2$, Theorem 8.9 gives

$$\int_{X_1} \Big(\int_{X_2} f_{x_1} \, d\overline{\mu_2} \Big) \, d\overline{\mu_1} = \int_{X_2} \Big(\int_{X_1} f_{x_2} \, d\overline{\mu_1} \Big) \, d\overline{\mu_2} = \int_{X_1 \times X_2} f \, d(\mu_1 \otimes \mu_2) < +\infty.$$

This implies $\int_{X_2} f_{x_1} \, d\overline{\mu_2} < +\infty$ for μ_1-a.e. $x_1 \in X_1$ and $\int_{X_1} f_{x_2} \, d\overline{\mu_1} < +\infty$ for μ_2-a.e. $x_2 \in X_2$. Thus, the conclusion of the theorem is true for non-negative functions.
(b) If $f : X_1 \times X_2 \to \overline{\mathbf{R}}$ is integrable with respect to $\mu_1 \otimes \mu_2$, the same is true for f^+ and f^- and, by the result of (a), the conclusion is true for these two functions. Since $f_{x_1} = (f^+)_{x_1} - (f^-)_{x_1}$ and $f_{x_2} = (f^+)_{x_2} - (f^-)_{x_2}$ for every $x_1 \in X_1$ and $x_2 \in X_2$, the conclusion is, by linearity, true also for f.
(c) If $f : X_1 \times X_2 \to \mathbf{C}$ is integrable with respect to $\mu_1 \otimes \mu_2$, the same is true

for $\Re(f)$ and $\Im(f)$. By the result of (b), the conclusion is true for $\Re(f)$ and $\Im(f)$ and, since $f_{x_1} = \Re(f)_{x_1} + i\Im(f)_{x_1}$ and $f_{x_2} = \Re(f)_{x_2} + i\Im(f)_{x_2}$ for every $x_1 \in X_1$ and $x_2 \in X_2$, the conclusion is, by linearity, true for f.
(d) Finally, let $f : X_1 \times X_2 \to \overline{\mathbf{C}}$ be integrable with respect to $\mu_1 \otimes \mu_2$. Then the set $E = f^{-1}(\{\infty\}) \in \Sigma_{\mu_1 \otimes \mu_2}$ has $(\mu_1 \otimes \mu_2)(E) = 0$. Theorem 8.7 implies that $\overline{\mu_2}(E_{x_1}) = 0$ for μ_1-a.e. $x_1 \in X_1$ and $\overline{\mu_1}(E_{x_2}) = 0$ for μ_2-a.e. $x_2 \in X_2$.

If we define $F = f\chi_{E^c}$, then $F : X_1 \times X_2 \to \mathbf{C}$ is integrable with respect to $\mu_1 \otimes \mu_2$ and, by (c), 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 $\int_{X_1 \times X_2} F \, d(\mu_1 \otimes \mu_2) = \int_{X_1 \times X_2} f \, d(\mu_1 \otimes \mu_2)$. We, also, have that $F_{x_1} = f_{x_1}$ on $X_2 \setminus E_{x_1}$ and, hence, $F_{x_1} = f_{x_1}$ holds $\overline{\mu_2}$-a.e. on X_2 for μ_1-a.e. $x_1 \in X_1$. Therefore, f_{x_1} is integrable with respect to $\overline{\mu_2}$ and $\int_{X_2} f_{x_1} \, d\overline{\mu_2} = \int_{X_2} F_{x_1} \, d\overline{\mu_2}$, for μ_1-a.e. $x_1 \in X_1$. This implies $\int_{X_1} \Big(\int_{X_2} f_{x_1} \, d\overline{\mu_2} \Big) \, d\overline{\mu_1}(x_1) = \int_{X_1} \Big(\int_{X_2} F_{x_1} \, d\overline{\mu_2} \Big) \, d\overline{\mu_1}(x_1)$ and, equating the corresponding integrals of F, we find $\int_{X_1 \times X_2} f \, d(\mu_1 \otimes \mu_2) = \int_{X_1} \Big(\int_{X_2} f_{x_1} \, d\overline{\mu_2} \Big) \, d\overline{\mu_1}(x_1)$. The argument is the same for x_2-sections.

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 integrable with respect to the product measure $\mu_1 \otimes \mu_2$. The Theorem of Tonelli is applied to non-negative functions f which must be $\Sigma_{\mu_1 \otimes \mu_2}$-measurable and whose set $f^{-1}((0, +\infty])$ must be of σ-finite $\mu_1 \otimes \mu_2$-measure. Thus, the assumptions of 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 to interchange successive integrations, is first to prove that f is $\Sigma_{\mu_1 \otimes \mu_2}$-measurable and that the set $\{(x_1, x_2) \,|\, f(x_1, x_2) \neq 0\}$ is of σ-finite $\mu_1 \otimes \mu_2$-measure. We, then, apply the Theorem of Tonelli to $|f|$ and have

$$\int_{X_1 \times X_2} |f| \, d(\mu_1 \otimes \mu_2) = \int_{X_1} \Big(\int_{X_2} |f|_{x_1} \, d\overline{\mu_2} \Big) \, d\overline{\mu_1} = \int_{X_2} \Big(\int_{X_1} |f|_{x_2} \, d\overline{\mu_1} \Big) \, d\overline{\mu_2}.$$

By calculating either the second or the third term in this string of equalities, we calculate the $\int_{X_1 \times X_2} |f| \, d(\mu_1 \otimes \mu_2)$. If it is finite, then f is integrable with respect to the product measure $\mu_1 \otimes \mu_2$ and we may apply the Theorem of Fubini to find the desired

$$\begin{aligned}\int_{X_1 \times X_2} f(x_1, x_2) \, d(\mu_1 \otimes \mu_2)(x_1, x_2) &= \int_{X_1} \Big(\int_{X_2} f(x_1, x_2) \, d\overline{\mu_2}(x_2) \Big) \, d\overline{\mu_1}(x_1) \\ &= \int_{X_2} \Big(\int_{X_1} f(x_1, x_2) \, d\overline{\mu_1}(x_1) \Big) \, d\overline{\mu_2}(x_2).\end{aligned}$$

Of the two starting assumptions, the σ-finiteness of $\{(x_1, x_2) \,|\, f(x_1, x_2) \neq 0\}$ is usually easy to check. For example, if the measure spaces (X_1, Σ_1, μ_1) and (X_2, Σ_2, μ_2) are both σ-finite, then the measure space $(X_1 \times X_2, \Sigma_{\mu_1 \otimes \mu_2}, \mu_1 \otimes \mu_2)$

is also σ-finite and all subsets of $X_1 \times X_2$ are obviously of σ-finite $\mu_1 \otimes \mu_2$-measure.

The assumption of $\Sigma_{\mu_1\otimes\mu_2}$-measurability of f is more subtle and sometimes difficult to verify.

Theorem 8.9 *(Tonelli) Let (X_1, Σ_1, μ_1) and (X_2, Σ_2, μ_2) be σ-finite measure spaces and $(X_1 \times X_2, \Sigma_1 \otimes \Sigma_2, \mu_1 \otimes \mu_2)$ be their (restricted) product measure space.*

If $f : X_1 \times X_2 \to [0, +\infty]$ is $\Sigma_1 \otimes \Sigma_2$-measurable, then f_{x_1} is Σ_2-measurable for every $x_1 \in X_1$ and f_{x_2} is Σ_1-measurable for every $x_2 \in X_2$ and the functions

$$x_1 \mapsto \int_{X_2} f_{x_1}\, d\mu_2, \qquad x_2 \mapsto \int_{X_1} f_{x_2}\, d\mu_1$$

are Σ_1-measurable and, respectively, Σ_2-measurable. Also,

$$\int_{X_1\times X_2} f\, d(\mu_1\otimes\mu_2) = \int_{X_1} \Big(\int_{X_2} f_{x_1}\, d\mu_2\Big)\, d\mu_1(x_1) = \int_{X_2} \Big(\int_{X_1} f_{x_2}\, d\mu_1\Big)\, d\mu_2(x_2).$$

Proof The measurability of the sections is an immediate application of Theorem 8.2 and does not need the assumption about σ-finiteness. Otherwise, the proof results from Theorem 8.8 in exactly the same way in which the proof of Theorem 8.9 results from Theorem 8.7.

Theorem 8.10 *(Fubini) Let (X_1, Σ_1, μ_1) and (X_2, Σ_2, μ_2) be two σ-finite measure spaces and $(X_1 \times X_2, \Sigma_1 \otimes \Sigma_2, \mu_1 \otimes \mu_2)$ be their (restricted) product measure space.*

Let $f : X_1 \times X_2 \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$ be $\Sigma_1 \otimes \Sigma_2$-measurable and integrable with respect to $\mu_1 \otimes \mu_2$. Then f_{x_1} is Σ_2-measurable for every $x_1 \in X_1$ and integrable with respect to μ_2 for μ_1-a.e. $x_1 \in X_1$. Also, f_{x_2} is Σ_1-measurable for every $x_2 \in X_2$ and integrable with respect to μ_1 for μ_2-a.e. $x_2 \in X_2$. The a.e. defined functions

$$x_1 \mapsto \int_{X_2} f_{x_1}\, d\mu_2, \qquad x_2 \mapsto \int_{X_1} f_{x_2}\, d\mu_1$$

are integrable with respect to μ_1 and, respectively, integrable with respect to μ_2 and

$$\int_{X_1\times X_2} f\, d(\mu_1\otimes\mu_2) = \int_{X_1} \Big(\int_{X_2} f_{x_1}\, d\mu_2\Big)\, d\mu_1(x_1) = \int_{X_2} \Big(\int_{X_1} f_{x_2}\, d\mu_1\Big)\, d\mu_2(x_2).$$

Proof: Again, the measurability of the sections is an immediate application of Theorem 8.2 and does not need the assumption about σ-finiteness. Otherwise, the proof results from Theorem 8.11 in exactly the same way in which the proof of Theorem 8.10 results from Theorem 8.9.

8.4 Surface measure on S^{n-1}.

For every $x = (x_1, \dots, x_n) \in \mathbf{R}^n_* = \mathbf{R}^n \setminus \{0\}$ we write

$$r = |x| = \sqrt{x_1^2 + \cdots + x_n^2} \in \mathbf{R}^+ = (0, +\infty), \qquad y = \frac{x}{|x|} \in S^{n-1},$$

where $S^{n-1} = \{y \in \mathbf{R}^n \mid |y| = 1\}$ is the **unit spere of** $\mathbf{R}^n$.

The mapping $\Phi : \mathbf{R}^n_* \to \mathbf{R}^+ \times S^{n-1}$ defined by

$$\Phi(x) = (r, y) = \Big(|x|, \frac{x}{|x|}\Big)$$

is one-to-one and onto and its inverse $\Phi^{-1} : \mathbf{R}^+ \times S^{n-1} \to \mathbf{R}^n_*$ is given by

$$\Phi^{-1}(r, y) = x = ry.$$

The numbers $r = |x|$ and $y = \frac{x}{|x|}$ are called **the polar coordinates of** x and the mappings Φ and Φ^{-1} determine an *identification* of $\mathbf{R}^n_*$ with the cartesian product $\mathbf{R}^+ \times S^{n-1}$, where every point $x \neq 0$ is identified with the pair (r, y) of its polar coordinates.

As usual, we consider S^{n-1} as a metric subspace of $\mathbf{R}^n$. This means that the distance between points of S^{n-1} is their euclidean distance considered as points of the larger space $\mathbf{R}^n$. Namely

$$|y - y'| = \sqrt{(y_1 - y_1')^2 + \cdots + (y_n - y_n')^2},$$

for every $y = (y_1, \dots, y_n), y' = (y_1', \dots, y_n') \in S^{n-1}$. No two points of S^{n-1} have distance greater that 2 and, if two points have distance 2, then they are opposite or, equivalently, anti-diametric. The open ball in S^{n-1} with center $y \in S^{n-1}$ and radius $r > 0$ is the *spherical cap* $S(y; r) = \{y' \in S^{n-1} \mid |y' - y| < r\}$, which is the intersection of the euclidean ball $B(y; r) = \{x \in \mathbf{R}^n \mid |x - y| < r\}$ with S^{n-1}. In fact, the intersection of an arbitrary euclidean open ball in $\mathbf{R}^n$ with S^{n-1} is, if non-empty, a spherical cap of 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 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 S^{n-1} of all open balls in $\mathbf{R}^n$ with rational centers and rational radii: if U is an arbitrary open subset of S^{n-1} and we take arbitrary $y \in U$, we can find r so that $B(y; r) \cap S^{n-1} \subseteq U$. Then, we can find an open ball $B(x'; r')$ with rational x' and rational r' so that $y \in B(x'; r') \subseteq B(y; r)$. Now, y belongs to the spherical cap $B(x'; r') \cap S^{n-1} \subseteq U$.

If we equip $\mathbf{R}^+ \times S^{n-1}$ with the *product topology* through the *product metric*

$$d\big((r, y), (r', y')\big) = \max\{|r - r'|, |y - y'|\},$$

then the mappings Φ and Φ^{-1} are both continuous. In fact, it is clear that the convergence $(r_k, y_k) \to (r, y)$ in the product metric of $\mathbf{R}^+ \times S^{n-1}$ is equivalent to the simultaneous $r_k \to r$ and $y_k \to y$. Therefore, if $x_k \to x$ in $\mathbf{R}^n_*$, then $r_k = |x_k| \to |x| = r$ and $y_k = \frac{x_k}{|x_k|} \to \frac{x}{|x|} = y$ and hence $\Phi(x_k) = (r_k, y_k) \to (r, y) = \Phi(x)$ in $\mathbf{R}^+ \times S^{n-1}$. Conversely, if $(r_k, y_k) \to (r, y)$ in $\mathbf{R}^+ \times S^{n-1}$, then $r_k \to r$ and $y_k \to y$ and hence $\Phi^{-1}(r_k, y_k) = r_k y_k \to ry = \Phi^{-1}(r, y)$ in $\mathbf{R}^n_*$.

We may observe that the *open balls* in the product topology of $\mathbf{R}^+ \times S^{n-1}$ are exactly all the cartesian products $(a, b) \times S(y; r)$ of open subintervals of $\mathbf{R}^+$ with spherical caps of S^{n-1}.

The next proposition contains information about the Borel structures of $\mathbf{R}^n_*$ and of $\mathbf{R}^+$, S^{n-1} and their product $\mathbf{R}^+ \times S^{n-1}$.

Proposition 8.6 *(i)* $\mathcal{B}_{\mathbf{R}^n_*} = \{E \in \mathcal{B}_{\mathbf{R}^n} \,|\, E \subseteq \mathbf{R}^n_*\}$.
(ii) $\mathcal{B}_{\mathbf{R}^+} = \{E \in \mathcal{B}_{\mathbf{R}} \,|\, E \subseteq \mathbf{R}^+\}$ *and* $\mathcal{B}_{\mathbf{R}^+}$ *is generated by the collection of all open subintervals of* $\mathbf{R}^+$ *and, also, by the collection of all open-closed subintervals of* $\mathbf{R}^+$*.*
(iii) $\mathcal{B}_{S^{n-1}} = \{E \in \mathcal{B}_{\mathbf{R}^n} \,|\, E \subseteq S^{n-1}\}$ *and* $\mathcal{B}_{S^{n-1}}$ *is generated by the collection of all spherical caps.*
(iv) $\mathcal{B}_{\mathbf{R}^+ \times S^{n-1}} = \mathcal{B}_{\mathbf{R}^+} \otimes \mathcal{B}_{S^{n-1}}$*.*
(v) $\Phi(E)$ *is a Borel set in* $\mathbf{R}^+ \times S^{n-1}$ *for every Borel set* E *in* $\mathbf{R}^n_*$ *and* $\Phi^{-1}(E)$ *is a Borel set in* $\mathbf{R}^n_*$ *for every Borel set* E *in* $\mathbf{R}^+ \times S^{n-1}$*.*
(vi) $M \cdot A = \{ry \,|\, r \in M, y \in A\}$ *is a Borel set in* $\mathbf{R}^n_*$ *for every Borel set* A *in* S^{n-1} *and every Borel set* M *in* $\mathbf{R}^+$*.*

Proof: The equalities of (i),(ii) and (iii) are simple consequences of Theorem 1.3 or, more directly, of Exercise 1.6.6. That $\mathcal{B}_{\mathbf{R}^+}$ is generated by the collection of all open or of all open-closed subintervals of $\mathbf{R}^+$ is due to the fact that every open subset of $\mathbf{R}^+$ is a countable union of such intervals. Also, that $\mathcal{B}_{S^{n-1}}$ is generated by the collection of all spherical caps is due to the fact that every open subset of S^{n-1} is a countable union of spherical caps.
(iv) Both $\mathcal{B}_{\mathbf{R}^+ \times S^{n-1}}$ and $\mathcal{B}_{\mathbf{R}^+} \otimes \mathcal{B}_{S^{n-1}}$ are σ-algebras of subsets of the space $\mathbf{R}^+ \times S^{n-1}$. The second is generated by the collection of all cartesian products of open subsets of $\mathbf{R}^+$ with open subsets of S^{n-1} and all these sets are open subsets of $\mathbf{R}^+ \times S^{n-1}$ and, hence, belong to the first σ-algebra. Therefore, the second σ-algebra is included in the first. Conversely, the first σ-algebra is generated by the collection of all open subsets of $\mathbf{R}^+ \times S^{n-1}$ and every such set is a *countable* union of open balls, i.e. of cartesian products of open subintervals of $\mathbf{R}^+$ with spherical caps of S^{n-1}. Thus, every open subset of $\mathbf{R}^+ \times S^{n-1}$ is contained in the second σ-algebra and, hence, the first σ-algebra is included in the second.
(v) Since Φ is continuous, it is $(\mathcal{B}_{\mathbf{R}^n_*}, \mathcal{B}_{\mathbf{R}^+ \times S^{n-1}})$-measurable and, thus, $\Phi^{-1}(E)$ is a Borel set in $\mathbf{R}^n_*$ for every Borel set E in $\mathbf{R}^+ \times S^{n-1}$. The other statement is, similarly, a consequence of the continuity of Φ^{-1}.
(vi) $M \times A$ is a Borel set (measurable interval) in $\mathbf{R}^+ \times S^{n-1}$. Since Φ is continuous, $M \cdot A = \Phi^{-1}(M \times A)$ is a Borel set in $\mathbf{R}^n_*$.

A set $\Gamma \subseteq \mathbf{R}^n_*$ is called a **positive cone** if $rx \in \Gamma$ for every $r \in \mathbf{R}^+$ 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 $B \subseteq \mathbf{R}^n_*$, then the set $\mathbf{R}^+ \cdot B = \{rb \,|\, r \in \mathbf{R}^+, b \in B\}$ is, obviously, a positive cone and it is called **the positive cone determined by** B. It is trivial to see that, if Γ is a positive cone and $A = \Gamma \cap S^{n-1}$, then Γ is the positive cone determined by A and, conversely, that, if $A \subseteq S^{n-1}$ and Γ is the positive cone determined by A, then $\Gamma \cap S^{n-1} = A$. This means that there is a one-to-one correspondence between the subsets of S^{n-1} and the positive cones of $\mathbf{R}^n$.

The next result expresses a simple characterization of open and of Borel subsets of S^{n-1} in terms of the corresponding positive cones.

Proposition 8.7 *Let $A \subseteq S^{n-1}$.*
(i) A is open in S^{n-1} if and only if the cone $\mathbf{R}^+ \cdot A$ is open in $\mathbf{R}^n$.
(ii) A is a Borel set in S^{n-1} if and only if $\mathbf{R}^+ \cdot A$ is a Borel set in $\mathbf{R}^n$.

Proof: (i) By the definition of the product topology, A is open in S^{n-1} if and only if $\mathbf{R}^+ \times A$ is open in $\mathbf{R}^+ \times S^{n-1}$. By the continuity of Φ and Φ^{-1}, this last one is true if and only if $\mathbf{R}^+ \cdot A = \Phi^{-1}(\mathbf{R}^+ \times A)$ is open in $\mathbf{R}^n_*$ if and only if $\mathbf{R}^+ \cdot A$ is open in $\mathbf{R}^n$.
(ii) If A is a Borel set in S^{n-1} then, as a measurable interval, $\mathbf{R}^+ \times A$ is a Borel set in $\mathbf{R}^+ \times S^{n-1}$. Conversely, if $\mathbf{R}^+ \times A$ is a Borel set in $\mathbf{R}^+ \times S^{n-1}$, then all its r-sections, and in particular A, are Borel sets in S^{n-1}. Therefore, A is a Borel set in S^{n-1} if and only if $\mathbf{R}^+ \times A$ is a Borel set in $\mathbf{R}^+ \times S^{n-1}$. Proposition 8.6 implies that this is true if and only if $\mathbf{R}^+ \cdot A = \Phi^{-1}(\mathbf{R}^+ \times A)$ is a Borel set in $\mathbf{R}^n_*$ if and only if $\mathbf{R}^+ \cdot A$ is a Borel set in $\mathbf{R}^n$.

Proposition 8.8 *If we define*

$$\sigma_{n-1}(A) = n \cdot m_n\big((0,1] \cdot A\big)$$

for every $A \in \mathcal{B}_{S^{n-1}}$, then σ_{n-1} is a measure on $(S^{n-1}, \mathcal{B}_{S^{n-1}})$.

Proof: By the last statement of Proposition 8.6, $(0,1] \cdot A$ is a Borel set in $\mathbf{R}^n_*$ and thus $\sigma_{n-1}(A)$ is well defined. We have $\sigma_{n-1}(\emptyset) = n \cdot m_n\big((0,1] \cdot \emptyset\big) = n \cdot m_n(\emptyset) = 0$. Moreover, if $A_1, A_2, \ldots \in \mathcal{B}_{S^{n-1}}$ are pairwise disjoint, then the sets $(0,1] \cdot A_1, (0,1] \cdot A_2, \ldots$ are also pairwise disjoint. Hence, $\sigma_{n-1}(\cup_{j=1}^{+\infty} A_j) = n \cdot m_n\big((0,1] \cdot \cup_{j=1}^{+\infty} A_j\big) = n \cdot m_n\big(\cup_{j=1}^{+\infty} ((0,1] \cdot A_j)\big) = \sum_{j=1}^{+\infty} n \cdot m_n\big((0,1] \cdot A_j\big) = \sum_{j=1}^{+\infty} \sigma_{n-1}(A_j)$.

Definition 8.5 *The measure σ_{n-1} on $(S^{n-1}, \mathcal{B}_{S^{n-1}})$, which is defined in Proposition 8.8, is called* ***the** $(n-1)$**-dimensional surface measure on** S^{n-1}.*

Lemma 8.1 *If we define*

$$\rho(N) = \int_N r^{n-1}\, dr$$

for every $N \in \mathcal{B}_{\mathbf{R}^+}$, then ρ is a measure on $(\mathbf{R}^+, \mathcal{B}_{\mathbf{R}^+})$.

Proof: A simple consequence of Theorem 7.13.

Lemma 8.2 *If we define*

$$\widetilde{m_n}(E) = m_n(\Phi^{-1}(E))$$

for every Borel set E in $\mathbf{R}^+ \times S^{n-1}$, then $\widetilde{m_n}$ is a measure on the measurable space $(\mathbf{R}^+ \times S^{n-1}, \mathcal{B}_{\mathbf{R}^+ \times S^{n-1}})$.

Proof: $\Phi^{-1}(E)$ is a Borel set in $\mathbf{R}^n_*$ for every Borel set E in $\mathbf{R}^+ \times S^{n-1}$ and, hence, $\widetilde{m_n}(E)$ is well defined. Clearly, $\widetilde{m_n}(\emptyset) = m_n(\Phi^{-1}(\emptyset)) = m_n(\emptyset) = 0$. If $E_1, E_2, \ldots$ are pairwise disjoint, then $\Phi^{-1}(E_1), \Phi^{-1}(E_2), \ldots$ are also pairwise disjoint and $\widetilde{m_n}(\cup_{j=1}^{+\infty} E_j) = m_n(\Phi^{-1}(\cup_{j=1}^{+\infty} E_j)) = m_n(\cup_{j=1}^{+\infty} \Phi^{-1}(E_j)) = \sum_{j=1}^{+\infty} m_n(\Phi^{-1}(E_j)) = \sum_{j=1}^{+\infty} \widetilde{m_n}(E_j)$.

Lemma 8.3 *The measures $\widetilde{m_n}$ and $\rho \otimes \sigma_{n-1}$ are identical on the measurable space $(\mathbf{R}^+ \times S^{n-1}, \mathcal{B}_{\mathbf{R}^+ \times S^{n-1}}) = (\mathbf{R}^+ \times S^{n-1}, \mathcal{B}_{\mathbf{R}^+} \otimes \mathcal{B}_{S^{n-1}})$.*

Proof: The equality $\mathcal{B}_{\mathbf{R}^+ \times S^{n-1}} = \mathcal{B}_{\mathbf{R}^+} \otimes \mathcal{B}_{S^{n-1}}$ is in Proposition 8.6.

If A is a Borel set in S^{n-1}, then the sets $(0, b] \cdot A$ and $(0, 1] \cdot A$ are both Borel sets in $\mathbf{R}^n$ and the first is a dilate of the second by the factor $b > 0$. By Theorem 4.7, $m_n((0, b] \cdot A) = b^n m_n((0, 1] \cdot A)$ for every $b > 0$. By a simple subtraction we find that $m_n((a, b] \cdot A) = (b^n - a^n) m_n((0, 1] \cdot A)$ for every a, b with $0 \leq a < b < +\infty$.

Therefore, if A is a Borel set in S^{n-1}, then

$$\begin{aligned}
\widetilde{m_n}((a, b] \times A) &= m_n(\Phi^{-1}((a, b] \times A)) = m_n((a, b] \cdot A) \\
&= (b^n - a^n) m_n((0, 1] \cdot A) = \frac{b^n - a^n}{n} \sigma_{n-1}(A) \\
&= \int_{(a,b]} r^{n-1}\, dr \; \sigma_{n-1}(A) = \rho((a, b]) \; \sigma_{n-1}(A) \\
&= (\rho \otimes \sigma_{n-1}) \; ((a, b] \times A).
\end{aligned}$$

If we define

$$\mu(N) = \widetilde{m_n}(N \times A), \qquad \nu(N) = (\rho \otimes \sigma_{n-1})(N \times A)$$

for every Borel set N in $\mathbf{R}^+$, it is easy to see that both μ and ν are Borel measures on $\mathbf{R}^+$ and, by what we just proved, they satisfy $\mu((a, b]) = \nu((a, b])$ for every interval in $\mathbf{R}^+$. This, obviously, extends to all finite unions of pairwise disjoint open-closed intervals. Theorem 2.4 implies, now, that the two measures are equal on the σ-algebra generated by the collection of all these sets, which, by Proposition 8.6, is $\mathcal{B}_{\mathbf{R}^+}$. Therefore,

$$\widetilde{m_n}(N \times A) = (\rho \otimes \sigma_{n-1})(N \times A)$$

for every Borel set N in $\mathbf{R}^+$ and every Borel set A in S^{n-1}.

Theorem 8.4 implies now the equality of the two measures, because both measures ρ and σ_{n-1} are σ-finite.

If $E \subseteq \mathbf{R}^n_*$, we consider the set $\Phi(E) \subseteq \mathbf{R}^+ \times S^{n-1}$. We also consider the r-sections $\Phi(E)_r = \{y \in S^{n-1} \,|\, (r,y) \in \Phi(E)\} = \{y \in S^{n-1} \,|\, ry \in E\}$ and the y-sections $\Phi(E)_y = \{r \in \mathbf{R}^+ \,|\, (r,y) \in \Phi(E)\} = \{r \in \mathbf{R}^+ \,|\, ry \in E\}$ of $\Phi(E)$. We extend the notation as follows.

Definition 8.6 *If $E \subseteq \mathbf{R}^n$, we define, for every $r \in \mathbf{R}^+$ and every $y \in S^{n-1}$,*

$$E_r = \{y \in S^{n-1} \,|\, ry \in E\}, \qquad E_y = \{r \in \mathbf{R}^+ \,|\, ry \in E\}$$

*and call them **the r-sections** and **the y-sections** of E, respectively.*

Observe that E may contain 0, but this plays no role. 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 8.11 *Let E be any Borel set in $\mathbf{R}^n$. Then E_r is a Borel set in S^{n-1} for every $r \in \mathbf{R}^+$ and E_y is a Borel set in $\mathbf{R}^+$ for every $y \in S^{n-1}$ and the functions*

$$r \mapsto \sigma_{n-1}(E_r), \qquad y \mapsto \int_{E_y} r^{n-1}\, dr$$

are $\mathcal{B}_{\mathbf{R}^+}$-measurable and, respectively, $\mathcal{B}_{S^{n-1}}$-measurable. Also,

$$m_n(E) = \int_0^{+\infty} \sigma_{n-1}(E_r) r^{n-1}\, dr = \int_{S^{n-1}} \Big(\int_{E_y} r^{n-1}\, dr \Big)\, d\sigma_{n-1}(y).$$

Proof: The set $E \setminus \{0\}$ is a Borel set in $\mathbf{R}^n_*$, while $E_r = \Phi(E \setminus \{0\})_r$ and $E_y = \Phi(E \setminus \{0\})_y$.

Lemmas 8.2 and 8.3 imply that $m_n(E) = m_n(E \setminus \{0\}) = \widetilde{m_n}\big(\Phi(E \setminus \{0\})\big) = (\rho \otimes \sigma_{n-1})\big(\Phi(E \setminus \{0\})\big)$. Proposition 8.6 says that $\Phi(E \setminus \{0\})$ is a Borel set in $\mathbf{R}^+ \times S^{n-1}$ and the rest is a consequence of Theorem 8.8.

The next result gives a simple description of the completion of the measure space $(S^{n-1}, \mathcal{B}_{S^{n-1}}, \sigma_{n-1})$ in terms of positive cones.

Definition 8.7 *We denote $(S^{n-1}, \mathcal{S}_{n-1}, \sigma_{n-1})$ the completion of the measure space $(S^{n-1}, \mathcal{B}_{S^{n-1}}, \sigma_{n-1})$.*

Proposition 8.9 *If $A \subseteq S^{n-1}$, then*
(i) $A \in \mathcal{S}_{n-1}$ if and only if $\mathbf{R}^+ \cdot A \in \mathcal{L}_n$ if and only if $(0,1] \cdot A \in \mathcal{L}_n$,
(ii) $\sigma_{n-1}(A) = n \cdot m_n\big((0,1] \cdot A\big)$ for every $A \in \mathcal{S}_{n-1}$.

Proof: (i) If $A \in \mathcal{S}_{n-1}$, there exist $A_1, A_2 \in \mathcal{B}_{S^{n-1}}$ with $\sigma_{n-1}(A_2) = 0$ so that $A_1 \subseteq A$ and $A \setminus A_1 \subseteq A_2$. Proposition 8.7 implies that the positive cones $\mathbf{R}^+ \cdot A_1$ and $\mathbf{R}^+ \cdot A_2$ are Borel sets in $\mathbf{R}^n$ with $\mathbf{R}^+ \cdot A_1 \subseteq \mathbf{R}^+ \cdot A$ and $(\mathbf{R}^+ \cdot A) \setminus (\mathbf{R}^+ \cdot A_1) \subseteq \mathbf{R}^+ \cdot A_2$. Lemmas 8.2 and 8.3 or Theorem 8.13 imply $m_n(\mathbf{R}^+ \cdot A_2) = \rho(\mathbf{R}^+)\sigma_{n-1}(A_2) = 0$. Hence, $\mathbf{R}^+ \cdot A \in \mathcal{L}_n$.

Conversely, let $\mathbf{R}^+ \cdot A \in \mathcal{L}_n$. Then, there are Borel sets $B_1, B_2 \subseteq \mathbf{R}^n$ with $m_n(B_2) = 0$, so that $B_1 \subseteq \mathbf{R}^+ \cdot A$ and $(\mathbf{R}^+ \cdot A) \setminus B_1 \subseteq B_2$. For every

$r \in \mathbf{R}^+$ we have that $(B_1)_r \subseteq A$ and $A \setminus (B_1)_r \subseteq (B_2)_r$. From Theorem 8.13, $\int_0^{+\infty} \sigma_{n-1}((B_2)_r) r^{n-1}\, dr = m_n(B_2) = 0$, implying that $\sigma_{n-1}((B_2)_r) = 0$ 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 S^{n-1}, we conclude that $A \in \mathcal{S}_{n-1}$.

If $\mathbf{R}^+ \cdot A \in \mathcal{L}_n$, then $(0,1] \cdot A = (\mathbf{R}^+ \cdot A) \cap B_n \in \mathcal{L}_n$. Conversely, if $(0,1] \cdot A \in \mathcal{L}_n$, then $\mathbf{R}^+ \cdot A = \bigcup_{k=1}^{+\infty} k \cdot \big((0,1] \cdot A\big) \in \mathcal{L}_n$.

(ii) We take $A \in \mathcal{S}_{n-1}$ and $A_1, A_2 \in \mathcal{B}_{S^{n-1}}$ with $\sigma_{n-1}(A_2) = 0$ so that $A_1 \subseteq A$ and $A \setminus A_1 \subseteq A_2$. Then the sets $(0,1] \cdot A_1$ and $(0,1] \cdot A_2$ are Borel sets in $\mathbf{R}^n$ with $(0,1] \cdot A_1 \subseteq (0,1] \cdot A$ and $(0,1] \cdot A \setminus (0,1] \cdot A_1 \subseteq (0,1] \cdot A_2$. Since $m_n((0,1] \cdot A_2) = \frac{1}{n}\, \sigma_{n-1}(A_2) = 0$, we conclude that $\sigma_{n-1}(A) = \sigma_{n-1}(A_1) = n \cdot m_n((0,1] \cdot A_1) = n \cdot m_n((0,1] \cdot A)$.

The next result is an extension of Theorem 8.13 to Lebesgue sets.

Theorem 8.12 *Let $E \in \mathcal{L}_n$. Then $E_r \in \mathcal{S}_{n-1}$ for m_1-a.e. $r \in \mathbf{R}^+$ and $E_y \in \mathcal{L}_1$ for σ_{n-1}-a.e. $y \in S^{n-1}$ and the a.e. defined functions*

$$r \mapsto \sigma_{n-1}(E_r), \qquad y \mapsto \int_{E_y} r^{n-1}\, dr$$

are $\mathcal{L}_1$-measurable and, respectively, $\mathcal{S}_{n-1}$-measurable. Also,

$$m_n(E) = \int_0^{+\infty} \sigma_{n-1}(E_r) r^{n-1}\, dr = \int_{S^{n-1}} \Big(\int_{E_y} r^{n-1}\, dr \Big)\, d\sigma_{n-1}(y).$$

Proof: We consider Borel sets B_1, B_2 in $\mathbf{R}^n$ with $m_n(B_2) = 0$, so that $B_1 \subseteq E$ and $E \setminus B_1 \subseteq B_2$.

Theorem 8.13 implies that, for every $r \in \mathbf{R}^+$, $(B_1)_r$ and $(B_2)_r$ are Borel sets in S^{n-1} with $(B_1)_r \subseteq E_r$ and $E_r \setminus (B_1)_r \subseteq (B_2)_r$. From Theorem 8.13 again, $\int_0^{+\infty} \sigma_{n-1}((B_2)_r) r^{n-1}\, dr = m_n(B_2) = 0$ and we get that $\sigma_{n-1}((B_2)_r) = 0$ for m_1-a.e. $r \in \mathbf{R}^+$. Therefore, $E_r \in \mathcal{S}_{n-1}$ and $\sigma_{n-1}(E_r) = \sigma_{n-1}((B_1)_r)$ for m_1-a.e. $r \in \mathbf{R}^+$.

Similarly, for every $y \in S^{n-1}$, $(B_1)_y$ and $(B_2)_y$ are Borel sets in $\mathbf{R}^+$ with $(B_1)_y \subseteq E_y$ and $E_y \setminus (B_1)_y \subseteq (B_2)_y$. From $\int_{S^{n-1}} \big(\int_{(B_2)_y} r^{n-1}\, dr \big)\, d\sigma_{n-1}(y) = m_n(B_2) = 0$, we get that $\int_{(B_2)_y} r^{n-1}\, dr = 0$ for σ_{n-1}-a.e. $y \in S^{n-1}$. This implies $m_1((B_2)_y) = 0$ for σ_{n-1}-a.e. $y \in S^{n-1}$ and, hence, $E_y \in \mathcal{L}_1$ and $\int_{E_y} r^{n-1}\, dr = \int_{(B_1)_y} r^{n-1}\, dr$ for σ_{n-1}-a.e. $y \in S^{n-1}$. Theorem 8.13 implies $m_n(E) = m_n(B_1) = \int_0^{+\infty} \sigma_{n-1}((B_1)_r) r^{n-1}\, dr = \int_0^{+\infty} \sigma_{n-1}(E_r) r^{n-1}\, dr$ and, also, $= \int_{S^{n-1}} \big(\int_{(B_1)_y} r^{n-1}\, dr \big)\, d\sigma_{n-1}(y) = \int_{S^{n-1}} \big(\int_{E_y} r^{n-1}\, dr \big)\, d\sigma_{n-1}(y)$.

The rest of this section consists of a series of theorems which describe the so-called method of **integration by polar coordinates**.

Definition 8.8 *Let $f : \mathbf{R}^n \to Y$. For every $r \in \mathbf{R}^+$ and every $y \in S^{n-1}$ we define the functions $f_r : S^{n-1} \to Y$ and $f_y : \mathbf{R}^+ \to Y$ by the formulas*

$$f_r(y) = f_y(r) = f(ry).$$

*f_r is called **the r-section of** f and f_y is called **the y-section of** f.*

The next two theorems cover integration by polar coordinates for Borel measurable functions.

Theorem 8.13 *Let $f : \mathbf{R}^n \to [0,+\infty]$ be $\mathcal{B}_{\mathbf{R}^n}$-measurable. Then every f_r is $\mathcal{B}_{S^{n-1}}$-measurable and every f_y is $\mathcal{B}_{\mathbf{R}^+}$-measurable. The functions*

$$r \mapsto \int_{S^{n-1}} f(ry)\, d\sigma_{n-1}, \qquad y \mapsto \int_0^{+\infty} f(ry) r^{n-1}\, dr$$

are $\mathcal{B}_{\mathbf{R}^+}$-measurable and, respectively, $\mathcal{B}_{S^{n-1}}$-measurable. Moreover

$$\begin{aligned} \int_{\mathbf{R}^n} f(x)\, dm_n(x) &= \int_0^{+\infty} \Big(\int_{S^{n-1}} f(ry)\, d\sigma_{n-1}(y) \Big) r^{n-1}\, dr \\ &= \int_{S^{n-1}} \Big(\int_0^{+\infty} f(ry) r^{n-1}\, dr \Big)\, d\sigma_{n-1}(y). \end{aligned}$$

Proof: The results of this theorem and of Theorem 8.13 are the same in case $f = \chi_E$. Using the linearity of the integrals, we prove the theorem in the case of a simple function $\phi : \mathbf{R}^n \to [0,+\infty]$. Finally, applying the Monotone Convergence Theorem to an increasing sequence of simple functions, we prove the theorem in the general case.

Theorem 8.14 *Let $f : \mathbf{R}^n \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$ be $\mathcal{B}_{\mathbf{R}^n}$-measurable and integrable with respect to m_n. Then every f_r is $\mathcal{B}_{S^{n-1}}$-measurable and, for m_1-a.e. $r \in \mathbf{R}^+$, f_r is integrable with respect to σ_{n-1}. Also, every f_y is $\mathcal{B}_{\mathbf{R}^+}$-measurable, and for σ_{n-1}-a.e. $y \in S^{n-1}$, f_y is integrable with respect to m_1. The a.e. defined functions*

$$r \mapsto \int_{S^{n-1}} f(ry)\, d\sigma_{n-1}(y), \qquad y \mapsto \int_0^{+\infty} f(ry) r^{n-1}\, dr$$

are integrable with respect to m_1 and, respectively, with respect to σ_{n-1}. Also

$$\begin{aligned} \int_{\mathbf{R}^n} f(x)\, dm_n(x) &= \int_0^{+\infty} \Big(\int_{S^{n-1}} f(ry)\, d\sigma_{n-1}(y) \Big) r^{n-1}\, dr \\ &= \int_{S^{n-1}} \Big(\int_0^{+\infty} f(ry) r^{n-1}\, dr \Big)\, d\sigma_{n-1}(y). \end{aligned}$$

Proof: We use Theorem 8.15 to pass to the case of functions $f : \mathbf{R}^n \to \overline{\mathbf{R}}$, by writing them as $f = f^+ - f^-$. We next treat the case of $f : \mathbf{R}^n \to \overline{\mathbf{C}}$, by writing $f = \Re(f) + i\Im(f)$, after we exclude, in the usual manner, the set $f^{-1}(\{\infty\})$.

The next two theorems treat integration by polar coordinates in the case of Lebesgue measurable functions. They are proved, one after the other, using Theorem 8.14 exactly as Theorems 8.15 and 8.16 were proved with the use of Theorem 8.13.

Theorem 8.15 *Let $f : \mathbf{R}^n \to [0, +\infty]$ be $\mathcal{L}_n$-measurable. Then, for m_1-a.e. $r \in \mathbf{R}^+$, the function f_r is $\mathcal{S}_{n-1}$-measurable and, for σ_{n-1}-a.e. $y \in S^{n-1}$, the function f_y is $\mathcal{L}_1$-measurable. The a.e. defined functions*

$$r \mapsto \int_{S^{n-1}} f(ry)\, d\sigma_{n-1}(y), \qquad y \mapsto \int_0^{+\infty} f(ry) r^{n-1}\, dr$$

are $\mathcal{L}_1$-measurable and, respectively, $\mathcal{S}_{n-1}$-measurable. Moreover

$$\begin{aligned} \int_{\mathbf{R}^n} f(x)\, dm_n(x) &= \int_0^{+\infty} \Big(\int_{S^{n-1}} f(ry)\, d\sigma_{n-1}(y) \Big) r^{n-1}\, dr \\ &= \int_{S^{n-1}} \Big(\int_0^{+\infty} f(ry) r^{n-1}\, dr \Big)\, d\sigma_{n-1}(y). \end{aligned}$$

Theorem 8.16 *Let $f : \mathbf{R}^n \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$ be $\mathcal{L}_n$-measurable and integrable with respect to m_n. Then, for m_1-a.e. $r \in \mathbf{R}^+$, f_r is integrable with respect to σ_{n-1} and, for σ_{n-1}-a.e. $y \in S^{n-1}$, f_y is integrable with respect to m_1. The a.e. defined functions*

$$r \mapsto \int_{S^{n-1}} f(ry)\, d\sigma_{n-1}(y), \qquad y \mapsto \int_0^{+\infty} f(ry) r^{n-1}\, dr$$

are integrable with respect to m_1 and, respectively, with respect to σ_{n-1}. Also

$$\begin{aligned} \int_{\mathbf{R}^n} f(x)\, dm_n(x) &= \int_0^{+\infty} \Big(\int_{S^{n-1}} f(ry)\, d\sigma_{n-1}(y) \Big) r^{n-1}\, dr \\ &= \int_{S^{n-1}} \Big(\int_0^{+\infty} f(ry) r^{n-1}\, dr \Big)\, d\sigma_{n-1}(y). \end{aligned}$$

Definition 8.9 *A set $E \subseteq \mathbf{R}^n$ is called* ***radial*** *if $x \in E$ implies that $x' \in E$ for all x' with $|x'| = |x|$.*

A function $f : \mathbf{R}^n \to Y$ 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 is radial, we may define the **radial projection of** E as

$$\widetilde{E} = \{ r \in \mathbf{R}^+ \mid x \in E \text{ when } |x| = r \}.$$

Also, if f is radial, we may define the **radial projection of** f as the function $\widetilde{f} : \mathbf{R}^+ \to Y$ by

$$\widetilde{f}(r) = f(x)$$

for every $x \in \mathbf{R}^n$ 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 8.10 *(i) The radial set $E \subseteq \mathbf{R}^n$ is in $\mathcal{B}_{\mathbf{R}^n}$ or in $\mathcal{L}_n$ if and only if its radial projection is in $\mathcal{B}_{\mathbf{R}^+}$ or, respectively, in $\mathcal{L}_1$. In any case we have*

$$m_n(E) = \sigma_{n-1}(S^{n-1}) \int_{\widetilde{E}} r^{n-1}\, dr.$$

(ii) If (Y, Σ') is a measurable space, then the radial function $f : \mathbf{R}^n \to Y$ is $(\mathcal{B}_{\mathbf{R}^n}, \Sigma')$-measurable or $(\mathcal{L}_n, \Sigma')$-measurable if and only if its radial projection is $(\mathcal{B}_{\mathbf{R}^+}, \Sigma')$-measurable or, respectively, $(\mathcal{L}_1, \Sigma')$-measurable.

If $f : \mathbf{R}^n \to [0, +\infty]$ is Borel or Lebesgue measurable or if $f : \mathbf{R}^n \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$ is Borel or Lebesgue measurable and integrable with respect to m_n, then

$$\int_{\mathbf{R}^n} f(x)\, dm_n(x) = \sigma_{n-1}(S^{n-1}) \int_0^{+\infty} \widetilde{f}(r) r^{n-1}\, dr.$$

Proof: (i) If $E \in \mathcal{B}_{\mathbf{R}^n}$ or $E \in \mathcal{L}_n$ is radial, then, for every $y \in S^{n-1}$, we have $E_y = \widetilde{E}$ and, hence, the result is a consequence of Theorems 8.13 and 8.14.

For the converse we may argue as follows: we consider the collection of all subsets of $\mathbf{R}^+$ which are radial projections of radial Borel sets in $\mathbf{R}^n$, we then prove easily that this collection is a σ-algebra which contains all open subsets of $\mathbf{R}^+$ and we conclude that it contains all Borel sets in $\mathbf{R}^+$.

Now, if E is radial and $\widetilde{E} \in \mathcal{L}_1$, we take Borel sets M_1, M_2 in $\mathbf{R}^+$ 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 $E_1, E_2 \subseteq \mathbf{R}^n$ so that $\widetilde{E_1} = M_1$ and $\widetilde{E_2} = M_2$, which are Borel sets, by the result of the previous paragraph. Then we have $E_1 \subseteq E$ and $E \setminus E_1 \subseteq E_2$. Since $0 = m_n(E_2) = \int_{S^{n-1}} \left(\int_{(E_2)_y} r^{n-1}\, dr \right) d\sigma_{n-1} = \sigma_{n-1}(S^{n-1}) \int_{\widetilde{E_2}} r^{n-1}\, dr$, we have $\int_{\widetilde{E_2}} r^{n-1}\, dr$ and, hence, $m_1(\widetilde{E_2}) = 0$. This implies that $E \in \mathcal{L}_1$.
(ii) The statement about measurability is a trivial consequence of the definition of measurability and the result of part (i). The integral formulas are consequences of Theorems 8.15 up to 8.18.

8.5 Exercises.

1. If B is open in $\mathbf{R}^n_*$, prove that $\mathbf{R}^+ \cdot B$ is open in $\mathbf{R}^n_*$.

   If B is a Borel set in $\mathbf{R}^n_*$, prove that $\mathbf{R}^+ \cdot B$ is a Borel set in $\mathbf{R}^n_*$.

2. Consider the measure spaces $(\mathbf{R}, \mathcal{B}_{\mathbf{R}}, m_1)$ and $(\mathbf{R}, \mathcal{P}(\mathbf{R}), \sharp)$, where $\sharp$ is the counting measure. If $E = \{(x_1, x_2) \,|\, 0 \leq x_1 = x_2 \leq 1\}$, prove that all numbers $(m_1 \otimes \sharp)(E)$, $\int_{\mathbf{R}} \sharp(E_{x_1})\, dm_1(x_1)$ and $\int_{\mathbf{R}} m_1(E_{x_2})\, d\sharp(x_2)$ are different.

3. 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 $\sum_{n=1}^{+\infty} \left(\sum_{m=1}^{+\infty} a_{m,n} \right) \neq \sum_{m=1}^{+\infty} \left(\sum_{n=1}^{+\infty} a_{m,n} \right)$. Explain, through the Theorem of Fubini.

4. *The graph and the area under the graph of a function.*

Suppose that (X, Σ, μ) is a measure space and $f : X \to [0, +\infty]$ is Σ-measurable. If

$$A_f = \{(x, y) \in X \times \mathbf{R} \,|\, 0 \le y < f(x)\}$$

and

$$G_f = \{(x, y) \in X \times \mathbf{R} \,|\, y = f(x)\},$$

prove that both A_f and G_f are $\Sigma \otimes \mathcal{B}_{\mathbf{R}}$-measurable. If, moreover, μ is σ-finite, prove that

$$(\mu \otimes m_1)(A_f) = \int_X f \, d\mu, \qquad (\mu \otimes m_1)(G_f) = 0.$$

5. *The distribution function.*

Suppose that (X, Σ, μ) is a σ-finite measure space and $f : X \to [0, +\infty]$ is Σ-measurable. Calculating the measure $\mu \otimes \mu_G$ of the set $A_f = \{(x, y) \in X \times \mathbf{R} \,|\, 0 \le y < f(x)\}$, prove Proposition 7.14.

6. Consider measure spaces (X_1, Σ_1, μ_1) and (X_2, Σ_2, μ_2), a Σ_1-measurable $f_1 : X_1 \to \mathbf{C}$ and a Σ_2-measurable $f_2 : X_2 \to \mathbf{C}$. Consider the function $f : X_1 \times X_2 \to \mathbf{C}$ defined by $f(x_1, x_2) = f_1(x_1) f_2(x_2)$.

Prove that f is $\Sigma_1 \otimes \Sigma_2$-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

$$\int_{X_1 \times X_2} f \, d(\mu_1 \otimes \mu_2) = \int_{X_1} f_1 \, d\mu_1 \int_{X_2} f_2 \, d\mu_2.$$

7. *The volume of the unit ball in $\mathbf{R}^n$ and the surface measure of S^{n-1}.*

(i) If $v_n = m_n(B_n)$ is the Lebesgue measure of the unit ball of $\mathbf{R}^n$, prove that

$$v_n = 2 v_{n-1} \int_0^1 (1 - t^2)^{\frac{n-1}{2}} \, dt.$$

(ii) Set $J_n = \int_0^1 (1 - t^2)^{\frac{n-1}{2}} \, dt$ for $n \ge 0$ and prove the inductive formula $J_n = \frac{n-1}{n} J_{n-2}$, $n \ge 2$.
(iii) Prove that the gamma-function (defined in Exercise 7.9.38) satisfies the inductive formula

$$\Gamma(z + 1) = z \Gamma(z)$$

for every $z \in H_+$, and that $\Gamma(1) = 1$, $\Gamma(\frac{1}{2}) = \sqrt{\pi}$.
(iii) Prove that

$$v_n = \frac{\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2} + 1)}\,, \qquad \sigma_{n-1}(S^{n-1}) = \frac{2\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2})}\,.$$

8. *The integral of Gauss and the measures of B_n and of S^{n-1}.*

Define
$$I_n = \int_{\mathbf{R}^n} e^{-\frac{|x|^2}{2}}\, dx.$$
(i) Prove that $I_n = I_1^n$ for every $n \in \mathbf{N}$.
(ii) Use integration by polar coordinates to prove that $I_2 = 2\pi$ and, hence, that
$$\int_{\mathbf{R}^n} e^{-\frac{|x|^2}{2}}\, dx = (2\pi)^{\frac{n}{2}}.$$
(iii) Use integration by polar coordinates to prove that
$$(2\pi)^{\frac{n}{2}} = \sigma_{n-1}(S^{n-1}) \int_0^{+\infty} e^{-\frac{r^2}{2}} r^{n-1}\, dr$$
and, hence,
$$\sigma_{n-1}(S^{n-1}) = \frac{2\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2})}\,, \qquad v_n = m_n(B_n) = \frac{\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2}+1)}\,.$$

9. 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
$$\int_0^{\to+\infty} \frac{\sin x}{x}\, dx = \frac{\pi}{2}\,.$$

10. *Convolution.*

Let $f, g : \mathbf{R}^n \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$ be $\mathcal{L}_n$-measurable.

(i) Prove that the function $H : \mathbf{R}^n \times \mathbf{R}^n \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$, which is defined by the formula
$$H(x, y) = f(x-y)g(y),$$
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
$$\int_{\mathbf{R}^{2n}} |H|\, dm_{2n} \le \int_{\mathbf{R}^n} |f|\, dm_n \int_{\mathbf{R}^n} |g|\, dm_n.$$
(iii) Prove that for m_n-a.e. $x \in \mathbf{R}^n$ the function $f(x-\cdot)g(\cdot)$ is integrable with respect to m_n.

The a.e. defined function $f * g : \mathbf{R}^n \to \mathbf{R}$ or $\mathbf{C}$ by the formula
$$(f * g)(x) = \int_{\mathbf{R}^n} f(x-y)g(y)\, dm_n(y)$$
is called **the convolution of f and g.**

(iv) Prove that $f * g$ is integrable with respect to m_n, that

$$\int_{\mathbf{R}^n} (f * g)\, dm_n = \int_{\mathbf{R}^n} f\, dm_n \int_{\mathbf{R}^n} g\, dm_n$$

and

$$\int_{\mathbf{R}^n} |f * g|\, dm_n \le \int_{\mathbf{R}^n} |f|\, dm_n \int_{\mathbf{R}^n} |g|\, dm_n.$$

(v) Prove that, for every f, g, h, f_1, f_2 which are Lebesgue integrable, we have m_n-a.e. on $\mathbf{R}^n$ that $f*g = g*f$, $(f*g)*h = f*(g*h)$, $(\lambda f)*g = \lambda(f*g)$ and $(f_1 + f_2) * g = f_1 * g + f_2 * g$.

11. *The Fourier transforms of Lebesgue integrable functions.*

 Let $f : \mathbf{R}^n \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$ be Lebesgue integrable over $\mathbf{R}^n$. We define the function $\widehat{f} : \mathbf{R}^n \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$ by the formula

$$\widehat{f}(\xi) = \int_{\mathbf{R}^n} e^{-2\pi i x\cdot\xi} f(x)\, dm_n(x),$$

 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) Prove that $\widehat{f * g} = \widehat{f}\,\widehat{g}$, where $f*g$ is the convolution defined in Exercise 8.5.10.
 (iii) If $g(x) = f(x - a)$ for a.e. $x \in \mathbf{R}^n$, prove that $\widehat{g}(\xi) = e^{-2\pi i a\cdot\xi}\widehat{f}(\xi)$ for all $\xi \in \mathbf{R}^n$.
 (iv) If $g(x) = e^{-2\pi i a\cdot x} f(x)$ for a.e. $x \in \mathbf{R}^n$, prove that $\widehat{g}(\xi) = \widehat{f}(\xi + a)$ for all $\xi \in \mathbf{R}^n$.
 (v) If $g(x) = \overline{f(x)}$ for a.e. $x \in \mathbf{R}^n$, prove that $\widehat{g}(\xi) = \overline{\widehat{f}(-\xi)}$ for all $\xi \in \mathbf{R}^n$.
 (vi) If $T : \mathbf{R}^n \to \mathbf{R}^n$ is a linear transformation with $\det(T) \neq 0$ and $g(x) = f(Tx)$ for a.e. $x \in \mathbf{R}^n$, prove that $\widehat{g}(\xi) = \frac{1}{\det(T)}\,\widehat{f}\big((T^*)^{-1}(\xi)\big)$ for all $\xi \in \mathbf{R}^n$, where T^* is the adjoint of T.
 (vii) Prove that $\widehat{f}$ is continuous on $\mathbf{R}^n$.
 (viii) Prove that $|\widehat{f}(\xi)| \le \int_{\mathbf{R}^n} |f(x)|\, dm_n(x)$ for every $\xi \in \mathbf{R}^n$.

12. Let K be a Cantor-type set in $[0, 1]$ of the type considered in Exercise 4.6.16 with $m_1(K) > 0$. Prove that $\{(x, y) \in [0, 1] \times [0, 1] \,|\, x - y \in K\}$ is a compact subset of $\mathbf{R}^2$ with positive m_2-measure, which does not contain any measurable interval of positive m_2-measure.

13. *Uniqueness of Lebesgue measure.*

 Let μ and ν be two locally finite Borel measures on $\mathbf{R}^n$, which are translation invariant. Namely: $\mu(A+x) = \mu(A)$ and $\nu(A+x) = \nu(A)$ for every $x \in \mathbf{R}^n$ and every $A \in \mathcal{B}_{\mathbf{R}^n}$.

 Working with $\int_{\mathbf{R}^n\times\mathbf{R}^n} \chi_A(x)\chi_B(x + y)\, d(\mu \otimes \nu)(x, y)$, prove that either $\mu = \lambda\nu$ or $\nu = \lambda\mu$ for some $\lambda \in [0, +\infty)$.

Conclude that the only locally finite Borel measure on $\mathbf{R}^n$ which has value 1 at the unit cube $[0,1]^n$ is the Lebesgue measure m_n.

14. 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 $[0,1] \setminus E_y$. Prove that E is not Lebesgue measurable.

15. Let (X, Σ, μ) be a measure space and (Y, Σ') be a measurable space. Suppose that for every $x \in X$ there exists a measure ν_x on (Y, Σ') so that for every $B \in \Sigma'$ the function $x \mapsto \nu_x(B)$ is Σ-measurable.

    We define $\nu(B) = \int_X \nu_x(B)\, d\mu(x)$ for every $B \in \Sigma'$.
    (i) Prove that ν is a measure on (Y, Σ').
    (ii) If $g : Y \to [0, +\infty]$ is Σ'-measurable and if $f(x) = \int_Y g\, d\nu_x$ for every $x \in X$, prove that f is Σ-measurable and $\int_X f\, d\mu = \int_Y g\, d\nu$.

16. *Interchange of successive summations.*

    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

$$\sum_{i_1 \in I_1, i_2 \in I_2} c_{i_1,i_2} = \sum_{i_1 \in I_1} \Big(\sum_{i_2 \in I_2} c_{i_1,i_2} \Big) = \sum_{i_2 \in I_2} \Big(\sum_{i_1 \in I_1} c_{i_1,i_2} \Big).$$

17. Consider, for every $p \in (0, +\infty)$, the function $f : \mathbf{R}^n \to [0, +\infty]$, defined by $f(x) = \frac{1}{|x|^p}$.

    (i) Prove that f is not Lebesgue integrable over $\mathbf{R}^n$.
    (ii) Prove that f is integrable over the set $A_\delta = \{x \in \mathbf{R}^n \,|\, 0 < \delta \le |x|\}$ if and only if $p > 1$.
    (iii) Prove that f is integrable over the set $B_R = \{x \in \mathbf{R}^n \,|\, |x| \le R < +\infty\}$ if and only if $p < 1$.

18. Suppose that (Y, Σ) and (X_i, Σ_i) are measurable spaces for all $i \in I$ and that $g : X_{i_0} \to Y$ is (Σ_{i_0}, Σ)-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 $(\otimes_{i \in I} \Sigma_i, \Sigma)$-measurable.

19. *Integration by parts.*

    Consider the interval $\tilde{R} = (a, b] \times (a, b]$ and partition it into the two sets $\Delta_1 = \{(t, s) \in \tilde{R} \,|\, t \le s\}$ and $\Delta_2 = \{(t, s) \in \tilde{R} \,|\, s < t\}$. Writing $(\mu_G \otimes \mu_F)(\tilde{R}) = (\mu_G \otimes \mu_F)(\Delta_1) + (\mu_G \otimes \mu_F)(\Delta_2)$, prove Proposition 7.11.

Chapter 9

Convergence of functions

9.1 a.e. convergence and uniformly 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, briefly, recall their definitions and simple properties.

Suppose A is an arbitrary set and $f, f_n : A \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$ for every $n \in \mathbf{N}$. We say that (f_n) *converges 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: $|f_n(x) - f(x)| \le \epsilon$ for every $n \ge n_0$.

Suppose A is an arbitrary set and $f, f_n : A \to \mathbf{C}$ for every $n \in \mathbf{N}$. We say that (f_n) *converges to f uniformly on A* if for every $\epsilon > 0$ there is an $n_0 = n_0(\epsilon)$ so that: $|f_n(x) - f(x)| \le \epsilon$ for every $x \in A$ and every $n \ge n_0$ or, equivalently, $\sup_{x \in A} |f_n(x) - f(x)| \le \epsilon$ for every $n \ge n_0$. In other words, (f_n) converges 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 of (f_n) to f implies pointwise convergence on A. The converse is not true in general. As a counter-example, if $f_n = \chi_{(0, \frac{1}{n})}$ for every n, then (f_n) converges to $f = 0$ 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 and, hence, the same is true for the uniform limit.

Assume that $f, g, f_n, g_n : A \to \mathbf{C}$ for all n. If (f_n) converges to f and (g_n) converges to g pointwise on A, then $(f_n + g_n)$ converges to $f + g$ and $(f_n g_n)$ converges to fg pointwise on A. The same is true for uniform convergence, provided that in the case of the product we also assume that the two sequences are uniformly bounded: this means that there is an $M < +\infty$ so that $|f_n(x)|, |g_n(x)| \le M$ for every $x \in A$ and every $n \in \mathbf{N}$.

Another well-known fact is that, if $f_n : A \to \mathbf{C}$ for all n and (f_n) is Cauchy uniformly on A, then there is an $f : A \to \mathbf{C}$ so that (f_n) converges to f

uniformly on A. Indeed, suppose 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, the sequence $(f_n(x))$ is a Cauchy sequence of complex numbers and, hence, it converges to some complex number. If we define $f : A \to \mathbf{C}$ by $f(x) = \lim_{n\to+\infty} f_n(x)$ and 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 $n \geq n_0$. Hence, (f_n) converges to f uniformly on A.

It is almost straightforward to extend these two notions of convergence to measure spaces.

Suppose that (X, Σ, μ) is an arbitrary measure space.

We have already seen the notion of a.e. convergence. If $f, f_n : X \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$ for every n, we say that (f_n) converges to f **(pointwise) a.e. on** $A \in \Sigma$ if there is a set $B \in \Sigma$, $B \subseteq A$, so that $\mu(A \setminus B) = 0$ and (f_n) converges to f pointwise on B.

If $f, f_n : X \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$ for every n, we say that (f_n) converges to f **uniformly a.e. on** $A \in \Sigma$ if there is a set $B \in \Sigma$, $B \subseteq A$, so that $\mu(A \setminus B) = 0$, f and f_n are finite on B for all n and (f_n) converges to f uniformly on B.

It is clear that uniform convergence a.e. on A implies convergence a.e. on A. The converse is not true in general and the counter-example is the same as above.

If (f_n) converges to both f and f' a.e. on A, then $f = f'$ a.e. on A. Indeed, there are $B, B' \in \Sigma$ with $B, B' \subseteq A$ so that $\mu(A \setminus B) = \mu(A \setminus B') = 0$ and (f_n) converges to f pointwise on B and to f' pointwise on B'. Therefore, (f_n) converges to both f and f' pointwise on $B \cap B'$ and, hence, $f = f'$ on $B \cap B'$. Since $\mu(A \setminus (B \cap B')) = 0$, 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) converges to both f and f' uniformly a.e. on A, then $f = f'$ a.e. on A.

Moreover, if $f, g, f_n, g_n : A \to \mathbf{C}$ a.e. on A for every n and (f_n) converges to f and (g_n) converges to g a.e. on A, then $(f_n + g_n)$ converges to $f + g$ and $(f_n g_n)$ converges to fg a.e. on A. The same is true for uniform convergence a.e., provided that in the case of the product we also assume that the two sequences are *uniformly bounded a.e.*: namely, that there is an $M < +\infty$ so that $|f_n|, |g_n| \leq M$ a.e. on A for every $n \in \mathbf{N}$.

9.2 Convergence in the mean.

Assume that (X, Σ, μ) is a measure space.

Definition 9.1 *Let $f, f_n : X \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$ be measurable for all n. We say that (f_n)* ***converges to f in the mean on*** *$A \in \Sigma$ if f and f_n are finite a.e. on A*

for all n and

$$\int_A |f_n - f| \, d\mu \to 0$$

as $n \to +\infty$.

We say that (f_n) ***is Cauchy in the mean on*** *$A \in \Sigma$ if f_n is finite a.e. on A for all n and*

$$\int_A |f_n - f_m| \, d\mu \to 0$$

as $m, n \to +\infty$.

It is necessary to make a comment regarding the definition. The functions $|f_n - f|$ and $|f_n - f_m|$ are defined only a.e. on A. In fact, if all f, f_n are *finite* on $B \in \Sigma$ with $B \subseteq A$ and $\mu(A \setminus B) = 0$, then $|f_n - f|$ and $|f_n - f_m|$ are all defined on B and are $\Sigma \rceil B$-measurable. Therefore, only the integrals $\int_B |f_n - f| \, d\mu$ and $\int_B |f_n - f_m| \, d\mu$ are well-defined. If we want to be able to write the integrals $\int_A |f_n - f| \, d\mu$ and $\int_A |f_n - f_m| \, d\mu$, we must extend the functions $|f_n - f|$ and $|f_n - f_m|$ on X so that they are Σ-measurable and, after that, the integrals $\int_A |f_n - f| \, d\mu$ and $\int_A |f_n - f_m| \, d\mu$ will be defined and *equal to* $\int_B |f_n - f| \, d\mu$ and $\int_B |f_n - f_m| \, d\mu$, respectively. Since the values of the extensions outside B do not affect the resulting values of the integrals over A, it is simple and enough to extend all f, f_n as 0 on $X \setminus B$.

Thus, *the replacement of all f, f_n by 0 on $X \setminus B$ makes all functions finite everywhere on A without affecting the fact that (f_n) converges to f in the mean on A or that (f_n) is Cauchy in the mean on A.*

Proposition 9.1 *If (f_n) converges to both f and f' in the mean on A, then $f = f'$ a.e. on A.*

Proof: By the comment of the previous paragraph, we may assume that all f, f' and f_n are finite on A. This does not affect either the hypothesis or the result of the statement.

We write $\int_A |f - f'| \, d\mu \leq \int_A |f_n - f| \, d\mu + \int_A |f_n - f'| \, d\mu \to 0$ as $n \to +\infty$. Hence, $\int_A |f - f'| \, d\mu = 0$, implying that $f = f'$ a.e. on A.

Proposition 9.2 *Suppose (f_n) converges to f and (g_n) converges to g in the mean on A and $\lambda \in \mathbf{C}$. Then*
(i) $(f_n + g_n)$ converges to $f + g$ in the mean on A.
(ii) (λf_n) converges to λf in the mean on A.

Proof: We may assume that all f, g, f_n, g_n are finite on A.

Then, $\int_A |(f_n + g_n) - (f + g)| \, d\mu \leq \int_A |f_n - f| \, d\mu + \int_A |g_n - g| \, d\mu \to 0$ as $n \to +\infty$, and $\int_A |\lambda f_n - \lambda f| \, d\mu = |\lambda| \int_A |f_n - f| \, d\mu \to 0$ as $n \to +\infty$.

It is trivial to prove that, if (f_n) converges to f in the mean on A, then (f_n) is Cauchy in the mean on A. Indeed, assuming all f, f_n are finite on A, $\int_A |f_n - f_m| \, d\mu \leq \int_A |f_n - f| \, d\mu + \int_A |f_m - f| \, d\mu \to 0$ as $n, m \to +\infty$. The following basic theorem expresses the converse.

Theorem 9.1 *If (f_n) is Cauchy in the mean on A, then there is $f : X \to \mathbf{C}$ so that (f_n) converges to f in the mean on A. Moreover, there is a subsequence (f_{n_k}) which converges to f a.e. on A.*

As a corollary: if (f_n) converges to f in the mean on A, there is a subsequence (f_{n_k}) which converges to f a.e. on A.

Proof: As usual, we assume that all f, f_n are finite on A.

We have that, for every k, there is n_k so that $\int_A |f_n - f_m| \, d\mu < \frac{1}{2^k}$ for every $n, m \geq n_k$. Since we may assume that each n_k is as large as we like, we 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

$$\int_A |f_{n_{k+1}} - f_{n_k}| \, d\mu < \frac{1}{2^k}$$

for every k. Then, the measurable function $G : X \to [0, +\infty]$ defined by

$$G = \begin{cases} \sum_{k=1}^{+\infty} |f_{n_{k+1}} - f_{n_k}|, & \text{on } A \\ 0, & \text{on } A^c \end{cases}$$

satisfies $\int_X G \, d\mu = \sum_{k=1}^{+\infty} \int_A |f_{n_{k+1}} - f_{n_k}| \, d\mu = 1 < +\infty$. Thus, $G < +\infty$ a.e. on A and, hence, the series $\sum_{k=1}^{+\infty} (f_{n_{k+1}}(x) - f_{n_k}(x))$ converges for a.e. $x \in A$. Therefore, there is a $B \in \Sigma$, $B \subseteq A$ so that $\mu(A \setminus B) = 0$ and $\sum_{k=1}^{+\infty} (f_{n_{k+1}}(x) - f_{n_k}(x))$ converges for every $x \in B$. We define the measurable $f : X \to \mathbf{C}$ 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, hence, (f_{n_k}) converges to f a.e. on A.

We, also, have on B that $|f_{n_K} - f| = |f_{n_K} - f_{n_1} - \sum_{k=1}^{+\infty} (f_{n_{k+1}} - f_{n_k})| = |\sum_{k=1}^{K-1} (f_{n_{k+1}} - f_{n_k}) - \sum_{k=1}^{+\infty} (f_{n_{k+1}} - f_{n_k})| \leq \sum_{k=K}^{+\infty} |f_{n_{k+1}} - f_{n_k}|$ for all K. Hence,

$$\int_A |f_{n_K} - f| \, d\mu \leq \sum_{k=K}^{+\infty} \int_A |f_{n_{k+1}} - f_{n_k}| \, d\mu < \sum_{k=K}^{+\infty} \frac{1}{2^k} = \frac{1}{2^{K-1}} \to 0$$

as $K \to +\infty$.

From $n_k \to +\infty$, we get $\int_A |f_k - f| \, d\mu \leq \int_A |f_k - f_{n_k}| \, d\mu + \int_A |f_{n_k} - f| \, d\mu \to 0$ as $k \to +\infty$ and we conclude that (f_n) converges to f in the mean on A.

Example

Consider the sequence $f_1 = \chi_{(0,1)}, f_2 = \chi_{(0,\frac{1}{2})}, f_3 = \chi_{(\frac{1}{2},1)}, f_4 = \chi_{(0,\frac{1}{3})}, f_5 = \chi_{(\frac{1}{3},\frac{2}{3})}, f_6 = \chi_{(\frac{2}{3},1)}, f_7 = \chi_{(0,\frac{1}{4})}, f_8 = \chi_{(\frac{1}{4},\frac{2}{4})}, f_9 = \chi_{(\frac{2}{4},\frac{3}{4})}, f_{10} = \chi_{(\frac{3}{4},1)}$ and so on.

It is clear that $\int_{(0,1)} |f_n(x)| \, dm_1(x) \to 0$ as $n \to +\infty$ (the sequence of integrals 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, hence, (f_n) converges to 0 in the mean on $(0,1)$. By Theorem 9.1, there exists a subsequence converging to 0 a.e. on $(0,1)$ and it is easy to find many such subsequences: indeed, $f_1 = \chi_{(0,1)}, f_2 = \chi_{(0,\frac{1}{2})}, f_4 = \chi_{(0,\frac{1}{3})}, f_7 = \chi_{(0,\frac{1}{4})}$ and so on, is one such subsequence.

But, it is *not true* that (f_n) itself converges 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, thus, $(f_n(x))$ 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 9.2 as follows.

Proposition 9.3 *Suppose (f_n) converges to f and (g_n) converges to g in the mean on A.*
(i) If there is $M < +\infty$ so that $|f_n| \leq M$ a.e. on A, then $|f| \leq M$ a.e. on A.
(ii) If there is an $M < +\infty$ so that $|f_n|, |g_n| \leq M$ a.e. on A, then $(f_n g_n)$ converges to fg in the mean on A.

Proof: (i) Theorem 9.1 implies that there is a subsequence (f_{n_k}) which converges to f a.e. on A. Therefore, $|f_{n_k}| \to |f|$ a.e. on A and, hence, $|f| \leq M$ a.e. on A.
(ii) Assuming that all f, g, f_n, g_n are finite on A and using the result of (i), $\int_A |f_n g_n - fg|\, d\mu \leq \int_A |f_n g_n - f g_n|\, d\mu + \int_A |f g_n - fg|\, d\mu \leq M \int_A |f_n - f|\, d\mu + M \int_A |g_n - g|\, d\mu \to 0$ as $n \to +\infty$.

9.3 Convergence in measure.

Assume that (X, Σ, μ) is a measure space.

Definition 9.2 *Let $f, f_n : X \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$ be measurable for all n. We say that (f_n)* ***converges to f in (μ-)measure on*** *$A \in \Sigma$ if all f, f_n are finite a.e. on A and if for every $\epsilon > 0$ we have*

$$\mu(\{x \in A \,|\, |f_n(x) - f(x)| \geq \epsilon\}) \to 0$$

as $n \to +\infty$.

We say that (f_n) ***is Cauchy in (μ-)measure on*** *$A \in \Sigma$ if all f_n are finite a.e. on A and if for every $\epsilon > 0$ we have*

$$\mu(\{x \in A \,|\, |f_n(x) - f_m(x)| \geq \epsilon\}) \to 0$$

as $n, m \to +\infty$.

We make a comment similar to the comment following Definition 9.1. If we want to be able to write the values $\mu(\{x \in A \,|\, |f_n(x) - f(x)| \geq \epsilon\})$ and $\mu(\{x \in A \,|\, |f_n(x) - f(x)| \geq \epsilon\})$, we first extend the functions $|f_n - f|$ and $|f_n - f_m|$ outside the set $B \subseteq A$, where all f, f_n are finite, as functions defined on X and measurable. Then, since $\mu(A \setminus B) = 0$, we get that the above values are equal to the values $\mu(\{x \in B \,|\, |f_n(x) - f(x)| \geq \epsilon\})$ and, respectively, $\mu(\{x \in$

$B \,|\, |f_n(x) - f(x)| \geq \epsilon\})$. Therefore, the actual extensions play no role and, hence, we may for simplicity extend all f, f_n as 0 on $X \setminus B$.

Thus *the replacement of all f, f_n by 0 on $X \setminus B$ makes all functions finite everywhere on A and does not affect the fact that (f_n) converges to f in measure on A or that (f_n) is Cauchy in measure on A.*

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 9.4 *If (f_n) converges to both f and f' in measure on A, then $f = f'$ a.e. on A.*

Proof: We may assume that all f, f', f_n are finite on A.

Applying the above trick we find that $\mu(\{x \in A \,|\, |f(x) - f'(x)| \geq \epsilon\}) \leq \mu(\{x \in A \,|\, |f_n(x) - f(x)| \geq \frac{\epsilon}{2}\}) + \mu(\{x \in A \,|\, |f_n(x) - f'(x)| \geq \frac{\epsilon}{2}\}) \to 0$ as $n \to +\infty$. This implies $\mu(\{x \in A \,|\, |f(x) - f'(x)| \geq \epsilon\}) = 0$ for every $\epsilon > 0$.

We, now, write $\{x \in A \,|\, f(x) \neq f'(x)\} = \bigcup_{k=1}^{+\infty} \{x \in A \,|\, |f(x) - f'(x)| \geq \frac{1}{k}\}$. Since all terms in the union are null sets, we get $\mu(\{x \in A \,|\, f(x) \neq f'(x)\}) = 0$ and conclude that $f = f'$ a.e. on A.

Proposition 9.5 *Suppose (f_n) converges to f and (g_n) converges to g in measure on A and $\lambda \in \mathbf{C}$. Then*
(i) $(f_n + g_n)$ converges to $f + g$ in measure on A.
(ii) (λf_n) converges to λf in measure on A.
(iii) If there is $M < +\infty$ so that $|f_n| \leq M$ a.e. on A, then $|f| \leq M$ a.e. on A.
(iv) If there is $M < +\infty$ so that $|f_n|, |g_n| \leq M$ a.e. on A, then $(f_n g_n)$ converges to fg in measure on A.

Proof: We may assume that all f, f_n are finite on A, since all hypotheses and all results to be proved are not affected by any change of the functions on a subset of A of zero measure.
(i) We apply the usual trick and $\mu(\{x \in A \,|\, |(f_n + g_n)(x) - (f + g)(x)| \geq \epsilon\}) \leq \mu(\{x \in A \,|\, |f_n(x) - f(x)| \geq \frac{\epsilon}{2}\}) + \mu(\{x \in A \,|\, |g_n(x) - g(x)| \geq \frac{\epsilon}{2}\}) \to 0$ as $n \to +\infty$.
(ii) Also $\mu(\{x \in A \,|\, |\lambda f_n(x) - \lambda f(x)| \geq \epsilon\}) = \mu(\{x \in A \,|\, |f_n(x) - f(x)| \geq \frac{\epsilon}{|\lambda|}\}) \to 0$ as $n \to +\infty$.
(iii) We write $\mu(\{x \in A \,|\, |f(x)| \geq M + \epsilon\}) \leq \mu(\{x \in A \,|\, |f_n(x)| \geq M + \frac{\epsilon}{2}\}) + \mu(\{x \in A \,|\, |f_n(x) - f(x)| \geq \frac{\epsilon}{2}\}) = \mu(\{x \in A \,|\, |f_n(x) - f(x)| \geq \frac{\epsilon}{2}\}) \to 0$ as $n \to +\infty$. Hence, $\mu(\{x \in A \,|\, |f(x)| \geq M + \epsilon\}) = 0$ for every $\epsilon > 0$.

We have $\{x \in A \,|\, |f(x)| > M\} \subseteq \bigcup_{k=1}^{+\infty} \{x \in A \,|\, |f(x)| \geq M + \frac{1}{k}\}$ 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), $\mu(\{x \in A \,|\, |f_n(x) g_n(x) - f(x) g(x)| \geq \epsilon\}) \leq$

$\mu(\{x \in A \,|\, |f_n(x)g_n(x) - f_n(x)g(x)| \geq \frac{\epsilon}{2}\}) + \mu(\{x \in A \,|\, |f_n(x)g(x) - f(x)g(x)| \geq \frac{\epsilon}{2}\}) \leq \mu(\{x \in A \,|\, |g_n(x) - g(x)| \geq \frac{\epsilon}{2M}\}) + \mu(\{x \in A \,|\, |f_n(x) - f(x)| \geq \frac{\epsilon}{2M}\}) \to 0$ as $n \to +\infty$.

If (f_n) converges to f in measure on A, then (f_n) is Cauchy in measure on A. Indeed, taking all f, f_n finite on A, $\mu(\{x \in A \,|\, |f_n(x) - f_m(x)| \geq \epsilon\}) \leq \mu(\{x \in A \,|\, |f_n(x) - f(x)| \geq \frac{\epsilon}{2}\}) + \mu(\{x \in A \,|\, |f_m(x) - f(x)| \geq \frac{\epsilon}{2}\}) \to 0$ as $n, m \to +\infty$.

Theorem 9.2 *If (f_n) is Cauchy in measure on A, then there is $f : X \to \mathbf{C}$ so that (f_n) converges to f in measure on A. Moreover, there is a subsequence (f_{n_k}) which converges to f a.e. on A.*

As a corollary: if (f_n) converges to f in measure on A, there is a subsequence (f_{n_k}) which converges to f a.e. on A.

Proof: As usual, we assume that all f_n are finite on A.

We have, for all k, $\mu(\{x \in A \,|\, |f_n(x) - f_m(x)| \geq \frac{1}{2^k}\}) \to 0$ as $n, m \to +\infty$. Therefore, there is n_k so that $\mu(\{x \in A \,|\, |f_n(x) - f_m(x)| \geq \frac{1}{2^k}\}) < \frac{1}{2^k}$ 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. Hence, (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 that

$$\mu\Big(\Big\{x \in A \,|\, |f_{n_{k+1}}(x) - f_{n_k}(x)| \geq \frac{1}{2^k}\Big\}\Big) < \frac{1}{2^k}$$

for every k. For simplicity, we write

$$E_k = \Big\{x \in A \,|\, |f_{n_{k+1}}(x) - f_{n_k}(x)| \geq \frac{1}{2^k}\Big\}$$

and, hence, $\mu(E_k) < \frac{1}{2^k}$ for all k. We also define the *subsets* of A:

$$F_m = \cup_{k=m}^{+\infty} E_k \,, \qquad F = \cap_{m=1}^{+\infty} F_m = \limsup E_k \,.$$

Now, $\mu(F_m) \leq \sum_{k=m}^{+\infty} \mu(E_k) < \sum_{k=m}^{+\infty} \frac{1}{2^k} = \frac{1}{2^{m-1}}$ and, hence, $\mu(F) \leq \mu(F_m) < \frac{1}{2^{m-1}}$ for every m. This implies

$$\mu(F) = 0.$$

If $x \in A \setminus F$, 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 $k \geq m$, so that $\sum_{k=m}^{+\infty} |f_{n_{k+1}}(x) - f_{n_k}(x)| < \frac{1}{2^{m-1}}$. Thus, the series $\sum_{k=m}^{+\infty} (f_{n_{k+1}}(x) - f_{n_k}(x))$ converges and we may define $f : X \to \mathbf{C}$ by

$$f = \begin{cases} f_{n_1}(x) + \sum_{k=1}^{+\infty} (f_{n_{k+1}} - f_{n_k}), & \text{on } A \setminus F \\ 0, & \text{on } A^c \cup F. \end{cases}$$

By $f(x) = f_{n_1}(x) + \lim_{K \to +\infty} \sum_{k=1}^{K-1} (f_{n_{k+1}}(x) - f_{n_k}(x)) = \lim_{K \to +\infty} f_{n_K}(x)$ for every $x \in A \setminus F$ 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 $|f_{n_m} - f| = |f_{n_m} - f_{n_1} - \sum_{k=1}^{+\infty}(f_{n_{k+1}} - f_{n_k})| = |\sum_{k=1}^{m-1}(f_{n_{k+1}} - f_{n_k}) - \sum_{k=1}^{+\infty}(f_{n_{k+1}} - f_{n_k})| \leq \sum_{k=m}^{+\infty} |f_{n_{k+1}} - f_{n_k}| < \frac{1}{2^{m-1}}$. Therefore, $\{x \in A \,|\, |f_{n_m}(x) - f(x)| \geq \frac{1}{2^{m-1}}\} \subseteq F_m$ and, hence,

$$\mu\Big(\Big\{x \in A \,|\, |f_{n_m}(x) - f(x)| \geq \frac{1}{2^{m-1}}\Big\}\Big) \leq \mu(F_m) < \frac{1}{2^{m-1}}.$$

Take an arbitrary $\epsilon > 0$ and m_0 large enough so that $\frac{1}{2^{m_0-1}} \leq \epsilon$. If $m \geq m_0$, $\{x \in A \,|\, |f_{n_m}(x) - f(x)| \geq \epsilon\} \subseteq \{x \in A \,|\, |f_{n_m}(x) - f(x)| \geq \frac{1}{2^{m-1}}\}$ and, hence,

$$\mu(\{x \in A \,|\, |f_{n_m}(x) - f(x)| \geq \epsilon\}) < \frac{1}{2^{m-1}} \to 0$$

as $m \to +\infty$. This means that (f_{n_k}) converges to f in measure on A.

Since $n_k \to +\infty$ as $k \to +\infty$, we get $\mu(\{x \in A \,|\, |f_k(x) - f(x)| \geq \epsilon\}) \leq \mu(\{x \in A \,|\, |f_k(x) - f_{n_k}(x)| \geq \frac{\epsilon}{2}\}) + \mu(\{x \in A \,|\, |f_{n_k}(x) - f(x)| \geq \frac{\epsilon}{2}\}) \to 0$ as $k \to +\infty$ and we conclude that (f_n) converges to f in measure on A.

Example
We consider the example just after Theorem 9.1. If $0 < \epsilon \leq 1$, the sequence of the values $m_1(\{x \in (0,1) \,|\, |f_n(x)| \geq \epsilon\})$ 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, hence, converges to 0. Therefore, (f_n) converges to 0 in measure on $(0,1)$. But, as we have seen, it is not true that (f_n) converges to 0 a.e. on $(0,1)$.

9.4 Almost uniform convergence.

Assume that (X, Σ, μ) is a measure space.

Definition 9.3 *Let $f, f_n : X \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$ be measurable for all $n \in \mathbf{N}$. We say that (f_n)* ***converges to f (μ-)almost uniformly on*** *$A \in \Sigma$ if for every $\delta > 0$ there is $B \in \Sigma$, $B \subseteq A$, so that $\mu(A \setminus B) < \delta$ and (f_n) converges to f uniformly on B.*

We say that (f_n) ***is Cauchy (μ-)almost uniformly on*** *$A \in \Sigma$ if for every $\delta > 0$ there is $B \in \Sigma$, $B \subseteq A$, so that $\mu(A \setminus B) < \delta$ and (f_n) is Cauchy uniformly on B.*

Suppose that some $g : X \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$ is measurable and that, for every k, there is a $B_k \in \Sigma$, $B_k \subseteq A$, with $\mu(A \setminus B_k) < \frac{1}{k}$ so that g is finite on B_k. Now, it is clear that g is finite on the set $F = \cup_{k=1}^{+\infty} B_k$ and that $\mu(A \setminus F) \leq \mu(A \setminus B_k) < \frac{1}{k}$ for all k. This implies that $\mu(A \setminus F) = 0$ and, hence, g is finite a.e. on A.

From the statement of Definition 9.3. it is implied by the uniform convergence that all functions f, f_n are finite on sets $B \in \Sigma$, $B \subseteq A$ with $\mu(A \setminus B) < \delta$. Since δ is arbitrary, by the discussion in the previous paragraph, we conclude that, *if (f_n) converges to f almost uniformly on A or if it is Cauchy almost uniformly on A, then all f, f_n are finite a.e. on A.* Now, if $F \in \Sigma$, $F \subseteq A$ with $\mu(A \setminus F) = 0$ is the set where all f, f_n are finite, then, *if we replace all f, f_n by 0 on $X \setminus F$, the resulting functions f, f_n are all finite on A and the fact that (f_n)*

converges to f almost uniformly on A or that it is Cauchy almost uniformly on A is not affected.

Proposition 9.6 *If (f_n) converges to both f and f' almost uniformly on A, then $f = f'$ a.e. on A.*

Proof: Suppose that $\mu(\{x \in A \,|\, f(x) \neq f'(x)\}) > 0$. For simplicity, we set $E = \{x \in A \,|\, f(x) \neq f'(x)\}$.

We find $B \in \Sigma$, $B \subseteq A$, with $\mu(A \setminus B) < \frac{\mu(E)}{2}$ so that (f_n) converges to f uniformly on B. We, also, find $B' \in \Sigma$, $B' \subseteq A$, with $\mu(A \setminus B') < \frac{\mu(E)}{2}$ so that (f_n) converges to f' uniformly on B'. We, then, set $D = B \cap B'$ and have that $\mu(A \setminus D) < \mu(E)$ and (f_n) converges to both f and f' uniformly on D. This, of course, implies that $f = f'$ on D and, hence, that $D \cap E = \emptyset$.

But, then, $E \subseteq A \setminus D$ and, hence, $\mu(E) \leq \mu(A \setminus D) < \mu(E)$ and we arrive at a contradiction.

Proposition 9.7 *Suppose (f_n) converges to f and (g_n) converges to g almost uniformly on A. Then*
(i) $(f_n + g_n)$ converges to $f + g$ almost uniformly on A.
(ii) (λf_n) converges to λf almost uniformly on A.
(iii) If there is $M < +\infty$ so that $|f_n| \leq M$ a.e. on A, then $|f| \leq M$ a.e. on A.
(iv) If there is $M < +\infty$ so that $|f_n|, |g_n| \leq M$ a.e. on A, then $(f_n g_n)$ converges to fg almost uniformly on A.

Proof: We may assume that all f, f_n are finite on A.
(i) For arbitrary $\delta > 0$, there is $B' \in \Sigma$, $B' \subseteq A$, with $\mu(A \setminus B') < \frac{\delta}{2}$ so that (f_n) converges to f uniformly on B' and there is $B'' \in \Sigma$, $B'' \subseteq A$, with $\mu(A \setminus B'') < \frac{\delta}{2}$ so that (g_n) converges to g uniformly on B''. We take $B = B' \cap B''$ and have that $\mu(A \setminus B) < \delta$ and that (f_n) and (g_n) converge to f and, respectively, g uniformly on B. Then $(f_n + g_n)$ converges to $f + g$ uniformly on B and, since δ is arbitrary, we conclude that $(f_n + g_n)$ converges to $f + g$ almost uniformly on A.
(ii) This is easier, since, if (f_n) converges to f uniformly on B, then (λf_n) converges to λf uniformly on B.
(iii) Suppose $\mu(\{x \in A \,|\, |f(x)| > M\}) > 0$ and set $E = \{x \in A \,|\, |f(x)| > M\}$.

We find $B \in \Sigma$, $B \subseteq A$, with $\mu(A \setminus B) < \mu(E)$ so that (f_n) converges to f uniformly on B. Then we have $|f| \leq M$ a.e. on B and, hence, $\mu(B \cap E) = 0$. 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 $\delta > 0$ we find $B_1 \in \Sigma$, $B_1 \subseteq A$, with $\mu(A \setminus B_1) < \delta$ so that (f_n) and (g_n) converge to f and, respectively, g uniformly on B_1. By the result of (iii), $|f| \leq M$ a.e. on A and, hence, there is a $B_2 \in \Sigma$, $B_2 \subseteq A$ with $\mu(A \setminus B_2) = 0$ so that $|f_n|, |g_n|, |f| \leq M$ on B_2. We set $B = B_1 \cap B_2$, so that $\mu(A \setminus B) = \mu(A \setminus B_1) < \delta$. 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, thus, $(f_n g_n)$ converges to fg uniformly on B. We conclude that $(f_n g_n)$ converges to fg almost uniformly on A.

One should notice the difference between the next result and the corresponding Theorems 9.1 and 9.2 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, hence, for *every* subsequence).

Before the next result, let us consider a simple general fact.

Assume that there is a collection of functions $g_i : B_i \to \mathbf{C}$, indexed by the set I of indices, where $B_i \subseteq X$ for every $i \in I$, and that (f_n) converges 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 $B = \cup_{i \in I} B_i$ of the domains of definition. Hence, we can define a single function $f : B \to \mathbf{C}$ by

$$f(x) = g_i(x),$$

where $i \in I$ is any index for which $x \in B_i$, and it is clear that (f_n) converges to f pointwise on B.

Theorem 9.3 *If (f_n) is Cauchy almost uniformly on A, then there is an $f : X \to \mathbf{C}$ so that (f_n) converges to f almost uniformly on A. Moreover, (f_n) converges to f a.e. on A.*

As a corollary: if (f_n) converges to f almost uniformly on A, then (f_n) converges to f a.e. on A.

Proof: For each k, there exists $B_k \in \Sigma$, $B_k \subseteq A$, with $\mu(A \setminus B_k) < \frac{1}{k}$ so that (f_n) is Cauchy uniformly on B_k. Therefore, there is a function $g_k : B_k \to \mathbf{C}$ so that (f_n) converges to g_k uniformly and, hence, pointwise on B_k.

By the general result of the paragraph just before this theorem, there is an $f : B \to \mathbf{C}$, where $B = \cup_{k=1}^{+\infty} B_k$, so that (f_n) converges to f pointwise on B. But, $\mu(A \setminus B) \leq \mu(A \setminus B_k) < \frac{1}{k}$ for every k and, thus, $\mu(A \setminus B) = 0$. If we extend $f : X \to \mathbf{C}$, by defining $f = 0$ on B^c, we conclude that (f_n) converges to f a.e. on A.

By the general construction of f, we have that $g_k = f$ on B_k and, hence, (f_n) converges 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 $\mu(A \setminus B_k) < \delta$. Hence, (f_n) converges to f almost uniformly on A.

9.5 Relations between 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, Σ, μ) be a measure space.

Theorem 9.4 *If (f_n) converges to f almost uniformly on A, then (f_n) converges to f a.e. on A.*

The converse is true under the additional assumption that either
(i) (Egoroff) all f, f_n *are finite a.e. on* A *and* $\mu(A) < +\infty$
or
(ii) there is a $g : A \to [0, +\infty]$ *with* $\int_A g\, d\mu < +\infty$ *and* $|f_n| \leq g$ *a.e. on* A *for every* n.

Proof: The first statement is inluded in Theorem 9.3.
(i) Assume (f_n) converges to f a.e. on A, all f, f_n are finite a.e. on A and $\mu(A) < +\infty$. We may assume that all f, f_n are finite on A and, for each k, n, we define

$$E_n(k) = \cup_{m=n}^{+\infty} \Big\{ x \in A \,|\, |f_m(x) - f(x)| > \frac{1}{k} \Big\}.$$

If $C = \{x \in A \,|\, f_n(x) \to f(x)\}$, then it is easy to see that $\cap_{n=1}^{+\infty} E_n(k) \subseteq A \setminus C$. Since $\mu(A \setminus C) = 0$, we get $\mu(\cap_{n=1}^{+\infty} E_n(k)) = 0$ for every k. From $E_n(k) \downarrow \cap_{n=1}^{+\infty} E_n(k)$, from $\mu(A) < +\infty$ and from the continuity of μ from above, we find that $\mu(E_n(k)) \to 0$ as $n \to +\infty$. Hence, for an arbitrary $\delta > 0$, there is n_k so that

$$\mu(E_{n_k}(k)) < \frac{\delta}{2^k}.$$

We define

$$E = \cup_{k=1}^{+\infty} E_{n_k}(k), \qquad B = A \setminus E$$

and have $\mu(E) \leq \sum_{k=1}^{+\infty} \mu(E_{n_k}(k)) < \delta$. Also, for every $x \in B$ we have that, for every $k \geq 1$, $|f_m(x) - f(x)| \leq \frac{1}{k}$ for all $m \geq n_k$. Equivalently, for every $k \geq 1$,

$$\sup_{x \in B} |f_m(x) - f(x)| \leq \frac{1}{k}$$

for every $m \geq n_k$. This implies, of course, that (f_n) converges to f uniformly on B. Since $\mu(A \setminus B) = \mu(E) < \delta$, we conclude that (f_n) converges to f almost uniformly on A.
(ii) If $|f_n| \leq g$ a.e. on A for all n, then also $|f| \leq g$ a.e. on A and, since $\int_A g\, d\mu < +\infty$, all f, f_n are finite a.e. on A. Assuming, as we may, that all f, f_n are finite on A, we get $|f_n - f| \leq 2g$ a.e. on A for all n. Using the same notation as in the proof of (i), this implies that $E_n(k) \subseteq \{x \in A \,|\, g(x) > \frac{1}{2k}\}$ except for a null set. Therefore

$$\mu(E_n(k)) \leq \mu\Big(\Big\{ x \in A \,|\, g(x) > \frac{1}{2k} \Big\} \Big)$$

for every n, k. It is clear that the assumption $\int_A g\, d\mu < +\infty$ implies

$$\mu\Big(\Big\{ x \in A \,|\, g(x) > \frac{1}{2k} \Big\} \Big) < +\infty.$$

Therefore, we may, again, apply 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.

Example
If $f_n = \chi_{(n,n+1)}$ for every $n \geq 1$, then (f_n) converges to 0 everywhere on $\mathbf{R}$, but (f_n) does not converge to 0 almost uniformly on $\mathbf{R}$. In fact, if $0 < \delta \leq 1$, then every Lebesgue measurable $B \subseteq \mathbf{R}$ with $m_1(\mathbf{R} \setminus B) < \delta$ must have non-empty intersection with every interval $(n, n+1)$ and, hence, $\sup_{x \in B} |f_n(x)| \geq 1$ for every n.

In this example, of course, $m_1(\mathbf{R}) = +\infty$ and it is easy to see that there is *no* $g : \mathbf{R} \to [0, +\infty]$ with $\int_{\mathbf{R}} g(x)\, dm_1(x) < +\infty$ satisfying $f_n \leq g$ a.e. on $\mathbf{R}$ for every n. Otherwise, $g \geq 1$ a.e. on $(1, +\infty)$.

Theorem 9.5 *If (f_n) converges to f almost uniformly on A, then (f_n) converges to f in measure on A.*

Conversely, if (f_n) converges to f in measure on A, then there is a subsequence (f_{n_k}) which converges to f almost uniformly on A.

Proof: Suppose that (f_n) converges to f almost uniformly on A and take an arbitrary $\epsilon > 0$. For every $\delta > 0$ there is a $B \in \Sigma$, $B \subseteq A$, with $\mu(A \setminus B) < \delta$ so that (f_n) converges to f uniformly on B.

Now, there exists an n_0 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 (f_n) converges to f in measure on A.

The idea for the converse is already in the proof of Theorem 9.2.

We assume that (f_n) converges to f in measure on A and, without loss of generality, that all f, f_n are finite on A. Then $\mu(\{x \in A \,|\, |f_n(x) - f(x)| \geq \frac{1}{2^k}\}) \to 0$ as $n \to +\infty$ and there is n_k so that $\mu(\{x \in A \,|\, |f_n(x) - f(x)| \geq \frac{1}{2^k}\}) < \frac{1}{2^k}$ for all $n \geq n_k$. We may, inductively, assume that $n_k < n_{k+1}$ for all k and, hence, that (f_{n_k}) is a subsequence of (f_n) for which

$$\mu\Big(\Big\{x \in A \,|\, |f_{n_k}(x) - f(x)| \geq \frac{1}{2^k}\Big\}\Big) < \frac{1}{2^k}$$

for every $k \geq 1$. We set

$$E_k = \Big\{x \in A \,|\, |f_n(x) - f(x)| \geq \frac{1}{2^k}\Big\}, \qquad F_m = \cup_{k=m}^{+\infty} E_k.$$

Then $\mu(F_m) < \sum_{k=m}^{+\infty} \frac{1}{2^k} = \frac{1}{2^{m-1}}$ for every m.

If $x \in A \setminus F_m$, then $x \in A \setminus E_k$ for every $k \geq m$ so that $|f_{n_k}(x) - f(x)| < \frac{1}{2^k}$ 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 hence $\sup_{x \in A \setminus F_m} |f_{n_k}(x) - f(x)| \to 0$ as $k \to +\infty$. Therefore, (f_{n_k}) converges to f uniformly on $A \setminus F_m$ and we conclude that (f_{n_k}) converges to f almost uniformly on A.

Example
We consider the example just after Theorem 9.1. The sequence (f_n) converges to 0 in measure on $(0,1)$ but it does not converge to 0 almost uniformly on $(0,1)$. In fact, if we take any δ with $0 < \delta \le 1$, then every $B \subseteq (0,1)$ 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 every value of m) and, hence, $\sup_{x\in B}|f_n(x)| \ge 1$ for infinitely many n.

The converse in Theorem 9.6 is a variant of the Dominated Convergence Theorem.

Theorem 9.6 *If (f_n) converges to f in the mean on A, then (f_n) converges 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 $|f_n| \le g$ a.e. on A.

Proof: It is clear that we may assume all f, f_n are finite on A.

Suppose that (f_n) converges to f in the mean on A. Then, for every $\epsilon > 0$ we have

$$\mu(\{x \in A \,|\, |f_n(x) - f(x)| \ge \epsilon\}) \le \frac{1}{\epsilon}\int_A |f_n - f|\,d\mu \to 0$$

as $n \to +\infty$. Therefore, (f_n) converges to f in measure on A.

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

$$\int_A |f_{n_k} - f|\,d\mu \ge \epsilon_0$$

for every $k \ge 1$. Since (f_{n_k}) converges to f in measure, Theorem 9.2 implies that there is a subsequence $(f_{n_{k_l}})$ which converges to f a.e. on A. From $|f_{n_{k_l}}| \le g$ a.e. on A, we find that $|f| \le g$ 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.

Example
Let $f_n = n\chi_{(0,\frac{1}{n})}$ for every n. If $0 < \epsilon \le 1$, then $\mu(\{x \in (0,1)\,|\,|f_n(x)| \ge \epsilon\}) = \frac{1}{n} \to 0$ as $n \to +\infty$ and, hence, (f_n) converges to 0 in measure on $(0,1)$. But $\int_0^1 |f_n(x)|\,dm_1(x) = 1$ and (f_n) does not converge to 0 in the mean on $(0,1)$.

If $g : (0,1) \to [0,+\infty]$ is such that $|f_n| \le g$ a.e. on $(0,1)$ for every n, then $g \ge n$ a.e. in each interval $[\frac{1}{n+1}, \frac{1}{n})$. Hence, $\int_0^1 g(x)\,dm_1(x) \ge \sum_{n=1}^{+\infty}\int_{\frac{1}{n+1}}^{\frac{1}{n}} n\,dm_1(x) = \sum_{n=1}^{+\infty} n(\frac{1}{n} - \frac{1}{n+1}) = \sum_{n=1}^{+\infty}\frac{1}{n+1} = +\infty$.

9.6 Exercises.

Except if specified otherwise, all exercises refer to a measure space (X, Σ, μ), all sets belong to Σ and all functions are Σ-measurable.

1. Let $\phi : \mathbf{C} \to \mathbf{C}$.
   (i) If ϕ is continuous and (f_n) converges to f a.e. on A, prove that $(\phi \circ f_n)$ converges to $\phi \circ f$ a.e. on A.
   (ii) If ϕ is uniformly continuous and (f_n) converges to f in measure or almost uniformly on A, prove that $(\phi \circ f_n)$ converges to $\phi \circ f$ in measure or, respectively, almost uniformly on A.

2. (i) If (f_n) converges to f with respect to any of the four types of convergence (a.e. or in the mean or in measure or almost uniformly) on A and (f_n) converges, also, to f' with respect to any other of the same four types of convergence, prove that $f = f'$ a.e. on A.

   (ii) If (f_n) converges to f with respect to any of the four types of convergence on A and $|f_n| \leq g$ a.e. on A for all n, prove that $|f| \leq g$ a.e. on A.

3. If $E_n \subseteq A$ for every n and (χ_{E_n}) converges to f in the mean or in measure or almost uniformly or a.e. on A, prove that there exists $E \subseteq A$ so that $f = \chi_E$ a.e. on A.

4. Suppose that $E_n \subseteq A$ for every n. Prove that (χ_{E_n}) is Cauchy in measure or in the mean or almost uniformly on A if and only if $\mu(E_n \triangle E_m) \to 0$ as $n, m \to +\infty$.

5. Let $\sharp$ be the counting measure on $(\mathbf{N}, \mathcal{P}(\mathbf{N}))$. Prove that (f_n) converges to f uniformly on $\mathbf{N}$ if and only if (f_n) converges to f in measure on $\mathbf{N}$.

6. *A variant of the Lemma of Fatou.*

   If $f_n \geq 0$ a.e. on A and (f_n) converges to f in measure on A, prove that $\int_A f \, d\mu \leq \liminf_{n \to +\infty} \int_A f_n \, d\mu$.

7. *The Dominated Convergence Theorem.*

   Prove the Dominated Convergence Theorem in two ways, using either the first converse or the second converse of Theorem 9.4.

8. *A variant of the Dominated Convergence Theorem.*

   Suppose that $|f_n| \leq g$ a.e. on A, that $\int_A g \, d\mu < +\infty$ and that (f_n) converges to f in measure on A. Prove that $\int_A f_n \, d\mu \to \int_A f \, d\mu$.

   One can follow three paths. One is to use the result of Exercise 9.6.2. 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.

9. Suppose that A is of σ-finite measure and (f_n) converges to f a.e. on A. Prove that, for each k, there exists $E_k \subseteq A$ so that (f_n) converges to f uniformly on E_k and $\mu(A \setminus \cup_{k=1}^{+\infty} E_k) = 0$.

10. Suppose that $E_k(\epsilon) = \{x \in A \,|\, |f_k(x) - f(x)| \geq \epsilon\}$ for every k and $\epsilon > 0$. If $\mu(A) < +\infty$, prove that (f_n) converges to f a.e. on A if and only if, for every $\epsilon > 0$, $\mu(\cup_{k=n}^{+\infty} E_k(\epsilon)) \to 0$ as $n \to +\infty$.

11. (i) Let (h_n) satisfy $\sup_{n\in\mathbf{N}} |h_n(x)| < \infty$ for a.e. $x \in A$. If $\mu(A) < +\infty$, prove that for every $\delta > 0$ there is a $B \subseteq A$ with $\mu(A \setminus B) < \delta$ so that $\sup_{x\in B, n\in\mathbf{N}} |h_n(x)| < +\infty$.
(ii) Let (f_n) converge to f in measure on A and (g_n) converge to g in measure on A. If $\mu(A) < +\infty$, prove that $(f_n g_n)$ converges to fg in measure on A.

12. Suppose that $\mu(A) < +\infty$ and every f_n is finite a.e. on A.
(i) Prove that there is a sequence (λ_n) of positive numbers so that $(\lambda_n f_n)$ converges to 0 a.e. on A.
(ii) Prove that there exists $g : A \to [0, +\infty]$ and a sequence (r_n) in $\mathbf{R}^+$ so that $|f_n| \leq r_n g$ a.e. on A for every n.

13. Suppose that $\mu(A) < +\infty$ and (f_n) converges to 0 a.e. on A.
(i) Prove that there exists a sequence (λ_n) in $\mathbf{R}^+$ with $\lambda_n \uparrow +\infty$ so that $(\lambda_n f_n)$ converges to 0 a.e. on A.
(ii) Prove that there exists $g : A \to [0, +\infty]$ and a sequence (ϵ_n) in $\mathbf{R}^+$ with $\epsilon_n \to 0$ so that $|f_n| \leq \epsilon_n g$ a.e. on A for every n.

14. *A characterisation of convergence in measure.*

If $\mu(A) < +\infty$, prove that (f_n) converges to f in measure on A if and only if $\int_A \frac{|f_n - f|}{1+|f_n - f|}\, d\mu \to 0$ as $n \to +\infty$.

In general, prove that (f_n) converges to f in measure on A if and only if

$$\inf_{\epsilon>0} \frac{\epsilon + \mu(\{x \in A \,|\, |f_n(x) - f(x)| \geq \epsilon\})}{1 + \epsilon + \mu(\{x \in A \,|\, |f_n(x) - f(x)| \geq \epsilon\})} \to 0$$

as $n \to +\infty$.

15. *A variant of Egoroff's Theorem for continuous parameter.*

Let $\mu(X) < +\infty$ and $f : X \times [0, 1] \to \mathbf{C}$ has the properties:
(a) $f(\cdot, y) : X \to \mathbf{C}$ is measurable for every $y \in [0, 1]$
(b) $f(x, \cdot) : [0, 1] \to \mathbf{C}$ is continuous for every $x \in X$.

(i) If $\epsilon, \eta > 0$, prove that $\{x \in X \,|\, |f(x, y) - f(x, 0)| \leq \epsilon$ for all $y < \eta\}$ belongs to Σ.
(ii) Prove that for every $\delta > 0$ there is $B \subseteq X$ with $\mu(X \setminus B) < \delta$ and $f(\cdot, y) \to f(\cdot, 0)$ uniformly on B as $y \to 0+$.

16. Let (f_n) converge 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.

17. Prove the converse part of Theorem 9.6 using the converse part of Theorem 9.5.

18. *The complete relation between convergence in the mean and convergence in measure: the Theorem of Vitali.*

    We say that **the indefinite integrals of** (f_n) **are uniformly absolutely continuous over** A if for every $\epsilon > 0$ there exists $\delta > 0$ so that $|\int_E f_n \, d\mu| < \epsilon$ for all $n \geq 1$ and all $E \subseteq A$ with $\mu(E) < \delta$.

    We say that **the indefinite integrals of** (f_n) **are equicontinuous from above at** $\emptyset$ **over** 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 $n \geq 1$.

    Prove that (f_n) converges to f in the mean on A if and only if (f_n) converges 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 9.6 related to this result?

19. *The Theorem of Lusin.*

    If f is Lebesgue measurable and finite a.e. on $\mathbf{R}^n$, then for every $\delta > 0$ there is a Lebesgue set $B \subseteq \mathbf{R}^n$ and a g, continuous on $\mathbf{R}^n$, so that $m_n(B^c) < \delta$ and $f = g$ on B.

    (i) Use Theorem 7.16 to find a sequence (ϕ_n) of functions continuous on $\mathbf{R}^n$ so that $\int_{\mathbf{R}^n} |f - \phi_n| \, dm_n \to 0$ as $n \to +\infty$. Theorem 9.1 implies that there is a subsequence (ϕ_{n_k}) which converges to f a.e. on $\mathbf{R}^n$.
    (ii) Consider the qubes $Q_{m_1,\ldots,m_n} = [m_1, m_1 + 1) \times \cdots \times [m_n, m_n + 1)$ for every choice of $m_1, \ldots, m_n \in \mathbf{Z}$ and enumerate them as $Q_1, Q_2, \ldots$. Then, these qubes are pairwise disjoint and they cover $\mathbf{R}^n$. Apply Egoroff's Theorem to prove that for each Q_k there is a closed set $B_k \subseteq Q_k$ with $m_n(Q_k \setminus B_k) < \frac{\delta}{2^k}$ so that (ϕ_{n_k}) converges to f uniformly on B_k. Conclude that the restriction $f \rceil B_k$ of f on B_k is continuous on B_k.
    (iii) Take $B = \cup_{k=1}^{+\infty} B_k$ and prove that $m_n(B^c) < \delta$, that B is closed and that the restriction $f \rceil B$ of f on B is continuous on B.
    (iv) Use the Extension Theorem of Tietze to prove that there is a g, continuous on $\mathbf{R}^n$, so that $g = f \rceil B$ on B.

20. If $f : \mathbf{R}^n \to \mathbf{C}$ is continuous in each variable separately, prove that f is Lebesgue measurable.

Chapter 10

Signed measures and complex measures

10.1 Signed measures.

Let (X, Σ) be a measurable space.

Definition 10.1 *A function $\nu : \Sigma \to \overline{\mathbf{R}}$ is called **a signed measure on** (X, Σ) if*
(i) either $\nu(A) \neq -\infty$ for all $A \in \Sigma$ or $\nu(A) \neq +\infty$ for all $A \in \Sigma$,
(ii) $\nu(\emptyset) = 0$,
(iii) $\nu(\cup_{j=1}^{+\infty} A_j) = \sum_{j=1}^{+\infty} \nu(A_j)$ for all pairwise disjoint $A_1, A_2, \ldots \in \Sigma$.

If ν is a signed measure on (X, Σ) and $\nu(A) \in \mathbf{R}$ for every $A \in \Sigma$, then ν is called **a real measure**. It is obvious that ν is **a non-negative signed measure** (i.e. with $\nu(A) \geq 0$ for every $A \in \Sigma$) if and only if ν is a measure. If $\nu(A) \leq 0$ for every $A \in \Sigma$, then ν is called **a non-positive signed measure**.

It is clear that, if ν is a non-negative signed measure, then $-\nu$ is a non-positive signed measure and conversely. Also, if ν and ν' are signed measures on (X, Σ) with either $\nu(A), \nu'(A) \neq -\infty$ for all $A \in \Sigma$ or $\nu(A), \nu'(A) \neq +\infty$ for all $A \in \Sigma$, then $\nu + \nu'$, well-defined by $(\nu + \nu')(A) = \nu(A) + \nu'(A)$ for all $A \in \Sigma$, is a signed measure. Similarly, the $\kappa\nu$, defined by $(\kappa\nu)(A) = \kappa\nu(A)$ for all $A \in \Sigma$, is a signed measure for every $\kappa \in \mathbf{R}$.

Examples
1. Let μ_1, μ_2 be two measures on (X, Σ). If $\mu_2(X) < +\infty$, then $\mu_2(A) \leq \mu_2(X) < +\infty$ for every $A \in \Sigma$. Then, $\nu = \mu_1 - \mu_2$ is well-defined and it is a signed measure on (X, Σ), because $\nu(A) = \mu_1(A) - \mu_2(A) \geq -\mu_2(A) > -\infty$ for all $A \in \Sigma$. Similarly, if $\mu_1(X) < +\infty$, then $\nu = \mu_1 - \mu_2$ is a signed measure on (X, Σ) with $\nu(A) < +\infty$ for all $A \in \Sigma$.

Hence, *the difference of two measures, at least one of which is finite, is a signed measure.*

2. Let μ be a measure on (X, Σ) and $f : X \to \overline{\mathbf{R}}$ be a measurable function such that the $\int_X f\, d\mu$ is defined. Lemma 7.10 says that the $\int_A f\, d\mu$ is defined for every $A \in \Sigma$. If we consider the function $\lambda : \Sigma \to \overline{\mathbf{R}}$ defined by

$$\lambda(A) = \int_A f\, d\mu$$

for all $A \in \Sigma$, then Proposition 7.6 and Theorem 7.13 imply that λ is a signed measure on (X, Σ).

Definition 10.2 *The signed measure λ which is defined in the previous paragraph is called **the indefinite integral of** f **with respect to** μ and it is denoted by $f\mu$. Thus, the defining relation for $f\mu$ is*

$$(f\mu)(A) = \int_A f\, d\mu, \qquad A \in \Sigma.$$

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 $A \in \Sigma$. 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 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. Then the signed measure $f^+\mu$ is a *finite* measure (because $(f^+\mu)(X) = \int_X f^+\, d\mu < +\infty$) and the signed measure $f^-\mu$ is a measure. Also, for every $A \in \Sigma$ 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 $f^- d\mu$, 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 $A, B \in \Sigma$ 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 10.1 *Let ν be a signed measure on (X, Σ).*
(i) Let $A, B \in \Sigma$ and $A \subseteq B$. If $\nu(B) < +\infty$, then $\nu(A) < +\infty$ and, if $\nu(B) > -\infty$, then $\nu(A) > -\infty$. In particular, if $\nu(B) \in \mathbf{R}$, then $\nu(A) \in \mathbf{R}$.
(ii) If $A, B \in \Sigma$, $A \subseteq B$ and $\nu(A) \in \mathbf{R}$, then $\nu(B \setminus A) = \nu(B) - \nu(A)$.
(iii) (Continuity from below) If $A_1, A_2, \ldots \in \Sigma$ and $A_n \subseteq A_{n+1}$ for all n, then $\nu(\cup_{n=1}^{+\infty} A_n) = \lim_{n \to +\infty} \nu(A_n)$.
(iv) (Continuity from above) If $A_1, A_2, \ldots \in \Sigma$, $\nu(A_1) \in \mathbf{R}$ and $A_n \supseteq A_{n+1}$ for all n, then $\nu(\cap_{n=1}^{+\infty} A_n) = \lim_{n \to +\infty} \nu(A_n)$.

Proof: (i) We have $\nu(B) = \nu(A) + \nu(B \setminus A)$.

If $\nu(A) = +\infty$, then $\nu(B \setminus A) > -\infty$ and, thus, $\nu(B) = +\infty$. Similarly, if $\nu(A) = -\infty$, then $\nu(B \setminus A) < +\infty$ and, thus, $\nu(B) = -\infty$.

The proofs of (ii), (iii) and (iv) are the same as the proofs of the corresponding parts of Theorem 2.1.

10.2 The Hahn and Jordan decompositions, I.

Let (X, Σ) be a measurable space.

Definition 10.3 *Let ν be a signed measure on (X, Σ).*
(i) $P \in \Sigma$ is called ***a positive set for*** *ν if $\nu(A) \geq 0$ for every $A \in \Sigma$, $A \subseteq P$.*
(ii) $N \in \Sigma$ is called ***a negative set for*** *ν if $\nu(A) \leq 0$ for every $A \in \Sigma$, $A \subseteq N$.*
(iii) $Q \in \Sigma$ is called ***a null set for*** *ν if $\nu(A) = 0$ for every $A \in \Sigma$, $A \subseteq Q$.*

It is obvious that an element of Σ which is both a positive and a negative set for ν is a null set for ν. It is also obvious that, if μ is a measure, then every $A \in \Sigma$ is a positive set for μ.

Proposition 10.1 *Let ν be a signed measure on (X, Σ).*
(i) If P is a positive set for ν, $P' \in \Sigma$, $P' \subseteq P$, then P' is a positive set for ν.
(ii) If $P_1, P_2, \ldots$ are positive sets for ν, then $\cup_{k=1}^{+\infty} P_k$ is a positive set for ν.
The same results are, also, true for negative sets and for null sets for ν.

Proof: (i) For every $A \in \Sigma$, $A \subseteq P'$ we have $A \subseteq P$ and, hence, $\nu(A) \geq 0$.
(ii) Take arbitrary $A \in \Sigma$, $A \subseteq \cup_{k=1}^{+\infty} P_k$. We can write $A = \cup_{k=1}^{+\infty} A_k$, where $A_1, A_2, \ldots \in \Sigma$ are pairwise disjoint and $A_k \subseteq P_k$ for every k. Indeed, we may set $A_1 = A \cap P_1$ and $A_k = A \cap \big(P_k \setminus (P_1 \cup \cdots \cup P_{k-1})\big)$ for all $k \geq 2$. By the result of (i), we then have $\nu(A) = \sum_{k=1}^{+\infty} \nu(A_k) \geq 0$.

Theorem 10.2 *Let ν be a signed measure on (X, Σ).*
(i) There exist a positive set P and a negative set N for ν so that $P \cup N = X$ and $P \cap N = \emptyset$.
(ii) $\nu(N) \leq \nu(A) \leq \nu(P)$ for every $A \in \Sigma$.
(iii) If $\nu(A) < +\infty$ for every $A \in \Sigma$, then ν is bounded from above, while if $-\infty < \nu(A)$ for every $A \in \Sigma$, then ν is bounded from below.
(iv) If P' is a positive set for ν and N' is a negative set for ν with $P' \cup N' = X$ and $P' \cap N' = \emptyset$, then $P \triangle P' = N \triangle N'$ is a null set for ν.

Proof: (i) We consider the case when $\nu(A) < +\infty$ for every $A \in \Sigma$.

We define the quantity

$$\kappa = \sup\{\nu(P) \,|\, P \text{ is a positive set for } \nu\}.$$

This set is non-empty since $\nu(\emptyset) = 0$ is one of its elements. Thus, $0 \leq \kappa$. We consider a sequence (P_k) of positive sets for ν so that $\nu(P_k) \to \kappa$ and form the set $P = \cup_{k=1}^{+\infty} P_k$ which, by Proposition 10.1, is a positive set for ν. This implies that $\nu(P \setminus P_k) \geq 0$ for every k and, hence, $\nu(P_k) \leq \nu(P) \leq \kappa$ for every k. Taking the limit, we find that

$$\kappa = \nu(P) < +\infty.$$

This P is *a positive set for ν of maximal ν-measure* and we shall prove that the set $N = X \setminus P$ is a negative set for ν.

Suppose that N is not a negative set for ν. Then there is $A_0 \in \Sigma$, $A_0 \subseteq N$, with $0 < \nu(A_0) < +\infty$. The set A_0 is not a positive set or, otherwise, the set $P \cup A_0$ would be a positive set with $\nu(P \cup A_0) = \nu(P) + \nu(A_0) > \nu(P)$, contradicting the maximality of P. Hence, there is at least one subset of A_0 in Σ having negative ν-measure. This means that

$$\tau_0 = \inf\{\nu(B) \,|\, B \in \Sigma, B \subseteq A_0\} < 0.$$

If $\tau_0 < -1$, there is $B_1 \in \Sigma$, $B_1 \subseteq A_0$ with $\nu(B_1) < -1$. If $-1 \leq \tau_0 < 0$, there is a $B_1 \in \Sigma$, $B_1 \subseteq A_0$ with $\nu(B_1) < \frac{\tau_0}{2}$. We set $A_1 = A_0 \setminus B_1$ and have $\nu(A_0) = \nu(A_1) + \nu(B_1) < \nu(A_1) < +\infty$. Observe that we are using Theorem 10.1 to imply $\nu(A_1), \nu(B_1) \in \mathbf{R}$ from $\nu(A_0) \in \mathbf{R}$.

Suppose that we have constructed sets $A_0, A_1, \ldots, A_n \in \Sigma$ and $B_1, \ldots, B_n \in \Sigma$ so that

- ⋄ $A_n \subseteq A_{n-1} \subseteq \cdots \subseteq A_1 \subseteq A_0 \subseteq N, \quad B_n = A_{n-1} \setminus A_n, \ldots, B_1 = A_0 \setminus A_1,$
- ⋄ $\tau_{k-1} = \inf\{\nu(B) \,|\, B \in \Sigma, B \subseteq A_{k-1}\} < 0,$
- ⋄ $\nu(B_k) < \begin{cases} -1, & \text{if } \tau_{k-1} < -1 \\ \frac{\tau_{k-1}}{2}, & \text{if } -1 \leq \tau_{k-1} < 0 \end{cases} \qquad \text{for all } k = 1, \ldots, n,$
- ⋄ $0 < \nu(A_0) < \nu(A_1) < \cdots < \nu(A_{n-1}) < \nu(A_n) < +\infty.$

Now, A_n is not a positive set for ν for the same reason that A_0 is not a positive set for ν. Hence, there is at least one subset of A_n in Σ having negative ν-measure. This means that

$$\tau_n = \inf\{\nu(B) \,|\, B \in \Sigma, B \subseteq A_n\} < 0.$$

If $\tau_n < -1$, there is $B_{n+1} \in \Sigma$, $B_{n+1} \subseteq A_n$ with $\nu(B_{n+1}) < -1$. If $-1 \leq \tau_n < 0$, there is a $B_{n+1} \in \Sigma$, $B_{n+1} \subseteq A_n$ with $\nu(B_{n+1}) < \frac{\tau_n}{2}$. We set $A_{n+1} = A_n \setminus B_{n+1}$ and have $\nu(A_n) = \nu(A_{n+1}) + \nu(B_{n+1}) < \nu(A_{n+1}) < +\infty$. This means that we have, inductively, constructed two sequences (A_n), (B_n) satisfying all the properties ⋄.

Now, the sets $B_1, B_2, \ldots$ and $\cap_{n=1}^{+\infty} A_n$ are pairwise disjoint and we have $A_0 = (\cap_{n=1}^{+\infty} A_n) \cup (\cup_{n=1}^{+\infty} B_n)$. Therefore, $\nu(A_0) = \nu(\cap_{n=1}^{+\infty} A_n) + \sum_{n=1}^{+\infty} \nu(B_n)$, from which we find

$$\sum_{n=1}^{+\infty} \nu(B_n) > -\infty.$$

This implies that $\nu(B_n) \to 0$ as $n \to +\infty$ and, by the third property ⋄,

$$\tau_{n-1} \to 0$$

as $n \to +\infty$. Now the set $A = \cap_{n=1}^{+\infty} A_n \in \Sigma$, by continuity from above of ν, has

$$\nu(A) = \lim_{n \to +\infty} \nu(A_n) > 0.$$

Moreover, A is not a positive set for ν for the same reason that A_0 is not a positive set for ν. Hence, there is some $B \in \Sigma$, $B \subseteq A$ with $\nu(B) < 0$. But then $B \subseteq A_{n-1}$ for all n and, hence, $\tau_{n-1} \leq \nu(B) < 0$ for all n. We, thus, arrive at a contradiction with the limit $\tau_{n-1} \to 0$.

In the same way, we can prove that, if $-\infty < \nu(A)$ for every $A \in \Sigma$, then there is a negative set N for ν of minimal ν-measure so that the set $P = X \setminus N$ is a positive set for ν.

Thus, in any case we have a positive set P and a negative set N for ν so that $P \cup N = X$ and $P \cap N = \emptyset$.

(ii) If $A \in \Sigma$, then $\nu(P \setminus A) \geq 0$, because $P \setminus A \subseteq P$. This implies $\nu(P) = \nu(P \cap A) + \nu(P \setminus A) \geq \nu(P \cap A)$ and, similarly, $\nu(N) \leq \nu(N \cap A)$. Therefore, $\nu(A) = \nu(P \cap A) + \nu(N \cap A) \leq \nu(P \cap A) \leq \nu(P)$ and $\nu(A) = \nu(P \cap A) + \nu(N \cap A) \geq \nu(N \cap A) \geq \nu(N)$.

(iii) This a consequence of the result of (ii).

(iv) Now, let P' be a positive set and N' be a negative set for ν with $P' \cup N' = X$ and $P' \cap N' = \emptyset$. Then, since $P \setminus P' = N' \setminus N \subseteq P \cap N'$, the set $P \setminus P' = N' \setminus N$ is both a positive set and a negative set for ν and, hence, a null set for ν. Similarly, $P' \setminus P = N \setminus N'$ is a null set for ν and we conclude that their union $P \triangle P' = N \triangle N'$ is a null set for ν.

Definition 10.4 *Let ν be a signed measure on (X, Σ). Every partition of X into a positive and a negative set for ν is called* ***a Hahn decomposition of** X **for** ν.*

It is clear from Theorem 10.2 that if P, N is a Hahn decomposition of X for ν, then

$$\nu(P) = \max\{\nu(A) \mid A \in \Sigma\}, \qquad \nu(N) = \min\{\nu(A) \mid A \in \Sigma\}.$$

Definition 10.5 *Let ν_1, ν_2 be two signed measures on (X, Σ). We say that they are* ***mutually singular*** *(or that ν_1 is singular to ν_2 or that ν_2 is singular to ν_1) if there exist $A_1 \in \Sigma$ which is null for ν_2 and $A_2 \in \Sigma$ which is null for ν_1 so that $A_1 \cup A_2 = X$ and $A_1 \cap A_2 = \emptyset$.*

We use the symbol $\nu_1 \perp \nu_2$ to denote that ν_1, ν_2 are mutually singular.

In other words, *two signed measures are mutually singular if there is a set in Σ which is null for one of them and its complement is null for the other.*

If ν_1, ν_2 are mutually singular and A_1, A_2 are as in the Definition 10.5, then it is clear that

$$\nu_1(A) = \nu_1(A \cap A_1), \qquad \nu_2(A) = \nu_2(A \cap A_2)$$

for every $A \in \Sigma$. Thus, in a free language, we may say that ν_1 *is concentrated on* A_1 and ν_2 *is concentrated on* A_2.

Proposition 10.2 *Let ν, ν_1, ν_2 be signed measures on (X, Σ). If $\nu_1, \nu_2 \perp \nu$ and $\nu_1 + \nu_2$ is defined, then $\nu_1 + \nu_2 \perp \nu$.*

Proof: Take $A_1, B_1, A_2, B_2 \in \Sigma$ 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 $\nu_1 + \nu_2$. 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 $\nu_1 + \nu_2 \perp \nu$.

Theorem 10.3 *Let ν be a signed measure on (X, Σ). There exist two non-negative signed measures (i.e. measures) ν^+ and ν^-, at least one of which is finite, so that*

$$\nu = \nu^+ - \nu^- , \qquad \nu^+ \perp \nu^- .$$

If μ_1, μ_2 are two measures on (X, Σ), at least one of which is finite, so that $\nu = \mu_1 - \mu_2$ and $\mu_1 \perp \mu_2$, then $\mu_1 = \nu^+$ and $\mu_2 = \nu^-$.

Proof: We consider any Hahn decomposition of X for ν: P is a positive set and N a negative set for ν so that $P \cup N = X$ and $P \cap N = \emptyset$.

We define $\nu^+, \nu^- : \Sigma \to [0, +\infty]$ by

$$\nu^+(A) = \nu(A \cap P), \qquad \nu^-(A) = -\nu(A \cap N)$$

for every $A \in \Sigma$. It is trivial to see that ν^+, ν^- are non-negative signed measures on (X, Σ). If $\nu(A) < +\infty$ for every $A \in \Sigma$, then $\nu^+(X) = \nu(P) < +\infty$ and, hence, ν^+ is a finite measure. Similarly, if $-\infty < \nu(A)$ for every $A \in \Sigma$, then $\nu^-(X) = -\nu(N) < +\infty$ and, hence, ν^- is a finite measure.

Also, $\nu(A) = \nu(A \cap P) + \nu(A \cap N) = \nu^+(A) - \nu^-(A)$ for all $A \in \Sigma$ and, thus, $\nu = \nu^+ - \nu^-$.

If $A \in \Sigma$ 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, hence, $\nu^+ \perp \nu^-$.

Now, let μ_1, μ_2 be two measures on (X, Σ), at least one of which is finite, so that $\nu = \mu_1 - \mu_2$ and $\mu_1 \perp \mu_2$. Consider $A_1, A_2 \in \Sigma$, with $A_1 \cup A_2 = X$ and $A_1 \cap A_2 = \emptyset$, so that A_2 is a null set for μ_1 and A_1 is a null set for μ_2.

If $A \in \Sigma$, $A \subseteq A_2$, then $\nu(A) = \mu_1(A) - \mu_2(A) = -\mu_2(A) \leq 0$ and, if $A \subseteq A_1$, then $\nu(A) = \mu_1(A) - \mu_2(A) = \mu_1(A) \geq 0$. Hence, A_1, A_2 is a Hahn decomposition of X for ν. Theorem 10.2 implies that $A_1 \triangle P = A_2 \triangle N$ is a null set for ν. Therefore, for every $A \in \Sigma$, we have $\mu_1(A) = \mu_1(A \cap A_1) + \mu_1(A \cap A_2) = \mu_1(A \cap A_1) = \mu_1(A \cap A_1) - \mu_2(A \cap A_1) = \nu(A \cap A_1) = \nu(A \cap A_1 \cap P) + \nu(A \cap A_1 \cap N) = \nu(A \cap A_1 \cap P)$, since $A \cap A_1 \cap N \subseteq A_1 \triangle P$. On the other hand, $\nu^+(A) = \nu(A \cap P) = \nu(A \cap A_1 \cap P) + \nu(A \cap A_2 \cap P) = \nu(A \cap A_1 \cap P)$, since $A \cap A_2 \cap P \subseteq A_2 \triangle N$. From the two equalities we get $\mu_1(A) = \nu^+(A)$ for every $A \in \Sigma$ and, thus, $\mu_1 = \nu^+$. We, similarly, prove $\mu_2 = \nu^-$.

Definition 10.6 *Let ν be a signed measure on (X, Σ). We say that the pair of mutually singular measures ν^+, ν^-, whose existence and uniqueness is proved in Theorem 10.3, constitute* ***the Jordan decomposition of*** *ν.*

ν^+ is called ***the positive variation of*** *ν and ν^- is called* ***the negative variation of*** *ν.*

The measure $|\nu| = \nu^+ + \nu^-$ is called ***the absolute variation of*** *ν, while the quantity $|\nu|(X)$ is called* ***the total variation of*** *ν.*

Observe that the total variation of ν is equal to

$$|\nu|(X) = \nu^+(X) + \nu^-(X) = \nu(P) - \nu(N),$$

where the sets P, N constitute a Hahn decomposition of X for ν. Hence, *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 positive and the negative variations are finite measures if and only if the signed measure takes only finite values.*

Proposition 10.3 *Suppose μ is a measure on (X, Σ), $f : X \to \overline{\mathbf{R}}$ is measurable and $\int_X f\,d\mu$ is defined. Then the sets $P = \{x \in X \,|\, f(x) \geq 0\}$ and $N = \{x \in X \,|\, f(x) < 0\}$ constitute a Hahn decomposition of X for the signed measure $f\mu$. Also,*

$$(f\mu)^+ = f^+\mu, \qquad (f\mu)^- = f^-\mu$$

constitute the Jordan decomposition of $f\mu$ and

$$|f\mu| = |f|\mu.$$

Proof: If $A \in \Sigma$ and $A \subseteq P$, then $(f\mu)(A) = \int_A f\,d\mu \geq 0$, while, if $A \subseteq N$, then $(f\mu)(A) = \int_A f\,d\mu \leq 0$. Therefore, P is a positive set for $f\mu$ and N is a negative 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 $A \in \Sigma$.

Therefore, $(f\mu)^+ = f^+\mu$ and $(f\mu)^- = f^-\mu$.

Now, $|f\mu| = (f\mu)^+ + (f\mu)^- = f^+\mu + f^-\mu = |f|\mu$.

It is easy to see that another Hahn decomposition of X for $f\mu$ consists of the sets $P' = \{x \in X \,|\, f(x) > 0\}$ and $N' = \{x \in X \,|\, f(x) \leq 0\}$.

Proposition 10.4 *Suppose μ is a measure on (X, Σ), $f : X \to \overline{\mathbf{R}}$ is measurable and $\int_X f\,d\mu$ is defined. Let $E \in \Sigma$.*
(i) E is a positive set for $f\mu$ if and only if $f \geq 0$ a.e. on E.
(ii) E is a negative set for $f\mu$ if and only if $f \leq 0$ 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 $f \geq 0$ a.e. on E and take any $A \in \Sigma$, $A \subseteq E$. Then $f \geq 0$ a.e. on A and, hence, $(f\mu)(A) = \int_A f\,d\mu \geq 0$. Thus, E is a positive set for $f\mu$. Suppose, conversely, that E is a positive set for $f\mu$. If $n \in \mathbf{N}$ and $A_n = \{x \in E \,|\, f(x) \leq -\frac{1}{n}\}$, then $0 \leq (f\mu)(A_n) = \int_{A_n} f\,d\mu \leq -\frac{1}{n}\mu(A_n)$. This implies that $\mu(A_n) = 0$ and, since $\{x \in E \,|\, f(x) < 0\} = \cup_{n=1}^{+\infty} A_n$, we conclude that $\mu(\{x \in E \,|\, f(x) < 0\}) = 0$. This means that $f \geq 0$ a.e. on E.

The proof of (ii) is identical to the proof of (i), and (iii) is a consequence of the results of (i) and (ii).

We recall that, for every $a \in \overline{\mathbf{R}}$, the positive part of a and the negative part of a are defined as

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

and, hence,

$$a = a^+ - a^-, \qquad |a| = a^+ + a^-.$$

It is trivial to prove that

$$(a+b)^+ \le a^+ + b^+, \qquad (a+b)^- \le a^- + b^-$$

for every $a, b \in \overline{\mathbf{R}}$ for which $a + b$ is defined.

Definition 10.7 *Let $A \in \Sigma$. If $A_1, \ldots, A_n \in \Sigma$ are pairwise disjoint and $A = \cup_{k=1}^n A_k$, then $\{A_1, \ldots, A_n\}$ is called* ***a (finite) measurable partition of*** *A.*

Theorem 10.4 *Let ν be a signed measure on (X, Σ) and let $|\nu|, \nu^+$ and ν^- be the absolute, the positive and the negative variation of ν, respectively. Then, for every $A \in \Sigma$,*

$$|\nu|(A) = \sup\Big\{ \sum_{k=1}^n |\nu(A_k)| \,|\, n \in \mathbf{N}, \{A_1, \ldots, A_n\} \text{ measurable partition of } A \Big\},$$

$$\nu^+(A) = \sup\Big\{ \sum_{k=1}^n \nu(A_k)^+ \,|\, n \in \mathbf{N}, \{A_1, \ldots, A_n\} \text{ measurable partition of } A \Big\},$$

$$\nu^-(A) = \sup\Big\{ \sum_{k=1}^n \nu(A_k)^- \,|\, n \in \mathbf{N}, \{A_1, \ldots, A_n\} \text{ measurable partition of } A \Big\}.$$

Proof: We let P, N be a Hahn decomposition of X for ν. For every pairwise disjoint $A_1, \ldots, A_n \in \Sigma$ with $\cup_{k=1}^n A_k = A$ we have that

$$\begin{aligned} \sum_{k=1}^n |\nu(A_k)| &= \sum_{k=1}^n |\nu^+(A_k) - \nu^-(A_k)| \le \sum_{k=1}^n \nu^+(A_k) + \sum_{k=1}^n \nu^-(A_k) \\ &= \nu^+(A) + \nu^-(A) = |\nu|(A). \end{aligned}$$

Therefore, the supremum of the left side is $\le |\nu|(A)$. On the other hand, $\{A \cap P, A \cap N\}$ is a particular measurable partition of A for which $|\nu(A \cap P)| + |\nu(A \cap N)| = \nu(A \cap P) - \nu(A \cap N) = \nu^+(A) + \nu^-(A) = |\nu|(A)$ and, hence, the supremum is equal to $|\nu|(A)$.

The proofs of the other two equalities are identical.

Lemma 10.1 *Let ν be a signed measure on (X, Σ) and $A \in \Sigma$. Then, A is a null set for ν if and only if it is a null set for both ν^+, ν^- if and only if it is a null set for $|\nu|$.*

Proof Since $|\nu| = \nu^+ + \nu^-$, the second equivalence is trivial.

Let A be null for $|\nu|$. For every $B \in \Sigma$, $B \subseteq A$, we have that $|\nu(B)| = |\nu^+(B) - \nu^-(B)| \leq \nu^+(B) + \nu^-(B) = |\nu|(B) = 0$. Hence, $\nu(B) = 0$ and A is null for ν.

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, hence, $\sum_{k=1}^n |\nu(A_k)| = 0$. Taking the supremum of the left side, Theorem 10.4 implies that $|\nu|(A) = 0$ and, thus, A is null for $|\nu|$.

Proposition 10.5 *Let ν_1 and ν_2 be two signed measures on (X, Σ). 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.*

Proof: The proof is a trivial consequence of Lemma 10.1.

Proposition 10.6 *Let ν, ν_1, ν_2 be signed measures on (X, Σ) and $\kappa \in \mathbf{R}$. If $\nu_1 + \nu_2$ is defined, we have*

$$|\nu_1 + \nu_2| \leq |\nu_1| + |\nu_2|, \qquad |\kappa\nu| = |\kappa||\nu|.$$

Proof: We take an arbitrary measurable partition $\{A_1, \ldots, A_n\}$ of $A \in \Sigma$ and we have $\sum_{k=1}^n |(\nu_1+\nu_2)(A_k)| \leq \sum_{k=1}^n |\nu_1(A_k)| + \sum_{k=1}^n |\nu_2(A_k)| \leq |\nu_1|(A) + |\nu_2|(A)$. 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, $\sum_{k=1}^n |(\kappa\nu)(A_k)| = |\kappa| \sum_{k=1}^n |\nu(A_k)|$. This equality implies $\sum_{k=1}^n |(\kappa\nu)(A_k)| \leq |\kappa||\nu|(A)$ and, taking supremum of the left side, $|\kappa\nu|(A) \leq |\kappa||\nu|(A)$. The same equality, also, implies $|\kappa\nu|(A) \geq |\kappa| \sum_{k=1}^n |\nu(A_k)|$ and, taking supremum of the right side, $|\kappa\nu|(A) \geq |\kappa||\nu|(A)$.

10.3 The Hahn and Jordan decompositions, II.

In this section we shall describe another method of constructing the Hahn and Jordan decompositions of a signed measure. In the previous section we derived the Hahn decomposition first and, based on it, we derived the Jordan decomposition. We shall, now, follow the reverse procedure.

Let (X, Σ) be a measurable space.

Definition 10.8 *Let ν be a signed measure on (X, Σ). For every $A \in \Sigma$ we define*

$$|\nu|(A) = \sup\Big\{\sum_{k=1}^n |\nu(A_k)| \,|\, n \in \mathbf{N}, \{A_1, \ldots, A_n\} \text{ measurable partition of } A\Big\},$$

$$\nu^+(A) = \sup\Big\{\sum_{k=1}^n \nu(A_k)^+ \,|\, n \in \mathbf{N}, \{A_1, \ldots, A_n\} \text{ measurable partition of } A\Big\},$$

$$\nu^-(A) = \sup\Big\{\sum_{k=1}^n \nu(A_k)^- \,|\, n \in \mathbf{N}, \{A_1, \ldots, A_n\} \text{ measurable partition of } A\Big\}.$$

Lemma 10.2 *Let ν be a signed measure on (X, Σ). Then,*

$$\nu^+(A) + \nu^-(A) = |\nu|(A)$$

and

$$\nu^+(A) = \sup\{\nu(B) \,|\, B \in \Sigma, B \subseteq A\}, \quad \nu^-(A) = -\inf\{\nu(B) \,|\, B \in \Sigma, B \subseteq A\}$$

for every $A \in \Sigma$.

Proof: (a) Take any $A \in \Sigma$ and any measurable partition $\{A_1, \ldots, A_n\}$ of A. Then,

$$\sum_{k=1}^n |\nu(A_k)| = \sum_{k=1}^n \nu(A_k)^+ + \sum_{k=1}^n \nu(A_k)^- \leq \nu^+(A) + \nu^-(A).$$

Taking the supremum of the left side, we get $|\nu|(A) \leq \nu^+(A) + \nu^-(A)$.

Now take arbitrary partitions $\{A_1, \ldots, A_n\}$ and $\{A'_1, \ldots, A'_{n'}\}$ of A. Then

$$\sum_{k=1}^n \nu(A_k)^+ \leq \sum_{k=1}^n \Big(\sum_{k'=1}^{n'} \nu(A_k \cap A'_{k'})^+ \Big),$$

$$\sum_{k'=1}^{n'} \nu(A'_{k'})^- \leq \sum_{k'=1}^{n'} \Big(\sum_{k=1}^{n} \nu(A_k \cap A'_{k'})^- \Big)$$

and, adding,

$$\sum_{k=1}^n \nu(A_k)^+ + \sum_{k'=1}^{n'} \nu(A'_{k'})^- \leq \sum_{1 \leq k \leq n, 1 \leq k' \leq n'} |\nu(A_k \cap A'_{k'})|.$$

Since $\{A_k \cap A'_{k'} \,|\, 1 \leq k \leq n, 1 \leq k' \leq n'\}$ is a measurable partition of A, we get

$$\sum_{k=1}^n \nu(A_k)^+ + \sum_{k'=1}^{n'} \nu(A'_{k'})^- \leq |\nu|(A).$$

Finally, taking the supremum of the left side, we find $\nu^+(A) + \nu^-(A) \leq |\nu|(A)$.
(b) If $B \in \Sigma$ and $B \subseteq A$, then $\{B, A \setminus B\}$ is a measurable partition of A and, hence, $\nu(B) \leq \nu(B)^+ \leq \nu(B)^+ + \nu(A \setminus B)^+ \leq \nu^+(A)$. This proves that $\sup\{\nu(B) \,|\, B \in \Sigma, B \subseteq A\} \leq \nu^+(A)$.

Let $\{A_1, \ldots, A_n\}$ be any measurable partition of A. If $A_{i_1}, \ldots, A_{i_m}$ are exactly the sets with non-negative ν-measure and if $B_0 = \cup_{l=1}^m A_{i_l} \subseteq A$, then $\sum_{k=1}^n \nu(A_k)^+ = \sum_{l=1}^m \nu(A_{i_l}) = \nu(B_0)$. This implies that $\sum_{k=1}^n \nu(A_k)^+ \leq \sup\{\nu(B) \,|\, B \in \Sigma, B \subseteq A\}$ and, hence, $\nu^+(A) \leq \sup\{\nu(B) \,|\, B \in \Sigma, B \subseteq A\}$.

We conclude that $\nu^+(A) = \sup\{\nu(B) \,|\, B \in \Sigma, B \subseteq A\}$ and a similar argument proves the last equality.

Theorem 10.5 *Let ν be a signed measure on (X, Σ). Then, the functions $|\nu|, \nu^+, \nu^- : \Sigma \to [0, +\infty]$, which were defined in Definition 10.8, are measures on (X, Σ).*

At least one of ν^+, ν^- is finite and

$$\nu^+ - \nu^- = \nu, \qquad \nu^+ + \nu^- = |\nu|, \qquad \nu^+ \perp \nu^-.$$

Proof: (a) We shall first prove that $|\nu|$ is a measure.

It is obvious that $|\nu|(\emptyset) = 0$ and take arbitrary pairwise disjoint $A^1, A^2, \ldots \in \Sigma$ and $A = \cup_{j=1}^{+\infty} A^j$.

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, $\sum_{k=1}^n |\nu(A_k)| = \sum_{k=1}^n |\sum_{j=1}^{+\infty} \nu(A_k \cap A^j)| \le \sum_{k=1}^n \big(\sum_{j=1}^{+\infty} |\nu(A_k \cap A^j)|\big) = \sum_{j=1}^{+\infty} \big(\sum_{k=1}^n |\nu(A_k \cap A^j)|\big) \le \sum_{j=1}^{+\infty} |\nu|(A^j)$ and, taking the supremum of the left side, $|\nu|(A) \le \sum_{j=1}^{+\infty} |\nu|(A^j)$.

Fix arbitrary $N \in \mathbf{N}$ and for every $j = 1, \ldots, N$ take any measurable partition $\{A_1^j, \ldots, A_{n_j}^j\}$ of A^j. Then $\{A_1^1, \ldots, A_{n_1}^1, \ldots, A_1^N, \ldots, A_{n_N}^N, \cup_{j=N+1}^{+\infty} A^j\}$ is a measurable partition of A and, hence, $|\nu|(A) \ge \sum_{j=1}^N \big(\sum_{k=1}^{n_j} |\nu(A_k^j)|\big) + |\nu(\cup_{j=N+1}^{+\infty} A^j)| \ge \sum_{j=1}^N \big(\sum_{k=1}^{n_j} |\nu(A_k^j)|\big)$. Taking the supremum of the right side, we get $|\nu|(A) \ge \sum_{j=1}^N |\nu|(A^j)$ and, taking the limit as $N \to +\infty$, we find $|\nu|(A) \ge \sum_{j=1}^{+\infty} |\nu|(A^j)$.

Hence, $|\nu|(A) = \sum_{j=1}^{+\infty} |\nu|(A^j)$.

The proofs that ν^+ and ν^- are measures are identical to the proof we have just seen.

(b) In case $\nu(A) < +\infty$ for every $A \in \Sigma$, we shall prove that $\nu^+(X) < +\infty$.

We claim that *for every $A \in \Sigma$ with $\nu^+(A) = +\infty$ and every $M > 0$, there exists $B \in \Sigma$, $B \subseteq A$, so that $\nu^+(B) = +\infty$ and $\nu(B) \ge M$.*

Suppose that the claim is not true. Then, there is $A \in \Sigma$ with $\nu^+(A) = +\infty$ and an $M > 0$ so that, if $B \in \Sigma$, $B \subseteq A$, has $\nu(B) \ge M$, then $\nu^+(B) < +\infty$. Now, by Lemma 10.2, there is $B_1 \in \Sigma$, $B_1 \subseteq A$ with $\nu(B_1) \ge M$ and, hence, $\nu^+(B_1) < +\infty$. Suppose that we have constructed pairwise disjoint $B_1, \ldots, B_m \in \Sigma$ subsets of A with $\nu(B_j) \ge M$ and $\nu^+(B_j) < +\infty$ for every $j = 1, \ldots, m$. Since ν^+ is a measure, we have $\sum_{j=1}^m \nu^+(B_j) + \nu^+(A \setminus \cup_{j=1}^m B_j) = \nu^+(A) = +\infty$ and, thus, $\nu^+(A \setminus \cup_{j=1}^m B_j) = +\infty$. Lemma 10.2 implies that there is $B_{m+1} \in \Sigma$, $B_{m+1} \subseteq A \setminus \cup_{j=1}^m B_j$ with $\nu(B_{m+1}) \ge M$ and, hence, $\nu^+(B_{m+1}) < +\infty$.

We, thus, inductively construct a sequence (B_m) in Σ of pairwise disjoint subsets of A with $\nu(B_m) \ge M$. But, then, $\nu(\cup_{m=1}^{+\infty} B_m) = \sum_{m=1}^{+\infty} \nu(B_m) = +\infty$ and we arrive at a contradiction.

Using the claimed result and assuming that $\nu^+(X) = +\infty$, we find $B^1 \in \Sigma$ with $\nu(B^1) \ge 1$ and $\nu^+(B^1) = +\infty$. We, similarly, find $B^2 \in \Sigma$, $B^2 \subseteq B^1$, with $\nu(B^2) \ge 2$ and $\nu^+(B^2) = +\infty$. Continuing inductively, a decreasing sequence (B^m) is constructed in Σ with $\nu(B^m) \ge m$ for every m. Then, $\nu(\cap_{l=1}^{+\infty} B^l) = \lim_{m \to +\infty} \nu(B^m) = +\infty$ and we arrive at a contradiction.

Therefore, $\nu^+(X) < +\infty$.

If $-\infty < \nu(A)$ for every $A \in \Sigma$, we prove in the same way that $\nu^-(X) < +\infty$.
(c) Suppose that $\nu(A) < +\infty$ for every $A \in \Sigma$ and, hence, $\nu^+(X) < +\infty$, by the result of (b).

We take any $A \in \Sigma$ and any $B \in \Sigma$, $B \subseteq A$. Then $\nu(A \setminus B) \leq \nu^+(A)$ and, hence, $\nu(A) \leq \nu^+(A) + \nu(B)$. Taking the infimum over B and using the $\nu^+(A) < +\infty$, we get $\nu(A) \leq \nu^+(A) - \nu^-(A)$.

To prove the opposite inequality, we first assume $\nu^-(A) < +\infty$. For every $B \in \Sigma$, $B \subseteq A$, we have $-\nu^-(A) \leq \nu(A \setminus B)$ and, hence, $\nu(B) - \nu^-(A) \leq \nu(A)$. Taking the supremum over B we find $\nu^+(A) - \nu^-(A) \leq \nu(A)$. If $\nu^-(A) = +\infty$, then, since $\nu^+(A) < +\infty$, the $\nu^+(A) - \nu^-(A) \leq \nu(A)$ is clearly true.

We conclude that $\nu(A) = \nu^+(A) - \nu^-(A)$ for every $A \in \Sigma$ and the same can be proved if we assume that $-\infty < \nu(A)$ for every $A \in \Sigma$.

Therefore, $\nu = \nu^+ - \nu^-$.
(d) The equality $|\nu| = \nu^+ + \nu^-$ is contained in Lemma 10.2.
(e) We, again, assume $\nu(A) < +\infty$ for every $A \in \Sigma$ and, hence, $\nu^+(X) < +\infty$.

Using Lemma 10.2, we take a sequence (B_n) in Σ so that $\nu(B_n) \to \nu^+(X)$ as $n \to +\infty$. Since $\nu(B_n) \leq \nu^+(B_n) \leq \nu^+(X)$, we have that $\nu^+(B_n) \to \nu^+(X)$ as $n \to +\infty$. From $\nu(B_n) = \nu^+(B_n) - \nu^-(B_n)$, we get $\nu^-(B_n) \to 0$ as $n \to +\infty$.

We find a strictly increasing (n_k) so that $\nu^-(B_{n_k}) < \frac{1}{2^k}$ for all k. If we set $F_k = \cup_{l=k}^{+\infty} B_{n_l}$, then $\nu^-(F_k) \leq \sum_{l=k}^{+\infty} \nu^-(B_{n_l}) < \frac{1}{2^{k-1}}$ for every k and (F_k) is decreasing. Therefore, the set $F = \cap_{k=1}^{+\infty} F_k$ has $\nu^-(F) = 0$. We, also, have that $\nu^+(B_{n_k}) \leq \nu^+(F_k) \leq \nu^+(X)$ and, hence, $\nu^+(F_k) \to \nu^+(X)$ as $k \to +\infty$. Therefore, $\nu^+(F) = \nu^+(X)$.

We have constructed a set $F \in \Sigma$ so that $\nu^-(F) = 0$ and $\nu^+(F) = \nu^+(X)$. Since $\nu^+(X) < +\infty$, we find $\nu^+(X \setminus F) = 0$ and we conclude that $\nu^+ \perp \nu^-$.

The decomposition $\nu = \nu^+ - \nu^-$ of the signed measure ν on (X, Σ), which is given in Theorem 10.5, is the same as the Jordan decomposition of ν, which was defined in the previous section 10.2. This is justified both by the uniqueness of the Jordan decomposition of a signed measure and by the result of Theorem 10.4. Using, now, the Jordan decomposition, we shall produce the Hahn decomposition of a signed measure.

Theorem 10.6 *Let ν be a signed measure on (X, Σ) and ν^+, ν^- be the measures of Definition 10.8. Then, there exist $P, N \in \Sigma$ so that $P \cup N = X$, $P \cap N = \emptyset$, P is a positive set for ν, N is a negative set for ν and $\nu^+(N) = 0, \nu^-(P) = 0$.*

Proof: Theorem 10.5 implies that $\nu^+ \perp \nu^-$ and, hence, there are $P, N \in \Sigma$ so that $P \cup N = X$, $P \cap N = \emptyset$ and $\nu^+(N) = 0 = \nu^-(P)$.

If $A \in \Sigma$, $A \subseteq P$, then $\nu(A) = \nu^+(A) - \nu^-(A) = \nu^+(A) \geq 0$. Similarly, if $A \in \Sigma$, $A \subseteq N$, then $\nu(A) = \nu^+(A) - \nu^-(A) = -\nu^-(A) \leq 0$. Hence, P is a positive set for ν and N is a negative set for ν.

10.4 Complex measures.

Let (X, Σ) be a measurable space.

Definition 10.9 *A function $\nu : \Sigma \to \mathbf{C}$ is called* ***a complex measure on*** *(X, Σ) if*
(i) $\nu(\emptyset) = 0$,
(ii) $\nu(\cup_{j=1}^{+\infty} A_j) = \sum_{j=1}^{+\infty} \nu(A_j)$ for every pairwise disjoint $A_1, A_2, \ldots \in \Sigma$.

It is trivial to prove, taking real and imaginary parts, that the functions $\Re(\nu), \Im(\nu) : \Sigma \to \mathbf{R}$, which are defined by $\Re(\nu)(A) = \Re(\nu(A))$ and $\Im(\nu)(A) = \Im(\nu(A))$ for every $A \in \Sigma$, are real measures on (X, Σ) and, hence, they are bounded. That is, there is an $M < +\infty$ so that $|\Re(\nu)(A)|, |\Im(\nu)(A)| \leq M$ for every $A \in \Sigma$. This implies that $|\nu(A)| \leq 2M$ for every $A \in \Sigma$ and we have proved the

Proposition 10.7 *Let ν be a complex measure on (X, Σ). Then ν is bounded, i.e. there is an $M < +\infty$ so that $|\nu(A)| \leq M$ for every $A \in \Sigma$.*

If ν_1 and ν_2 are complex measures on (X, Σ) and $\kappa_1, \kappa_2 \in \mathbf{C}$, then $\kappa_1\nu_1 + \kappa_2\nu_2$, defined by $(\kappa_1\nu_1 + \kappa_2\nu_2)(A) = \kappa_1\nu_1(A) + \kappa_2\nu_2(A)$ for all $A \in \Sigma$, is a complex measure on (X, Σ).

The following are straightforward extensions of Definitions 10.3 and 10.5.

Definition 10.10 *Let ν be a complex measure on (X, Σ) and $A \in \Sigma$. We say that A* ***is a null set for*** *ν if $\nu(B) = 0$ for every $B \in \Sigma$, $B \subseteq A$.*

Definition 10.11 *Let ν_1 and ν_2 be complex or signed measures on (X, Σ). We say that ν_1* ***and*** *ν_2* ***are mutually singular****, and denote this by $\nu_1 \perp \nu_2$, if there are $A_1, A_2 \in \Sigma$ so that A_2 is null for ν_1, A_1 is null for ν_2 and $A_1 \cup A_2 = X$, $A_1 \cap A_2 = \emptyset$.*

Proposition 10.8 *Let ν be a complex measure on (X, Σ). If for every $A \in \Sigma$ we define*

$$|\nu|(A) = \sup\Big\{ \sum_{k=1}^{n} |\nu(A_k)| \,|\, n \in \mathbf{N}, \{A_1, \ldots, A_n\} \text{ measurable partition of } A \Big\},$$

then the function $|\nu| : \Sigma \to [0, +\infty]$ is a finite measure on (X, Σ).

Proof: The proof that $|\nu|$ is a measure is exactly the same as in part (a) of the proof of Theorem 10.5.

We take an arbitrary measurable partition $\{A_1, \ldots, A_n\}$ of X and have $\sum_{k=1}^{n} |\nu(A_k)| \leq \sum_{k=1}^{n} |\Re(\nu)(A_k)| + \sum_{k=1}^{n} |\Im(\nu)(A_k)| \leq |\Re(\nu)|(X) + |\Im(\nu)|(X)$. Taking the supremum of the left side, $|\nu|(X) \leq |\Re(\nu)|(X) + |\Im(\nu)|(X) < +\infty$, because the signed measures $\Re(\nu)$ and $\Im(\nu)$ have finite values.

Definition 10.12 *Let ν be a complex measure on (X, Σ). The measure $|\nu|$ defined in Proposition 10.8 is called* ***the absolute variation of*** *ν and the number $|\nu|(X)$ is called* ***the total variation of*** *ν.*

Proposition 10.9 *Let ν, ν_1, ν_2 be complex measures on (X, Σ) and $\kappa \in \mathbf{C}$. Then*
(i) $|\nu_1 + \nu_2| \leq |\nu_1| + |\nu_2|$ *and* $|\kappa\nu| = |\kappa||\nu|$
(ii) $|\Re(\nu)|, |\Im(\nu)| \leq |\nu| \leq |\Re(\nu)| + |\Im(\nu)|$.

Proof: (i) The proof is identical to the proof of Proposition 10.6.
(ii) In the same manner, if $\{A_1, \ldots, A_n\}$ is any measurable partition of $A \in \Sigma$, we have $\sum_{k=1}^n |\Re(\nu)(A_k)| \leq \sum_{k=1}^n |\nu(A_k)| \leq |\nu|(A)$ and also $\sum_{k=1}^n |\Im(\nu)(A_k)| \leq \sum_{k=1}^n |\nu(A_k)| \leq |\nu|(A)$. Taking supremum of the left sides of these two inequalities, we find $|\Re(\nu)|(A), |\Im(\nu)|(A) \leq |\nu|(A)$.

The last inequality is a consequence of the result of (i).

Lemma 10.3 *Let ν be a complex measure on (X, Σ) and $A \in \Sigma$. Then A is null for ν if and only if A is null for both $\Re(\nu)$ and $\Im(\nu)$ if and only if A is null for $|\nu|$.*

Proof: The first equivalence is trivial. The proof that A is null for ν if and only if A is null for $|\nu|$ is a repetition of the proof of the same result for a signed measure ν. See Lemma 10.1.

Proposition 10.10 *Let ν_1 and ν_2 be complex or signed measures on (X, Σ). Then, $\nu_1 \perp \nu_2$ if and only if each of $\Re(\nu_1), \Im(\nu_1)$ and each of $\Re(\nu_2), \Im(\nu_2)$ are mutually singular if and only if $|\nu_1| \perp |\nu_2|$.*

Proof: Trivial after Lemma 10.3.

Example
We take a measure μ on (X, Σ) and a measurable function $f : X \to \overline{\mathbf{C}}$ which is integrable over X. Then, $\int_A f \, d\mu$ is, by Lemma 7.10, a complex number for every $A \in \Sigma$, and Theorem 7.13 implies that the function $\lambda : \Sigma \to \mathbf{C}$, which is defined by

$$\lambda(A) = \int_A f \, d\mu$$

for every $A \in \Sigma$, is a complex measure on (X, Σ).

Definition 10.13 *Let μ be a measure on (X, Σ) and $f : X \to \overline{\mathbf{C}}$ be integrable. The complex measure λ defined in the previous paragraph is called* ***the indefinite integral of f with respect to μ*** *and it is denoted by $f\mu$. Thus,*

$$(f\mu)(A) = \int_A f \, d\mu, \qquad A \in \Sigma.$$

The next result is the analogue of Proposition 10.3.

Proposition 10.11 *Let μ be a measure on (X, Σ) and $f : X \to \overline{\mathbf{C}}$ be integrable with respect to μ. Then*

$$|f\mu|(A) = \int_A |f| \, d\mu$$

for every $A \in \Sigma$. Hence,

$$|f\mu| = |f|\mu.$$

Proof: If $\{A_1, \dots, A_n\}$ is an arbitrary measurable partition of $A \in \Sigma$, then $\sum_{k=1}^n |(f\mu)(A_k)| = \sum_{k=1}^n |\int_{A_k} f\,d\mu| \le \sum_{k=1}^n \int_{A_k} |f|\,d\mu = \int_A |f|\,d\mu$. Therefore, taking the supremum of the left side, $|f\mu|(A) \le \int_A |f|\,d\mu$.

Since f is integrable, it is finite a.e. on X. If $N = \{x \in X \,|\, f(x) \ne \infty\}$, then $\mu(N^c) = 0$ and Theorem 6.1 implies that there is a sequence (ϕ_m) of measurable simple functions with $\phi_m \to \overline{sign(f)}$ on N and $|\phi_m| \uparrow |\overline{sign(f)}| \le 1$ on N. Defining each ϕ_m as 0 on N^c, we have that all these properties hold a.e. on X.

If $\phi_m = \sum_{k=1}^{n_m} \kappa_k^m \chi_{E_k^m}$ is the standard representation of ϕ_m, then $|\kappa_k^m| \le 1$ for all $k = 1, \dots, n_m$ and, hence, $|\int_A f\phi_m\,d\mu| = |\sum_{k=1}^{n_m} \kappa_k^m \int_{A\cap E_k^m} f\,d\mu| \le \sum_{k=1}^{n_m} |(f\mu)(A \cap E_k^m)| \le |f\mu|(A)$, where the last inequality is true because $\{A \cap E_1^m, \dots, A \cap E_{n_m}^m\}$ is a measurable partition of A. By the Dominated Convergence Theorem, we get that $\int_A |f|\,d\mu = \int_A f\,\overline{sign(f)}\,d\mu \le |f\mu|(A)$.

We conclude that $|f\mu|(A) = \int_A |f|\,d\mu$ for every $A \in \Sigma$.

10.5 Integration.

Let (X, Σ) be a measurable space.

The next definition covers only the case when both f and ν have their values in $\overline{\mathbf{R}}$.

Definition 10.14 *Let ν be a signed measure on (X, Σ). If $f : X \to \overline{\mathbf{R}}$ is 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*

$$\int_X f\,d\nu = \int_X f\,d\nu^+ - \int_X f\,d\nu^-.$$

We say that f ***is integrable over*** X ***(with respect to*** ν***)*** *if the $\int_X f\,d\nu$ is finite.*

Proposition 10.12 *Let ν be a signed measure on (X, Σ) and $f : X \to \overline{\mathbf{R}}$ 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 $|\nu|$.

Lemma 10.4 *Let μ_1, μ_2 be two measures on (X, Σ) with $\mu_1 \le \mu_2$. Then $\int_X f\,d\mu_1 \le \int_X f\,d\mu_2$ for every measurable $f : X \to [0, +\infty]$.*

Proof: If $\phi = \sum_{j=1}^m \kappa_j \chi_{E_j}$ is a measurable non-negative simple function with its standard representation, then $\int_X \phi\,d\mu_1 = \sum_{j=1}^m \kappa_j \mu_1(E_j) \le \sum_{j=1}^m \kappa_j \mu_2(E_j) = \int_X \phi\,d\mu_2$. For the general f we take a sequence (ϕ_n) of measurable non-negative simple functions with $\phi_n \uparrow f$ on X. We write the inequality for each ϕ_n and the Monotone Convergence Theorem implies $\int_X f\,d\mu_1 \leq \int_X f\,d\mu_2$.

Now, suppose that ν is a signed measure or a complex measure on (X, Σ) and the function $f : X \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$ is measurable. If $\int_X |f|\,d|\nu| < +\infty$, then f is finite $|\nu|$-a.e. on X and the $|\nu|$-a.e. defined functions $\Re(f)$ and $\Im(f)$ satisfy $\int_X |\Re(f)|\,d|\nu| < +\infty$ and $\int_X |\Im(f)|\,d|\nu| < +\infty$. Since, by Proposition 10.9, $|\Re(\nu)| \leq |\nu|$ and $|\Im(\nu)| \leq |\nu|$, Lemma 10.4 implies that all integrals $\int_X |\Re(f)|\,d|\Re(\nu)|$, $\int_X |\Re(f)|\,d|\Im(\nu)|$, $\int_X |\Im(f)|\,d|\Re(\nu)|$ and $\int_X |\Im(f)|\,d|\Im(\nu)|$ are finite. Proposition 10.12 implies that all integrals $\int_X \Re(f)\,d\Re(\nu)$, $\int_X \Re(f)\,d\Im(\nu)$, $\int_X \Im(f)\,d\Re(\nu)$ and $\int_X \Im(f)\,d\Im(\nu)$ are defined and they all are real numbers.

Therefore, the following definition is valid.

Definition 10.15 *Let ν be a signed measure or a complex measure on (X, Σ) and $f : X \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$ be measurable. We say that f* ***is integrable over*** *X* ***(with respect to*** *ν***)*** *if f is integrable with respect to $|\nu|$. In such a case we say that* ***the integral*** *$\int_X f\,d\nu$* ***of*** *f* ***over*** *X* ***(with respect to*** *ν***) is defined*** *and it is given by*

$$\int_X f\,d\nu = \int_X \Re(f)\,d\Re(\nu) - \int_X \Im(f)\,d\Im(\nu) + i\int_X \Re(f)\,d\Im(\nu) + i\int_X \Im(f)\,d\Re(\nu).$$

Of course, when $f : X \to \overline{\mathbf{C}}$ and ν is signed, we have

$$\int_X f\,d\nu = \int_X \Re(f)\,d\nu + i\int_X \Im(f)\,d\nu,$$

and when $f : X \to \overline{\mathbf{R}}$ and ν is complex, we have

$$\int_X f\,d\nu = \int_X f\,d\Re(\nu) + i\int_X f\,d\Im(\nu),$$

all under the assumption that $\int_X |f|\,d|\nu| < +\infty$.

We shall not bother 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 positive and negative variations or to real and imaginary parts.*

For completeness, we shall only see a few most necessary properties, like the linearity properties and the appropriate version of the Dominated Convergence Theorem.

Proposition 10.13 *Let ν, ν_1, ν_2 be signed or complex measures on (X, Σ) and $f, f_1, f_2 : X \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$ be all integrable with respect to these measures. For every $\kappa_1, \kappa_2 \in \mathbf{C}$,*

$$\int_X (\kappa_1 f_1 + \kappa_2 f_2)\,d\nu = \kappa_1 \int_X f_1\,d\nu + \kappa_2 \int_X f_2\,d\nu\,,$$

$$\int_X f\,d(\kappa_1\nu_1 + \kappa_2\nu_2) = \kappa_1 \int_X f\,d\nu_1 + \kappa_2 \int_X f\,d\nu_2\,.$$

Proof: The proof is straightforward when we reduce everything to real functions and signed measures.

Theorem 10.7 ***(Dominated Convergence Theorem)*** *Let ν be a signed or complex measure on (X,Σ), $f, f_n : X \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$ and $g : X \to [0,+\infty]$ be 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 $\Re(\nu)$ and $\Im(\nu)$, if ν is complex. Moreover, by Lemma 10.4, $\int_X g\, d\nu^+, \int_X g\, d\nu^- < +\infty$, if ν is signed, and $\int_X g\, d|\Re(\nu)|, \int_X g\, d|\Im(\nu)| < +\infty$, if ν is complex.

Therefore, the proof reduces to the usual Dominated Convergence Theorem for measures.

Theorem 10.8 *Let ν be a signed or complex measure on (X,Σ) and $f : X \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$ 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 ν is a signed measure, $|\int_X f\, d\nu| = |\int_X f\, d\nu^+ - \int_X f\, d\nu^-| \leq |\int_X f\, d\nu^+| + |\int_X f\, d\nu^-| \leq \int_X |f|\, d\nu^+ + \int_X |f|\, d\nu^- = \int_X |f|\, d|\nu|$.

If ν is complex, we shall see a proof which is valid in all cases anyway.

Let $\phi : X \to \mathbf{C}$ be a measurable simple function with its 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 $|\int_X \phi\, d\nu| = |\sum_{k=1}^n \kappa_k \nu(E_k)| \leq \sum_{k=1}^n |\kappa_k||\nu(E_k)| \leq \sum_{k=1}^n |\kappa_k||\nu|(E_k) = \int_X |\phi|\, d|\nu|$.

Consider a sequence (ϕ_n) of measurable simple functions so that $\phi_n \to f$ on X and $|\phi_n| \uparrow |f|$ on X. The Monotone Convergence Theorem implies $\int_X |\phi_n|\, d|\nu| \to \int_X |f|\, d|\nu|$ and Theorem 10.7, together with $\int_X |f|\, d|\nu| < +\infty$, implies that $\int_X \phi_n\, d\nu \to \int_X f\, d\nu$. Taking the limit in $|\int_X \phi_n\, d\nu| \leq \int_X |\phi_n|\, d|\nu|$ we prove the $\left|\int_X f\, d\nu\right| \leq \int_X |f|\, d|\nu|$.

A companion to the previous theorem is

Theorem 10.9 *Let ν be a signed or complex measure on (X,Σ). Then*

$$|\nu|(A) = \sup\Big\{\Big|\int_A f\, d\nu\Big| \,\Big|\, f \text{ is measurable, } |f| \leq 1 \text{ on } A\Big\},$$

for every $A \in \Sigma$, where the functions f have real values, if ν is signed, and complex values, if ν is complex.

Proof: If f is measurable and $|f| \leq 1$ on A, then $|f\chi_A| \leq \chi_A$ on X and Theorem 10.8 implies $|\int_A f\,d\nu| = |\int_X f\chi_A\,d\nu| \leq \int_X |f\chi_A|\,d|\nu| \leq \int_X \chi_A\,d|\nu| = |\nu|(A)$. Therefore the supremum of the left side is $\leq |\nu|(A)$.

If ν is signed, we take a Hahn decomposition of X for ν. There are $P, N \in \Sigma$ so that $P \cup N = X$, $P \cap N = \emptyset$, P is a positive set and N a negative set for ν. We consider the function f with values $f = 1$ on P and $f = -1$ on N. Then $|\int_A f\,d\nu| = |\nu(A \cap P) - \nu(A \cap N)| = \nu(A \cap P) - \nu(A \cap N) = \nu^+(A) + \nu^-(A) = |\nu|(A)$. Therefore, the supremum is equal to $|\nu|(A)$.

If ν is complex, we find a measurable partition $\{A_1, \ldots, A_n\}$ of A so that $|\nu|(A) - \epsilon \leq \sum_{k=1}^n |\nu(A_k)|$. We, then, define the function $f = \sum_{k=1}^n \kappa_k \chi_{A_k}$, where $\kappa_k = \overline{sign(\nu(A_k))}$ for all k. Then, $|f| \leq 1$ on A and $|\int_A f\,d\nu| = |\sum_{k=1}^n \kappa_k \nu(A_k)| = \sum_{k=1}^n |\nu(A_k)| \geq |\nu|(A) - \epsilon$. This proves that the supremum is equal to $|\nu|(A)$.

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 10.10 *Let μ be a measure on (X, Σ) and $f : X \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$ be measurable so that $\int_X f\,d\mu$ is defined. Consider the signed measure or complex measure $f\mu$, the indefinite integral of f with respect to μ.*

A measurable function $g : X \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$ 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, without any restriction, if $f, g : X \to [0, +\infty]$ are measurable.

Proof: We consider first the case when $g, f : X \to [0, +\infty]$.

If $g = \chi_A$ for some $A \in \Sigma$, then $\int_X \chi_A\,d(f\mu) = (f\mu)(A) = \int_A f\,d\mu = \int_X \chi_A f\,d\mu$. Hence, the equality $\int_X g\,d(f\mu) = \int_X gf\,d\mu$ is true for characteristic functions. This extends, by linearity, to measurable non-negative simple functions $g = \phi$ and, by the Monotone Convergence Theorem, to the general 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.

10.6 Lebesgue decomposition, Radon-Nikodym derivative.

Let (X, Σ) be a measurable space.

Definition 10.16 *Let μ be a measure and ν a signed or complex measure on (X, Σ). We say that ν* ***is absolutely continuous with respect to*** *μ when $\nu(A) = 0$ for every $A \in \Sigma$ with $\mu(A) = 0$ and we denote by*

$$\nu \ll \mu.$$

Example
Let $f : X \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$ be measurable so that the $\int_X f\,d\mu$ is defined (recall that, in the case of $\overline{\mathbf{C}}$, this means that f is integrable). Then *the indefinite integral $f\mu$ is absolutely continuous with respect to μ.*

This is obvious: if $A \in \Sigma$ has $\mu(A) = 0$, then $(f\mu)(A) = \int_A f\,d\mu = 0$.

Proposition 10.14 *Let μ be a measure and ν, ν_1, ν_2 be signed or complex measures on (X, Σ).*
(i) If ν is complex, then $\nu \ll \mu$ if and only if $\Re(\nu) \ll \mu$ and $\Im(\nu) \ll \mu$ if and only if $|\nu| \ll \mu$.
(ii) If ν is signed, then $\nu \ll \mu$ if and only if $\nu^+ \ll \mu$ and $\nu^- \ll \mu$ if and only if $|\nu| \ll \mu$.
(iii) If $\nu \ll \mu$ and $\nu \perp \mu$, then $\nu = 0$.
(iv) If $\nu_1, \nu_2 \ll \mu$ and $\nu_1 + \nu_2$ is defined, then $\nu_1 + \nu_2 \ll \mu$.

Proof: (i) Since $\nu(A) = 0$ is equivalent to $\Re(\nu)(A) = \Im(\nu)(A) = 0$, the first equivalence is obvious.

Let $\nu \ll \mu$ and take any $A \in \Sigma$ 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, thus, $\sum_{k=1}^n |\nu(A_k)| = 0$. Taking the supremum of the left side we get $|\nu|(A) = 0$. Hence, $|\nu| \ll \mu$.

If $|\nu| \ll \mu$ and we take any $A \in \Sigma$ with $\mu(A) = 0$, then $|\nu(A)| \leq |\nu|(A) = 0$. Therefore, $\nu(A) = 0$ and $\nu \ll \mu$.
(ii) The argument of part (i) applies without change to prove that $\nu \ll \mu$ if and only if $|\nu| \ll \mu$. Since $|\nu| = \nu^+ + \nu^-$, it is obvious that $\nu^+ \ll \mu$ and $\nu^- \ll \mu$ if and only if $|\nu| \ll \mu$.
(iii) Take sets $M, N \in \Sigma$ 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, hence, $\nu = 0$.
(iv) If $A \in \Sigma$ has $\mu(A) = 0$, then $\nu_1(A) = \nu_2(A) = 0$ and, hence, $(\nu_1 + \nu_2)(A) = 0$.

The next result justifies the term *absolutely continuous* at least in the special case of a finite ν.

Proposition 10.15 *Let μ be a measure and ν a real or a complex measure on (X, Σ). Then $\nu \ll \mu$ if and only if for every $\epsilon > 0$ there is a $\delta > 0$ so that $|\nu(A)| < \epsilon$ for every $A \in \Sigma$ with $\mu(A) < \delta$.*

Proof: Suppose that for every $\epsilon > 0$ there is a $\delta > 0$ so that $|\nu(A)| < \epsilon$ for every $A \in \Sigma$ with $\mu(A) < \delta$. If $\mu(A) = 0$, then $\mu(A) < \delta$ for every $\delta > 0$ and, hence, $|\nu(A)| < \epsilon$ for every $\epsilon > 0$. Therefore, $\nu(A) = 0$ and $\nu \ll \mu$.

Suppose that $\nu \ll \mu$ but there is some $\epsilon_0 > 0$ so that, for every $\delta > 0$, there is $A \in \Sigma$ with $\mu(A) < \delta$ and $|\nu(A)| \geq \epsilon_0$. Then, for every k, there is $A_k \in \Sigma$ with

$\mu(A_k) < \frac{1}{2^k}$ and $|\nu|(A_k) \geq |\nu(A_k)| \geq \epsilon_0$. We define $B_k = \cup_{l=k}^{+\infty} A_l$ and, then, $\mu(B_k) < \frac{1}{2^{k-1}}$ and $|\nu|(B_k) \geq |\nu|(A_k) \geq \epsilon_0$ for every k. If we set $B = \cap_{k=1}^{+\infty} B_k$, then $B_k \downarrow B$ and, by the continuity of $|\nu|$ from above, we get $\mu(B) = 0$ and $|\nu|(B) \geq \epsilon_0$. This says that $|\nu|$ is not absolutely continuous with respect to μ and, by Proposition 10.14, we arrive at a contradiction.

Theorem 10.11 *Let μ be a measure on (X, Σ).*
(i) If $\lambda, \lambda', \rho, \rho'$ are signed or complex measures on (X, Σ) so that $\lambda, \lambda' \ll \mu$ and $\rho, \rho' \perp \mu$ and $\lambda + \rho = \lambda' + \rho'$, then $\lambda = \lambda'$ and $\rho = \rho'$.
(ii) If $f, f' : X \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$ are integrable over X with respect to μ and $f\mu = f'\mu$, then $f = f'$ μ-a.e. on X.
(iii) If $f, f' : X \to \overline{\mathbf{R}}$ are measurable and $\int_X f\, d\mu, \int_X f'\, d\mu$ are defined and $f\mu = f'\mu$, then $f = f'$ μ-a.e. on X, provided that μ, restricted to the set $\{x \in X \,|\, f(x) \neq f'(x)\}$, is semifinite.

Proof: (i) There exist sets $M, M', N, N' \in \Sigma$ with $M \cup N = X = M' \cup N'$, $M \cap N = \emptyset = M' \cap N'$ so that N, N' are null for μ, M is null for ρ and M' is null for ρ'. If we set $K = N \cup N'$, then K is null for μ and $K^c = M \cap M'$ is null for both ρ and ρ'. Since $\lambda, \lambda' \ll \mu$, we have that K is null for both λ and λ'.

If $A \in \Sigma$, $A \subseteq K$, then $\rho(A) = \rho(A) + \lambda(A) = \rho'(A) + \lambda'(A) = \rho'(A)$. If $A \in \Sigma$, $A \subseteq K^c$, then $\rho(A) = 0 = \rho'(A)$. Therefore, for every $A \in \Sigma$ we have $\rho(A) = \rho(A \cap K) + \rho(A \cap K^c) = \rho'(A \cap K) + \rho'(A \cap K^c) = \rho'(A)$ and, hence, $\rho = \rho'$.

A symmetric argument implies that $\lambda = \lambda'$.
(ii) We have $\int_A (f - f')\, d\mu = \int_A f\, d\mu - \int_A f'\, d\mu = (f\mu)(A) - (f'\mu)(A) = 0$ for all $A \in \Sigma$. Theorem 7.5 implies $f = f'$ μ-a.e. on X.
(iii) Let $t, s \in \mathbf{R}$ with $t < s$ and $A_{t,s} = \{x \in X \,|\, f(x) \leq t, s \leq f'(x)\}$. If $0 < \mu(A_{t,s}) < +\infty$, we define $B = A_{t,s}$. If $\mu(A_{t,s}) = +\infty$, we take $B \in \Sigma$ so that $B \subseteq A_{t,s}$ and $0 < \mu(B) < +\infty$. In any case, $(f\mu)(B) = \int_B f\, d\mu \leq t\mu(B)$ and $(f'\mu)(B) = \int_B f'\, d\mu \geq s\mu(B)$ and, thus, $s\mu(B) \leq t\mu(B)$. This implies $\mu(B) = 0$, which is false. The only remaining case is $\mu(A_{t,s}) = 0$.

Now, we observe that $\{x \in X \,|\, f(x) < f'(x)\} = \cup_{t,s \in \mathbf{Q}, t<s} A_{t,s}$, which implies $\mu(\{x \in X \,|\, f(x) < f'(x)\}) = 0$. Similarly, $\mu(\{x \in X \,|\, f(x) > f'(x)\}) = 0$ and we conclude that $f = f'$ μ-a.e. on X.

Lemma 10.5 *Let μ, ν be finite measures on (X, Σ). If μ and ν are not mutually singular, then there exists an $\epsilon_0 > 0$ and an $A_0 \in \Sigma$ with $\mu(A_0) > 0$ so that*

$$\frac{\nu(A)}{\mu(A)} \geq \epsilon_0$$

for every $A \in \Sigma$, $A \subseteq A_0$ with $\mu(A) > 0$.

Proof: We consider, for every n, a Hahn decomposition of the signed measure $\nu - \frac{1}{n}\mu$. There are sets $P_n, N_n \in \Sigma$ so that $P_n \cup N_n = X$, $P_n \cap N_n = \emptyset$ and P_n is a positive set and N_n is a negative set for $\nu - \frac{1}{n}\mu$.

We set $N = \cap_{n=1}^{+\infty} N_n$ and, since $N \subseteq N_n$ for all n, we get $(\nu - \frac{1}{n}\mu)(N) \leq 0$ for all n. Then $\nu(N) \leq \frac{1}{n}\mu(N)$ for all n and, since $\mu(N) < +\infty$, $\nu(N) = 0$. We set $P = \cup_{n=1}^{+\infty} P_n$ and have $P \cup N = X$ 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, $\mu(A_0) > 0$ and, if $A \in \Sigma$, $A \subseteq A_0$, then, since A_0 is a positive set for $\nu - \epsilon_0 \mu$, we get $\nu(A) - \epsilon_0 \mu(A) \geq 0$. If also $\mu(A) > 0$, then $\frac{\nu(A)}{\mu(A)} \geq \epsilon_0$.

Theorem 10.12 ***(Lebesgue-Radon-Nikodym Theorem.*** *Signed case.)* *Let ν be a σ-finite signed measure and μ be a σ-finite measure on (X, Σ). Then there exist unique σ-finite signed measures λ and ρ on (X, Σ) so that*

$$\nu = \lambda + \rho, \qquad \lambda \ll \mu, \qquad \rho \perp \mu.$$

Moreover, there exists a measurable $f : X \to \overline{\mathbf{R}}$ so that the $\int_X f\, d\mu$ is defined and

$$\lambda = f\mu.$$

If f' is another such function, then $f' = f$ 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 integrable over X with respect to μ.

Proof: The uniqueness part of the statement is a consequence of Theorem 10.11. Observe that μ is σ-finite and, hence, 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, Σ).

We define $\mathcal{C}$ to be the collection of all measurable $f : X \to [0, +\infty]$ with the property

$$\int_A f\, d\mu \leq \nu(A), \qquad A \in \Sigma.$$

The function 0, obviously, belongs to $\mathcal{C}$ and, if f_1, f_2 belong to $\mathcal{C}$, then the function $f = \max\{f_1, f_2\}$ also belongs to $\mathcal{C}$. Indeed, if $A \in \Sigma$ we consider $A_1 = \{x \in A \,|\, f_2(x) \leq f_1(x)\}$ and $A_2 = \{x \in A \,|\, f_1(x) < f_2(x)\}$ and we have $\int_A f\, d\mu = \int_{A_1} f\, d\mu + \int_{A_2} f\, d\mu = \int_{A_1} f_1\, d\mu + \int_{A_2} f_2\, d\mu \leq \nu(A_1) + \nu(A_2) = \nu(A)$.

We define

$$\kappa = \sup\Big\{ \int_X f\, d\mu \,|\, f \in \mathcal{C} \Big\}.$$

Since $0 \in \mathcal{C}$ and $\int_X f\, d\mu \leq \nu(X)$ for all $f \in \mathcal{C}$, we have $0 \leq \kappa \leq \nu(X) < +\infty$.

We take a sequence (f_n) in $\mathcal{C}$ so that $\int_X f_n\, d\mu \to \kappa$ and define $g_1 = f_1$ and, inductively, $g_n = \max\{g_{n-1}, f_n\}$ for all $n \geq 2$. Then all g_n belong to $\mathcal{C}$. If we set $f = \lim_{n\to+\infty} g_n$, then $g_n \uparrow f$ and, by the Monotone Convergence Theorem,

$$\int_A f\, d\mu \leq \nu(A), \qquad A \in \Sigma$$

and

$$\int_X f\, d\mu = \kappa < +\infty.$$

Since $(\nu - f\mu)(A) = \nu(A) - \int_A f\,d\mu \geq 0$ for all $A \in \Sigma$, the signed measure $\nu - f\mu$ is a finite measure. If $\nu - f\mu$ and μ are not mutually singular, then, by Lemma 10.5, there is $A_0 \in \Sigma$ and $\epsilon_0 > 0$ so that

$$\frac{\nu(A)}{\mu(A)} - \frac{1}{\mu(A)} \int_A f\,d\mu = \frac{(\nu - f\mu)(A)}{\mu(A)} \geq \epsilon_0$$

for all $A \in \Sigma$, $A \subseteq A_0$ with $\mu(A) > 0$. From this we get $\int_A (f + \epsilon_0 \chi_{A_0})\,d\mu \leq \nu(A)$ for all $A \in \Sigma$, $A \subseteq A_0$. Now for any $A \in \Sigma$ we have $\int_A (f + \epsilon_0 \chi_{A_0})\,d\mu = \int_{A\cap A_0} (f + \epsilon_0 \chi_{A_0})\,d\mu + \int_{A\setminus A_0} (f + \epsilon_0 \chi_{A_0})\,d\mu \leq \nu(A \cap A_0) + \int_{A\setminus A_0} (f + \epsilon_0 \chi_{A_0})\,d\mu = \nu(A \cap A_0) + \int_{A\setminus A_0} f\,d\mu \leq \nu(A \cap A_0) + \nu(A \setminus A_0) = \nu(A)$. This implies that $f + \epsilon_0 \chi_{A_0}$ belongs to $\mathcal{C}$ and hence $\kappa + \epsilon_0 \mu(A_0) = \int_X (f + \epsilon_0 \chi_{A_0})\,d\mu \leq \kappa$. 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, because $\lambda(X) = \int_X f\,d\mu = \kappa < +\infty$ and $\rho(X) = \nu(X) - \int_X f\,d\mu = \nu(X) - \kappa < +\infty$.

B. We, now, suppose that both μ, ν are σ-finite measures on (X, Σ).

Then, there are pairwise disjoint $F_1, F_2, \ldots \in \Sigma$ so that $X = \cup_{k=1}^{+\infty} F_k$ and $\mu(F_k) < +\infty$ for all k and pairwise disjoint $G_1, G_2, \ldots \in \Sigma$ so that $X = \cup_{l=1}^{+\infty} G_l$ and $\nu(G_l) < +\infty$ for all l. The sets $F_k \cap G_l$ are pairwise disjoint, they cover X and $\mu(F_k \cap G_l), \nu(F_k \cap G_l) < +\infty$ for all k, l. We enumerate them as $E_1, E_2, \ldots$ and have $X = \cup_{n=1}^{+\infty} E_n$ and $\mu(E_n), \nu(E_n) < +\infty$ for all n.

We define μ_n and ν_n by

$$\mu_n(A) = \mu(A \cap E_n), \qquad \nu_n(A) = \nu(A \cap E_n)$$

for all $A \in \Sigma$ and all n and we see that all μ_n, ν_n are finite measures on (X, Σ). We also have

$$\mu(A) = \sum_{n=1}^{+\infty} \mu_n(A), \qquad \nu(A) = \sum_{n=1}^{+\infty} \nu_n(A)$$

for all $A \in \Sigma$.

Applying the results of part A, we see that there exist finite measures λ_n, ρ_n on (X, Σ) and $f_n : X \to [0, +\infty]$ integrable with respect to μ_n so that

$$\nu_n = \lambda_n + \rho_n, \quad \lambda_n \ll \mu_n, \quad \rho_n \perp \mu_n, \quad \lambda_n(A) = \int_A f_n\,d\mu_n$$

for all n and all $A \in \Sigma$. 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 $A \in \Sigma$, $A \subseteq E_n^c$, the relation $\lambda_n(A) = \int_A f_n\,d\mu_n$ remains true for all $A \in \Sigma$ if we change f_n and make it 0 on E_n^c. We, therefore, assume that

$$f_n = 0 \quad \text{on } E_n^c, \qquad \lambda_n(A) = \int_{A\cap E_n} f_n\,d\mu_n$$

for all n and all $A \in \Sigma$.

We define $\lambda, \rho : \Sigma \to [0, +\infty]$ and $f : X \to [0, +\infty]$ by

$$\lambda(A) = \sum_{n=1}^{+\infty} \lambda_n(A), \quad \rho(A) = \sum_{n=1}^{+\infty} \rho_n(A), \quad f(x) = \sum_{n=1}^{+\infty} f_n(x)$$

for every $A \in \Sigma$ and every $x \in X$. It is trivial to see that λ and ρ are measures on (X, Σ) and that f is measurable.

The equality $\nu = \lambda + \rho$ is obvious.

If $A \in \Sigma$ has $\mu(A) = 0$, then $\mu_n(A) = \mu(A \cap E_n) = 0$ and, hence, $\lambda_n(A) = 0$ for all n. Thus, $\lambda(A) = 0$ and, thus, $\lambda \ll \mu$.

Since $\rho_n \perp \mu_n$, there is $R_n \in \Sigma$ 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 $R = \cup_{n=1}^{+\infty} R_n'$ is null for μ and $R^c = \cap_{n=1}^{+\infty} R_n'^c$ is null for all ρ_n and, hence, for ρ. We conclude that $\rho \perp \mu$.

The λ and ρ are σ-finite, because $\lambda(E_n) = \lambda_n(E_n) < +\infty$ and $\rho(E_n) = \rho_n(E_n) < +\infty$ for all n.

Finally, for every $A \in \Sigma$, $\lambda(A) = \sum_{n=1}^{+\infty} \lambda_n(A) = \sum_{n=1}^{+\infty} \int_{A \cap E_n} f_n \, d\mu_n = \sum_{n=1}^{+\infty} \int_{A \cap E_n} f \, d\mu_n = \sum_{n=1}^{+\infty} \int_{A \cap E_n} f \, d\mu = \int_A f \, d\mu$. The fourth equality is true because $\int_{E_n} f \, d\mu_n = \int_{E_n} f \, d\mu$ for all measurable $f : X \to [0, +\infty]$. This is justified as follows: if $f = \chi_A$, 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 measurable simple functions and, by the Monotone Convergence Theorem, it holds for all measurable $f : X \to [0, +\infty]$. Now, from $\lambda(A) = \int_A f \, d\mu$, we conclude that $\lambda = f\mu$ and that $\lambda \ll \mu$.

C. In the general case we write $\nu = \nu^+ - \nu^-$ and both ν^+, ν^- are σ-finite measures on (X, Σ). We apply the result of part B and 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 $\lambda_1, \lambda_2 \ll \mu$, $\rho_1, \rho_2 \perp \mu$. 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. We then write $\lambda = \lambda_1 - \lambda_2$ and $\rho = \rho_1 - \rho_2$ and have that $\nu = \lambda + \rho$ and $\lambda \ll \mu$, $\rho \perp \mu$.

We also have 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, hence, either $f_1 < +\infty$ μ-a.e. on X or $f_2 < +\infty$ μ-a.e. on X. The function $f = f_1 - f_2$ is defined μ-a.e. on X and the $\int_X f \, d\mu = \int_X f_1 \, d\mu - \int_X f_2 \, d\mu$ 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 $A \in \Sigma$ and, thus, $\lambda = f\mu$.

Theorem 10.13 ***(Lebesgue-Radon-Nikodym Theorem.*** `Complex case.`*)* *Let ν be a complex measure and μ be a σ-finite measure on (X, Σ). Then there exist unique complex measures λ and ρ on (X, Σ) so that*

$$\nu = \lambda + \rho, \qquad \lambda \ll \mu, \qquad \rho \perp \mu.$$

Moreover, there exists a measurable $f : X \to \overline{\mathbf{C}}$ so that f is integrable over X with respect to μ and

$$\lambda = f\mu.$$

If f' is another such function, then $f' = f$ 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 $\Re(\nu)$ and $\Im(\nu)$ are real measures and, by Theorem 10.12, there exist real measures $\lambda_1, \lambda_2, \rho_1, \rho_2$ on (X, Σ) so that $\Re(\nu) = \lambda_1 + \rho_1$, $\Im(\nu) = \lambda_2 + \rho_2$ and $\lambda_1, \lambda_2 \ll \mu$ and $\rho_1, \rho_2 \perp \mu$. We set $\lambda = \lambda_1 + i\lambda_2$ and $\rho = \rho_1 + i\rho_2$ and, then, $\nu = \lambda + \rho$ and, clearly, $\lambda \ll \mu$ and $\rho \perp \mu$. There are, also, $f_1, f_2 : X \to \overline{\mathbf{R}}$, which are integrable over X with respect to μ, so that $\lambda_1 = f_1\mu$ and $\lambda_2 = f_2\mu$. The function $f = f_1 + if_2 : X \to \mathbf{C}$ is μ-a.e. defined, it is integrable over X with respect to μ and we have $(f\mu)(A) = \int_A f\, d\mu = \int_A f_1\, d\mu + i\int_A f_2\, d\mu = \lambda_1(A) + i\lambda_2(A) = \lambda(A)$ for all $A \in \Sigma$. Hence, $\lambda = f\mu$.

The uniqueness is an easy consequence of Theorem 10.11.

Definition 10.17 *Let ν be a signed measure or a complex measure and μ a measure on (X, Σ). If there exist, necessarily unique, signed or complex measures λ and ρ, so that*

$$\nu = \lambda + \rho, \quad \lambda \ll \mu, \quad \rho \perp \mu,$$

we say that λ ***and*** *ρ* ***constitute the Lebesgue decomposition of*** *ν* ***with respect to*** *μ.*

λ is called ***the absolutely continuous part*** *and ρ is called* ***the singular part of*** *ν* ***with respect to*** *μ.*

Let ν be a signed or complex measure and μ a measure on (X, Σ) so that $\nu \ll \mu$. If there exists a measurable $f : X \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$ 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 by $\frac{d\nu}{d\mu}$.*

Theorems 10.12 and 10.13 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. Definition 10.17 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 $\nu(A) = \int_A f\, d\mu$, $A \in \Sigma$.
3. As we just observed, the real meaning of the symbol $\frac{d\nu}{d\mu}$ is through

$$\nu(A) = \int_A fracd\nu d\mu\, d\mu, \qquad A \in \Sigma,$$

which, after *formally* simplifying the fraction (!), changes into the *true* equality $\nu(A) = \int_A d\nu$.
4. Theorem 10.11 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 10.16 *Let ν_1, ν_2 be complex or σ-finite signed measures and μ a σ-finite measure on (X, Σ). If $\nu_1, \nu_2 \ll \mu$ and $\nu_1 + \nu_2$ is defined, then $\nu_1 + \nu_2 \ll \mu$ and*

$$\frac{d(\nu_1 + \nu_2)}{d\mu} = \frac{d\nu_1}{d\mu} + \frac{d\nu_2}{d\mu}\,, \quad \mu\text{-a.e. on } X.$$

Proof: We have $(\nu_1 + \nu_2)(A) = \int_A \frac{d\nu_1}{d\mu}\, d\mu + \int_A \frac{d\nu_2}{d\mu}\, d\mu = \int_A \left(\frac{d\nu_1}{d\mu} + \frac{d\nu_2}{d\mu}\right) d\mu$ for all $A \in \Sigma$ and, hence, $\frac{d(\nu_1+\nu_2)}{d\mu} = \frac{d\nu_1}{d\mu} + \frac{d\nu_2}{d\mu}$ μ-a.e. on X.

Proposition 10.17 *Let ν be a complex or a σ-finite signed measure and μ a σ-finite measure on (X, Σ). If $\nu \ll \mu$ and $\kappa \in \mathbf{C}$ or $\mathbf{R}$, then $\kappa\nu \ll \mu$ and*

$$\frac{d(\kappa\nu)}{d\mu} = \kappa \frac{d\nu}{d\mu}\,, \quad \mu\text{-a.e. on } X.$$

Proof: We have $(\kappa\nu)(A) = \kappa \int_A \frac{d\nu}{d\mu}\, d\mu = \int_A \left(\kappa \frac{d\nu}{d\mu}\right) d\mu$ for all $A \in \Sigma$ and, hence, $\frac{d(\kappa\nu)}{d\mu} = \kappa \frac{d\nu}{d\mu}$ μ-a.e. on X.

Proposition 10.18 *(Chain rule.) Let ν be a complex or σ-finite signed measure and μ', μ be σ-finite measures on (X, Σ). If $\nu \ll \mu'$ and $\mu' \ll \mu$, then $\nu \ll \mu$ and*

$$\frac{d\nu}{d\mu} = \frac{d\nu}{d\mu'} \frac{d\mu'}{d\mu}\,, \quad \mu\text{-a.e. on } X.$$

Proof: If $A \in \Sigma$ has $\mu(A) = 0$, then $\mu'(A) = 0$ and, hence, $\nu(A) = 0$. Therefore, $\nu \ll \mu$.

Theorem 10.10 implies that $\nu(A) = \int_A \frac{d\nu}{d\mu'}\, d\mu' = \int_A \frac{d\nu}{d\mu'} \frac{d\mu'}{d\mu}\, d\mu$ for every $A \in \Sigma$ and, hence, $\frac{d\nu}{d\mu} = \frac{d\nu}{d\mu'} \frac{d\mu'}{d\mu}$ μ-a.e. on X.

Proposition 10.19 *Let μ and μ' be two σ-finite measures on (X, Σ). If $\mu' \ll \mu$ and $\mu \ll \mu'$, then*

$$\frac{d\mu}{d\mu'} \frac{d\mu'}{d\mu} = 1\,, \quad \mu\text{-a.e. on } X.$$

Proof: We have $\mu(A) = \int_A d\mu$ for every $A \in \Sigma$ and, hence, $\frac{d\mu}{d\mu} = 1$ μ-a.e. on X. The result of this proposition is a trivial consequence of Proposition 10.18.

Proposition 10.20 *Let ν be a σ-finite measure on (X, Σ). Then $\nu \ll |\nu|$ and*

$$\left|\frac{d\nu}{d|\nu|}\right| = 1\,, \quad \nu\text{-a.e. on } X.$$

Proof: Proposition 10.11 implies that $\left|\frac{d\nu}{d|\nu|}\right||\nu| = \left|\frac{d\nu}{d|\nu|}|\nu|\right| = |\nu|$ and, hence, $\left|\frac{d\nu}{d|\nu|}\right| = 1$ $|\nu|$ -a.e. on X.

10.7 Differentiation of indefinite integrals in $\mathbf{R}^n$.

Let $f : [a, b] \to \mathbf{R}$ 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 $\frac{d}{dx} \int_a^x f(y)\, dy = f(x)$. This, of course, means that

$$\lim_{r\to 0+} \frac{\int_a^{x+r} f(y)\, dy - \int_a^x f(y)\, dy}{r} = \lim_{r\to 0+} \frac{\int_a^{x-r} f(y)\, dy - \int_a^x f(y)\, dy}{-r} = f(x).$$

Adding the two limits, we find

$$\lim_{r\to 0+} \frac{\int_{x-r}^{x+r} f(y)\, dy}{2r} = f(x).$$

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 in $\mathbf{R}^n$.*

Lemma 10.6 *(N. Wiener) Let $B_1, \ldots, B_m$ be open balls in $\mathbf{R}^n$. There exist pairwise disjoint $B_{1_1}, \ldots, B_{i_k}$ 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 $B_1, \ldots, B_m$ 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, we hav

$$m_n(C_2) \leq m_n(B^*_{i_2}) = 3^n m_n(B_{i_2}).$$

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 $B_{i_1}, \ldots, B_{i_{k-1}}$ and from them we choose a ball B_{i_k} with largest radius. This B_{i_k} does not intersect $B_{i_1}, \ldots, B_{i_{k-1}}$. 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,

$$m_n(C_k) \leq m_n(B^*_{i_k}) = 3^n m_n(B_{i_k}).$$

It is clear that each of the original balls $B_1, \ldots, B_m$ is either chosen as one of $B_{i_1}, \ldots, B_{i_k}$ or is a satellite of one of $B_{i_1}, \ldots, B_{i_k}$. Therefore, $B_1 \cup \cdots \cup B_m = C_1 \cup \cdots \cup C_k$ and, hence,

$$\begin{aligned} m_n(B_1 \cup \cdots \cup B_m) &= m_n(C_1 \cup \cdots \cup C_k) \leq m_n(C_1) + \cdots + m_n(C_k) \\ &\leq 3^n \big(m_n(B_{i_1}) + \cdots + m_n(B_{i_k})\big). \end{aligned}$$

Definition 10.18 *Let $f : \mathbf{R}^n \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$ be Lebesgue measurable. We say that f is* ***locally Lebesgue integrable*** *if for every $x \in \mathbf{R}^n$ there is an open neighborhood U_x of x so that $\int_{U_x} |f(y)|\, dm_n(y) < +\infty$.*

Lemma 10.7 *Let $f : \mathbf{R}^n \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$ be Lebesgue measurable. Then f is locally Lebesgue integrable if and only if $\int_M |f(y)|\, dm_n(y) < +\infty$ for every bounded set $M \in \mathcal{L}_n$.*

Proof: Let f be locally Lebesgue integrable and $M \subseteq \mathbf{R}^n$ be bounded. We consider a compact set $K \supseteq M$. (Such a K is the closure of M or just a closed ball or a closed cube including M.) For each $x \in K$ we take an open neighborhood U_x of x so that $\int_{U_x} |f(y)|\, dm_n(y) < +\infty$. We, then, take finitely many $x_1, \ldots, x_m$ 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 $\int_{B(x;1)} |f(y)|\, dm_n(y) < +\infty$ for every x and, hence, f is locally Lebesgue integrable.

Proposition 10.21 *Let $f, f_1, f_2 : \mathbf{R}^n \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$ be locally Lebesgue integrable and $\kappa \in \mathbf{C}$. Then*
(i) f is finite a.e. on $\mathbf{R}^n$,

(ii) $f_1 + f_2$ *is defined a.e. on* $\mathbf{R}^n$ *and any Lebesgue measurable definition of* $f_1 + f_2$ *is locally Lebesgue integrable,*
(iii) κf *is locally Lebesgue integrable.*

Proof: (i) Lemma 10.7 implies $\int_{B(0;k)} |f(y)| \, dm_n(y) < +\infty$ and, hence, f is finite a.e. in $B(0;k)$ for every k. Since $\mathbf{R}^n = \cup_{k=1}^{+\infty} B(0;k)$, we find that f is finite a.e. in $\mathbf{R}^n$.
(ii) By (i), both f_1, f_2 are finite and, hence, $f_1 + f_2$ is defined a.e. on $\mathbf{R}^n$. 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 $M \subseteq \mathbf{R}^n$ and, by Lemma 10.7, $f_1 + f_2$ 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 $M \subseteq \mathbf{R}^n$ and, hence, κf is locally Lebesgue integrable.

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 10.19 *Let* $f : \mathbf{R}^n \to \overline{\mathbf{R}}$ *or* $\overline{\mathbf{C}}$ *be locally Lebesgue integrable. The function* $M(f) : \mathbf{R}^n \to [0, +\infty]$*, 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 $x \in \mathbf{R}^n$*, is called* ***the Hardy-Littlewood maximal function of*** f*.*

Proposition 10.22 *Let* $f, f_1, f_2 : \mathbf{R}^n \to \overline{\mathbf{R}}$ *or* $\overline{\mathbf{C}}$ *be locally Lebesgue integrable and* $\kappa \in \mathbf{C}$*. 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 $B \ni x$, $\frac{1}{m_n(B)} \int_B |f_1(y) + f_2(y)| \, dm_n(y) \leq \frac{1}{m_n(B)} \int_B |f_1(y)| \, dm_n(y) + \frac{1}{m_n(B)} \int_B |f_2(y)| \, dm_n(y) \leq M(f_1)(x) + M(f_2)(x)$. Taking supremum of the left side, we get $M(f_1 + f_2)(x) \leq M(f_1)(x) + M(f_2)(x)$.
(ii) Also, $\frac{1}{m_n(B)} \int_B |\kappa f(y)| \, dm_n(y) = |\kappa| \frac{1}{m_n(B)} \int_B |f(y)| \, dm_n(y) \leq |\kappa| M(f)(x)$ and, taking the supremum of the left side, $M(\kappa f)(x) \leq |\kappa| M(f)(x)$. Conversely, $M(\kappa f)(x) \geq \frac{1}{m_n(B)} \int_B |\kappa f(y)| \, dm_n(y) = |\kappa| \frac{1}{m_n(B)} \int_B |f(y)| \, dm_n(y)$ and, taking the supremum of the right side, $M(\kappa f)(x) \geq |\kappa| M(f)(x)$.

Lemma 10.8 *Let* $f : \mathbf{R}^n \to \overline{\mathbf{R}}$ *or* $\overline{\mathbf{C}}$ *be locally Lebesgue integrable. Then, for every* $t > 0$*, the set* $\{x \in \mathbf{R}^n \,|\, t < M(f)(x)\}$ *is open in* $\mathbf{R}^n$*.*

Proof: Let $U = \{x \in \mathbf{R}^n \,|\, t < M(f)(x)\}$ and $x \in U$. Then $t < M(f)(x)$ and, hence, there exists 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 $\frac{1}{m_n(B)} \int_B |f(y)| \, dm_n(y) \leq M(f)(x')$ and, thus, $t < M(f)(x')$. Therefore, $B \subseteq U$ and U is open in $\mathbf{R}^n$.

Since $\{x \in \mathbf{R}^n \,|\, t < M(f)(x)\}$ is open, it is also a Lebesgue set.

Theorem 10.14 *(Hardy, Littlewood) Let $f : \mathbf{R}^n \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$ be Lebesgue integrable. Then, for every $t > 0$, we have*

$$m_n(\{x \in \mathbf{R}^n \,|\, t < M(f)(x)\}) \leq \frac{3^n}{t} \int_{\mathbf{R}^n} |f(y)| \, dm_n(y).$$

Proof: We take arbitrary compact $K \subseteq U = \{x \in \mathbf{R}^n \,|\, t < M(f)(x)\}$ and for each $x \in K$ we find an open ball $B_x \ni x$ with $t < \frac{1}{m_n(B_x)} \int_{B_x} |f(y)| \, dm_n(y)$. Since K is compact, there exist $x_1, \ldots, x_m$ so that $K \subseteq B_{x_1} \cup \cdots \cup B_{x_m}$. By Lemma 10.6, there exist *pairwise disjoint* $B_{x_{i_1}}, \ldots, B_{x_{i_k}}$ so that

$$m_n(B_{x_1} \cup \cdots \cup B_{x_m}) \leq 3^n \big(m_n(B_{x_{i_1}}) + \cdots + m_n(B_{x_{i_k}}) \big).$$

Then

$$\begin{aligned} m_n(K) &\leq m_n(B_{x_1} \cup \cdots \cup B_{x_m}) \\ &\leq \frac{3^n}{t} \Big(\int_{B_{x_{i_1}}} |f(y)| \, dm_n(y) + \cdots + \int_{B_{x_{i_k}}} |f(y)| \, dm_n(y) \Big) \\ &\leq \frac{3^n}{t} \int_{\mathbf{R}^n} |f(y)| \, dm_n(y). \end{aligned}$$

By the regularity of m_n, $m_n(U) = \sup\{m_n(K) \,|\, K \text{ is compact } \subseteq U\}$ and we conclude that $m_n(U) \leq \frac{3^n}{t} \int_{\mathbf{R}^n} |f(y)| \, dm_n(y)$.

Observe that the quantity $m_n(\{x \in \mathbf{R}^n \,|\, t < M(f)(x)\})$ 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 Theorem 10.14 is

$$\lambda_{M(f)}(t) \leq \frac{3^n}{t} \int_{\mathbf{R}^n} |f(y)| \, dm_n(y).$$

Definition 10.20 *Let (X, Σ, μ) be a measure space and $g : X \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$ be measurable. We say that g **is weakly integrable over** X **(with respect to** μ**)** if there is a $c < +\infty$ so that $\lambda_{|g|}(t) \leq \frac{c}{t}$ for every $t > 0$.*

Another way to state Theorem 10.14 is: *if f is Lebesgue integrable, then $M(f)$ is weakly Lebesgue integrable.*

Proposition 10.23 *Let (X, Σ, μ) be a measure space, $g, g_1, g_2 : X \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$ be weakly integrable over X and $\kappa \in \mathbf{C}$. Then*
(i) g is finite a.e. on X,
(ii) $g_1 + g_2$ is defined a.e. on X and any measurable definition of $g_1 + g_2$ is weakly integrable over X,
(iii) κg is weakly integrable over X.

Proof: (i) $\lambda_{|g|}(t) \leq \frac{c}{t}$ for all $t > 0$ implies that $\mu(\{x \in X \,|\, g(x) \text{ is infinite}\}) \leq \mu(\{x \in X \,|\, n < |g(x)|\}) \leq \frac{c}{n}$ for all n and, thus, $\mu(\{x \in X \,|\, g(x) \text{ is infinite}\}) = 0$.
(ii) By (i) both g_1 and g_2 are finite a.e. on X and, hence, $g_1 + g_2$ 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 any measurable definition of $g_1 + g_2$ satisfies, for every $t > 0$, the estimate: $\mu(\{x \in X \,|\, t < |g_1(x) + g_2(x)|\}) \leq \mu(\{x \in X \,|\, \frac{t}{2} < |g_1(x)|\}) + \mu(\{x \in X \,|\, \frac{t}{2} < |g_2(x)|\}) \leq \frac{2c_1+2c_2}{t}$.
(iii) If $\mu(\{x \in X \,|\, t < |g(x)|\}) \leq \frac{c}{t}$ for all $t > 0$, then $\mu(\{x \in X \,|\, t < |\kappa g(x)|\}) = \mu(\{x \in X \,|\, \frac{t}{|\kappa|} < |g(x)|\}) \leq \frac{c|\kappa|}{t}$ for all $t > 0$.

Proposition 10.24 *Let (X, Σ, μ) be a measure space and $g : X \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$ be integrable over X. Then g is also weakly integrable over X.*

Proof: We have $\lambda_{|g|}(t) = \mu(\{x \in X \,|\, t < |g(x)|\}) \leq \frac{1}{t} \int_{\{x \in X \,|\, t < |g(x)|\}} |g| \, d\mu \leq \frac{1}{t} \int_X |g| \, d\mu$ 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 10.24 is not true. Consider, for example, the function $g(x) = \frac{1}{|x|^n}, x \in \mathbf{R}^n$.

By Proposition 8.12, $\int_{\mathbf{R}^n} |g(x)| \, dm_n(x) = \sigma_{n-1}(S^{n-1}) \int_0^{+\infty} \frac{1}{r^n} r^{n-1} \, dr = \sigma_{n-1}(S^{n-1}) \int_0^{+\infty} \frac{1}{r} \, dr = +\infty$. But, $\{x \in \mathbf{R}^n \,|\, t < |g(x)|\} = B(0; t^{-\frac{1}{n}})$ and, hence, $\lambda_{|g|}(t) = v_n \cdot (t^{-\frac{1}{n}})^n = \frac{v_n}{t}$ for every $t > 0$, where $v_n = m_n(B(0; 1))$.

The next result says that the Hardy-Littlewood maximal function is never (except if the function is zero) Lebesgue integrable.

Proposition 10.25 *Let $f : \mathbf{R}^n \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$ be Lebesgue integrable. Then $M(f)$ is locally Lebesgue integrable. If $M(f)$ is Lebesgue integrable, then $f = 0$ a.e. on $\mathbf{R}^n$.*

Proof: Let $A = \{x \in \mathbf{R}^n \,|\, f(x) \neq 0\}$ and assume that $m_n(A) > 0$. Since $A = \cup_{k=1}^{+\infty} (A \cap B(0; k))$, 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 set M 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 take any x with $|x| \geq k$ and observe that there is an open ball B of diameter $|x| + k + 1$ containing x and including M. If $v_n = m_n(B(0; 1))$, then $M(f)(x) \geq \frac{1}{m_n(B)} \int_B |f(y)| \, dm_n(y) \geq \frac{2^n}{v_n \cdot (|x|+k+1)^n} \int_M |f(y)| \, dm_n(y) \geq \frac{c}{|x|^n}$, with $c = \frac{2^n}{v_n 3^n} \int_M |f(y)| \, dm_n(y) > 0$. This implies $\int_{\mathbf{R}^n} |M(f)(x)| \, dm_n(x) \geq c \int_{\{x \in \mathbf{R}^n \,|\, |x| \geq k\}} \frac{1}{|x|^n} \, dm_n(x) = c\sigma_{n-1}(S^{n-1}) \int_k^{+\infty} \frac{1}{r^n} r^{n-1} \, dr = +\infty$.

The next result is a direct generalization of the fundamental theorem of calculus and the proofs are identical.

Lemma 10.9 *Let $g : \mathbf{R}^n \to \mathbf{C}$ be continuous on $\mathbf{R}^n$. Then*

$$\lim_{r \to 0+} \frac{1}{m_n(B(x; r))} \int_{B(x;r)} |g(y) - g(x)| \, dm_n(y) = 0$$

for every $x \in \mathbf{R}^n$.

Proof: Let $\epsilon > 0$ and take $\delta > 0$ so that $|g(y) - g(x)| \leq \epsilon$ for every $y \in \mathbf{R}^n$ with $|y - x| < \delta$. Then, for every $r < \delta$, $\frac{1}{m_n(B(x;r))} \int_{B(x;r)} |g(y) - g(x)| \, dm_n(y) \leq \frac{1}{m_n(B(x;r))} \int_{B(x;r)} \epsilon \, dm_n(y) = \epsilon$.

Theorem 10.15 *(Lebesgue) Let $f : \mathbf{R}^n \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$ 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 a.e. $x \in \mathbf{R}^n$.

Proof: We first assume that f is integrable.

We take an arbitrary $\epsilon > 0$ and, through Theorem 7.16, we find $g : \mathbf{R}^n \to \mathbf{C}$, continuous on $\mathbf{R}^n$, so that $\int_{\mathbf{R}^n} |g - f| \, dm_n < \epsilon$. For all $x \in \mathbf{R}^n$ and $r > 0$ we get $\frac{1}{m_n(B(x;r))} \int_{B(x;r)} |f(y) - f(x)| \, dm_n(y) \leq \frac{1}{m_n(B(x;r))} \int_{B(x;r)} |f(y) - g(y)| \, dm_n(y) + \frac{1}{m_n(B(x;r))} \int_{B(x;r)} |g(y) - g(x)| \, dm_n(y) + \frac{1}{m_n(B(x;r))} \int_{B(x;r)} |g(x) - f(x)| \, dm_n(y) \leq M(f - g)(x) + \frac{1}{m_n(B(x;r))} \int_{B(x;r)} |g(y) - g(x)| \, dm_n(y) + |g(x) - f(x)|$.

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 10.9, implies

$$\limsup_{r \to 0+} A(f)(x;r) \leq M(f - g)(x) + 0 + |g(x) - f(x)|.$$

Now, for every $t > 0$, we get $m_n^*(\{x \in \mathbf{R}^n \mid t < \limsup_{r \to 0+} A(f)(x;r)\}) \leq m_n(\{x \in \mathbf{R}^n \mid \frac{t}{2} < M(f - g)(x)\}) + m_n(\{x \in \mathbf{R}^n \mid \frac{t}{2} < |g(x) - f(x)|\}) \leq \frac{2 \cdot 3^n}{t} \int_{\mathbf{R}^n} |f - g| \, dm_n + \frac{2}{t} \int_{\mathbf{R}^n} |f - g| \, dm_n \leq \frac{2 \cdot 3^n + 2}{t} \epsilon$, where the second inequality is a consequence of Theorem 10.14. Since ϵ is arbitrary, we find, for all $t > 0$,

$$m_n^*(\{x \in \mathbf{R}^n \mid t < \limsup_{r \to 0+} A(f)(x;r)\}) = 0.$$

By the subadditivity of m_n^*, $m_n^*(\{x \in \mathbf{R}^n \mid \limsup_{r \to 0+} A(f)(x;r) \neq 0\}) \leq \sum_{k=1}^{+\infty} m_n^*(\{x \in \mathbf{R}^n \mid \frac{1}{k} < \limsup_{r \to 0+} A(f)(x;r)\}) = 0$ and, hence,

$$m_n^*(\{x \in \mathbf{R}^n \mid \limsup_{r \to 0+} A(f)(x;r) \neq 0\}) = 0.$$

This implies that $\limsup_{r \to 0+} A(f)(x;r) = 0$ for a.e. $x \in \mathbf{R}^n$ and, since $\liminf_{r \to 0+} A(f)(x;r) \geq 0$ for every $x \in \mathbf{R}^n$, we conclude that

$$\lim_{r \to 0+} A(f)(x;r) = 0$$

for a.e. $x \in \mathbf{R}^n$.

Now, let f be locally Lebesgue integrable. We fix an arbitrary $k \geq 2$ 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 $r \leq 1$, 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 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 a.e. $x \in \mathbf{R}^n$.

Definition 10.21 *Let $f : \mathbf{R}^n \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$ be locally Lebesgue integrable. The set L_f of all $x \in \mathbf{R}^n$ 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 10.9.

Theorem 10.16 *Let $f : \mathbf{R}^n \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$ 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 $\left| \frac{1}{m_n(B(x;r))} \int_{B(x;r)} f(y) \, dm_n(y) - f(x) \right| \leq \frac{1}{m_n(B(x;r))} \int_{B(x;r)} |f(y) - f(x)| \, dm_n(y) \to 0$.

Definition 10.22 *Let $x \in \mathbf{R}^n$ and $\mathcal{E}$ be a collection of sets in $\mathcal{L}_n$ with the property that there is a $c > 0$ so that for every $E \in \mathcal{E}$ there is a ball $B(x;r)$ with $E \subseteq B(x;r)$ and $m_n(E) \geq c m_n(B(x;r))$. Then the collection $\mathcal{E}$ is called* ***a thick family of sets at*** *x.*

Examples
1. Any collection of qubes containing x and any collection of balls containing x is a thick family of sets at x.
2. Consider any collection $\mathcal{E}$ all elements of which are intervals S containing x. Let A_S be the length of the largest side and a_S be the length of the smallest side of S. If there is a constant $c > 0$ so that $\frac{a_S}{A_S} \geq c$ for every $S \in \mathcal{E}$, then $\mathcal{E}$ is a thick family of sets at x.

Theorem 10.17 *Let $f : \mathbf{R}^n \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$ be locally Lebesgue integrable. Then, for every x in the Lebesgue set of f and for every thick family $\mathcal{E}$ of sets at x, we have*

$$\lim_{E\in\mathcal{E}, m_n(E)\to 0+} \frac{1}{m_n(E)} \int_E |f(y) - f(x)| \, dm_n(y) = 0$$

$$\lim_{E\in\mathcal{E}, m_n(E)\to 0+} \frac{1}{m_n(E)} \int_E f(y) \, dm_n(y) = f(x).$$

Proof: There is a $c > 0$ so that for every $E \in \mathcal{E}$ there is a ball $B(x;r_E)$ with $E \subseteq B(x;r_E)$ and $m_n(E) \geq c m_n(B(x;r_E))$. 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 $m_n(E) < c v_n \delta^n$, where $v_n = m_n(B(0;1))$, then $r_E < \delta$ and, hence, $\frac{1}{m_n(E)} \int_E |f(y) - f(x)| \, dm_n(y) \leq \frac{1}{c m_n(B(x;r_E))} \int_{B(x;r_E)} |f(y) - f(x)| \, dm_n(y) < \epsilon$. This means that $\lim_{E\in\mathcal{E}, m_n(E)\to 0+} \frac{1}{m_n(E)} \int_E |f(y) - f(x)| \, dm_n(y) = 0$.

By $\left| \frac{1}{m_n(E)} \int_E f(y) \, dm_n(y) - f(x) \right| \leq \frac{1}{m_n(E)} \int_E |f(y) - f(x)| \, dm_n(y)$ and by the first limit, we prove the second.

10.8 Differentiation of Borel measures in $\mathbf{R}^n$.

Definition 10.23 *Any signed or complex measure on* $(\mathbf{R}^n, \mathcal{B}_{\mathbf{R}^n})$ *is called* ***a Borel signed or complex measure*** *on* $\mathbf{R}^n$.

Definition 10.24 *Let* ν *be a Borel signed measure in* $\mathbf{R}^n$. *We say that* ν ***is locally finite*** *if for every* $x \in \mathbf{R}^n$ *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 10.26 *Let* ν *be a Borel signed measure in* $\mathbf{R}^n$. *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 $x \in \mathbf{R}^n$ we take 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, hence, $|\nu|(U_x)$ are finite. Therefore, $|\nu|$ is locally finite.

Proposition 10.27 *Let* ν *be a Borel signed measure in* $\mathbf{R}^n$. *Then,* ν *is locally finite if and only if* $\nu(M)$ *is finite for all bounded Borel sets* $M \subseteq \mathbf{R}^n$.

Proof: One direction is easy, since every open ball is a bounded set. For the other direction, we suppose that ν is locally finite and, by Proposition 10.26, that $|\nu|$ is also locally finite. Lemma 5.7 implies that $|\nu(M)| \leq |\nu|(M) < +\infty$ for all bounded Borel sets $M \subseteq \mathbf{R}^n$.

Theorem 10.18 *Let* ρ *be a locally finite Borel signed measure or a complex measure on* $\mathbf{R}^n$ *with* $\rho \perp m_n$. *Then,*

$$\lim_{r \to 0+} \frac{\rho(B(x;r))}{m_n(B(x;r))} = 0$$

for m_n*-a.e.* $x \in \mathbf{R}^n$.

Proof: If ρ is complex, then $|\rho|$ is a finite Borel measure on $\mathbf{R}^n$. Proposition 10.26 implies that, if ρ is signed, then $|\rho|$ is a locally finite Borel measure on $\mathbf{R}^n$. Moreover, Proposition 10.10 implies that $|\rho| \perp m_n$. Hence, there exist sets $R, M \in \mathcal{B}_{\mathbf{R}^n}$ with $M \cup R = \mathbf{R}^n$, $M \cap R = \emptyset$ so that R is null for m_n and M is null for $|\rho|$.

We define $A(|\rho|)(x;r) = \frac{|\rho|(B(x;r))}{m_n(B(x;r))}$, take an arbitrary $t > 0$ and consider the set $M_t = \{x \in M \,|\, t < \limsup_{r \to 0+} A(|\rho|)(x;r)\}$.

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 $t < A(|\rho|)(x;r_x) = \frac{|\rho|(B(x;r_x))}{m_n(B(x;r_x))}$ and $B(x;r_x) \subseteq U$.

We form the open set $V = \cup_{x \in M_t} B(x; r_x)$ and take an arbitrary compact set $K \subseteq V$. Now, there exist finitely many $x_1, \ldots, x_m \in M_t$ so that $K \subseteq B(x_1; r_{x_1}) \cup \cdots \cup B(x_m; r_{x_m})$. Lemma 10.6 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

$$m_n(K) \leq \frac{3^n}{t} \big(|\rho|(B(x_{i_1}; r_{x_{i_1}})) + \cdots + |\rho|(B(x_{i_k}; r_{x_{i_k}}))\big) \leq \frac{3^n}{t} |\rho|(U) < \frac{3^n}{t} \epsilon.$$

By the regularity of m_n and since K is an arbitrary compact subset of V, we find that $m_n(V) \leq \frac{3^n}{t} \epsilon$. Since $M_t \subseteq V$, we have that $m_n^*(M_t) \leq \frac{3^n}{t} \epsilon$ and, since ϵ is arbitrary, we conclude that M_t is a Lebesgue set and $m_n(M_t) = 0$. Finally, since $\{x \in M \mid \limsup_{r \to 0+} A(|\rho|)(x; r) \neq 0\} = \cup_{k=1}^{+\infty} M_{\frac{1}{k}}$, we get that $\limsup_{r \to 0+} A(|\rho|)(x; r) = 0$ for m_n-a.e. $x \in \mathbf{R}^n$. Now, from $0 \leq \liminf_{r \to 0+} A(|\rho|)(x; r)$, we conclude that $\lim_{r \to 0+} A(|\rho|)(x; r) = 0$ for m_n-a.e. $x \in \mathbf{R}^n$.

Lemma 10.10 *Let ν be a locally finite Borel signed measure on $\mathbf{R}^n$. 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 $\mathbf{R}^n = \cup_{k=1}^{+\infty} B(0; k)$ and $\nu(B(0; k))$ is finite for every k, we find that ν is σ-finite and Theorem 10.12 implies the existence of the Lebesgue decomposition of ν.

Since $\rho \perp m_n$, there exist Borel sets R, 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, take any bounded Borel set M. Since $\nu(M)$ is finite, Theorem 10.1 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, hence, $\lambda(M)$ 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.

Take, again, any bounded Borel set M. Then $\int_M f(x)\, dm_n(x) = \lambda(M)$ is finite and, hence, $\int_X |f(x)|\, dm_n(x) < +\infty$. This implies that f is locally Lebesgue integrable.

Theorem 10.19 *Let ν be a locally finite Borel signed measure or a Borel complex measure on $\mathbf{R}^n$. 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. $x \in \mathbf{R}^n$.

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 10.10 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. Theorems 10.16 and 10.18 imply

$$\begin{aligned}\lim_{r\to 0+} \frac{\nu(B(x;r))}{m_n(B(x;r))} &= \lim_{r\to 0+} \frac{1}{m_n(B(x;r))} \int_{B(x;r)} f(y)\, dm_n(y) \\ &\quad + \lim_{r\to 0+} \frac{\rho(B(x;r))}{m_n(B(x;r))} = f(x) + 0 \\ &= f(x)\end{aligned}$$

for m_n-a.e. $x \in \mathbf{R}^n$.

Theorem 10.20 *Let ν be a locally finite Borel signed measure or a Borel complex measure on $\mathbf{R}^n$. 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. $x \in \mathbf{R}^n$,*

$$\lim_{E\in\mathcal{E}, m_n(E)\to 0+} \frac{\nu(E)}{m_n(E)} = f(x)$$

for every thick family $\mathcal{E}$ of sets at x.

Proof: If ρ is the singular part of ν with respect to m_n, then $|\rho| \perp m_n$ and, by Theorem 10.18, $\lim_{r\to 0+} \frac{|\rho|(B(x;r))}{m_n(B(x;r))} = 0$ for m_n-a.e. $x \in \mathbf{R}^n$.

We, now, take any x for which $\lim_{r\to 0+} \frac{|\rho|(B(x;r))}{m_n(B(x;r))} = 0$ and any thick family $\mathcal{E}$ of sets at x. This means that there is a $c > 0$ so that for every $E \in \mathcal{E}$ 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 $\frac{|\rho|(B(x;r))}{m_n(B(x;r))} < c\epsilon$. Therefore, if $m_n(E) < c v_n \delta^n$, where $v_n = m_n(B(0;1))$, then $r_E < \delta$ and, hence, $\left|\frac{\rho(E)}{m_n(E)}\right| \leq \frac{|\rho|(E)}{m_n(E)} \leq \frac{1}{c} \frac{|\rho|(B(x;r_E))}{m_n(B(x;r_E))} < \epsilon$. This means that, for m_n-a.e. $x \in \mathbf{R}^n$,

$$\lim_{E\in\mathcal{E}, m_n(E)\to 0+} \frac{\rho(E)}{m_n(E)} = 0$$

for every thick family $\mathcal{E}$ of sets at x.

We combine this with Theorem 10.17 to complete the proof.

10.9 Exercises.

1. Let ν be a signed measure on (X, Σ) and let μ_1, μ_2 be two measures on (X, Σ) at least one of which is finite. If $\nu = \mu_1 - \mu_2$, prove that $\nu^+ \leq \mu_1$ and $\nu^- \leq \mu_2$.

2. Let $\sharp$ be the counting measure on $(\mathbf{N}, \mathcal{P}(\mathbf{N}))$ and μ be the point-mass distribution on $\mathbf{N}$ induced by the function $a_n = \frac{1}{2^n}$, $n \in \mathbf{N}$. Prove that there is an $\epsilon_0 > 0$ and a sequence (E_k) of subsets of $\mathbf{N}$, so that $\mu(E_k) \to 0$ and $\sharp(E_k) \geq \epsilon_0$ for all k. On the other hand, prove that $\sharp \ll \mu$.

3. Let ν_1, μ_1 be σ-finite measures on (X_1, Σ_1) and ν_2, μ_2 be σ-finite measures on (X_2, Σ_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. $(x_1, x_2) \in X_1 \times X_2$.

4. Let $\sharp$ be the counting measure on $(\mathbf{R}, \mathcal{B}_{\mathbf{R}})$.
(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 ?

5. *Generalization of the Radon-Nikodym Theorem.*

   Let ν be a signed measure and μ be a σ-finite measure on (X, Σ) so that $\nu \ll \mu$. Prove that there is a measurable $f : X \to \overline{\mathbf{R}}$, so that $\int_X f\, d\mu$ exists and $\nu = f\mu$.

6. *Generalization of the Lebesgue Decomposition Theorem.*

   Let ν be a σ-finite signed measure and μ a measure on (X, Σ). Prove that there are unique σ-finite signed measures λ, ρ on (X, Σ) so that $\lambda \ll \mu$, $\rho \perp \mu$ and $\nu = \lambda + \rho$.

7. Let ν, μ be two measures on (X, Σ) 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 $\nu = \frac{f}{1-f}\mu$.

8. *Conditional Expectation.*

   Let μ be a σ-finite measure on (X, Σ), Σ_0 be a σ-algebra with $\Sigma_0 \subseteq \Sigma$ and μ be the restriction of the measure on (X, Σ_0).
   (i) If $f : X \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$ is Σ-measurable and $\int_X f\, d\mu$ exists, prove that there is a Σ_0-measurable $f_0 : X \to \overline{\mathbf{R}}$ or, respectively, $\overline{\mathbf{C}}$ so that $\int_X f_0\, d\mu$ exists and
$$\int_A f_0\, d\mu = \int_A f\, d\mu\,, \qquad A \in \Sigma_0.$$
   If f_0' has the same properties as f_0, prove that $f_0' = f_0$ μ-a.e. on X.

   Any f_0 with the above properies is called **a conditional expectation of** f **with respect to** Σ_0 and it is denoted by
$$E(f|\Sigma_0).$$
   (ii) Prove that

   (a) $E(f|\Sigma) = f$ μ-a.e. on X,
   (b) $E(f + g|\Sigma_0) = E(f|\Sigma_0) + E(g|\Sigma_0)$ μ-a.e. on X,
   (c) $E(\kappa f|\Sigma_0) = \kappa E(f|\Sigma_0)$ μ-a.e. on X,
   (d) if g is Σ_0-measurable, then $E(gf|\Sigma_0) = gE(f|\Sigma_0)$ μ-a.e. on X,

(e) if $\Sigma_1 \subseteq \Sigma_0 \subseteq \Sigma$, then $E(f|\Sigma_1) = E(E(f|\Sigma_0)|\Sigma_1)$ μ-a.e. on X.

9. Let ν be a real or complex measure on (X, Σ). If $\nu(X) = |\nu|(X)$, prove that $\nu = |\nu|$.

10. Let ν be a signed or complex measure on (X, Σ). We say that $\{A_1, A_2, \ldots\}$ is **a (countable) measurable partition of** $A \in \Sigma$, if $A_k \in \Sigma$ for all k, the sets $A_1, A_2, \ldots$ are pairwise disjoint and $A = A_1 \cup A_2 \cup \cdots$. Prove that

$$|\nu|(A) = \sup\Big\{ \sum_{k=1}^{+\infty} |\nu(A_k)| \,\Big|\, \{A_1, A_2, \ldots\} \text{ is a measurable partition of } A \Big\}$$

for every $A \in \Sigma$.

11. *A variant of the Hardy-Littlewood maximal function.*

Let $f : \mathbf{R}^n \to \overline{\mathbf{R}}$ or $\overline{\mathbf{C}}$ be locally Lebesgue integrable. We define

$$H(f)(x) = \sup_{r>0} \frac{1}{m_n(B(x;r))} \int_{B(x;r)} |f(y)| \, dm_n(y)$$

for every $x \in \mathbf{R}^n$.
(i) Prove that the set $\{x \in \mathbf{R}^n \,|\, t < H(f)(x)\}$ is open for every $t > 0$.
(ii) Prove that $\frac{1}{2^n} M(f)(x) \leq H(f)(x) \leq M(f)(x)$ for every $x \in \mathbf{R}^n$.

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.

12. *The Vitali Covering Theorem.*

Let $E \subseteq \mathbf{R}^n$ 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 $m_n^*(E \setminus \cup_{k=1}^{+\infty} B_k) = 0$.

13. *Points of density.*

Let $E \in \mathcal{L}_n$. If $x \in \mathbf{R}^n$, 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. 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 a.e. $x \in E$ is a density point of E.
(iii) For any $\alpha \in (0, 1)$ find $x \in \mathbf{R}$ and $E \in \mathcal{L}_1$ so that $D_E(x) = \alpha$. Also, find $x \in \mathbf{R}$ and $E \in \mathcal{L}_1$ so that $D_E(x)$ does not exist.

14. Let f be the Cantor function on $[0, 1]$ (see Exercise 4.6.10) extended as 0 on $(-\infty, 0)$ and as 1 on $(1, +\infty)$ and let μ_f be the Lebesgue-Stieltjes measure on $(\mathbf{R}, \mathcal{B}_{\mathbf{R}})$ induced by f. Prove that $\mu_f \perp m_1$.

15. Let ν be a signed measure on (X, Σ). Prove that $\nu^+, \nu^- \ll |\nu|$ and find formulas for Radon-Nikodym derivatives $\frac{d\nu^+}{d|\nu|}$ and $\frac{d\nu^-}{d|\nu|}$.

16. Let μ be a finite measure on (X, Σ). We define

$$d(A, B) = \mu(A \triangle B), \qquad A, B \in \Sigma.$$

(i) Prove that (Σ, d) is a complete metric space.
(ii) If ν is a real or a complex measure on (X, Σ), prove that ν is continuous on Σ (with respect to d) if and only if ν is continuous at $\emptyset$ (with respect to d) if and only if $\nu \ll \mu$.

Chapter 11

The classical Banach spaces

11.1 Normed spaces.

Definition 11.1 *Let Z be a linear space over the field $F = \mathbf{R}$ or over the field $F = \mathbf{C}$ and let $\| \cdot \| : Z \to \mathbf{R}$ 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 $u = o$, where o is the zero element of Z.
*Then, $\| \cdot \|$ is called **a norm on** Z and $(Z, \| \cdot \|)$ is called **a 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 11.1 *If $\| \cdot \|$ is a norm on the linear space Z, then*
(i) $\|o\| = 0$, where o is the zero element of Z,
(ii) $\| - u\| = \|u\|$, for all $u \in Z$,
(iii) $\|u\| \geq 0$, for all $u \in Z$.

Proof: (i) $\|o\| = \|0 \cdot o\| = |0| \|o\| = 0$.
(ii) $\| - u\| = \|(-1)u\| = | - 1| \|u\| = \|u\|$.
(iii) $0 = \|o\| = \|u + (-u)\| \leq \|u\| + \| - u\| = 2\|u\|$ and, hence, $0 \leq \|u\|$.

Proposition 11.2 *Let $(Z, \| \cdot \|)$ be a normed space. If we define $d : Z \times Z \to \mathbf{R}$ by*

$$d(u, v) = \|u - v\|$$

for all $u, v \in Z$, then d is a metric on Z.

Proof: Using Proposition 11.1, we have
a. $d(u, v) = \|u - v\| \geq 0$ for all $u, v \in Z$ and, if $d(u, v) = 0$, then $\|u - v\| = 0$ and, hence, $u - v = o$ or, equivalently, $u = v$.
b. $d(u, v) = \|u - v\| = \| - (v - u)\| = \|v - u\| = d(v, u)$.
c. $d(u, v) = \|u-v\| = \|(u-w)+(w-v)\| \leq \|u-w\|+\|w-v\| = d(u, w)+d(w, v)$.

Definition 11.2 *Let $(Z, \|\cdot\|)$ be a normed space. If d is the metric defined in Proposition 11.2, 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 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: $u_n \to u$ in Z or $\lim_{n\to+\infty} u_n = u$ in Z.

A set $U \subseteq Z$ is open in Z if for every $u \in U$ there is an $r > 0$ so that $B(u; r) \subseteq U$. Any union of open sets in Z is open in Z and any finite intersection of open sets in Z is open in Z. The sets $\emptyset$ and Z are open in Z.

A set $K \subseteq Z$ is closed in Z if its complement $Z \setminus K$ is open in Z or, equivalently, if the limit of every sequence in K (which has a limit) belongs to K. Any intersection of closed sets in Z is closed in Z and any finite union of closed sets in Z is closed in Z. The sets $\emptyset$ and Z are closed in Z.

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 in Z is convergent, then Z is a complete metric space.

Definition 11.3 *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.

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,

$$u + A = \{u + v \mid v \in A\}, \qquad \kappa A = \{\kappa v \mid v \in A\}$$

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

$$\tau_u(v) = v + u, \qquad l_\kappa(v) = \kappa v$$

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 $l_\kappa^{-1} = l_{\frac{1}{\kappa}}$. It is obvious that $u + A = \tau_u(A)$ and $\kappa A = l_\kappa(A)$.

Proposition 11.3 *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) Translations and dilations are homeomorphisms. This means that they, together with their inverses, are continuous on Z.

(vi) If A is open or closed in Z and $u \in Z$, then $u + A$ is open or, respectively, closed in Z.
(vii) If A is open or closed in Z and $\kappa \in F \setminus \{0\}$, then κA is open or, respectively, closed in Z.

Proof: (i) $w \in u + B(v; r)$ if and only if $w - u \in B(v; r)$ if and only if $\|w - u - v\| < r$ if and only if $w \in B(u + v; r)$.
(ii) $w \in \kappa B(v; r)$ if and only if $\frac{1}{\kappa} w \in B(v; r)$ if and only if $\|\frac{1}{\kappa} w - v\| < r$ if and only if $\|w - \kappa v\| < |\kappa| r$ if and only if $w \in B(\kappa v; |\kappa| r)$.
(iii) $\|(u_n + v_n) - (u + v)\| \leq \|u_n - u\| + \|v_n - v\| \to 0$ as $n \to +\infty$.
(iv) $\|\kappa_n u_n - \kappa u\| \leq |\kappa_n| \|u_n - u\| + |\kappa_n - \kappa| \|u\| \to 0$ as $n \to +\infty$, because (κ_n) is bounded in F.
(v) If $v_n \to v$ in Z, then $\tau_u(v_n) = u + v_n \to u + v = \tau_u(v)$, by (iii). Also, $l_\kappa(v_n) = \kappa v_n \to \kappa v = l_\kappa(v)$, by (iv). Therefore, τ_u and l_κ are continuous on Z. Their inverses are also continuous, because they are also a translation, τ_{-u}, and a dilation, $l_{\frac{1}{\kappa}}$, respectively.
(vi) $u + A = \tau_{-u}^{-1}(A)$ is the inverse image of A under the continuous τ_{-u}.
(vii) $\kappa A = l_{\frac{1}{\kappa}}^{-1}(A)$ is the inverse image of A under the continuous $l_{\frac{1}{\kappa}}$.

As in any linear space, we define **a linear functional on** Z to be a function $l : Z \to F$ which satisfies

$$l(u + v) = l(u) + l(v), \qquad l(\kappa u) = \kappa l(u)$$

for every $u, v \in Z$ and $\kappa \in F$. If l is a linear functional on Z, then $l(o) = l(0o) = 0l(o) = 0$ and $l(-u) = l((-1)u) = (-1)l(u) = -l(u)$ for all $u \in Z$. We define the **sum** $l_1 + l_2 : Z \to F$ of two linear functionals l_1, l_2 on Z by

$$(l_1 + l_2)(u) = l_1(u) + l_2(u), \qquad u \in Z$$

and the **product** $\kappa l : Z \to F$ of a linear functional l on Z and a $\kappa \in F$ by

$$(\kappa l)(u) = \kappa l(u), \qquad u \in Z.$$

It is trivial to prove that $l_1 + l_2$ and κl are linear functionals on Z and that the set Z' whose elements are all the linear functionals on Z,

$$Z' = \{l \,|\, l \text{ is a linear functional on } Z\},$$

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 $o : Z \to F$ with $o(u) = 0$ for every $u \in Z$ and the **opposite** of a linear functional l on Z is the linear functional $-l : Z \to F$ with $(-l)(u) = -l(u)$ for every $u \in Z$.

Definition 11.4 *Let $(Z, \|\cdot\|)$ be a normed space and $l \in Z'$ a linear functional on Z. Then l is called **a bounded linear functional on** Z if there is an $M < +\infty$ so that*

$$|l(u)| \leq M \|u\|$$

for every $u \in Z$.

Theorem 11.1 *Let $(Z, \|\cdot\|)$ be a normed space and $l \in Z'$. The following are equivalent.*
(i) l is bounded.
(ii) $l : Z \to F$ is continuous on Z.
(iii) $l : Z \to F$ is continuous at $o \in Z$.

Proof: Suppose that l is bounded and, hence, there is an $M < +\infty$ so that $|l(u)| \leq M\|u\|$ for every $u \in Z$. If $u_n \to u$ in Z, then $|l(u_n) - l(u)| = |l(u_n - u)| \leq M\|u_n - u\| \to 0$ as $n \to +\infty$ and, thus, $l(u_n) \to l(u)$ in F as $n \to +\infty$. This says that l is continuous on Z.

If l is continuous on Z, then it is certainly continuous at $o \in Z$.

Suppose that l is continuous at $o \in Z$. Then, for $\epsilon = 1$ there exists a $\delta > 0$ so that $|l(u)| = |l(u) - l(o)| < 1$ for every $u \in Z$ with $\|u\| = \|u - o\| < \delta$. We take an arbitrary $u \in Z \setminus \{o\}$ and an arbitrary $t > 1$ and have that $\left\|\frac{\delta}{t\|u\|}\, u\right\| = \frac{\delta}{t} < \delta$. Therefore, $\left|l\left(\frac{\delta}{t\|u\|}\, u\right)\right| < 1$, implying that $|l(u)| \leq \frac{t}{\delta}\,\|u\|$. This is trivially true also for $u = o$ and we conclude that $|l(u)| \leq \frac{t}{\delta}\,\|u\|$ for every $u \in Z$. For the arbitrary $u \in Z$, letting $t \to 1+$, we get $|l(u)| \leq M\|u\|$, where $M = \frac{1}{\delta}$. This says that l is bounded.

Definition 11.5 *Let $(Z, \|\cdot\|)$ 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.*

Proposition 11.4 *Let $(Z, \|\cdot\|)$ be a normed space and l a bounded linear functional on Z. Then there is a smallest M 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 $|l(u)| > m\|u\|$.

Proof: We consider

$$M_0 = \inf\{M \mid |l(u)| \leq M\|u\| \text{ for every } u \in Z\}.$$

The set $L = \{M \mid |l(u)| \leq M\|u\| \text{ for every } u \in Z\}$ 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 $|l(u)| \leq M_n\|u\|$ for every $u \in Z$, we find $|l(u)| \leq M_0\|u\|$ for every $u \in Z$.

If $m < M_0$, then $m \notin L$ and, hence, there is a $u \in Z$ so that $|l(u)| > m\|u\|$.

Definition 11.6 *Let $(Z, \|\cdot\|)$ be a normed space and l a bounded linear functional on Z. The smallest M with the property that $|l(u)| \leq M\|u\|$ for every $u \in Z$ is called **the norm of** l and it is denoted by $\|l\|_*$.*

Proposition 11.4, which proves the existence of $\|l\|_*$, states also its characterizing properties:

1. $|l(u)| \leq \|l\|_* \|u\|$ for every $u \in Z$,

2. for every $m < \|l\|_*$ there is a $u \in Z$ so that $|l(u)| > m\|u\|$.

The zero linear functional $o : Z \to F$ is bounded and, since $|o(u)| = 0 \leq 0\|u\|$ for every $u \in Z$, we have that

$$\|o\|_* = 0.$$

On the other hand, if $l \in Z^*$ has $\|l\|_* = 0$, then $|l(u)| \leq 0\|u\| = 0$ for every $u \in Z$ and, hence, $l = o$ is the zero linear functional on Z.

Proposition 11.5 *Let $(Z, \|\cdot\|)$ be a normed space and $l \in Z^*$. Then*

$$\|l\|_* = \sup_{u \in Z, u \neq o} \frac{|l(u)|}{\|u\|} = \sup_{u \in Z, \|u\|=1} |l(u)| = \sup_{u \in Z, \|u\| \leq 1} |l(u)| \,.$$

Proof: Every u with $\|u\| = 1$ satisfies $\|u\| \leq 1$. Therefore, $\sup_{u \in Z, \|u\|=1} |l(u)| \leq \sup_{u \in Z, \|u\| \leq 1} |l(u)|$.

Writing $v = \frac{u}{\|u\|}$ for every $u \in Z \setminus \{o\}$, we have that $\|v\| = 1$. Therefore, $\sup_{u \in Z, u \neq o} \frac{|l(u)|}{\|u\|} = \sup_{u \in Z, u \neq o} \left| l\left(\frac{u}{\|u\|}\right)\right| \leq \sup_{u \in Z, \|u\|=1} |l(u)|$.

For every u with $\|u\| \leq 1$, we have $|l(u)| \leq \|l\|_* \|u\| \leq \|l\|_*$ and, thus, $\sup_{u \in Z, \|u\| \leq 1} |l(u)| \leq \|l\|_*$.

If we set $M = \sup_{u \in Z, u \neq o} \frac{|l(u)|}{\|u\|}$, then $\frac{|l(u)|}{\|u\|} \leq M$ and, hence, $|l(u)| \leq M\|u\|$ for all $u \neq o$. Since this is obviously true for $u = o$, we have that $\|l\|_* \leq M$ and this finishes the proof.

Proposition 11.6 *Let $(Z, \|\cdot\|)$ 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*

$$\|l_1 + l_2\|_* \leq \|l_1\|_* + \|l_2\|_* \,, \qquad \|\kappa l\|_* = |\kappa| \|l\|_* \,.$$

Proof: We have that $|(l_1 + l_2)(u)| \leq |l_1(u)| + |l_2(u)| \leq \|l_1\|_* \|u\| + \|l_2\|_* \|u\| = (\|l_1\|_* + \|l_2\|_*)\|u\|$ for every $u \in Z$. This implies that $l_1 + l_2$ is bounded and that $\|l_1 + l_2\|_* \leq \|l_1\|_* + \|l_2\|_*$.

Similarly, $|(\kappa l)(u)| = |\kappa||l(u)| \leq |\kappa|\|l\|_* \|u\|$ for every $u \in Z$. This implies that κl is bounded and that $\|\kappa l\|_* \leq |\kappa|\|l\|_*$. If $\kappa = 0$, then the equality is obvious. If $\kappa \neq 0$, to get the opposite inequality, we write $|\kappa||l(u)| = |(\kappa l)(u)| \leq \|\kappa l\|_* \|u\|$. This implies that $|l(u)| \leq \frac{\|\kappa l\|_*}{|\kappa|} \|u\|$ for every $u \in Z$ and, hence, that $\|l\|_* \leq \frac{\|\kappa l\|_*}{|\kappa|}$.

Proposition 11.6 together with the remarks about the norm of the zero functional imply that Z^* is a linear subspace of Z' and that $\|\cdot\|_* : Z^* \to \mathbf{R}$ is a norm on Z^*.

Theorem 11.2 *If $(Z, \|\cdot\|)$ is a normed space, then $(Z^*, \|\cdot\|_*)$ is a Banach space.*

Proof: Let (l_n) be a Cauchy sequence in Z^*. For all $u \in Z$, $|l_n(u) - l_m(u)| = |(l_n - l_m)(u)| \leq \|l_n - l_m\|_* \|u\| \to 0$ 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

$$l(u) = \lim_{n \to +\infty} l_n(u)$$

for every $u \in Z$.

For every $u, v \in Z$ and $\kappa \in F$ we have $l(u+v) = \lim_{n \to +\infty} l_n(u+v) = \lim_{n \to +\infty} l_n(u) + \lim_{n \to +\infty} l_n(v) = l(u) + l(v)$ and $l(\kappa u) = \lim_{n \to +\infty} l_n(\kappa u) = \kappa \lim_{n \to +\infty} l_n(u) = \kappa l(u)$. Therefore, $l \in Z'$.

There is N so that $\|l_n - l_m\|_* \leq 1$ for all $n, m \geq N$. This implies that $|l_n(u) - l_m(u)| \leq \|l_n - l_m\|_* \|u\| \leq \|u\|$ for all $u \in Z$ and all $n, m \geq N$ and, taking the limit as $n \to +\infty$ and, taking $m = N$, we find $|l(u) - l_N(u)| \leq \|u\|$ for all $u \in Z$. Therefore, $|l(u)| \leq |l_N(u)| + \|u\| \leq (\|l_N\|_* + 1)\|u\|$ for every $u \in Z$ and, hence, $l \in Z^*$.

For an arbitrary $\epsilon > 0$ there is N so that $\|l_n - l_m\|_* \leq \epsilon$ for all $n, m \geq N$. This implies $|l_n(u) - l_m(u)| \leq \|l_n - l_m\|_* \|u\| \leq \epsilon \|u\|$ for all $u \in Z$ and all $n, m \geq N$ and, taking the limit as $m \to +\infty$, we find $|l_n(u) - l(u)| \leq \epsilon \|u\|$ for all $u \in Z$ and all $n \geq N$. Therefore, $\|l_n - l\|_* \leq \epsilon$ for all $n \geq N$ and, hence, $l_n \to l$ in Z^*.

Definition 11.7 *Let Z and W be two linear spaces over the same F and a function $T : Z \to W$. T is called* ***a linear transformation*** *or* ***a linear operator from Z to W*** *if*

$$T(u+v) = T(u) + T(v), \qquad T(\kappa u) = \kappa T(u)$$

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 transformation. T is one-to-one if and only if $T(u) = o$ (the zero element of W) implies $u = o$ (the zero element of Z). The subset $N(T) = \{u \in Z \,|\, T(u) = o\}$ 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.

The linear transformation $T : Z \to W$ is one-to-one if and only if $N(T) = \{o\}$ and T is onto if and only if $R(T) = W$.

If the linear transformation $T : Z \to W$ is one-to-one and onto, then the inverse function $T^{-1} : W \to Z$ is also a linear transformation. 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))*: 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 the linear transformation $T : Z \to W$ is one-to-one but not onto, then we may consider the restriction $T : Z \to R(T)$. This is a linear transformation which is one-to-one and onto and, thus, 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 11.8 *Let $(Z, \|\cdot\|_Z)$ and $(W, \|\cdot\|_W)$ be two normed spaces and a linear transformation $T : Z \to W$. We say that T is **a bounded linear transformation from Z to W** if there exists an $M < +\infty$ so that*

$$\|T(u)\|_W \leq M\|u\|_Z$$

for all $u \in Z$.

Theorem 11.3 *Let $(Z, \|\cdot\|_Z)$ and $(W, \|\cdot\|_W)$ be two normed spaces and a linear transformation $T : Z \to W$. The following are equivalent.*
(i) T is bounded.
(ii) $T : Z \to W$ is continuous on Z.
(iii) $T : Z \to W$ is continuous at $o \in Z$.

Proof: Suppose that T is bounded and, hence, there is an $M < +\infty$ so that $\|T(u)\|_W \leq M\|u\|_Z$ for every $u \in Z$. If $u_n \to u$ in Z, then $\|T(u_n) - T(u)\|_W = \|T(u_n - u)\|_W \leq M\|u_n - u\|_Z \to 0$ as $n \to +\infty$ and, thus, $T(u_n) \to T(u)$ in W as $n \to +\infty$. This says that T is continuous on Z.

If T is continuous on Z, then it is certainly continuous at $o \in Z$.

Suppose that T is continuous at $o \in Z$. Then, for $\epsilon = 1$ there exists a $\delta > 0$ so that $\|T(u)\|_W = \|T(u) - T(o)\|_W < 1$ for every $u \in Z$ with $\|u\|_Z = \|u - o\|_Z < \delta$. We take an arbitrary $u \in Z \setminus \{o\}$ and an arbitrary $t > 1$ and have that $\big\|\frac{\delta}{t\|u\|_Z}\, u\big\|_Z = \frac{\delta}{t} < \delta$. Therefore, $\big\|T\big(\frac{\delta}{t\|u\|_Z}\, u\big)\big\|_W < 1$, implying that $\|T(u)\|_W \leq \frac{t}{\delta}\,\|u\|_Z$. This is trivially true also for $u = o$ and, hence, $\|T(u)\|_W \leq \frac{t}{\delta}\,\|u\|_Z$ for all $u \in Z$. Letting $t \to 1+$, we find $\|T(u)\|_W \leq M\|u\|_Z$, where $M = \frac{1}{\delta}$. This says that T is bounded.

Proposition 11.7 *Let $(Z, \|\cdot\|_Z)$ and $(W, \|\cdot\|_W)$ be two normed spaces and a bounded linear transformation $T : Z \to W$. Then there is a smallest M with the property: $\|T(u)\|_W \leq M\|u\|_Z$ for every $u \in Z$. This M_0 is characterized by the two properties:*
(i) $\|T(u)\|_W \leq M_0\|u\|_Z$ for every $u \in Z$,
(ii) for every $m < M_0$ there is a $u \in Z$ so that $\|T(u)\|_W > m\|u\|_Z$.

Proof: We consider

$$M_0 = \inf\{M \mid \|T(u)\|_W \leq M\|u\|_Z \text{ for every } u \in Z\}.$$

The set $L = \{M \mid \|T(u)\|_W \leq M\|u\|_Z \text{ for every } u \in Z\}$ 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 $\|T(u)\|_W \leq M_n\|u\|_Z$ for every $u \in Z$, we find $\|T(u)\|_W \leq M_0\|u\|_Z$ for every $u \in Z$.

If $m < M_0$, then $m \notin L$ and, hence, there is a $u \in Z$ so that $\|T(u)\|_W > m\|u\|_Z$.

Definition 11.9 *Let $(Z, \|\cdot\|_Z)$ and $(W, \|\cdot\|_W)$ be two normed spaces and a bounded linear transformation $T : Z \to W$. The smallest M with the property that $\|T(u)\|_W \leq M\|u\|_Z$ for every $u \in Z$ is called* ***the norm of*** *T and it is denoted by $\|T\|$.*

By Proposition 11.7, which proves the existence of $\|T\|$, we have:

1. $\|T(u)\|_W \leq \|T\|\|u\|_Z$ for every $u \in Z$,
2. for every $m < \|T\|$ there is a $u \in Z$ so that $\|T(u)\|_W > m\|u\|_Z$.

The zero linear transformation $o : Z \to W$ is bounded and, since $\|o(u)\|_W = 0 \leq 0\|u\|_Z$ for every $u \in Z$, we have that

$$\|o\| = 0.$$

On the other hand, if T is a bounded linear transformation with $\|T\| = 0$, then $\|T(u)\|_W \leq 0\|u\|_Z = 0$ for every $u \in Z$ and, hence, $T = o$ is the zero linear transformation.

Proposition 11.8 *Let $(Z, \|\cdot\|_Z)$ and $(W, \|\cdot\|_W)$ be two normed spaces and a bounded linear transformation $T : Z \to W$. Then*

$$\|T\| = \sup_{u\in Z, u\neq o} \frac{\|T(u)\|_W}{\|u\|_Z} = \sup_{u\in Z, \|u\|_Z=1} \|T(u)\|_W = \sup_{u\in Z, \|u\|_Z\leq 1} \|T(u)\|_W .$$

Proof: Every u with $\|u\|_Z = 1$ satisfies $\|u\|_Z \leq 1$. This, clearly, implies that $\sup_{u\in Z,\|u\|_Z=1} \|T(u)\|_W \leq \sup_{u\in Z,\|u\|_Z\leq 1} \|T(u)\|_W$.

Writing $v = \frac{u}{\|u\|_Z}$ for every $u \in Z \setminus \{o\}$, we have that $\|v\|_Z = 1$. Therefore, $\sup_{u\in Z,u\neq o} \frac{\|T(u)\|_W}{\|u\|_Z} = \sup_{u\in Z,u\neq o} \big\|T\big(\frac{u}{\|u\|_Z}\big)\big\|_W \leq \sup_{u\in Z,\|u\|_Z=1} \|T(u)\|_W$.

For every u with $\|u\|_Z \leq 1$, we have $\|T(u)\|_W \leq \|T\|\|u\|_Z \leq \|T\|$ and, thus, $\sup_{u\in Z,\|u\|_Z\leq 1} \|T(u)\|_W \leq \|T\|$.

If we set $M = \sup_{u\in Z,u\neq o} \frac{\|T(u)\|_W}{\|u\|_Z}$, then $\frac{\|T(u)\|_W}{\|u\|_Z} \leq M$ and this implies $\|T(u)\|_W \leq M\|u\|_Z$ for all $u \neq o$. Since this is obviously true for $u = o$, we have that $\|T\| \leq M$ and this finishes the proof.

Definition 11.10 *Let $(Z, \|\cdot\|_Z)$ and $(W, \|\cdot\|_W)$ be two normed spaces and a bounded linear transformation $T : Z \to W$.*

If T is onto W and $\|T(u)\|_W = \|u\|_Z$ for every $u \in Z$, then we say that T is an ***isometry from Z onto W*** *or an* ***isometry between Z and W.***

If $\|T(u)\|_W = \|u\|_Z$ for every $u \in Z$ (but T is not necessarily onto W), we say that T is an ***isometry from Z into W.***

Proposition 11.9 *Let $(Z, \|\cdot\|_Z)$ and $(W, \|\cdot\|_W)$ 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 also an isometry from W onto Z.

Proof: (i) If $T(u) = T(v)$, then $0 = \|T(u) - T(v)\|_W = \|T(u-v)\|_W = \|u-v\|_Z$ and, hence, $u = v$.
(ii) From (i) we have that T is one-to-one and, thus, the inverse mapping $T^{-1} : W \to Z$ exists. If $w, w_1, w_2 \in W$ and $\kappa \in F$, we take the (unique) $u, u_1, u_2 \in Z$ so that $T(u) = w, T(u_1) = w_1$ and $T(u_2) = w_2$. Then $T(u_1 + u_2) = T(u_1) + T(u_2) = w_1 + w_2$ and, hence, $T^{-1}(w_1 + w_2) = u_1 + u_2 = T^{-1}(w_1) + T^{-1}(w_2)$. Also, $T(\kappa u) = \kappa T(u) = \kappa w$ and, hence, $T^{-1}(\kappa w) = \kappa u = \kappa T^{-1}(w)$. These imply that $T^{-1} : W \to Z$ is a linear transformation.

Moreover, $\|T^{-1}(w)\|_Z = \|u\|_Z = \|T(u)\|_W = \|w\|_W$. Therefore, T^{-1} is an isometry from W onto Z.

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 after Definition 11.7) but we may also identify them as *metric* spaces: the distances between elements are unaffected by changing their ((names)): $\|T(u) - T(v)\|_W = \|T(u-v)\|_W = \|u-v\|_Z$.

If T is an isometry from Z into W, then, clearly, T is an isometry from Z onto $R(T)$ and now we may identify Z with the subspace $R(T)$ of W or we may view $R(T)$ as a ((copy)) of Z inside W.

11.2 The spaces $L^p(X, \Sigma, \mu)$.

In this whole section and the next, (X, Σ, μ) will be a fixed measure space.

Definition 11.11 *If $0 < p < +\infty$, we define the space $\mathcal{L}^p(X, \Sigma, \mu)$ to be the set of all measurable functions $f : X \to \overline{F}$, where $F = \mathbf{R}$ or $F = \mathbf{C}$, with*

$$\int_X |f|^p \, d\mu < +\infty.$$

Observe that the space $\mathcal{L}^1(X, \Sigma, \mu)$ is the set of all functions which are integrable over X with respect to μ.

Whenever any of X, Σ, μ is uniquelly 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 11.10 *$\mathcal{L}^p$ is a linear space over F.*

Proof: We shall use the trivial inequality

$$(a + b)^p \leq 2^p(a^p + b^p), \qquad 0 \leq a, b.$$

This can be proved by $(a+b)^p \leq (2 \max\{a, b\})^p = 2^p \max\{a^p, b^p\} \leq 2^p(a^p + b^p)$.

Suppose that $f_1, f_2 \in \mathcal{L}^p$. Then both f_1 and f_2 are finite a.e. on X and, hence, $f_1 + f_2$ is defined a.e. on X. If $f_1 + f_2$ is any measurable definition of $f_1 + f_2$, then, using the above elementary inequality, $|(f_1 + f_2)(x)|^p \leq 2^p(|f_1(x)|^p + |f_2(x)|^p)$ for a.e. $x \in X$ and, hence,

$$\int_X |f_1 + f_2|^p \, d\mu \leq 2^p \int_X |f_1|^p \, d\mu + 2^p \int_X |f_2|^p \, d\mu < +\infty.$$

Therefore $f_1 + f_2 \in \mathcal{L}^p$.

If $f \in \mathcal{L}^p$ and $\kappa \in F$, then

$$\int_X |\kappa f|^p \, d\mu = |\kappa|^p \int_X |f|^p \, d\mu < +\infty.$$

Therefore, $\kappa f \in \mathcal{L}^p$.

Definition 11.12 *Let $f : X \to \overline{F}$ be measurable. We say that f is **essentially bounded over X (with respect to μ)** if there is $M < +\infty$ so that $|f| \leq M$ a.e. on X.*

Proposition 11.11 *Let $f : X \to \overline{F}$ be measurable. If f is essentially bounded over 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 $A = \{M \,|\, |f| \leq M \text{ a.e. on } X\}$ and the

$$M_0 = \inf\{M \,|\, |f| \leq M \text{ a.e. on } X\}.$$

The set A is non-empty by assumption and is included in $[0, +\infty)$ and, hence, M_0 exists.

We take a sequence (M_n) in A with $M_n \to M_0$. From $M_n \in A$, we find $\mu(\{x \in X \,|\, |f(x)| > M_n\}) = 0$ for every n and, since $\{x \in X \,|\, |f(x)| > M_0\} = \cup_{n=1}^{+\infty}\{x \in X \,|\, |f(x)| > M_n\}$, we conclude that $\mu(\{x \in X \,|\, |f(x)| > M_0\}) = 0$. Therefore, $|f| \leq M_0$ a.e. on X.

If $m < M_0$, then $m \notin A$ and, hence, $\mu(\{x \in X \,|\, |f(x)| > m\}) > 0$.

Definition 11.13 *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) instead of ess-sup$_{X,\mu}(f)$.

The ess-sup(f) is characterized by the properties:

1. $|f| \leq$ ess-sup(f) a.e. on X,

2. for every $m <$ ess-sup(f), we have $\mu(\{x \in X \,|\, |f(x)| > m\}) > 0$.

Definition 11.14 *We define $\mathcal{L}^\infty(X, \Sigma, \mu)$ to be the set of all measurable functions $f : X \to \overline{F}$ which are essentially bounded over X.*

Proposition 11.12 *$\mathcal{L}^\infty$ is a linear space over F.*

Proof: If $f_1, f_2 \in \mathcal{L}^\infty$, then there are sets $A_1, A_2 \in \Sigma$ 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 $|f_1 + f_2| \le |f_1| + |f_2| \le$ ess-sup(f_1)+ess-sup(f_2) on A. Hence $f_1 + f_2$ is essentially bounded over X and

$$\text{ess-sup}(f_1 + f_2) \le \text{ ess-sup}(f_1) + \text{ess-sup}(f_2).$$

If $f \in \mathcal{L}^\infty$ and $\kappa \in F$, then there is $A \in \Sigma$ with $\mu(A^c) = 0$ so that $|f| \le$ ess-sup(f) on A. We, now, have $|\kappa f| \le |\kappa|$ess-sup(f) on A. Hence κf is essentially bounded over X and ess-sup$(\kappa f) \le |\kappa|$ess-sup(f). If $\kappa = 0$, this inequality, obviously, becomes equality. If $\kappa \ne 0$, we apply the same inequality to $\frac{1}{\kappa}$ and κf and get ess-sup$(f) =$ ess-sup$(\frac{1}{\kappa}(\kappa f)) \le \frac{1}{|\kappa|}$ ess-sup(κf). Therefore

$$\text{ess-sup}(\kappa f) = |\kappa|\text{ess-sup}(f).$$

Definition 11.15 *Let $1 \le p \le +\infty$. We define*

$$p' = \begin{cases} \frac{p}{p-1}\,, & \text{if } 1 < p < +\infty \\ +\infty\,, & \text{if } p = 1 \\ 1\,, & \text{if } p = +\infty. \end{cases}$$

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 $1 \le p' \le +\infty$ and p is the conjugate of p'. Moreover, p, p' are related by the symmetric equality

$$\frac{1}{p} + \frac{1}{p'} = 1.$$

Lemma 11.1 *Let $0 < t < 1$. For every $a, b \ge 0$ we have*

$$a^t b^{1-t} \le ta + (1-t)b.$$

Proof: If $b = 0$ the inequality is obviously true: $0 \le ta$.

If $b > 0$, the inequality is equivalent to $(\frac{a}{b})^t \le t\frac{a}{b} + 1 - t$ and, setting $x = \frac{a}{b}$, it is equivalent to $x^t \le tx + 1 - t$, $0 \le x$. To prove it we form the function $f(x) = x^t - tx$ on $[0, +\infty)$ and we easily see that it is increasing in $[0, 1]$ and decreasing in $[1, +\infty)$. Therefore, $f(x) \le f(1) = 1 - t$ for all $x \in [0, +\infty)$.

Theorem 11.4 ***(Hölder's inequalities)*** *Let $1 \le p, p' \le +\infty$ 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*

$$\int_X |fg|\, d\mu \le \Big(\int_X |f|^p\, d\mu\Big)^{\frac{1}{p}} \Big(\int_X |g|^{p'}\, d\mu\Big)^{\frac{1}{p'}}, \qquad 1 < p, p' < +\infty,$$

$$\int_X |fg|\, d\mu \le \int_X |f|\, d\mu \;\cdot \textit{ess-sup}(g)\,, \qquad p = 1, p' = +\infty,$$

$$\int_X |fg|\, d\mu \le \textit{ess-sup}(f) \int_X |g|\, d\mu\,, \qquad p = +\infty, p' = 1.$$

Proof: (a) 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. It becomes 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 11.1 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 have that

$$\frac{|fg|}{A^{\frac{1}{p}} B^{\frac{1}{p'}}} \leq \frac{1}{p} \frac{|f|^p}{A} + \frac{1}{p'} \frac{|g|^{p'}}{B}$$

a.e. on X. Integrating, we find

$$\frac{1}{A^{\frac{1}{p}} B^{\frac{1}{p'}}} \int_X |fg| \, d\mu \leq \frac{1}{p} + \frac{1}{p'} = 1$$

and this implies the inequality we wanted to prove.

(b) 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.

(c) The proof in the case $p = +\infty$, $p' = 1$ is the same as in (b).

Theorem 11.5 ***(Minkowski's inequalities)*** *Let $1 \leq p \leq +\infty$. If $f_1, f_2 \in \mathcal{L}^p$, then*

$$\Big(\int_X |f_1 + f_2|^p \, d\mu \Big)^{\frac{1}{p}} \leq \Big(\int_X |f_1|^p \, d\mu \Big)^{\frac{1}{p}} + \Big(\int_X |f_2|^p \, d\mu \Big)^{\frac{1}{p}}, \qquad 1 \leq p < +\infty,$$

$$\textit{ess-sup}(f_1 + f_2) \leq \textit{ess-sup}(f_1) + \textit{ess-sup}(f_2), \qquad p = +\infty.$$

Proof: The case $p = +\infty$ is included in the proof of Proposition 11.12. 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

$$\begin{aligned} \int_X |f_1 + f_2|^p \, d\mu \; &\leq \; \Big(\int_X |f_1|^p \, d\mu \Big)^{\frac{1}{p}} \Big(\int_X |f_1 + f_2|^{(p-1)p'} \, d\mu \Big)^{\frac{1}{p'}} \\ &\quad + \Big(\int_X |f_2|^p \, d\mu \Big)^{\frac{1}{p}} \Big(\int_X |f_1 + f_2|^{(p-1)p'} \, d\mu \Big)^{\frac{1}{p'}} \\ &= \Big(\int_X |f_1|^p \, d\mu \Big)^{\frac{1}{p}} \Big(\int_X |f_1 + f_2|^p \, d\mu \Big)^{\frac{1}{p'}} \\ &\quad + \Big(\int_X |f_2|^p \, d\mu \Big)^{\frac{1}{p}} \Big(\int_X |f_1 + f_2|^p \, d\mu \Big)^{\frac{1}{p'}}. \end{aligned}$$

Simplifying, we get the inequality we want to prove.

Definition 11.16 *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)* ***converges to f in the p-mean*** *if*

$$\int_X |f_n - f|^p \, d\mu \to 0, \qquad 1 \leq p < +\infty,$$

$$\text{ess-sup}(f_n - f) \to 0, \qquad p = +\infty$$

as $n \to +\infty$. We say that (f_n) is Cauchy in the p-mean if

$$\int_X |f_n - f_m|^p \, d\mu \to 0, \qquad 1 \leq p < +\infty,$$

$$\text{ess-sup}(f_n - f_m) \to 0, \qquad p = +\infty$$

as $n, m \to +\infty$.

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, $\big(\int_X |f_n - f_m|^p \, d\mu\big)^{\frac{1}{p}} \leq \big(\int_X |f_n - f|^p \, d\mu\big)^{\frac{1}{p}} + \big(\int_X |f_m - f|^p \, d\mu\big)^{\frac{1}{p}} \to 0$ 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 11.6 is an extension of Theorem 9.1.

Theorem 11.6 *If (f_n) is Cauchy in the p-mean, then there is $f \in \mathcal{L}^p$ so that (f_n) converges to f in the p-mean. Moreover, there is a subsequence (f_{n_k}) which converges to f a.e. on X.*

As a corollary: if (f_n) converges to f in the p-mean, there is a subsequence (f_{n_k}) which converges to f a.e. on X.

Proof: (a) We consider first the case $1 \leq p < +\infty$.
First proof. Since each f_n is finite a.e. on X, there is $A \in \Sigma$ so that $\mu(A^c) = 0$ and all f_n are finite on A.

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 $n, m \geq n_k$. 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

$$G = \begin{cases} \sum_{k=1}^{+\infty} |f_{n_{k+1}} - f_{n_k}|, & \text{on } A \\ 0, & \text{on } A^c \end{cases}.$$

If

$$G_K = \begin{cases} \sum_{k=1}^{K-1} |f_{n_{k+1}} - f_{n_k}|, & \text{on } A \\ 0, & \text{on } A^c \end{cases},$$

then $\left(\int_X G_K^p\,d\mu\right)^{\frac{1}{p}} \leq \sum_{k=1}^{K-1}\left(\int_X |f_{n_{k+1}} - f_{n_k}|^p\,d\mu\right)^{\frac{1}{p}} < 1$, by Minkowski's inequality. Since $G_K \uparrow G$ on X, we find that $\int_X G^p\,d\mu \leq 1$ 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. $x \in A$. Therefore, there is a $B \in \Sigma$, $B \subseteq A$ so that $\mu(A \setminus B) = 0$ and $\sum_{k=1}^{+\infty}(f_{n_{k+1}}(x) - f_{n_k}(x))$ converges for every $x \in B$. 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}) converges to f a.e. on X.

We, also, have on B that $|f_{n_K} - f| = |f_{n_K} - f_{n_1} - \sum_{k=1}^{+\infty}(f_{n_{k+1}} - f_{n_k})| = |\sum_{k=1}^{K-1}(f_{n_{k+1}} - f_{n_k}) - \sum_{k=1}^{+\infty}(f_{n_{k+1}} - f_{n_k})| \leq \sum_{k=K}^{+\infty}|f_{n_{k+1}} - f_{n_k}| \leq G$ for every K and, hence, $|f_{n_K} - f|^p \leq G^p$ 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 $n_k \to +\infty$, we get $\left(\int_X |f_k - f|^p\,d\mu\right)^{\frac{1}{p}} \leq \left(\int_X |f_k - f_{n_k}|^p\,d\mu\right)^{\frac{1}{p}} + \left(\int_X |f_{n_k} - f|^p\,d\mu\right)^{\frac{1}{p}} \to 0$ as $k \to +\infty$ and we conclude that (f_n) converges to f in the p-mean.

Second proof. For every $\epsilon > 0$ we have that $\mu(\{x \in X \mid |f_n(x) - f_m(x)| \geq \epsilon\}) \leq \frac{1}{\epsilon}\left(\int_X |f_n - f_m|^p\,d\mu\right)^{\frac{1}{p}}$ and, hence, (f_n) is Cauchy in measure on X. Theorem 9.2 implies that there is a subsequence (f_{n_k}) which converges to some f a.e. on X.

Now, for every $\epsilon > 0$ there is an N so that $\int_X |f_n - f_m|^p\,d\mu \leq \epsilon$ for all $n, m \geq N$. 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

$$\int_X |f_n - f|^p\,d\mu \leq \liminf_{k\to+\infty}\int_X |f_n - f_{n_k}|^p\,d\mu \leq \epsilon$$

for all $n \geq N$. This, of course, says that (f_n) converges to f in the p-mean.

(b) Now, let $p = +\infty$.

For each n, m we have a set $A_{n,m} \in \Sigma$ with $\mu(A_{n,m}^c) = 0$ and $|f_n - f_m| \leq$ ess-sup$(f_n - f_m)$ on $A_{n,m}$. We form the set $A = \cap_{1\leq n,m} A_{n,m}$ and have that $\mu(A^c) = 0$ and $|f_n - f_m| \leq$ 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) converges to f uniformly on A. Now,

$$\text{ess-sup}(f_n - f) \leq \sup_{x\in A}|f_n(x) - f(x)| \to 0$$

as $n \to +\infty$.

If, for every $f \in \mathcal{L}^p$, we set

$$N_p(f) = \begin{cases} \left(\int_X |f|^p \, d\mu \right)^{\frac{1}{p}}, & \text{if } 1 \leq p < +\infty \\ \text{ess-sup}(f), & \text{if } p = +\infty, \end{cases}$$

then, Propositions 11.10 and 11.12 and Theorem 11.5 imply that the function $N_p : \mathcal{L}^p \to \mathbf{R}$ satisfies

1. $N_p(f_1 + f_2) \leq N_p(f_1) + N_p(f_2)$,
2. $N_p(\kappa f) = |\kappa| N_p(f)$

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 11.17 *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 11.13 *The relation $\sim$ on $\mathcal{L}^p$ is an equivalence relation.*

Proof: It is obvious that $f \sim f$ and that, if $f_1 \sim f_2$, then $f_2 \sim f_1$. Now, if $f_1 \sim f_2$ and $f_2 \sim f_3$, then there are $A, B \in \Sigma$ with $\mu(A^c) = \mu(B^c) = 0$ so that $f_1 = f_2$ on A and $f_2 = f_3$ on B. This implies that $\mu((A \cap B)^c) = 0$ and $f_1 = f_3$ on $A \cap B$ and, hence, $f_1 \sim f_3$.

As with 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 $f' \in \mathcal{L}^p$ which are equivalent to f:

$$[f] = \{f' \in \mathcal{L}^p \,|\, f' \sim f\}.$$

Proposition 11.14 *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: (i) Assume $f_1 \sim f_2$. If $f \in [f_1]$, then $f \sim f_1$. Therefore, $f \sim f_2$ and, hence, $f \in [f_2]$. Symmetrically, if $f \in [f_2]$, then $f \in [f_1]$ and, thus, $[f_1] = [f_2]$.

If $[f_1] = [f_2]$, then $f_1 \in [f_1]$ and, hence, $f_1 \in [f_2]$. Therefore, $f_1 \sim f_2$.
(ii) If $f \in [f_1]$ and $f \in [f_2]$, then $f \sim f_1$ and $f \sim f_2$ and, hence, $f_1 \sim f_2$. This, by the result of (i), implies $[f_1] = [f_2]$.

For the last statement, we observe that every $f \in \mathcal{L}^p$ belongs to $[f]$.

Proposition 11.14 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 11.18 *We define*

$$L^p(X,\Sigma,\mu)=\mathcal{L}^p(X,\Sigma,\mu)/_\sim=\{[f]\,|\,f\in\mathcal{L}^p(X,\Sigma,\mu)\}.$$

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 11.15 *Let $f, f_1, f_2, f', f_1', f_2' \in \mathcal{L}^p$ and $\kappa \in F$.*
(i) If $f_1 \sim f_1'$ and $f_2 \sim f_2'$, then $f_1 + f_2 \sim f_1' + f_2'$.
(ii) If $f \sim f'$, then $\kappa f \sim \kappa f'$.

Proof: (i) There are $A_1, A_2 \in \Sigma$ with $\mu(A_1^c) = \mu(A_2^c) = 0$ so that $f_1 = f_1'$ on A_1 and both f_1, f_1' are finite on A_1 and, also, $f_2 = f_2'$ on A_2 and both f_2, f_2' are finite on A_2. Then $\mu((A_1 \cap A_2)^c) = 0$ and $f_1 + f_2 = f_1' + f_2'$ on $A_1 \cap A_2$. Hence, $f_1 + f_2 \sim f_1' + f_2'$.
(ii) There is $A \in \Sigma$ with $\mu(A^c) = 0$ so that $f = f'$ on A. Then, $\kappa f = \kappa f'$ on A and, hence $\kappa f \sim \kappa f'$.

Because of Proposition 11.14, another way to state the results of Proposition 11.15 is:

1. $[f_1] = [f_1']$ and $[f_2] = [f_2']$ imply $[f_1 + f_1'] = [f_2 + f_2']$,
2. $[f] = [f']$ implies $[\kappa f] = [\kappa f']$.

These allow the following definition.

Definition 11.19 *We define addition and multiplication in L^p as follows:*

$$[f_1] + [f_2] = [f_1 + f_2], \qquad \kappa[f] = [\kappa f].$$

It is a matter of routine to prove, now, that the set L^p becomes a linear space under this addition and multiplication. Then L^p is a linear space over F.

The zero element of L^p is the equivalence class $[o]$ of the function o which is identically 0 on X. The opposite of an $[f]$ is the equivalence class $[-f]$.

The next task is to define a norm on L^p.

Proposition 11.16 *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*

$$\int_X |f_1|^p\,d\mu = \int_X |f_2|^p\,d\mu\,, \qquad 1 \le p < +\infty,$$
$$\textit{ess-sup}(f_1) = \textit{ess-sup}(f_2)\,, \qquad p = +\infty.$$

Proof: It is well known that $f_1 = f_2$ a.e. on X implies the first equality. Regarding the second equality, we have sets $B, A_1, A_2 \in \Sigma$ with $\mu(B^c) = \mu(A_1^c) = \mu(A_2^c) = 0$ so that $f_1 = f_2$ on B, $|f_1| \le \text{ess-sup}(f_1)$ on A_1 and $|f_2| \le \text{ess-sup}(f_2)$ on A_2. Then, the set $A = B \cap A_1 \cap A_2$ has $\mu(A^c) = 0$. Moreover, $|f_1| = |f_2| \le \text{ess-sup}(f_2)$ on A and, hence, $\text{ess-sup}(f_1) \le \text{ess-sup}(f_2)$. Also, $|f_2| = |f_1| \le \text{ess-sup}(f_1)$ on A and, hence, $\text{ess-sup}(f_2) \le \text{ess-sup}(f_1)$.

An equivalent way to state the result of Proposition 11.16 is

1. $[f_1] = [f_2]$ implies $\int_X |f_1|^p \, d\mu = \int_X |f_2|^p \, d\mu$, if $1 \le p < +\infty$,

2. $[f_1] = [f_2]$ implies ess-sup(f_1) = ess-sup(f_2), if $p = +\infty$.

These allow the

Definition 11.20 *We define, for every $[f] \in L^p$,*

$$\|[f]\|_p = N_p(f) = \begin{cases} \left(\int_X |f|^p \, d\mu \right)^{\frac{1}{p}}, & \text{if } 1 \le p < +\infty \\ \text{ess-sup}(f), & \text{if } p = +\infty. \end{cases}$$

Proposition 11.17 *The function $\| \cdot \|_p$ is a norm on L^p.*

Proof: We have $\|[f_1] + [f_2]\|_p = \|[f_1 + f_2]\|_p = N_p(f_1 + f_2) \le N_p(f_1) + N_p(f_2) = \|[f_1]\|_p + \|[f_2]\|_p$. 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, $f \sim o$ or, equivalently, $[f]$ is the zero element of L^p.

In order to simplify things and not have to carry 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 $\int_X fg \, d\mu$ for the element $f \in L^p$, we mean the integral $\int_X fg \, d\mu$ for the element-function $f \in \mathcal{L}^p$ *and*, at the same time, all integrals $\int_X f'g \, d\mu$ (equal to each other) for all functions $f' \in \mathcal{L}^p$ such that $f' = f$ 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 $\left(\int_X |f|^p \, d\mu \right)^{\frac{1}{p}}$, 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 $f' \in \mathcal{L}^p$ such that $f' = f$ 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 $g \in L^{p'}$.

Theorem 11.7 *All L^p are Banach spaces.*

Proof: Let (f_n) be a Cauchy sequence in L^p. Then $\|f_n - f_m\|_p \to 0$ and, hence, $\int_X |f_n - f_m|^p \, d\mu \to 0$, if $1 \leq p < +\infty$, and ess-sup$(f_n - f_m) \to 0$, if $p = +\infty$. Theorem 11.6 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 $1 \leq p < +\infty$, 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.

Definition 11.21 *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 $I = \mathbf{N}$, we denote $l^p = l^p(\mathbf{N})$.

If $1 \leq p < +\infty$, 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,

$$\sum_{i \in I} |b_i|^p < +\infty.$$

If $|b_i| = +\infty$ for at least one $i \in I$, then $\sum_{i \in I} |b_i|^p = +\infty$.

Definition 11.22 *Let I be an index set and $b : I \to F$. If $1 \leq p < +\infty$, we say that $b = \{b_i\}_{i \in I}$ is p-**summable** if $\sum_{i \in I} |b_i|^p < +\infty$.*

Hence, $b = \{b_i\}_{i \in I}$ is p-summable if and only if it belongs to $l^p(I)$. We also have

$$\|b\|_p = \big(\sum_{i \in I} |b_i|^p\big)^{\frac{1}{p}}.$$

When $1 \leq p < +\infty$, Minkowski's inequality becomes

$$\big(\sum_{i \in I} |b_i^{(1)} + b_i^{(2)}|^p\big)^{\frac{1}{p}} \leq \big(\sum_{i \in I} |b_i^{(1)}|^p\big)^{\frac{1}{p}} + \big(\sum_{i \in I} |b_i^{(2)}|^p\big)^{\frac{1}{p}}$$

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

$$\sum_{i \in I} |b_i c_i| \leq \big(\sum_{i \in I} |b_i|^p\big)^{\frac{1}{p}} \big(\sum_{i \in I} |c_i|^{p'}\big)^{\frac{1}{p'}}$$

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 over I with respect to $\sharp$ if and only if there is an $M < +\infty$ so that $|b_i| \leq M$ for all $i \in I$. It is obvious that the smallest M with the property that $|b_i| \leq M$ for all $i \in I$ is the $M_0 = \sup_{i \in I} |b_i|$.

Definition 11.23 *Let I be an index set and $b : I \to F$. We say that $b = \{b_i\}_{i \in I}$ is **bounded** if $\sup_{i \in I} |b_i| < +\infty$.*

Therefore, b is essentially bounded over 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

$$\sum_{i \in I} |b_i c_i| \le \sum_{i \in I} |b_i| \cdot \sup_{i \in I} |c_i|$$

for all summable $b = \{b_i\}_{i \in I}$ and all bounded $c = \{c_i\}_{i \in I}$.

The spaces $l^p(I)$ are all Banach spaces.

As we have already mentioned, a particular case is when $I = \mathbf{N}$. Then

$$l^p = \Big\{ x = (x_1, x_2, \dots) \mid \sum_{k=1}^{+\infty} |x_k|^p < +\infty \Big\}, \qquad 1 \le p < +\infty,$$

$$l^\infty = \big\{ x = (x_1, x_2, \dots) \mid \sup_{k \ge 1} |x_k| < +\infty \big\}, \qquad p = +\infty.$$

The corresponding norms are

$$\|x\|_p = \big(\sum_{k=1}^{+\infty} |x_k|^p \big)^{\frac{1}{p}}, \qquad 1 \le p < +\infty,$$

$$\|x\|_\infty = \sup_{k \ge 1} |x_k|, \qquad p = +\infty,$$

for every $x = (x_1, x_2, \dots) \in l^p$.

Another very special case is when $I = \{1, \dots, n\}$. In this case we have $l^p(I) = F^n$. The norms are

$$\|x\|_p = \Big(\sum_{k=1}^{n} |x_k|^p \Big)^{\frac{1}{p}}, \qquad 1 \le p < +\infty,$$

$$\|x\|_\infty = \max_{1 \le k \le n} |x_k|, \qquad p = +\infty,$$

for every $x = (x_1, \dots, x_n) \in F^n$.

11.3 The dual of $L^p(X, \Sigma, \mu)$.

In this section $p, p' \in [1, +\infty]$ are meant to be conjugate.

Theorem 11.8 *Let $g \in L^{p'}$. If $1 < p \leq +\infty$, then*

$$\|g\|_{p'} = \sup\Big\{\Big|\int_X fg\,d\mu\Big| \;\Big|\; f \in L^p, \|f\|_p \leq 1\Big\}.$$

If μ is semifinite, the same is true when $p = 1$.

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 have, by Hölder's inequality, that $|\int_X fg\,d\mu| \leq \|f\|_p\|g\|_{p'} \leq \|g\|_{p'}$. Therefore,

$$\sup\Big\{\Big|\int_X fg\,d\mu\Big| \;\Big|\; f \in L^p, \|f\|_p \leq 1\Big\} \leq \|g\|_{p'}.$$

If $\|g\|_{p'} = 0$, then the inequality between the sup and the $\|g\|_{p'}$, obviously, becomes equality. Anyway, we have $\int_X |g|^{p'}\,d\mu = 0$ and, hence, $g = 0$ a.e. on X. This implies that $\int_X fg\,d\mu = 0$ for every $f \in L^p$.

Now, let $\|g\|_{p'} > 0$. We consider the function f_0 defined by

$$f_0(x) = \begin{cases} \frac{|g(x)|^{p'-1}\overline{sign(g(x))}}{\|g\|_{p'}^{p'-1}}\,, & \text{if } g(x) \text{ is finite and } g(x) \neq 0, \\ 0\,, & \text{if } g(x) \text{ is infinite or } g(x) = 0. \end{cases}$$

Then,

$$f_0(x)g(x) = \begin{cases} \frac{|g(x)|^{p'}}{\|g\|_{p'}^{p'-1}}\,, & \text{if } g(x) \text{ is finite,} \\ 0\,, & \text{if } g(x) \text{ is infinite} \end{cases}$$

and, hence, $\int_X f_0 g\,d\mu = \frac{1}{\|g\|_{p'}^{p'-1}} \int_X |g|^{p'}\,d\mu = \|g\|_{p'}$.

If $1 < p, p' < +\infty$, then, since $p(p'-1) = p'$,

$$|f_0(x)|^p = \begin{cases} \frac{|g(x)|^{p'}}{\|g\|_{p'}^{p'}}\,, & \text{if } g(x) \text{ is finite,} \\ 0\,, & \text{if } g(x) \text{ is infinite} \end{cases}$$

and, hence, $\|f_0\|_p = \big(\int_X |f_0|^p\,d\mu\big)^{\frac{1}{p}} = 1$.

If $p = +\infty, p' = 1$, then

$$|f_0(x)| = \begin{cases} 1\,, & \text{if } g(x) \text{ is finite and } \neq 0, \\ 0\,, & \text{if } g(x) \text{ is infinite or } = 0 \end{cases}$$

and, hence, $\|f_0\|_\infty = \text{ess-sup}(f_0) = 1$.

We conclude that

$$\|g\|_{p'} = \max\Big\{\Big|\int_X fg\,d\mu\Big| \;\Big|\; f \in L^p, \|f\|_p \leq 1\Big\}.$$

(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,

$$\sup\Big\{\Big|\int_X fg\,d\mu\Big| \,|\, f \in L^1, \|f\|_1 \leq 1\Big\} \leq \|g\|_\infty.$$

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, thus, the inequality between the sup and the $\|g\|_\infty$ becomes equality.

Let $\|g\|_\infty > 0$. We consider an arbitrary ϵ with $0 < \epsilon < \|g\|_\infty$ and, then $\mu(\{x \in X \,|\, \|g\|_\infty - \epsilon < |g(x)| \leq \|g\|_\infty\}) > 0$. If μ is semifinite, there exists a $B \in \Sigma$ 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

$$f_0(x) = \begin{cases} \frac{\overline{sign(g(x))}\chi_B(x)}{\mu(B)}\,, & \text{if } g(x) \text{ is finite,} \\ 0\,, & \text{if } g(x) \text{ is infinite.} \end{cases}$$

Then,

$$f_0(x)g(x) = \begin{cases} \frac{|g(x)|\chi_B(x)}{\mu(B)}\,, & \text{if } g(x) \text{ is finite,} \\ 0\,, & \text{if } g(x) \text{ is infinite} \end{cases}$$

and, hence, $\int_X f_0 g\,d\mu = \frac{1}{\mu(B)}\int_B |g|\,d\mu \geq \|g\|_\infty - \epsilon$.

Also,

$$|f_0(x)| = \begin{cases} \frac{\chi_B(x)}{\mu(B)}\,, & \text{if } g(x) \text{ is finite,} \\ 0\,, & \text{if } g(x) \text{ is infinite} \end{cases}$$

and, hence, $\|f_0\|_1 = \int_X |f_0|\,d\mu = \frac{1}{\mu(B)}\int_B d\mu = 1$.

These imply

$$\sup\Big\{\Big|\int_X fg\,d\mu\Big| \,|\, f \in L^1, \|f\|_1 \leq 1\Big\} \geq \|g\|_\infty - \epsilon$$

for every ϵ with $0 < \epsilon < \|g\|_\infty$ and, taking the limit as $\epsilon \to 0+$, we conclude that

$$\|g\|_\infty = \sup\Big\{\Big|\int_X fg\,d\mu\Big| \,|\, f \in L^1, \|f\|_1 \leq 1\Big\}.$$

Definition 11.24 *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 11.18 *Let $1 \leq p \leq +\infty$. For every $g \in L^{p'}$, the function l_g of Definition 11.24 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 11.8 together with Proposition 11.5 imply that, if $1 < p \leq +\infty$, then $\|l_g\|_* = \|g\|_{p'}$. If μ is semifinite, the same is true, also, for $p = 1$.

If $p = 1$, for every $f \in L^1$ we have $|l_g(f)| = \left|\int_X fg\,d\mu\right| \leq \|g\|_\infty \|f\|_1$. Therefore, $\|l_g\|_* \leq \|g\|_\infty$.

Definition 11.25 *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 11.19 *The function J of Definition 11.25 is a bounded linear transformation. If $1 < p \leq +\infty$, J is an isometry from $L^{p'}$ into $(L^p)^*$. This is true, also, when $p = 1$, if μ is semifinite.*

Proof: For every $f \in L^p$ we have $l_{g_1+g_2}(f) = \int_X f(g_1 + g_2)\,d\mu = \int_X fg_1\,d\mu + \int_X fg_2\,d\mu = l_{g_1}(f) + l_{g_2}(f) = (l_{g_1} + l_{g_2})(f)$ and, hence, $J(g_1 + g_2) = l_{g_1+g_2} = l_{g_1} + l_{g_2} = J(g_1) + J(g_2)$.

Moreover, $l_{\kappa g}(f) = \int_X f(\kappa g)\,d\mu = \kappa \int_X fg\,d\mu = \kappa l_g(f) = (\kappa l_g)(f)$ and, hence, $J(\kappa g) = l_{\kappa g} = \kappa l_g = \kappa J(g)$.

Now, $\|J(g)\|_* = \|l_g\|_* \leq \|g\|_{p'}$ and J is bounded. That J is an isometry is a consequence of Proposition 11.18.

Lemma 11.2 *Let $l \in (L^p(X, \Sigma, \mu))^*$. If $E \in \Sigma$, $\Sigma\rceil E = \{A \in \Sigma \,|\, A \subseteq E\}$ is the restriction of Σ on E and $\mu\rceil E$ is the restricted measure on $(E, \Sigma\rceil E)$, we define $l\rceil E$ by*

$$(l\rceil E)(h) = l(\widetilde{h}), \qquad h \in L^p(E, \Sigma\rceil E, \mu\rceil E),$$

where $\widetilde{h}$ is the extension of h as 0 on $X \setminus E$.

Then, $l\rceil E \in (L^p(E, \Sigma\rceil E, \mu\rceil E))^*$ and $\|l\rceil E\|_* \leq \|l\|_*$. Moreover,

$$l(f\chi_E) = (l\rceil E)(f\rceil E), \qquad f \in L^p(X, \Sigma, \mu),$$

where $f\rceil E$ is the restriction of f on E.

Proof: For all $h, h_1, h_2 \in L^p(E, \Sigma\rceil E, \mu\rceil E)$ we consider the corresponding extensions $\widetilde{h}, \widetilde{h_1}, \widetilde{h_2} \in L^p(X, \Sigma, \mu)$. Since $\widetilde{h_1} + \widetilde{h_2}$ and $\kappa\widetilde{h}$ are the extensions of $h_1 + h_2$ and κh, respectively, we have $(l\rceil E)(h_1 + h_2) = l(\widetilde{h_1} + \widetilde{h_2}) = l(\widetilde{h_1}) + l(\widetilde{h_2}) = (l\rceil E)(h_1) + (l\rceil E)(h_2)$ and $(l\rceil E)(\kappa h) = l(\kappa\widetilde{h}) = \kappa l(\widetilde{h}) = \kappa(l\rceil E)(h)$. 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 $f \in L^p(X, \Sigma, \mu)$, 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 11.26 *The $l\rceil E$ defined in Lemma 11.2 is called* ***the restriction of*** *$l \in (L^p(X, \Sigma, \mu))^*$* ***on*** *$L^p(E, \Sigma\rceil E, \mu\rceil E)$.*

Theorem 11.9 *Let $1 < p < +\infty$.*
(i) For every $l \in (L^p)^*$ there exists a unique $g \in L^{p'}$ so that

$$l(f) = \int_X fg \, d\mu$$

for every $f \in L^p$.
(ii) The function J of Definition 11.25 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 $A \in \Sigma$. We define the function $\nu : \Sigma \to F$ by

$$\nu(A) = l(\chi_A), \qquad A \in \Sigma.$$

We have $\nu(\emptyset) = l(\chi_\emptyset) = l(o) = 0$. If $A_1, A_2, \ldots \in \Sigma$ are pairwise disjoint and $A = \cup_{j=1}^{+\infty} A_j$, then $\chi_A = \sum_{j=1}^{+\infty} \chi_{A_j}$. Therefore, $\| \sum_{j=1}^{n} \chi_{A_j} - \chi_A \|_p^p = \int_X | \sum_{j=n+1}^{+\infty} \chi_{A_j} |^p \, d\mu = \int_X |\chi_{\cup_{j=n+1}^{+\infty} A_j}|^p \, d\mu = \mu(\cup_{j=n+1}^{+\infty} A_j) \to \mu(\emptyset) = 0$, by the continuity of μ from above. The linearity and the continuity of l imply, now, that $\sum_{j=1}^{n} \nu(A_j) = \sum_{j=1}^{n} l(\chi_{A_j}) = l(\sum_{j=1}^{n} \chi_{A_j}) \to l(\chi_A) = \nu(A)$ or, equivalently, that $\sum_{j=1}^{+\infty} \nu(A_j) = \nu(A)$.

Hence, ν is a real or complex measure (depending on whether $F = \mathbf{R}$ or $F = \mathbf{C}$) on (X, Σ).

We observe that, if $A \in \Sigma$ has $\mu(A) = 0$, then $\nu(A) = l(\chi_A) = l(o) = 0$, because the function χ_A is the zero element o of L^p. Therefore, $\nu \ll \mu$ and, by Theorems 10.12 and 10.13, 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 $A \in \Sigma$. By the 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 which are bounded on X. Indeed, let $f \in L^p$ be such that $|f| \leq M$ 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 $|\phi_n g| \leq |fg| \leq M|g|$ 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 $\int_X |\phi_n - f|^p \, d\mu \to 0$ as $n \to +\infty$ and, hence, $\phi_n \to f$ in L^p. By the continuity of l, we get $\int_X \phi_n g \, d\mu = l(\phi_n) \to l(f)$ and, hence,

$$\diamond \qquad l(f) = \int_X fg \, d\mu$$

for every $f \in L^p$ which is bounded on X.

Our first task, now, 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

$$\phi_n(x) = \begin{cases} \psi_n(x)\overline{sign(g(x))}, & \text{if } g(x) \text{ is finite} \\ 0, & \text{if } g(x) \text{ is infinite.} \end{cases}$$

Then, $0 \leq \phi_n g = \psi_n |g| \uparrow |g|^{p'}$ a.e. on X and each ϕ_n is bounded on X. Hence, $\|\psi_n\|_p^p = \int_X \psi_n^p \, d\mu \leq \int_X \psi_n |g| \, d\mu = \int_X \phi_n g \, d\mu = l(\phi_n) \leq \|l\|_* \|\phi_n\|_p \leq \|l\|_* \|\psi_n\|_p$, where the last equality is justified by $\diamond$. 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 $A = \{x \in X \,|\, t < |g(x)|\}$ has $\mu(A) > 0$. We define the function

$$f(x) = \begin{cases} \chi_A(x)\overline{sign(g(x))}, & \text{if } g(x) \text{ is finite} \\ 0, & \text{if } g(x) \text{ is infinite.} \end{cases}$$

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 $\diamond$. 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, $|\int_X \phi_n g \, d\mu - \int_X fg \, d\mu| \leq \int_X |\phi_n - f||g| \, d\mu \leq \|\phi_n - f\|_p \|g\|_{p'} \to 0$, since $\|g\|_{p'} < +\infty$. From $l(\phi_n) = \int_X \phi_n g \, d\mu$, we conclude that

$$l(f) = \int_X fg \, d\mu, \qquad f \in L^p.$$

This implies, of course, 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 let $g' \in L^{p'}$ also satisfies $l = l_{g'}$. Then $J(g') = l = J(g)$ and, since J is an isometry (and, hence, one-to-one) we get that $g' = g$ a.e. on X.

B. We suppose, now, that μ is σ-finite and consider an increasing sequence (E_k) in Σ so that $E_k \uparrow X$ and $\mu(E_k) < +\infty$ for all k.

Let $l \in (L^p(X, \Sigma, \mu))^*$.

For each k, we consider the restriction $l\rceil E_k$ of l on $L^p(E_k,\Sigma\rceil E_k,\mu\rceil E_k)$, which is defined in Lemma 11.2. Since $l\rceil E_k \in (L^p(E_k,\Sigma\rceil E_k,\mu\rceil E_k))^*$ 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 $g_k\in L^{p'}(E_k,\Sigma\rceil E_k,\mu\rceil E_k)$ so that $\|g_k\|_{p'}\leq \|l\rceil E_k\|_*\leq \|l\|_*$ and

$$(l\rceil E_k)(h)=\int_{E_k} hg_k\,d(\mu\rceil E_k)$$

for every $h\in L^p(E_k,\Sigma\rceil E_k,\mu\rceil E_k)$. In particular,

$$l(f\chi_{E_k})=(l\rceil E_k)(f\rceil E_k)=\int_{E_k}(f\rceil E_k)g_k\,d(\mu\rceil E_k)$$

for every $f\in L^p(X,\Sigma,\mu)$.

For $h\in L^p(E_k,\Sigma\rceil E_k,\mu\rceil E_k)$, take its extension h' on E_{k+1} as 0 on $E_{k+1}\setminus E_k$. Since $\widetilde{h}=\widetilde{h'}$ on X, we get

$$\begin{aligned}\int_{E_k} hg_k\,d(\mu\rceil E_k)&=(l\rceil E_k)(h)=l(\widetilde{h})=l(\widetilde{h'})=(l\rceil E_{k+1})(h')\\&=\int_{E_{k+1}} h'g_{k+1}\,d(\mu\rceil E_{k+1})=\int_X \widetilde{h'g_{k+1}}\,d\mu\\&=\int_{E_k}(\widetilde{h'g_{k+1}})\rceil E_k\,d(\mu\rceil E_k)\\&=\int_{E_k} h(g_{k+1}\rceil E_k)\,d(\mu\rceil E_k).\end{aligned}$$

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.

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,

$$l(f\chi_{E_k})=\int_{E_k} fg\,d\mu\,,\qquad f\in L^p(X,\Sigma,\mu).$$

If $1<p'<+\infty$, then, since $|\widetilde{g_k}|\uparrow|g|$ on X, by the Monotone Convergence Theorem, $\int_X|g|^{p'}\,d\mu=\lim_{k\to+\infty}\int_X|\widetilde{g_k}|^{p'}\,d\mu=\lim_{k\to+\infty}\int_{E_k}|g_k|^{p'}\,d(\mu\rceil E_k)\leq \limsup_{k\to+\infty}\|l\rceil E_k\|_*^{p'}\leq\|l\|_*^{p'}<+\infty$. Hence, $g\in L^{p'}(X,\Sigma,\mu)$ and $\|g\|_{p'}\leq \|l\|_*$.

If $p'=+\infty$, we have that, for every k, $|g|=|g_k|\leq\|g_k\|_\infty\leq\|l\rceil E_k\|_*\leq\|l\|_*$ a.e. on E_k. This implies that $|g|\leq\|l\|_*$ a.e. on X and, thus, $g\in L^\infty(X,\Sigma,\mu)$ and $\|g\|_\infty\leq\|l\|_*$.

Hence, in all cases, $g\in L^{p'}(X,\Sigma,\mu)$ and $\|g\|_{p'}\leq\|l\|_*$.

For an arbitrary $f\in L^p(X,\Sigma,\mu)$, we have $\|f\chi_{E_k}-f\|_p^p=\int_X|f\chi_{E_k}-f|^p\,d\mu=\int_{E_k^c}|f|^p\,d\mu=\int_X\chi_{E_k^c}|f|^p\,d\mu\to 0$, by the Dominated Convergence Theorem. By the continuity of l, we get $l(f)=\lim_{k\to+\infty}l(f\chi_{E_k})=\lim_{k\to+\infty}\int_{E_k}fg\,d\mu=$

$\int_X fg\,d\mu$. The last equality holds since $|\int_{E_k} fg\,d\mu - \int_X fg\,d\mu| = |\int_{E_k^c} fg\,d\mu| \leq \big(\int_{E_k^c} |f|^p\,d\mu\big)^{\frac{1}{p}} \|g\|_{p'} \to 0$. We have proved that

$$l(f) = \int_X fg\,d\mu, \qquad f \in L^p(X,\Sigma,\mu)$$

and, hence, $l = l_g = J(g)$. Therefore, just as in part A, J is an isometry from $L^{p'}(X,\Sigma,\mu)$ *onto* $(L^p(X,\Sigma,\mu))^*$.

Again, if $g' \in L^{p'}(X,\Sigma,\mu)$ also satisfies $l = l_{g'}$, then $J(g') = l = J(g)$ and, since J is an isometry, we get that $g' = g$ a.e. on X.

C. Now, let $1 < p, p' < +\infty$ and μ be arbitrary.

Let $l \in (L^p(X,\Sigma,\mu))^*$.

We consider any $E \in \Sigma$ of σ-finite measure and the restriction $l\rceil E$ of l on $L^p(E,\Sigma\rceil E,\mu\rceil E)$, defined in Lemma 11.2. Since $l\rceil E \in (L^p(E,\Sigma\rceil E,\mu\rceil E))^*$ and $\|l\rceil E\|_* \leq \|l\|_*$, part B implies that there is a unique $g_E \in L^{p'}(E,\Sigma\rceil E,\mu\rceil E)$ so that $\|g_E\|_{p'} \leq \|l\rceil E\|_* \leq \|l\|_*$ and

$$(l\rceil E)(h) = \int_E hg_E\,d(\mu\rceil E)$$

for every $h \in L^p(E,\Sigma\rceil E,\mu\rceil E)$. In particular,

$$l(f\chi_E) = (l\rceil E)(f\rceil E) = \int_E (f\rceil E)g_E\,d(\mu\rceil E)$$

for every $f \in L^p(X,\Sigma,\mu)$.

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.

We define

$$M = \sup\Big\{ \int_E |g_E|^{p'}\,d(\mu\rceil E) \,\big|\, E \text{ of } \sigma\text{-finite measure}\Big\}$$

and, obviously, $M \leq \|l\|_*^{p'} < +\infty$. We take a sequence (E_n) in Σ, 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 $E = \cup_{n=1}^{+\infty} E_n$ and observe that E has σ-finite measure and, hence, $\int_E |g_E|^{p'}\,d(\mu\rceil E) \leq M$. 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, thus, $\int_{E_n} |g_{E_n}|^{p'}\,d(\mu\rceil E_n) \leq \int_E |g_E|^{p'}\,d(\mu\rceil E) \leq M$. 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 that

$$\int_X |g|^{p'}\,d\mu = \int_E |g_E|^{p'}\,d(\mu\rceil E) = M \leq \|l\|_*^{p'}.$$

Now consider an arbitrary $f \in L^p(X, \Sigma, \mu)$. The set

$$F = E \cup \{x \in X \mid f(x) \neq 0\}$$

has σ-finite measure. By $g_F \rceil E = g_E$ a.e. on E, we get $M = \int_E |g_E|^{p'} \, d(\mu \rceil E) = \int_E |g_F|^{p'} \, d(\mu \rceil F) \leq \int_E |g_F|^{p'} \, d(\mu \rceil F) + \int_{F \setminus E} |g_F|^{p'} \, d(\mu \rceil F) = \int_F |g_F|^{p'} \, d(\mu \rceil F) \leq M$. 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,

$$\begin{aligned} l(f) = l(f\chi_F) &= \int_F (f \rceil F) g_F \, d(\mu \rceil F) = \int_E (f \rceil F) g_F \, d(\mu \rceil F) \\ &= \int_E (f \rceil F) g_E \, d(\mu \rceil F) = \int_E (f \rceil E) g_E \, d(\mu \rceil E) \\ &= \int_X f g \, d\mu. \end{aligned}$$

Thus, $l = l_g = J(g)$ and, just as in parts A and B, J is an isometry from $L^{p'}(X, \Sigma, \mu)$ *onto* $(L^p(X, \Sigma, \mu))^*$.

Finally, if $g' \in L^{p'}(X, \Sigma, \mu)$ also satisfies $l = l_{g'}$, then $J(g') = l = J(g)$ and, since J is an isometry, we get that $g' = g$ a.e. on X.

11.4 The space $M(X, \Sigma)$.

Just as in the previous two sections, (X, Σ) will be a fixed measure space.

Definition 11.27 *Let (X, Σ) be a measurable space. The set of all real or complex (depending on whether $F = \mathbf{R}$ or $F = \mathbf{C}$) measures on (X, Σ) is denoted by $M(X, \Sigma)$.*

If there is no danger of confusion, we shall use the symbol M instead of $M(X, \Sigma)$.

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 $A \in \Sigma$. We, also, define $\kappa\nu \in M$ by $(\kappa\nu)(A) = \kappa\nu(A)$ for all $A \in \Sigma$ and $\kappa \in F$.

It is easy to show that M is a linear space over F. The zero element is the measure o defined by $o(A) = 0$ for all $A \in \Sigma$. The opposite to ν is $-\nu$ defined by $(-\nu)(A) = -\nu(A)$ for all $A \in \Sigma$.

Definition 11.28 *For every $\nu \in M$ we define*

$$\|\nu\| = |\nu|(X).$$

Thus, $\|\nu\|$ is just the total variation of ν.

Proposition 11.20 *$\| \cdot \|$ is a norm on M.*

Proof: Proposition 10.9 implies that $\|\nu_1 + \nu_2\| = |\nu_1 + \nu_2|(X) \leq |\nu_1|(X) + |\nu_2|(X) = \|\nu_1\| + \|\nu_2\|$ and $\|\kappa\nu\| = |\kappa\nu|(X) = |\kappa||\nu|(X) = |\kappa|\|\nu\|$.

If $\|\nu\| = 0$, then $|\nu|(X) = 0$. This implies that $|\nu(A)| \leq |\nu|(A) = 0$ for all $A \in \Sigma$ and, hence, $\nu = o$ is the zero measure.

Theorem 11.10 *M is a Banach space.*

Proof: Let (ν_n) be a Cauchy sequence in M. This means $|\nu_n - \nu_m|(X) = \|\nu_n - \nu_m\| \to 0$ 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 $A \in \Sigma$. Therefore, it converges to a finite number and we define

$$\nu(A) = \lim_{n \to +\infty} \nu_n(A)$$

for all $A \in \Sigma$.

It is clear that $\nu(\emptyset) = \lim_{n \to +\infty} \nu_n(\emptyset) = 0$.

Now, let $A_1, A_2, \ldots \in \Sigma$ be pairwise disjoint and $A = \cup_{j=1}^{+\infty} A_j$. 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 $n \geq N$,

$$\begin{aligned}
\sum_{j=J+1}^{+\infty} |\nu_n|(A_j) &\leq \sum_{j=J+1}^{+\infty} |\nu_n - \nu_N|(A_j) + \sum_{j=J+1}^{+\infty} |\nu_N|(A_j) \\
&\leq |\nu_n - \nu_N|(\cup_{j=J+1}^{+\infty} A_j) + \epsilon \\
&\leq |\nu_n - \nu_N|(X) + \epsilon \; = \; \|\nu_n - \nu_N\| + \epsilon \\
&\leq 2\epsilon.
\end{aligned}$$

This implies that, for arbitrary $K \geq J+1$ and every $n \geq N$, 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

$$\sum_{j=J+1}^{+\infty} |\nu(A_j)| \leq 2\epsilon.$$

We have $|\nu_n(A) - \sum_{j=1}^{J} \nu_n(A_j)| = |\sum_{j=J+1}^{+\infty} \nu_n(A_j)| \leq \sum_{j=J+1}^{+\infty} |\nu_n(A_j)| \leq \sum_{j=J+1}^{+\infty} |\nu_n|(A_j) \leq 2\epsilon$ for all $n \geq N$ and, taking the limit as $n \to +\infty$,

$$|\nu(A) - \sum_{j=1}^{J} \nu(A_j)| \leq 2\epsilon.$$

Altogether, 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 $\nu \in M$.

Consider an arbitrary measurable partition $\{A_1, \ldots, A_p\}$ of X. We have that $\sum_{k=1}^{p} |(\nu_n - \nu_m)(A_k)| \leq \|\nu_n - \nu_m\| \leq \epsilon$ 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, we get

$$\|\nu_n - \nu\| = |\nu_n - \nu|(X) \leq \epsilon.$$

This means that $\|\nu_n - \nu\| \to 0$ as $n \to +\infty$.

11.5 The space $C_0(X)$ and its dual.

Definition 11.29 *Let X be any non-empty set 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, the space B is a linear space over F.

Definition 11.30 *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, $\|o\|_u = \sup_{x \in X} 0 = 0$. If $\|f\|_u = 0$, then $\sup_{x \in X} |f(x)| = 0$ and, hence, $f(x) = 0$ for all $x \in X$. Moreover, $\|\kappa f\|_u = \sup_{x \in X} |\kappa f(x)| = |\kappa| \sup_{u \in X} |f(x)| = |\kappa| \|f\|_u$. 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.

Theorem 11.11 *B is a Banach space.*

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 $|f_n(x) - f(x)| \le \epsilon$ for all $x \in X$ and $n \ge N$. Therefore, $\|f_n - f\|_u \le \epsilon$ for all $n \ge N$ and (f_n) converges to f in B.

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 11.31 *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 11.32 $BC(X) = B(X) \cap C(X)$.

We may, again, 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 $f \in BC$.

Theorem 11.12 *BC is a Banach space.*

Proof: 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 $\|f_n - f\|_u \le \frac{\epsilon}{3}$ for all $n \ge N$ and, in particular, $\|f_N - f\|_u \le \frac{\epsilon}{3}$. By continuity of f_N, there is some open neighborhood U of x so that $|f_N(x') - f_N(x)| \le \frac{\epsilon}{3}$ for all $x' \in U$. Now, for all $x' \in U$ we have $|f(x') - f(x)| \le |f(x') - f_N(x')| + |f_N(x') - f_N(x)| + |f_N(x) - f(x)| \le \|f - f_N\|_u + \frac{\epsilon}{3} + \|f_N - f\|_u \le \epsilon$. Therefore f is continuous at x and, since x is arbitrary, f is continuous on X. Thus $f \in BC$.

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 $C = BC$.

Lemma 11.3 *Let μ be a real or complex (depending on whether $F = \mathbf{R}$ or $F = \mathbf{C}$) Borel measure on X. For every $f \in BC$ we have*

$$\Big| \int_X f \, d\mu \Big| \le \int_X |f| \, d|\mu| \le \|f\|_u \|\mu\|.$$

Proof: A consequence of Theorem 10.8.

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 11.33 *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 $\Re(\mu)$ and $\Im(\mu)$ are regular.
The space of all regular real or complex 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 11.21 *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 $|\Re(\mu)|, |\Im(\mu)| \leq |\mu|$ and $|\mu| \leq |\Re(\mu)| + |\Im(\mu)|$.

Theorem 11.13 *$M_{\mathcal{R}}$ 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. This means that $M_{\mathcal{R}}$ is closed in M.

We recall Theorem 5.7 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, $M_{\mathcal{R}} = M$.

Definition 11.34 *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 $\overline{V}$ is compact.*

Lemma 11.4 *Let X be locally compact Hausdorff. If $K \subseteq X$ is compact and $U \subseteq X$ is open and $K \subseteq U$, then there is an open V such that $K \subseteq V \subseteq \overline{V} \subseteq U$ and $\overline{V}$ is compact.*

The next result is a special case of a well-known more general Lemma of Urysohn.

Theorem 11.14 ***Urysohn's lemma***. *Let X be locally compact 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 supp(f) is a compact subset of U.*

Proof: Let $G = X \setminus L$ and denote $A_0 = K$ and $B_1 = G$. A_0 is closed and B_1 is open.

Then there is some open $B_{\frac{1}{2}}$ such that

$$A_0 \subseteq B_{\frac{1}{2}} \subseteq \overline{B_{\frac{1}{2}}} \subseteq B_1.$$

Similarly, there exist open $B_{\frac{1}{4}}$ and $B_{\frac{3}{4}}$ so that

$$B_0 \subseteq B_{\frac{1}{4}} \subseteq \overline{B_{\frac{1}{4}}} \subseteq B_{\frac{1}{2}} \subseteq \overline{B_{\frac{1}{2}}} \subseteq B_{\frac{3}{4}} \subseteq \overline{B_{\frac{3}{4}}} \subseteq B_1.$$

Continuing inductively, to every rational of the form $r = \frac{k}{2^n}$ with $0 < k \leq 2^n$ corresponds an open set B_r, with the property

$$A_0 \subseteq B_r \subseteq \overline{B_r} \subseteq B_s$$

for every two such rationals r, s with $r < s$. Let $\mathbf{Q}_d$ be the set of all these rational numbers.

We define $f(x) = \inf\{r \in \mathbf{Q}_d | x \in B_r\}$ if $x \in B_1$ and $f(x) = 1$ if $x \in X \setminus B_1$.

We see that $f = 0$ on K and $f = 1$ on L and that $f : X \to [0, 1]$ and it remains to prove that f is continuous on X.

Let $x \in X$ and $\epsilon > 0$. If $0 < f(x) < 1$, there are $r, r', s \in \mathbf{Q}_d$ so that $f(x) - \epsilon < r < r' < f(x) < s < f(x) + \epsilon$. If $y \in B_s$, then $f(y) \leq s < f(x) + \epsilon$. If $y \in X \setminus \overline{B_r}$, then $y \notin B_r$, hence $f(y) \geq r > f(x) - \epsilon$. Also, $x \in B_s$ and $x \notin B_{r'}$, therefore $x \in X \setminus \overline{B_r}$. Thus, the open set $V = B_s \cap (X \setminus \overline{B_r})$ contains x and $f(x) - \epsilon < f(y) < f(x) + \epsilon$ for every $y \in V$. Therefore, f is continuous at x.

If $f(x) = 1$, we take, like before, $r, r' \in \mathbf{Q}_d$ so that $1 - \epsilon < r < r' < 1$ and we see that the open set $V = X \setminus \overline{B_r}$ contains x and $1 - \epsilon < f(y) \leq 1 < 1 + \epsilon$ for every $y \in V$. Similarly, if $f(x) = 0$, we take $s \in \mathbf{Q}_d$ so that $0 < s < \epsilon$ and we get that the open set $V = B_s$ contains x and $-\epsilon < 0 \leq f(y) < \epsilon$ for every $y \in V$. Hence, in all cases f is continuous at x.

We have to say that Urysohn's Lemma holds, more generally, for the *normal* topological spaces, that is for Hausdorff topological spaces with the property that for any two disjoint closed subsets there exist two disjoint open subsets which contain them. This is the only property that was used in the proof of the Lemma. A class of normal spaces is, as we have seen, the compact Hausdorff spaces and another one is the metric spaces. Indeed, in the case of a metric space (X, d) the Lemma has a simple proof: we consider the function $f(x) = \frac{d(x,K)}{d(x,K)+d(x,L)}$ for all $x \in X$, where $d(x, A) = \inf_{y\in A} d(x, y)$ for any $A \subseteq X$.

Lemma 11.5 *(Partition of unity.) Let X be locally compact Hausdorff. If $K \subseteq X$ is compact and $U_1, \ldots, U_n \subseteq X$ are open so that $K \subseteq U_1 \cup \cdots \cup U_n$, then there exist $f_1, \ldots, f_n : X \to [0, 1]$ continuous on X so that supp(f_j) is a compact subset of U_j for all j and $f_1 + \cdots + f_n = 1$ 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 $K \setminus (U_2 \cup \cdots \cup U_n) \subseteq V_1 \subseteq \overline{V_1} \subseteq U_1$. 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 $K \setminus (V_1 \cup U_3 \cup \cdots \cup U_n) \subseteq V_2 \subseteq \overline{V_2} \subseteq U_2$. 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 $U_1, \ldots, U_n$ with open $V_1, \ldots, V_n$ so that $K \subseteq V_1 \cup \cdots \cup V_n$ and $\overline{V_j} \subseteq U_j$ for all j.

We repeat the process, so there exist open $W_1, \ldots, W_n$ so that $K \subseteq W_1 \cup \cdots \cup W_n$ and $\overline{W_j} \subseteq V_j \subseteq \overline{V_j} \subseteq U_j$ for all j.

By Urysohn's Lemma, there are $g_1, \ldots, g_n : X \to [0, 1]$ so that $g_j = 1$ on $\overline{W_j}$ and $g_j = 0$ out of V_j. Also, there exists $g_0 : X \to [0, 1]$ so that $g_0 = 0$ on K and $g_0 = 1$ out of $W_1 \cup \cdots \cup W_n$. We define $f_j = \frac{g_j}{g_0+g_1+\cdots+g_n}$ for every $j = 1, \ldots, n$.

If for any $x \in X$ the $g_0(x) = 1$ is not true, then $x \in W_1 \cup \cdots \cup W_n$ and then $g_j(x) = 1$ for some $j = 1, \ldots, n$. Therefore, $g_0 + g_1 + \cdots + g_n \geq 1$ on X, and the $f_1, \ldots, f_n : X \to [0, 1]$ are continuous on X.

If $x \notin V_j$, then $g_j(x) = 0$ and thus $f_j(x) = 0$. So supp$(f_j) \subseteq \overline{V_j} \subseteq U_j$. Also, $f_1 + \cdots + f_n = \frac{g_1+\cdots+g_n}{g_0+g_1+\cdots+g_n} = 1$ on K, because $g_0 = 0$ on K.

Definition 11.35 *Let K be compact and $U_1, \ldots, U_n$ be open subsets of the locally compact X and $K \subseteq U_1 \cup \cdots \cup U_n$. If the $f_1, \ldots, f_n : X \to [0, 1]$ are continuous on X so that supp(f_j) is a compact subset of U_j for all j and $f_1 + \cdots + f_n = 1$ on K, then the collection $\{f_1, \ldots, f_n\}$ is called* ***a partition of unity for** K* ***relative to its open cover*** *$\{U_1, \ldots, U_n\}$.*

Definition 11.36 *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

$$C_0(X) = \{f \in C(X) \,|\, f \text{ vanishes at infinity } \}.$$

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 $f \in C_0$.

If X is compact, then $C_0 = C = BC$.

Theorem 11.15 *C_0 is a Banach space.*

Theorem 11.16 *Let X be locally compact Hausdorff and $\mu \in M_{\mathcal{R}}$. Then*

$$\|\mu\| = \sup\Big\{\Big|\int_X f\,d\mu\Big| \mid f \in C_0, \|f\|_u \le 1\Big\}.$$

Proof: For all $f \in C(X)$ with $\|f\|_u \le 1$, Lemma 11.4 implies that $|\int_X f\,d\mu| \le \|f\|_u\|\mu\| \le \|\mu\|$. Therefore,

$$\sup\Big\{\Big|\int_X f\,d\mu\Big| \mid f \in C(X), \|f\|_u \le 1\Big\} \le \|\mu\|.$$

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, $|\mu|(U_j \setminus K_j) < \frac{1}{n}\,\epsilon$ for all j. Then, for all j there are $f_j : X \to [0,1]$ continuous on X so that $f_j = 1$ on K_j and $f_j = 0$ out of U_j. Finally, we define $\kappa_j = \overline{\mathrm{sign}\big(\int_{U_j} f_j\,d\mu\big)}$ and $f = \kappa_1 f_1 + \cdots + \kappa_n f_n$.

It is easy to see that $\|f\|_u \le 1$. Therefore,

$$\begin{aligned}\Big|\int_X f\,d\mu\Big| &= \Big|\sum_{j=1}^n \kappa_j \int_{U_j} f_j\,d\mu\Big| = \sum_{j=1}^n \Big|\int_{U_j} f_j\,d\mu\Big| \\ &\ge \sum_{j=1}^n |\mu(K_j)| - \sum_{j=1}^n \Big|\int_{U_j\setminus K_j} f_j\,d\mu\Big| \\ &> \|\mu\| - 2\epsilon - \sum_{j=1}^n |\mu|(U_j \setminus K_j) > \|\mu\| - 3\epsilon.\end{aligned}$$

Since $\epsilon > 0$ is arbitrary, we conclude that

$$\sup\Big\{\Big|\int_X f\,d\mu\Big| \mid f \in C(X), \|f\|_u \le 1\Big\} \ge \|\mu\|.$$

Definition 11.37 *Let X be locally compact Hausdorff. For every $\mu \in M_{\mathcal{R}}$ we define $l_\mu : C_0 \to F$ by*

$$l_\mu(f) = \int_X f\,d\mu, \qquad f \in C_0.$$

Proposition 11.22 *Let X be locally compact Hausdorff. For every $\mu \in M_{\mathcal{R}}$, the function l_μ of Definition 11.35 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 11.15 together with Proposition 11.5 imply that $\|l_\mu\|_*=\|\mu\|$.

Definition 11.38 *Let X be locally compact Hausdorff. We define the mapping $J : M_{\mathcal{R}}\to(C_0)^*$ by*

$$J(\mu)=l_\mu$$

for all $\mu\in M_{\mathcal{R}}$.

Proposition 11.23 *The function J of Definition 11.36 is an isometry from $M_{\mathcal{R}}$ into $(C_0)^*$*

Proof: For every $f\in C(X)$ we have $l_{\mu_1+\mu_2}(f)=\int_X f\,d(\mu_1+\mu_2)=\int_X f\,d\mu_1+\int_X f\,d\mu_2=l_{\mu_1}(f)+l_{\mu_2}(f)=(l_{\mu_1}+l_{\mu_2})(f)$ and, hence, $J(\mu_1+\mu_2)=l_{\mu_1+\mu_2}=l_{\mu_1}+l_{\mu_2}=J(\mu_1)+J(\mu_2)$.

Moreover, $l_{\kappa\mu}(f)=\int_X f\,d(\kappa\mu)=\kappa\int_X f\,d\mu=\kappa l_\mu(f)=(\kappa l_\mu)(f)$ and, hence, $J(\kappa\mu)=l_{\kappa\mu}=\kappa l_\mu=\kappa J(\mu)$.

Now, $\|J(\mu)\|_*=\|l_\mu\|_*\le\|\mu\|$ and J is an isometry.

Theorem 11.17 *(F. Riesz, Radon, Banach, Kakutani.) Let X be locally compact Hausdorff.*
(i) For every $l\in(C_0)^*$ there exists a unique regular (real or complex) Borel measure μ on X so that

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

for every $f\in C_0$.
If l is non-negative (in other words if $l(f)\ge 0$ for every non-negative $f\in C_0$), then μ is non-negative.
If l is real (in other words $l(f)\in\mathbf{R}$ for every real $f\in C_0$), then μ is real.
(ii) The function J of Definition 11.36 is an isometry from $M_{\mathcal{R}}$ onto $(C_0)^*$.

Proof: (A) Let $l\in C(X)^*$ be non-negative.

For each open $O\subseteq X$ and $f\in C(X)$ we denote $f\prec O$ whenever $f:X\to[0,1]$ and $supp(f)\subseteq O$.

For each open O 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,O' are open and $O\subseteq O'$, then $f\prec O$ implies $f\prec O'$ and, thus, $\mu(O)\le\mu(O')$. Hence, $\mu^*(O)=\mu(O)$ for each open O.

If $f\prec O$, then $l(f)\le\|l\|_*\|f\|_u\le\|l\|_*$. Hence, $\mu(O)\le\|l\|_*$ and $\mu^*(E)\le\|l\|_*$ for every $E\subseteq X$.

It is obvious that $\mu^*(\emptyset) = \mu(\emptyset) = 0$ and also that $\mu^*(E) \le \mu^*(E')$ for all E, E' with $E \subseteq E'$. 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 $K = supp(f) \subseteq O$. There is, then, N so that $K \subseteq O_1 \cup \cdots \cup O_N$ and we consider a partition of unity $\{f_1, \dots, f_N\}$ for K relative to $\{O_1, \dots, O_N\}$. Then $f = ff_1 + \cdots + ff_N$ and $supp(ff_j) \prec O_j$ for each j and, hence, $l(f) = l(ff_1) + \cdots + l(ff_N) \le \mu(O_1) + \cdots + \mu(O_N) \le \mu(O_1) + \mu(O_N) + \cdots$. This implies that $\mu(O) \le \mu(O_1) + \mu(O_N) + \cdots \le \mu^*(E_1) + \mu^*(E_2) + \cdots + \epsilon$ and, since $E \subseteq O$, we get $\mu^*(E) \le \mu^*(E_1) + \mu^*(E_2) + \cdots + \epsilon$. Since $\epsilon > 0$ is arbitrary, we get $\mu^*(E) \le \mu^*(E_1) + \mu^*(E_2) + \cdots$. 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 $O' \setminus supp(f)$ is open and we take $g \prec O' \setminus supp(f)$ so that $l(g) > \mu(O' \setminus supp(f)) - \epsilon$. We observe that $f + g \prec O'$, whence $\mu^*(E) + \epsilon > \mu(O') \ge l(f+g) = l(f) + l(g) > \mu(O' \cap O) + \mu(O' \setminus supp(f)) - 2\epsilon \ge \mu^*(E \cap O) + \mu^*(E \setminus O) - 2\epsilon$. Hence, $\mu^*(E) \ge \mu^*(E \cap O) + \mu^*(E \setminus O)$ 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

$$(\sharp) \qquad \mu(K) = \inf\{l(f) \,|\, f \in C(X) \text{ and } \chi_K \le f \text{ on } X\}$$

for all compact $K \subseteq X$.

We take any $f \in C(X)$ with $f \ge \chi_K$ (e.g. $f \ge 0$ on X and, in particular, $f \ge 1$ on K) and consider the open set $O = \{x \in X \,|\, f(x) > 1 - \epsilon\} \supseteq K$. If $g \prec O$, then $g \le \frac{1}{1-\epsilon} f$ on X and then $l(g) \le \frac{1}{1-\epsilon} l(f)$, since l is non-negative. Therefore, $\mu(O) \le \frac{1}{1-\epsilon} l(f)$, whence $\mu(K) \le \frac{1}{1-\epsilon} l(f)$. Since $\epsilon > 0$ is arbitrary, this implies that $\mu(K) \le l(f)$ and, thus, $\mu(K) \le \inf\{l(f) \,|\, f \in C(X) \text{ and } \chi_K \le f \text{ on } X\}$. We now take an open $O \supseteq K$ with $\mu(O) < \mu(K) + \epsilon$ and, then, an $f : X \to [0,1]$ continuous on X with $f = 1$ on K and $supp(f) \subseteq O$. Then $f \ge \chi_K$ and $f \prec O$ and, hence, $l(f) \le \mu(O) < \mu(K) + \epsilon$. Since ϵ is arbitrary, $\inf\{l(f) \,|\, f \in C(X) \text{ and } \chi_K \le f \text{ on } X\} \le \mu(K)$.

We shall next prove the regularity of μ.

For each Borel set E we have $\mu(E) = \mu^*(E) = \inf\{\mu(O) \,|\, O \text{ open } \supseteq E\}$ 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 $K = supp(g) \subseteq O$. For each $f \in C(X)$ with $f \ge \chi_K$, we get that $f \ge g$ and then $l(f) \ge l(g)$. From $(\sharp)$ it is implied that $\mu(K) \ge l(g)$. 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, $\mu(E) = \sup\{\mu(L) | L \text{ compact } \subseteq E\}$. This is the second regularity condition.

Finally, we shall prove that $l(f) = \int_X f\, d\mu$ for every $f \in C(X)$ and, by linearity, it is enough to prove it for real f. (Of course, if $F = \mathbf{R}$, then all functions are real anyway.) If f is real, we write $f^+ = \frac{1}{2}(|f| + f) \geq 0$ and $f^- = \frac{1}{2}(|f| - f) \geq 0$, whence $f = f^+ - f^-$. Therefore, it is enough to consider $f \geq 0$ and, multiplying with an appropriate positive constant, we may assume that $f \in C(X)$ and $0 \leq f \leq 1$ on X.

We take arbitrary N and define $K_k = \{x \in X |\, f(x) \geq \frac{k}{N}\}$ for $0 \leq k \leq N$. Every K_k is compact and, obviously, $K_0 = X$. Also, for each $j = 0, \ldots, N-1$ we define

$$f_j = \min\Big\{ \max\Big\{f, \frac{j}{N}\Big\}, \frac{j+1}{N}\Big\} - \frac{j}{N}\,.$$

Each f_j is continuous on X 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

$$\frac{1}{N}(\mu(K_1) + \cdots + \mu(K_N)) \leq \int_X f\, d\mu \leq \frac{1}{N}(\mu(K_0) + \cdots + \mu(K_{N-1})).$$

From $\chi_{K_{j+1}} \leq N f_j$ and ($\sharp$) 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,

$$\frac{1}{N}(\mu(K_1) + \cdots + \mu(K_N)) \leq l(f) \leq \frac{1}{N}(\mu(K_0) + \cdots + \mu(K_{N-1})).$$

Thus, $\big|\int_X f\, d\mu - l(f)\big| \leq \frac{1}{N}(\mu(K_0) + \cdots + \mu(K_{N-1})) - \frac{1}{N}(\mu(K_1) + \cdots + \mu(K_N)) = \frac{1}{N}\mu(K_0 \setminus K_N) \leq \frac{1}{N}\mu(X)$ and, since N is arbitrary,

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

and the case of non-negative l is finished.
(B) Let now l be real. For each non-negative $f \in C(X)$ we define

$$l^+(f) = \sup\{l(g) \,|\, g \in C(X), 0 \leq g \leq f \text{ on } X\}.$$

Obviously, $l^+(f) \geq l(0) = 0$ and $l^+(f) \geq l(f)$. Also, if $0 \leq g \leq f$, then $|l(g)| \leq \|l\|_* \|g\|_u \leq \|l\|_* \|f\|_u$ and, thus, $l^+(f) = |l^+(f)| \leq \|l\|_* \|f\|_u < +\infty$.

For every $\kappa > 0$ and non-negative $f \in C(X)$ we have $l^+(\kappa f) = \sup\{l(g) \,|\, g \in C(X), 0 \leq g \leq \kappa f \text{ on } X\} = \sup\{l(\kappa h) \,|\, h \in C(X), 0 \leq h \leq f \text{ on } X\} = \kappa \sup\{l(h) \,|\, h \in C(X), 0 \leq h \leq f \text{ on } X\} = \kappa l^+(f)$.

If $f_1, f_2 \in C(X)$ are non-negative, $0 \leq g_1 \leq f_1$ and $0 \leq g_2 \leq f_2$, then $l(g_1) + l(g_2) = l(g_1 + g_2)$ and, since $0 \leq g_1 + g_2 \leq f_1 + f_2$, it is implied that $l(g_1) + l(g_2) \leq l^+(f_1 + f_2)$. Taking supremum over g_1 and g_2, we find $l^+(f_1) + l^+(f_2) \leq l^+(f_1 + f_2)$. Now, let $0 \leq g \leq f_1 + f_2$. We set $g_1 = \min(f_1, g)$, from which $0 \leq g_1 \leq f_1$ and $g_1 \leq g$. If we set $g_2 = g - g_1$, then it is easy to see that $0 \leq g_2 \leq f_2$ and, of course, $g = g_1 + g_2$. Hence, $l(g) = l(g_1) + l(g_2) \leq l^+(f_1) + l^+(f_2)$, from which $l^+(f_1 + f_2) \leq l^+(f_1) + l^+(f_2)$. We conclude that $l^+(f_1 + f_2) = l^+(f_1) + l^+(f_2)$.

Until now, $l^+(f)$ is defined only for non-negative $f \in C(X)$. For an arbitrary real $f \in C(X)$ we write $f^+ = \frac{1}{2}(|f| + f) \geq 0$ and $f^- = \frac{1}{2}(|f| - f) \geq 0$, whence $f = f^+ - f^-$. We, then, define for each real $f \in C(X)$

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

Observe that, if $f = g - h$ for any non-negative $g, h \in C(X)$, then $f^+ + h = f^- + g$, whence $l^+(f^+) + l^+(h) = l^+(f^+ + h) = l^+(f^- + g) = l^+(f^-) + l^+(g)$. Hence, $l^+(f) = l^+(g) - l^+(h)$.

If $f_1, f_2 \in C(X)$ are real, then from the last identity we get $f_1 + f_2 = (f_1^+ + f_2^+) - (f_1^- + f_2^-)$, whence $l(f_1 + f_2) = l(f_1^+ + f_2^+) - l(f_1^- + f_2^-) = l(f_1^+) + l(f_2^+) - l(f_1^-) - l(f_2^-) = l(f_1) + l(f_2)$.

If $f \in C(X)$ is real and $\kappa \geq 0$, then $l^+(\kappa f) = l^+(\kappa f^+) - l^+(\kappa f^-) = \kappa l^+(f^+) - \kappa l^+(f^-) = \kappa l^+(f)$, while if $\kappa < 0$, then $l^+(\kappa f) = l^+(|\kappa| f^-) - l^+(|\kappa| f^+) = |\kappa| l^+(f^-) - |\kappa| l^+(f^+) = \kappa l^+(f)$.

If $F = \mathbf{R}$, we have already proved that $l^+ : C(X) \to \mathbf{R}$ is a linear functional.

If $F = \mathbf{C}$, for each complex $f \in C(X)$ we define

$$l^+(f) = l^+(\Re f) + i l^+(\Im f)$$

and it is easy to see that $l^+ : C(X) \to \mathbf{C}$ is a linear functional. If $f \in C(X)$ is real, then $|l^+(f)| = |l^+(f^+) - l^+(f^-)| \leq \max\{l^+(f^+), l^+(f^-)\} \leq \max\{\|l\|_* \|f^+\|_u, \|l\|_* \|f^-\|_u\} = \|l\|_* \|f\|_u$. While, if f is complex, then, with an appropriate $\kappa \in \mathbf{C}$ with $|\kappa| = 1$ we have $|l^+(f)| = \kappa l^+(f) = l^+(\kappa f) = \Re\big(l^+(\kappa f)\big) = l^+(\Re(\kappa f)) \leq \|l\|_* \|\Re(\kappa f)\|_u \leq \|l\|_* \|f\|_u$. Therefore, l^+ is a non-negative bounded linear functional of $C(X)$ with $\|l^+\|_* \leq \|l\|_*$.

We also define $l^- = l^+ - l : C(X) \to F$. Clearly, this is a bounded linear functional of $C(X)$ and it is non-negative, since for every non-negative $f \in C(X)$ we have $l^-(f) = l^+(f) - l(f) \geq 0$.

By part (A), there are two non-negative Borel measures μ_1 and μ_2 on X so that $l^+(f) = \int_X f \, d\mu_1$ and $l^-(f) = \int_X f \, d\mu_2$ for every $f \in C(X)$. Therefore, for the real Borel measure $\mu = \mu_1 - \mu_2$ we have $l(f) = l^+(f) - l^-(f) = \int_X f \, d\mu_1 - \int_X f \, d\mu_2 = \int_X f \, d\mu$ for every $f \in C(X)$.

At this point the proof is finished, if $F = \mathbf{C}$ and l is real or if $F = \mathbf{R}$ (whence l is automatically real).
(C) If $F = \mathbf{C}$ and l is complex, then $\Re(l)$ and $\Im(l)$ are bounded real $\mathbf{R}$-linear functionals of $C(X)$ and, hence, they are bounded $\mathbf{R}$-linear functionals of $C_r(X)$, the $\mathbf{R}$-linear space of real continuous functions on X. By the result of (B), there are two real Borel measures μ_1 and μ_2, so that $\Re(l(f)) = \int_X f\, d\mu_1$ and $\Im(l(f)) = \int_X f\, d\mu_2$ for every real $f \in C(X)$. Therefore, if we define $\mu = \mu_1 + i\mu_2$, then μ is a complex Borel measure on X and for every real $f \in C(X)$ we have $l(f) = \Re(l(f)) + i\Im(l(f)) = \int_X f\, d\mu_1 + i\int_X f\, d\mu_2 = \int_X f\, d\mu$. Therefore, for every $f \in C(X)$, $l(f) = l(\Re(f)) + il(\Im(f)) = \int_X \Re(f)\, d\mu + i\int_X \Im(f)\, d\mu = \int_X f\, d\mu$.

11.6 Exercises.

1. *Approximation*

   (i) Let $f \in L^p(X, \Sigma, \mu)$ and $\epsilon > 0$. Using Theorem 6.1, 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 $f \in L^p(\mathbf{R}^n, \mathcal{L}_n, m_n)$ and $\epsilon > 0$. If $p < +\infty$, prove that there exists a function g continuous on $\mathbf{R}^n$ and equal to 0 outside some bounded set so that $\|f - g\|_p < \epsilon$.

2. Let I be any index set and $0 < p < q \leq +\infty$. Prove that $l^p(I) \subseteq l^q(I)$ and that
$$\|b\|_q \leq \|b\|_p$$
   for every $b \in l^p(I)$.

3. Let $\mu(X) < +\infty$ and $0 < p < q \leq +\infty$. Prove that $L^q(X, \Sigma, \mu) \subseteq L^p(X, \Sigma, \mu)$ and that
$$\|f\|_p \leq \mu(X)^{\frac{1}{p} - \frac{1}{q}} \|f\|_q$$
   for every $f \in L^q(X, \Sigma, \mu)$.

4. Let $0 < p < q < r \leq +\infty$ and $f \in L^p(X, \Sigma, \mu) \cap L^r(X, \Sigma, \mu)$. Prove that $f \in L^q(X, \Sigma, \mu)$ and, if $\frac{1}{q} = \frac{t}{p} + \frac{1-t}{r}$, then
$$\|f\|_q \leq \|f\|_p^t \|f\|_r^{1-t}.$$

5. Let $1 \leq p < r \leq +\infty$. We set $Z = L^p(X, \Sigma, \mu) \cap L^r(X, \Sigma, \mu)$ and we 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 $T : Z \to L^q(X, \Sigma, \mu)$ with $T(f) = f$ for every $f \in Z$ (see Exercise 11.5.4). Prove that T is bounded.

6. Let $0 < p < q < r \leq +\infty$ and $f \in L^q(X, \Sigma, \mu)$. If $t > 0$ is arbitrary, consider the functions

$$g(x) = \begin{cases} f(x), & \text{if } |f(x)| > t \\ 0, & \text{if } |f(x)| \leq t \end{cases} \qquad h(x) = \begin{cases} 0, & \text{if } |f(x)| > t \\ f(x), & \text{if } |f(x)| \leq t \end{cases}.$$

Prove that $g \in L^p(X, \Sigma, \mu)$ and $h \in L^r(X, \Sigma, \mu)$ and that $f = g + h$ on X.

7. Let $1 \leq p < r \leq +\infty$. We define $W = L^p(X, \Sigma, \mu) + L^r(X, \Sigma, \mu) = \{g + h \mid g \in L^p(X, \Sigma, \mu), h \in L^r(X, \Sigma, \mu)\}$ and

$$\|f\| = \inf\big\{\|g\|_p + \|h\|_r \mid g \in L^p(X, \Sigma, \mu), h \in L^r(X, \Sigma, \mu), f = g + h\big\}$$

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 $T : L^q(X, \Sigma, \mu) \to W$ with $T(f) = f$ for every $f \in L^q(X, \Sigma, \mu)$ (see Exercise 11.5.6). Prove that T is bounded.

8. Let $0 < p < q < +\infty$. Prove that $L^p(X, \Sigma, \mu) \not\subseteq L^q(X, \Sigma, \mu)$ if and only if X includes sets of arbitrarily small positive measure and that $L^q(X, \Sigma, \mu) \not\subseteq L^p(X, \Sigma, \mu)$ if and only if X includes sets of arbitrarily large finite measure.

9. Let $1 \leq p < +\infty$ and (f_n) be a sequence in $L^p(X, \Sigma, \mu)$ so that $|f_n| \leq g$ a.e. on X for every n for some $g \in L^p(X, \Sigma, \mu)$. If (f_n) converges to f a.e. on X or in measure, prove that $\|f_n - f\|_p \to 0$.

10. Let $1 \leq p < +\infty$ and $f, f_n \in L^p(X, \Sigma, \mu)$ for all n. If $f_n \to f$ a.e. on X, prove that $\|f_n - f\|_p \to 0$ if and only if $\|f_n\|_p \to \|f\|_p$.

11. Let $1 \leq p \leq +\infty$ and $g \in L^\infty(X, \Sigma, \mu)$. We define the linear transformation $T : L^p(X, \Sigma, \mu) \to L^p(X, \Sigma, \mu)$ with $T(f) = gf$ for every $f \in L^p(X, \Sigma, \mu)$. Prove that T is bounded, that $\|T\| \leq \|g\|_\infty$ and that $\|T\| = \|g\|_\infty$ if μ is semifinite.

12. *The inequality of Chebychev.*

If $0 < p < +\infty$ and $f \in L^p(X, \Sigma, \mu)$, prove that

$$\lambda_{|f|}(t) \leq \frac{\|f\|_p^p}{t^p}, \qquad 0 < t < +\infty.$$

13. *The general Minkowski's Inequality.*

Let (X_1, Σ_1, μ_1) and (X_2, Σ_2, μ_2) be two σ-finite measure spaces and $1 \leq p < +\infty$.
(i) If $f : X_1 \times X_2 \to [0, +\infty]$ is $\Sigma_1 \otimes \Sigma_2$−measurable, prove that

$$\Big(\int_{X_1}\Big(\int_{X_2} f(\cdot,\cdot)\,d\mu_2\Big)^p d\mu_1\Big)^{\frac{1}{p}} \leq \int_{X_2}\Big(\int_{X_1} f(\cdot,\cdot)^p\,d\mu_1\Big)^{\frac{1}{p}} d\mu_2.$$

(ii) If $f(\cdot, x_2) \in L^p(X_1, \Sigma_1, \mu_1)$ for μ_2-a.e. $x_2 \in X_2$ and the function $x_2 \mapsto \|f(\cdot, x_2)\|_p$ is in $L^1(X_2, \Sigma_2, \mu_2)$, prove that $f(x_1, \cdot) \in L^1(X_2, \Sigma_2, \mu_2)$ 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 $L^p(X_1, \Sigma_1, \mu_1)$ and

$$\Big(\int_{X_1} \Big| \int_{X_2} f(\cdot, \cdot)\, d\mu_2 \Big|^p d\mu_1 \Big)^{\frac{1}{p}} \leq \int_{X_2} \Big(\int_{X_1} |f(\cdot, \cdot)|^p\, d\mu_1 \Big)^{\frac{1}{p}} d\mu_2.$$