Marko Kandić
marko.kandic@fmf.uni-lj.si

Lecture Notes 2021/22

Measure Theory

Contents

1 Introduction

The notion of a measure is a natural generalization of well-known geometrical notions: the length of an interval, the area of a plane figure or the area of a body in the space. Usually, if the boundary of a figure or a body is nice, then its area or volume can be calculated by the Riemann integral.

1.1 Question. *For each $n \in \mathbb{N}$ does there exist a measure μ_n which assigns to every subset $E \subseteq \mathbb{R}^n$ its measure $\mu_n(E) \in [0, \infty]$?*

The answer to this question is trivial since we do not have any requirements upon μ_n. What about if we require that μ_n also satisfies the following restrictions:

(i) If $K = [0, 1) \times \cdots [0, 1)$ is the unit cube in $\mathbb{R}^n$, then $\mu_n(K) = 1$.

(ii) If $F \subseteq \mathbb{R}^n$ is obtained from $E \subseteq \mathbb{R}^n$ after a finite number of translations, rotations and reflections, then $\mu(F) = \mu(E)$.

(iii) If $E_1, E_2, \ldots$ is a finite or an infinite sequence of pairwise disjoint sets in $\mathbb{R}^n$, then

$$\mu\Big(\bigcup_{n=1}^{\infty} E_n\Big) = \sum_{n=1}^{\infty} \mu(E_n).$$

Unfortunately such a map μ_n does not even exist in the case when $n = 1$. This will be proved later under the assumption of the axiom of choice. Since the first two requirements seem very reasonable, it makes sense to relax the third requirement. However, the additivity for countably additive families of disjoint sets guarantees the validity of very simple and important limit theorems. If one really wants to require that the third condition holds only for finite families of pairwise disjoint sets, then such mapping does not exist for all $n \geq 3$. This is a simple consequence of the famous Banach-Tarski paradox from 1924.

1.2 Banach-Tarski paradox. *Let $n \geq 3$. Suppose that U and V are open and bounded subsets of $\mathbb{R}^n$. There exists $k \in \mathbb{N}$, sets $E_1, \ldots, E_k$, $F_1, \ldots, F_k$ in $\mathbb{R}^n$ which satisfy*

(i) *$E_1 \cup \cdots \cup E_k = U$ and $E_i \cap E_j = \emptyset$ whenever $i \neq j$;*

(ii) *$F_1 \cup \cdots \cup F_k = V$ and $F_i \cap F_j = \emptyset$ whenever $i \neq j$:*

(iii) *For each $i = 1, \ldots, k$ the set F_i is obtained from E_i after a finite number of translations, rotations and reflections.*

The sets E_i and F_i are bizarre. We cannot visualize them as they are obtained by the axiom of choice. One cam apply Banach-Tarski paradox to prove that an additive measure μ_n with the requirements from above does not exist for $n \geq 3$. Indeed, first note that the Banach-Tarski paradox yields that $\mu_n(U) = \mu_n(V)$ for all nonempty open and bounded subsets of $\mathbb{R}^n$.

Consider the sets $U = (0,1) \times \cdots \times (0,1)$ and $V = (-1,1) \times \cdots \times (-1,1)$. Then $\mu_n(U) = \mu_n(V)$. If we take $K = [0,1) \times \cdots \times [0,1)$, by assumption we have $\mu_n(K) = 1$ so that by monotonicity of μ_n we have $\mu_n(U) = \mu_n(V) < \infty$. If W is an nonempty open subset of $V \setminus U$, then $\mu(V) \geq \mu(U \cup W) = \mu(U) + \mu(W)$ yields $\mu(W) \neq 0$. This contradicts the conclusion of the Banach-Tarski paradox.

The conclusion of the story from above is that focusing on measures that are defined on every subset is sometimes not fruitful. Instead we will focus on measures which are defined on a sufficiently big class of subsets.

2 Measures

2.1 σ-algebras

Let X be a nonempty set. A family $\mathcal{A}$ of subsets of X is said to be a ***σ-algebra*** if the following three conditions are satisfied:

(i) $X \in \mathcal{A}$;

(ii) if $A \in \mathcal{A}$, then $A^c \in \mathcal{A}$ as well;

(iii) for a countable family $(A_n) \subseteq \mathcal{A}$ the union $\bigcup_{n=1}^{\infty} A_n$ is in $\mathcal{A}$.

The sets of $\mathcal{A}$ are called ***measurable sets*** and the pair $(X, \mathcal{A})$ is called ***measurable space***. If in the definition of σ-algebra we replace (iii) with

(iii') for a finite family $A_1, \ldots, A_k \in \mathcal{A}$ the union $\bigcup_{n=1}^{k} A_n$ is in $\mathcal{A}$,

we obtain the definition of an ***algebra***.

2.1.1 Remark

(i) Since $X \in \mathcal{A}$, $\emptyset = X^c \in \mathcal{A}$.

(ii) Every σ-algebra is also an algebra.

(iii) The DeMorgan law implies that the countable intersection of members of a given σ-algebra again belongs to the σ-algebra.

2.1.2 Example Let X be a nonempty set.

(i) The family $\{\emptyset, X\}$ is a σ-algebra which is contained in every σ-algebra on X. This σ-algebra is called ***trivial***

(ii) The ***power σ-algebra*** $\mathbb{P}(X)$ is a σ-algebra on X that contains every σ-algebra on X.

(iii) The family $\mathcal{A}$ of all subsets A of X for which A or A^c is countable is a σ-algebra on X. It is easy to see $\mathcal{A} \neq \mathbb{P}(X)$ if X is uncountable.

2.1.3 Proposition *Let $\mathcal{B}$ be a family of subsets of a nonempty set X. Then the intersection of all σ-algebras on X which contains $\mathcal{B}$ is the smallest σ-algebra which contains $\mathcal{B}$.*

Proof. Let $\mathcal{C}$ be the family of all σ-algebras on X which contain $\mathcal{B}$. Then $\mathcal{C}$ is nonempty since $\mathbb{P}(X) \in \mathcal{C}$. Let $\mathcal{A}$ be the intersection of all members of $\mathcal{C}$. Then $\mathcal{A}$ is a σ-algebra and $\mathcal{B} \subseteq \mathcal{A}$. That $\mathcal{A}$ is the smallest σ-algebra which contains $\mathcal{B}$ is obvious by the definition. $\square$

The σ-algebra from Proposition 2.1.3 is called the σ-algebra generated by $\mathcal{B}$ and is denoted by $\sigma(\mathcal{B})$.

Borel Sets

Let (X, τ) be a topological space. By Proposition 2.1.3 there exists the smallest σ-algebra on X which contains all open sets of X. This σ-algebra is called the ***Borel σ-algebra*** on X and is denoted by $\mathcal{B}(X)$. Since each σ-algebra is closed under taking complements, it is obvious that Borel σ-algebra is also generated by the family of all closed subsets of X. Since every σ-algebra is closed under taking countable intersections and unions, a countable intersection of open sets is an element of $\mathcal{B}(X)$. Such sets are called ***G_δ-sets***. Also, the countable union of closed sets are again in $\mathcal{B}(X)$. This sets are called ***F_σ-sets***.

2.1.4 Example Since $[a, b) = \bigcup_{n=1}^{\infty}[a, b - \frac{1}{n}]$, we see that $[a, b)$ is a F_σ-set. Also, since $[a, b) = \bigcap_{n=1}^{\infty}(a - \frac{1}{n}, b)$, the set $[a, b)$ is also a G_δ-set.

2.1.5 Proposition *$\mathcal{B}(\mathbb{R})$ is generated by each of the following:*

(i) *the open intervals: $\mathcal{E}_1 = \{(a, b) : \ a < b\}$,*

(ii) *the closed intervals: $\mathcal{E}_2 = \{[a, b] : \ a < b\}$,*

(iii) *the half-open intervals: $\mathcal{E}_3 = \{(a, b] \ a < b\}$ or $\mathcal{E}_4\{[a, b) : \ a < b\}$,*

(iv) *the open rays: $\mathcal{E}_5 = \{(a, \infty)\}$ or $\mathcal{E}_6 = \{(-\infty, b)\}$,*

(v) *the closed rays: $\mathcal{E}_7 = \{[a, \infty)\}$ or $\mathcal{E}_8 = \{(-\infty, b]\}$,*

Proof. All sets in $\mathcal{E}_i$ for $i = 1, \dots, 8$ are Borel sets, so that $\sigma(\mathcal{E}_i) \subseteq \mathcal{B}(\mathbb{R})$. Since each open set in $\mathcal{R}$ is a countable union of sets in $\mathcal{E}_1$ we conclude $\sigma(\mathcal{E}_1) = \mathcal{B}(\mathbb{R})$.

Since $(a, b) = \bigcup_{n=1}^{\infty}[a + \delta_n, b - \delta_n]$ for an appropriate sequence (δ_n), we conclude $\sigma(\mathcal{E}_2) = \mathcal{B}(\mathbb{R})$. The reader should complete the proof.

□

2.2 Positive Measures

Examples of positive measures are the length of subsets of $\mathbb{R}$, the area of sets in the plane $\mathbb{R}^2$ or the volume of bodies in the space $\mathbb{R}^3$. We will generalize this notion to arbitrary measurable spaces.

Positive measure or just a ***measure*** on a measurable space $(X, \mathcal{A})$ is a set function $\mu \colon \mathcal{A} \to [0, \infty]$ which satisfies the following:

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

(ii) For every sequence of pairwise disjoint sets (A_n) in $\mathcal{A}$ we have

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

The triple $(X, \mathcal{A}, \mu)$ is called a ***space with measure***. The condition (ii) is called ***countable additivity*** or ***σ-additivity***. If for each $k \in \mathbb{N}$ in (ii) we take $A_{k+1} = A_{k+2} = \cdots = \emptyset$, we obtain

$$\mu\Big(\bigcup_{n=1}^{k} A_n\Big) = \sum_{n=1}^{k} \mu(A_n).$$

This condition is called ***finite additivity***.

2.2.1 Example Let X be a nonempty set.

(i) Pick $x \in X$. The set function $\delta_x \colon \mathbb{P}(X) \to [0, \infty)$ defined as

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

is a positive measure which is called the ***Dirac measure***.

(ii) The set function $\mu \colon \mathbb{P}(X) \to [0, \infty]$ defined as

$$\mu(A) = \begin{cases} |A|, & A \text{ is finite} \\ \infty, & \text{otherwise} \end{cases}$$

is a positive measure on $\mathbb{P}(X)$. This measure is called the ***counting measure***.

(iii) Suppose X is uncountable and $\mathcal{A}$ is the σ-algebra of all sets which are either countable or their complement is countable. Then the set function $\mu \colon \mathcal{A} \to [0, \infty]$ defined as

$$\mu(A) = \begin{cases} 0, & A \text{ is countable} \\ 1, & \text{otherwise} \end{cases}$$

is again a positive measure on $\mathcal{A}$.

(iv) Suppose X is infinite. Then the set function $\mu \colon \mathbb{P}(X) \to [0, \infty]$ defined as

$$\mu(A) = \begin{cases} 0, & A \text{ is finite} \\ \infty, & \text{otherwise} \end{cases}$$

is finitely additive which is not countably additive.

Suppose $(X, \mathcal{A}, \mu)$ is a space with measure. A measure μ is called ***finite*** if $\mu(X) < \infty$. If X can be written as a countable union of sets with finite measure, then μ is said to be ***σ-finite***.

2.2.2 Lemma *If $\mu \colon \mathcal{A} \to [0, \infty]$ is a finitely additive function on an algebra $\mathcal{A}$, then $\mu(A) \leq \mu(B)$ for all sets in $\mathcal{A}$ with $A \subseteq B$.*

Proof. Write $B = A \cup (B \setminus A)$. Then $\mu(B) = \mu(A) + \mu(B \setminus A) \geq \mu(A)$. □

Every measure is subadditive.

2.2.3 Proposition *Let $(X, \mathcal{A}, \mu)$ be a space with measure. Then for every sequence (A_n) in $\mathcal{A}$ we have*

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

Proof. Define $B_1 := A_1$ and $B_k = A_k \setminus (A_1 \cup \cdots \cup A_{k-1})$ for each $k > 1$. Since the sets $(B_n)_n$ are pairwise disjoint and for each $k \in \mathbb{N} \cup \{\infty\}$ we have $\bigcup_{n=1}^{k} A_n = \bigcup_{n=1}^{k} B_n$ we conclude

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

□

2.2.4 Proposition *A finitely additive set function $\mu \colon \mathcal{A} \to [0, \infty]$ on a measurable space $(X, \mathcal{A})$ is a measure iff for each increasing sequence (A_n) of sets in $\mathcal{A}$ we have*

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

Proof. Suppose first μ is a measure. Denote by A the union of sets $A_1, A_2, \ldots$. Define $B_1 = A_1$ and $B_n = A_n \setminus A_{n-1}$ for $n > 1$. The sets (B_n) are pairwise disjoint and their union is A. Since μ is a measure, we have

$$\begin{aligned}\mu(A) &= \sum_{k=1}^{\infty} \mu(B_k) = \lim_{n\to\infty} \sum_{k=1}^{n} \mu(B_k) \\ &= \mu(A_1) + \lim_{n\to\infty} \sum_{k=2}^{n} \mu(A_k \setminus A_{k-1}) = \lim_{n\to\infty} \mu(A_n).\end{aligned}$$

This works only if measures of sets A_n are finite. If some set A_n has infinite measure, then by monotonicity of the measure both sides from the proposition are infinite and we have an equality.

For the converse, let (A_n) be a sequence of pairwise disjoint measurable sets. Denote $B_n = A_1 \cup \cdots \cup A_n$. Since μ is finitely additive, we have $\mu(B_n) = \sum_{k=1}^{n} \mu(A_n)$. Since (B_n) is increasing with union $\bigcup_{n=1}^{\infty} A_n$, we have

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

□

2.2.5 Corollary *Let $(X, \mathcal{A}, \mu)$ be a space with measure and let (A_n) be a decreasing sequence in $\mathcal{A}$ with $\mu(A_1) < \infty$. Then*

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

Proof. Since (A_n) is decreasing, the sequence $(A_1 \setminus A_n)$ is an increasing sequence in $\mathcal{A}$ and it satisfies

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

Since $\mu(A_1) < \infty$, by Proposition 2.2.4 we have

$$\mu(A_1) - \mu\left(\bigcap_{n=1}^{\infty} A_n\right) = \mu(A_1) - \lim_{n\to\infty} \mu(A_n).$$

The conclusion of corollary now immediately follows. □

2.2.6 Example Let μ be the counting measure on $(\mathbb{N}, \mathbb{P}(\mathbb{N}))$. Define $A_n := \{k : \ k \geq n\}$. Then $\mu(A_n) = \infty$ while the intersection $A := \bigcap_{n=1}^{\infty} A_n = \emptyset$ and so $\mu(A) = 0$.

2.2.7 Example Later on we will see that there exists a measure m defined on a Borel σ-algebra $\mathcal{B}(\mathbb{R})$ which coincides with the usual length on all intervals of $\mathbb{R}$.

The generalized Cantor set C is a set that is obtained with the following procedure.

Pick $0 < \alpha_1 < 1$. Take the interval $I_0 = [0, 1]$ and delete the open middle interval centered at $\frac{1}{2}$ whose length is $\alpha_1 \cdot m(I_0)$. The resulting set is denoted by I_1. Pick $0 < \alpha_2 < 1$ and from each interval in I_1 delete the open middle interval with center at its midpoint and length $\frac{1}{2}\alpha_2 m(I_1)$. The total length of deleted intervals is $\alpha_2 \cdot m(I_1)$. The resulting set is denoted by I_2. On the n-th step we choose $0 < \alpha_n < 1$ and from each interval in I_{n-1} we delete the open middle interval centered at its midpoint with the length $\frac{1}{2^n}\alpha_n m(I_{n-1})$. The resulting set is denoted by I_n. The total length of deleted intervals is $\alpha_n \cdot m(I_{n-1})$. The set $\bigcap\limits_{n=1}^{\infty} I_n$ is called the ***generalized Cantor set***. By the construction we have $m(I_n) = (1 - \alpha_n)m(I_{n-1})$ and so

$$m(I_n) = (1 - \alpha_1) \cdot \dots \cdot (1 - \alpha_n).$$

By Corollary 2.2.5 we have

$$m(C) = \lim_{n\to\infty} m(I_n) = \prod_{n=1}^{\infty}(1 - \alpha_n).$$

If we take $\alpha_n = \frac{1}{3}$, we get the regular Cantor set whose Lebesgue measure is zero.

2.2.8 Proposition *The generalized Cantor set C is a metrizable, compact, nowhere dense, totally disconnected space without isolated points. The Lebesgue measure of C is $\prod_{n=1}^{\infty}(1-\alpha_n)$.*

2.2.9 Corollary *The generalized Cantor set has a positive measure iff the series $\sum_{n=1}^{\infty}\alpha_n$ converges.*

2.3 Completion of a Space with Measure

A space with measure $(X, \mathcal{A}, \mu)$ is said to be ***complete*** whenever $\mu(A) = 0$ and $B \subseteq A$ imply $B \in \mathcal{A}$. The sets with measure zero are said to be μ***-null***. The collection of all sets with measure zero is denoted by $\mathcal{N}$.

2.3.1 Lemma *Let $(X, \mathcal{A}, \mu)$ be a space with measure. Given a sequence $(N_n) \subseteq \mathcal{N}$, the set $\bigcup_{n=1}^{\infty} N_n$ is again in $\mathcal{N}$.*

Every measure space can be completed.

2.3.2 Theorem *Let $(X, \mathcal{A}, \mu)$ be a space with measure. Let $\overline{\mathcal{A}}$ be the set of all $B \subseteq X$ of the form $B = A \cup S$ with $A \in \mathcal{A}$ and $S \subseteq N$ for some $N \in \mathcal{N}$. Then $\overline{\mathcal{A}}$ is a σ-algebra on X and the set function $\overline{\mu}\colon \overline{\mathcal{A}} \to [0, \infty]$ defined as $\overline{\mu}(B) := \mu(A)$ is a positive measure on $(X, \overline{\mathcal{A}})$.*

The space with measure $(X\overline{\mathcal{A}}, \overline{\mu})$ is called the ***completion*** of the space with measure $(X, \mathcal{A}, \mu)$.

Proof. Since $X = X \cup \emptyset$, we have $X \in \overline{\mathcal{A}}$. If $B = A \cup S \in \overline{\mathcal{A}}$ for some $A \in \mathcal{A}$ and $S \subseteq N$ for some $N \in \mathcal{N}$, then

$$(A \cup S)^c = A^c \cap S^c = A^c \cap (N^c \cup N \setminus S) = (A^c \cap N^c) \cup (A^c \cap N \cap S^c)$$

where $(A^c \cap N^c) \in \mathcal{A}$ and $A^c \cap N \in \mathcal{N}$.

Pick a sequence (B_n) in $\overline{\mathcal{A}}$. Then for each $n \in \mathbb{N}$ there exist $A_n \in \mathcal{A}$, $N_n \in \mathcal{N}$ and $S_n \subseteq N_n$ with $B_n = A_n \cup S_n$. Then

$$\bigcup_{n=1}^{\infty} B_n = \bigcup_{n=1}^{\infty} A_n \cup \bigcup_{n=1}^{\infty} S_n$$

and since $\bigcup_{n=1}^{\infty} N_n \in \mathcal{N}$, we have $\bigcup_{n=1}^{\infty} B_n \in \overline{\mathcal{A}}$.

Define $\overline{\mu}\colon \mathcal{A} \to [0, \infty]$ as $\overline{\mu}(B) := \mu(A)$. If $B = A_1 \cup S_1 = A_2 \cup S_2$ where $S_1 \subseteq N_1$ and $S_2 \subseteq N_2$ for some $A_1, A_2 \in \mathcal{A}$ and $N_1, N_2 \in \mathcal{N}$, then

$$\overline{\mu}(B_1) = \mu(A_1) \leq \mu(A_2 \cup N_2) \leq \mu(A_2) = \overline{\mu}(B_2).$$

By symmetry we also have $\overline{\mu}(B_2) \leq \overline{\mu}(B_1)$, so that we have an equality. Obviously we have $\overline{\mu}(\emptyset) = 0$. If $(B_n) \subseteq \overline{\mathcal{A}}$ are pairwise disjoint, then $(A_n) \subseteq \mathcal{A}$ are pairwise disjoint, from where it follows

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

Suppose $\overline{\mu}(N) = 0$ and pick $S \subseteq N$. Then $S = \emptyset \cup S$ implies $S \in \overline{\mathcal{A}}$, so that the space with measure $(X, \overline{\mathcal{A}}, \overline{\mu})$ is complete. □

As we already mentioned, later will see that there exists a unique measure m defined on the Borel σ-algebra $\mathcal{B}(\mathbb{R})$ which satisfies $m([a,b]) = b-a$ for each interval $[a,b]$. This measure is ***invariant under translation*** which means $m(A+x) = m(A)$ for each $x \in R$ and each Borel set $A \subseteq \mathbb{R}$. The cardinality of $\mathcal{B}(\mathbb{R})$ is continuum $\mathfrak{c}$. The Cantor set C is a closed set with measure zero. It is well-known that the cardinality of the Cantor set is $\mathfrak{c}$. If $(\mathbb{R}, \mathcal{B}(\mathbb{R}), m)$ is complete, then every subset of C would be a Borel set, so that the cardinality of $\mathcal{B}(\mathbb{R})$ is at least $2^{\mathfrak{c}}$. Hence, $(\mathbb{R}, \mathcal{B}(\mathbb{R}), m)$ is not complete. Its completion is $(\mathbb{R}, \mathcal{L}(\mathbb{R}), \overline{m})$ where $\mathcal{L}(\mathbb{R})$ is called the ***Lebesgue σ-alegbra*** on $\mathbb{R}$ and its members are called ***Lebesgue measurable sets***. Hence, a set B is Lebesgue measurable whenever there exists a Borel set $A \in \mathcal{B}(\mathbb{R})$, a Borel set N with $m(N) = 0$ and $S \subseteq N$ such that $B = A \cup S$.

2.3.3 Theorem *The cardinality of the Lebesgue σ-algebra $\mathcal{L}(\mathbb{R})$ is $2^{\mathfrak{c}}$.*

Proof. We already proved that the cardinality of $\mathcal{L}(\mathbb{R})$ is at least $2^{\mathfrak{c}}$. On the other hand $\mathcal{L}(\mathbb{R}) \subseteq \mathbb{P}(\mathbb{R})$, so that the cardinality of $\mathcal{L}(\mathbb{R})$ is also smaller than or equal to $2^{\mathfrak{c}}$. Hence $|\mathcal{L}(\mathbb{R})| = 2^{\mathfrak{c}}$. □

The following example shows that there is no translation invariant measure defined on $\mathbb{P}(\mathbb{R})$ such that on intervals it coincides with the usual length.

2.3.4 Example On the set of all real numbers $\mathbb{R}$ we introduce the relation $\sim$ as $x \sim y$ iff $x - y \in \mathbb{Q}$. It is easy to see that $\sim$ is equivalence relation on $\mathbb{R}$. The equivalence classes are the elements of the quotient group $\mathbb{R}/\mathbb{Q}$. In each equivalence class $x + \mathbb{Q}$ where $x \in [-1,1]$ choose an element from $(x+\mathbb{Q}) \cap [-1,1]$ and let S be the set of all chosen elements. This can be made by Axiom of choice.

Denote $\mathbb{Q}_1 = \mathbb{Q} \cap [-2,2]$. The family of sets $\mathcal{F} := \{r + S : r \in \mathbb{Q}_1\}$ are disjoint and since for each $x \in [-1,1]$ there exists $y \in S$ such that $q := x - y \in \mathbb{Q}$ and $|q| \leq |x| + |y| \leq 2$, the union of the family $\mathcal{F}$ contains $[-1,1]$. For $q \in [-2,2]$ we have $q + S \subseteq [-3,3]$. If the set S would be measurable for some translation invariant measure which coincides with the length on the intervals, then we would have

$$2 = \mu([-1,1]) \leq \mu\left(\bigcup_{q \in \mathbb{Q}_1} (q+S)\right) = \sum_{q \in \mathbb{Q}_1} \mu(q+S) \leq 6.$$

Translation invariance would yield $\mu(q+S)=0$ for each $q \in \mathbb{Q}_1$ which is a contradiction.

Solovay [**?**] proved in 1964 that without the axiom of choice one cannot construct a non-Lebesgue measurable set. Even if one adds the countable version of the axiom of choice to the Zermelo-Frankel **ZF** axioms it is not possible to construct a non-Lebesgue measurable set. To be precise, recall first that the relation R on a nonempty set X is said to be ***entire*** whenever for any $x \in X$ there is $y \in X$ with xRy. On the set X we say to have an ***axiom*** of ***dependent choice*** **DC** whenever for any entire relation R there exists a sequence (x_n) in X such that $x_n R x_{n+1}$. It turns out that **DC** implies the ***axiom of countable choice***. Solovay's model satisfies **ZF+DC** and every subset of reals is Lebesgue measurable.

2.4 Outer Measure

An ***outer measure*** on a set X is a set function $\zeta\colon \mathbb{P}(X) \to [0,\infty]$ which satisfies the following

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

(ii) $\zeta(A) \le \zeta(B)$ whenever $A \subseteq B$;

(iii) $\zeta(\bigcup_{n=1}^{\infty} A_n) \le \sum_{n=1}^{\infty} \zeta(A_n)$ for any sequence of sets (A_n) in X.

The following proposition tells how to generate outer measures.

2.4.1 Proposition *Let $\mathcal{S}$ be a family of subsets of a nonempty set X. Suppose $\emptyset, X \in \mathcal{S}$ and let $\mu\colon \mathcal{S} \to [0,\infty]$ be a set function with $\mu(\emptyset)=0$. Then the set function $\mu^*\colon \mathbb{P}(X) \to [0,\infty]$ defined as*

$$\mu^*(Y) := \inf\left\{ \sum_{j=1}^{\infty} \mu(A_j) : \ (A_j) \subseteq S \text{ is a cover for } Y \right\}$$

is an outer measure.

Proof. Obviously $\mu^*(\emptyset)=\mu(\emptyset)=0$. Suppose $Y \subseteq Z$. Since each cover for Z is also a cover for Y, we have $\mu^*(Y) \le \mu^*(Z)$.

Pick $(Y_n) \subseteq \mathbb{P}(X)$. If $\mu^*(Y_n)=\infty$ for some $n \in \mathbb{N}$, then countable subadditivity follows from monotonicity of μ^*. Suppose now $\mu^*(Y_n) < \infty$ for each $n \in \mathbb{N}$ and pick an arbitrary $\epsilon > 0$. For each $n \in \mathbb{N}$ find a countable cover $(A_{n,j})_j$ for Y_n such that

$$\sum_{j=1}^{\infty} \mu(A_{n,j}) \le \mu^*(Y_n) + \frac{\epsilon}{2^n}.$$

Since the countable family $(A_{n,j})_{n,j}$ covers $Y := \bigcup_{n=1}^{\infty} Y_n$ we have

$$\mu^*(Y) \leq \sum_{n=1}^{\infty}\sum_{j=1}^{\infty} \mu(A_{n,j}) \leq \sum_{n=1}^{\infty}(\mu^*(Y_n) + \frac{\epsilon}{2^n}) = \sum_{n=1}^{\infty} \mu^*(Y_n) + \epsilon.$$

Since $\epsilon > 0$ was chosen arbitrarily, we have $\mu^*(Y) \leq \sum_{n=1}^{\infty} \mu^*(Y_n)$. □

Carathéodory's theorem will tell us how outer measures on X induce positive measures on σ-algebras.

Let ζ be an outer measure on X. A set $A \subseteq X$ is said to be ***ζ-measurable*** if

$$\zeta(Y) = \zeta(A \cap Y) + \zeta(A^c \cap Y)$$

for every subset Y of X. The family of all ζ-measurable subsets of X is denoted by $\mathcal{A}_\zeta$.

2.4.2 Remark Since every outer measure is countably subadditive, the inequality

$$\zeta(Y) \leq \zeta(A \cap Y) + \zeta(A^c \cap Y)$$

holds for all subsets A and Y of X. Therefore, to prove that a given set A is ζ-measurable it suffices to prove

$$\zeta(Y) \geq \zeta(A \cap Y) + \zeta(A^c \cap Y) \tag{2.1}$$

for each subset Y of X. Since the latter inequality holds whenever $\zeta(Y) = \infty$, we conclude that a set $A \subseteq X$ is ζ-measurable iff (2.1) holds for all subsets $Y \subseteq X$ with $\zeta(Y) < \infty$.

2.4.3 Carathéodory's Theorem *If ζ is an outer measure on X, the family $\mathcal{A}_\zeta$ is a σ-algebra on X, the restriction $\zeta|_{\mathcal{A}_\zeta}$ is a measure and $(X, \mathcal{A}_\zeta, \zeta|_{\mathcal{A}_\zeta})$ is a complete space with measure.*

Proof. We first prove that $\mathcal{A}_\zeta$ is an algebra on X. Since $X^c \cap Y = \emptyset$ and $X \cap Y = Y$, we immediately conclude $X \in \mathcal{A}_\zeta$. Also, since the requirement that a set A is ζ-measurable is symmetric in A and A^c, we conclude that $\mathcal{A}_\zeta$ is closed under taking complements.

Choose $A, B \in \mathcal{A}_\zeta$. We will prove $A \cup B \in \mathcal{A}_\zeta$. Pick any $Y \subseteq X$. We need to prove

$$\zeta((A \cup B) \cap Y) + \zeta((A \cup B)^c \cap Y) \leq \zeta(Y).$$

Since $A \cup B = A \cup (B \cap A^c)$ the left-hand side can be written as

$$\zeta((A \cup (B \cap A^c)) \cap Y) + \zeta(B^c \cap A^c \cap Y).$$

Since ζ is subadditive, the latter expression is less than or equal to

$$\zeta(A \cap Y) + \zeta(B \cap A^c \cap Y) + \zeta(B^c \cap A^c \cap Y).$$

Since $B \in \mathcal{A}_\zeta$, we have $\zeta(B \cap A^c \cap Y) + \zeta(B^c \cap A^c \cap Y) = \zeta(A^c \cap Y)$ and since $A \in \mathcal{A}_\zeta$, we have

$$\zeta((A \cup B) \cap Y) + \zeta((A \cup B)^c \cap Y) \leq \zeta(A \cap Y) + \zeta(A^c \cap Y) = \zeta(Y).$$

Since $\mathcal{A}_\zeta$ is an algebra on X, every countable union of sets from $\mathcal{A}_\zeta$ can be written as a union of sequence of pairwise disjoint sets $A_j \in \mathcal{A}_\zeta$. Define $B := \bigcup_{j=1}^\infty A_j$. If we prove $B \in \mathcal{A}_\zeta$ and $\zeta(B) = \sum_{j=1}^\infty \zeta(A_j)$, then $\mathcal{A}_\zeta$ is a σ-algebra on X and $\zeta|_{\mathcal{A}_\zeta}$ is a measure on $\mathcal{A}_\zeta$.

For each $n \in \mathbb{N}$ define $B_n = \bigcup_{j=1}^n A_j$. Since $A_n \in \mathcal{A}_\zeta$ for each $Y \subseteq X$ we have

$$\begin{aligned} \zeta(Y \cap B_n) &= \zeta((Y \cap B_n) \cap A_n) + \zeta((Y \cap B_n) \cap A_n^c) \\ &= \zeta(Y \cap A_n) + \zeta(Y \cap B_{n-1}). \end{aligned}$$

By easy induction we obtain

$$\zeta(Y \cap B_n) = \sum_{j=1}^n \zeta(Y \cap A_j). \tag{2.2}$$

By taking $Y = B_n$ we conclude ζ is finitely additive on $\mathcal{A}_\zeta$. Since $B_n \in \mathcal{A}_\zeta$ and ζ is monotone, for each $Y \subseteq X$ we have

$$\begin{aligned} \zeta(Y) &= \zeta(Y \cap B_n) + \zeta(Y \cap B_n^c) \geq \zeta(Y \cap B_n) + \zeta(Y \cap B^c) \\ &= \sum_{j=1}^n \zeta(Y \cap A_j) + \zeta(Y \cap B^c). \end{aligned}$$

Letting $n \to \infty$ and applying σ-subadditivity of ζ we conclude

$$\zeta(Y) \geq \sum_{j=1}^\infty \zeta(Y \cap A_j) + \zeta(Y \cap B^c) \geq \zeta(Y \cap B) + \zeta(Y \cap B^c)$$

which proves $B \in \mathcal{A}_\zeta$. If we take $Y = B_n$ in (2.2) and apply $\zeta(B) \geq \zeta(B_n)$ we conclude

$$\zeta(B) \geq \sum_{j=1}^n \zeta(A_j).$$

Letting $n \to \infty$ and applying σ-subadditivity we conclude $\zeta(B) = \sum_{j=1}^\infty \zeta(A_j)$.

The only thing left to prove is completeness: Suppose $N \in \mathcal{A}_\zeta$ with $\zeta(N) = 0$ and $A \subseteq N$. Pick $Y \subseteq X$. Since ζ is monotone, we have $\zeta(A \cap Y) = 0$, so that

$$\zeta(Y) \leq \zeta(A \cap Y) + \zeta(A^c \cap Y) = \zeta(A^c \cap Y) \leq \zeta(Y).$$

□

2.5 Measures on Algebras

A ***measure*** on an algebra $\mathcal{A}$ is a set function $\mu\colon \mathcal{A} \to [0,\infty]$ which satisfies

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

(ii) $\mu(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mu(A_n)$ for every sequence of disjoint sets (A_n) in $\mathcal{A}$ whose union $\bigcup_{n=1}^{\infty} A_n$ belongs to $\mathcal{A}$.

Formally it can happen that the union $\bigcup_{n=1}^{\infty} A_n$ does not belong to $\mathcal{A}$. Every measure on an algebra can be extended to an outer measure.

2.5.1 Theorem *Let μ be a measure on an algebra $\mathcal{A}$ on some set X. For $Y \subseteq X$ we define*

$$\mu^*(Y) = \inf\left\{ \sum_{n=1}^{\infty} \mu(A_n) : \ (A_n) \subseteq \mathcal{A} \text{ is a cover for } Y \right\}.$$

Then μ^* is an outer measure on X and $\mu^*(A) = \mu(A)$ for each $A \in \mathcal{A}$.

Proof. By Proposition 2.4.1 μ^* is an outer measure on X. Since the countable family $\{A, \emptyset, \emptyset, \ldots\}$ covers $A \in \mathcal{A}$, by definition we have $\mu^*(A) \leq \mu(A) + \mu(\emptyset) + \mu(\emptyset) + \cdots = \mu(A)$. For the reverse inequality pick any sequence (A_n) in $\mathcal{A}$ which covers A. Define $B_1 = A \cap A_1$ and $B_n = A \cap (A_n \setminus (A_1 \cup \cdots \cup A_{n-1}))$ for $n > 1$. Since μ is a measure on the algebra $\mathcal{A}$ we conclude

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

Since (A_n) was an arbitrary cover for A with elements in $\mathcal{A}$, by definition we have $\mu(A) \leq \mu^*(A)$. □

Let μ_0 be a measure on an algebra $\mathcal{A}$ on some set X. Then μ_0 induces an outer measure μ_0^* on X. By Carathéodory's theorem the family $\mathcal{A}_{\mu_0^*}$ of all μ_0^*-measurable sets is a σ-algebra and $\mu := \mu_0^*|_{\mathcal{A}_{\mu_0^*}}$ is a complete measure on $\mathcal{A}_{\mu_0^*}$.

2.5.2 Theorem *Let μ_0, $\mathcal{A}$, μ and $\mathcal{A}_{\mu_0^*}$ be as before.*

(i) *We have $\mathcal{A} \subseteq \mathcal{A}_{\mu_0^*}$.*

(ii) *If ν is any extension of the measure μ_0, then for any $A \in \mathcal{A}_{\mu_0^*}$ we have $\nu(A) \leq \mu(A)$.*

(iii) *If $\mu(A) < \infty$, then $\nu(A) = \mu(A)$.*

(iv) *If μ_0 is σ-finite, then μ is the only extension of μ_0 on the σ-algebra $\mathcal{A}_{\mu_0^*}$.*

Proof. (i) Let $A \in \mathcal{A}$. To prove $A \in \mathcal{A}_{\mu_0^*}$ it suffices to prove

$$\mu_0^*(A \cap Y) + \mu_0^*(A^c \cap Y) \leq \mu_0^*(Y)$$

for any subset $Y \subseteq X$ with $\mu_0^*(Y) < \infty$. Pick $\epsilon > 0$ and a cover $(A_n) \subseteq \mathcal{A}$ for Y such that

$$\sum_{n=1}^{\infty} \mu_0(A_n) \leq \mu_0^*(Y) + \epsilon.$$

Since $(A_n \cap A)$ and $(A_n \cap A^c)$ are coverings for $(Y \cap A)$ and $(Y \cap A^c)$, respectively, we have

$$\mu_0^*(Y \cap A) \leq \sum_{n=1}^{\infty} \mu_0(A_n \cap A) \qquad \text{in} \qquad \mu_0^*(Y \cap A^c) \leq \sum_{n=1}^{\infty} \mu_0(A_n \cap A^c).$$

By summing the inequalities and using the fact that μ_0 is a measure on the algebra $\mathcal{A}$ we obtain

$$\begin{aligned} \mu_0^*(Y \cap A) + \mu_0^*(Y \cap A^c) &\leq \sum_{n=1}^{\infty} \mu_0(A_n \cap A) + \sum_{n=1}^{\infty} \mu_0(A_n \cap A^c) \\ &= \sum_{n=1}^{\infty} \mu_0(A_n) < \mu_0^*(Y) + \epsilon. \end{aligned}$$

(ii) Pick another measure ν on $\mathcal{A}_{\mu_0^*}$ with $\nu|_{\mathcal{A}} = \mu$. Pick $A \in \mathcal{A}_{\mu_0^*}$. Then for any countable cover $(A_n) \subseteq \mathcal{A}$ of A we have

$$\nu(A) \leq \sum_{n=1}^{\infty} \nu(A_n) = \sum_{n=1}^{\infty} \mu_0(A_n).$$

This implies $\nu(A) \leq \mu_0^*(A) = \mu(A)$.

(iii) Suppose $A \in \mathcal{A}_{\mu_0^*}$ satisfies $\mu(A) < \infty$. Pick $\epsilon > 0$ and find a countable cover $(A_n) \subseteq \mathcal{A}_{\mu_0^*}$ for A with

$$\sum_{n=1}^{\infty} \mu_0(A_n) < \mu_0^*(A) + \epsilon.$$

Denote by B the set $\bigcup_{n=1}^{\infty} A_n$. Then

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

Since μ is a measure, we conclude

$$\mu(B \setminus A) = \mu(B) - \mu(A) \leq \sum_{n=1}^{\infty} \mu(A_n) - \mu(A) < \epsilon.$$

Hence,

$$\mu(A) \leq \mu(B) = \nu(B) = \nu(A) + \nu(B \setminus A) \leq \nu(A) + \mu(B \setminus A) \leq \nu(A) + \epsilon.$$

This yields $\mu(A) \leq \nu(A)$.

(iv) If μ_0 is σ-finite, then X can be written as a countable union of pairwise disjoint sets $X_n \in \mathcal{A}$ with $\mu_0(X_n) < \infty$. Then for any set $A \in \mathcal{A}_{\mu_0^*}$ we have $\mu(A \cap X_n) < \infty$, so that

$$\nu(A) = \sum_{n=1}^{\infty} \nu(A \cap X_n) = \sum_{n=1}^{\infty} \mu(A \cap X_n) = \mu(A).$$

□

2.6 Semi-algebras and Semi-measures

Lebesgue measure on $\mathbb{R}$ will be constructed from the length of the intervals. There is a problem since the family of all intervals of $\mathbb{R}$ is not even an algebra on $\mathbb{R}$. Let $\mathcal{S}(\mathbb{R}) \subseteq \mathbb{P}(\mathbb{R})$ be the family which contains $\emptyset$, all half-open intervals of the form $[a, b)$ and all intervals of the form $(-\infty, b)$ and $[a, \infty)$ for $a, b \in \mathbb{R}$. This family is not an algebra, yet it is a semi-algebra with respect to the following definition.

A ***semi-algebra*** on a set X is a family $\mathcal{S} \subseteq \mathbb{P}(X)$ which satisfies

(i) $\emptyset \in \mathcal{S}$;

(ii) if $A, B \in \mathcal{S}$, then $A \cap B \in \mathcal{S}$;

(iii) if $A \in \mathcal{S}$, then A^c can be written as a union of finitely many disjoint sets from $\mathcal{S}$.

A ***semi-measure*** on a semi-algebra $\mathcal{S}$ is a set function $\mu \colon \mathcal{S} \to [0, \infty]$ which satisfies

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

(ii) if $A_1, \dots, A_n \in \mathcal{S}$ are pairwise disjoint and their union is in $\mathcal{S}$, then $\mu(A) = \sum_{j=1}^{n} \mu(A_j)$.

(iii) for any sequence (A_n) of pairwise disjoint sets in $\mathcal{S}$ with union A in $\mathcal{A}$ we have $\mu(A) \leq \sum_{n=1}^{\infty} \mu(A_n)$.

2.6.1 Example On the semi-algebra $\mathcal{S}(\mathbb{R})$ we define the set function $\mu \colon \mathcal{S} \to [0, \infty]$ as follows:

$$\mu([a, b)) := b - a \ (b > a), \quad \mu((-\infty, b)) = \infty = \mu([a, \infty)), \quad \mu(\emptyset) = 0.$$

Then μ is a semi-measure on $\mathcal{S}(\mathbb{R})$.

2.6.2 Proposition *If $\mathcal{S}$ is a semi-algebra on a set X, then the set $\mathcal{A}$ of all finite unions of pairwise disjoint elements from $\mathcal{S}$ is an algebra on X which contains $\mathcal{S}$.*

Proof. Since $\emptyset \in \mathcal{S}$, we have $\mathcal{S} \subseteq \mathcal{A}$ and, in particular, $\emptyset \in \mathcal{A}$.

We need to show that $\mathcal{A}$ is closed under taking complements and finite unions. By de Morgan's law $\mathcal{A}$ is an algebra iff it is closed under taking complements and finite intersections. Suppose $A = \bigcup_{i=1}^{n} A_i$ and $B = \bigcup_{j=1}^{m} B_j$ with (A_i) and (B_j) pairwise disjoint in $\mathcal{S}$, respectively. Then

$$A \cap B = \bigcup_{i=1}^{n} \bigcup_{j=1}^{m} (A_i \cap B_j)$$

is a finite union of pairwise disjoint sets from $\mathcal{S}$. To see that $A^c \in \mathcal{A}$, note first that each A_i^c can be written as a finite union of pairwise disjoint sets from $\mathcal{S}$, so that $A_i^c \in \mathcal{A}$. Since $\mathcal{A}$ is closed under taking finite intersections, we conclude $A^c = \bigcap_{i=1}^{n} A_i^c \in \mathcal{A}$. □

It is obvious that $\mathcal{A}$ is the algebra generated by $\mathcal{S}$. Also, now it is clear that $\mathcal{A}$ contains all finite unions of sets from $\mathcal{S}$.

2.6.3 Theorem *Let $\mathcal{S}$ be a semi-algebra on X, μ a semi-measure on $\mathcal{S}$ and $\mathcal{A}$ the algebra generated by $\mathcal{S}$. If for each finite family $A_1, \dots, A_n$ of pairwise disjoint sets in $\mathcal{S}$ we define*

$$\widetilde{\mu}\Big(\bigcup_{j=1}^{n} A_j\Big) := \sum_{j=1}^{n} \mu(A_j),$$

then $\widetilde{\mu}$ is a measure on $\mathcal{A}$ which extends μ.

Proof. First we need to prove that $\widetilde{\mu}$ is well-defined. Suppose $A \in \mathcal{A}$ can be written as $A = \bigcup_{i=1}^{n} A_i = \bigcup_{j=1}^{m} B_j$ where (A_i) and (B_j) are pairwise disjoint, respectively. Since $B_j = \bigcup_{i=1}^{n} (B_j \cap A_i)$ and μ is a semi-measure on $\mathcal{S}$, we conclude

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

Similarly one can see

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

This proves $\widetilde{\mu}$ is well-defined. It is also clear that $\widetilde{\mu}$ extends μ.

The function $\widetilde{\mu}$ is obviously finitely-additive and therefore monotone. Pick a family (A_n) of pairwise disjoint sets in $\mathcal{A}$ and suppose their union $A := \bigcup_{n=1}^{\infty} A_n$ is in $\mathcal{A}$. Since $\bigcup_{i=1}^{n} A_i \subseteq A$ for each $n \in \mathbb{N}$, we have

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

Letting $n \to \infty$ we obtain

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

To prove the converse inequality, each A_j can be written as a finite union $A_j = \bigcup_i B_{ij}$ of pairwise disjoint sets $B_{ij} \in \mathcal{S}$. Since $A \in \mathcal{A}$, A can be also written as a finite union $A = \bigcup_k C_k$ of pairwise disjoint sets $C_k \in \mathcal{S}$. Since $C_k \cap B_{ij} \in \mathcal{S}$, since

$$C_k = \bigcup_{i,j} (C_k \cap B_{ij})$$

is a countable union of elements of $\mathcal{S}$ and since μ is a semi-measure, we conclude

$$\mu(C_k) \leq \sum_{i,j} \mu(C_k \cap B_{ij}).$$

Since

$$A_j = \left(\bigcup_i B_{ij}\right) \cap \left(\bigcup_k C_k\right) = \bigcup_{i,k} (C_k \cap B_{ij})$$

is a finite disjoint union of sets from $\mathcal{S}$, we have $\widetilde{\mu}(A_j) = \sum_{i,k} \mu(C_k \cap B_{ij})$ from where it follows

$$\widetilde{\mu}(A) = \sum_k \mu(C_k) \leq \sum_j \sum_{i,k} \mu(C_k \cap B_{ij}) = \sum_j \widetilde{\mu}(A_j).$$

□

2.7 Lebesgue-Stieltjes Measures

Let X be a topological space and $\mathcal{B}(X)$ the Borel σ-algebra on X. Any measure defined on $\mathcal{B}(X)$ is called the ***Borel measure*** on X.

Suppose μ is a Borel measure on $\mathbb{R}$. If μ is finite we can define its ***distribution function*** $F\colon \mathbb{R} \to \mathbb{R}$ by $F(x) = \mu((\infty, x))$. These functions are important in Probability theory since distributions of random variables are given by them.

2.7.1 Proposition *Let μ be a Borel measure on $\mathbb{R}$ which is finite on all bounded Borel sets. Then the function $F_\mu\colon \mathbb{R} \to \mathbb{R}$ given by*

$$F_\mu(x) = \begin{cases} \mu([0,x)) & : \ x > 0 \\ 0 & : \ x = 0 \\ -\mu([x,0)) & ; \ x < 0 \end{cases}$$

is increasing and left-continuous.

(i) *If $a < b$ then $\mu([a,b)) = F_\mu(b) - F_\mu(a)$.*

(ii) *If μ is finite, then $F = F_\mu + \mu((-\infty, 0))$.*

Proof. Since μ is monotone, the function F_μ is increasing. Suppose now that (x_n) is an increasing sequence of positive real numbers with limit $x > 0$. We need to prove $\lim_{n\to\infty} F_\mu(x_n) = F_\mu(x)$. To see this, note

$$F_\mu(x) = \mu([0,x)) = \mu\left(\bigcup_{n=1}^{\infty}[0,x_n)\right) = \lim_{n\to\infty}\mu([0,x_n)) = \lim_{n\to\infty}F_\mu(x_n).$$

(i) We consider only the case $0 < a < b$ and leave the remaining cases for the reader.

$$\mu([a,b)) = \mu([0,b)\setminus[0,a)) = \mu([0,b)) - \mu([0,a)) = F_\mu(b) - F_\mu(a).$$

(ii) is again left for the reader. □

By Proposition 2.7.1 every Borel measure on $\mathbb{R}$ which is finite on bounded Borel sets gives an increasing left-continuous function F_μ such that $\mu([a,b)) = F_\mu(b) - F_\mu(a)$. Do increasing left-continuous functions $f\colon \mathbb{R} \to \mathbb{R}$ induce Borel measures on $\mathbb{R}$ which are finite on bounded Borel sets? This is the essence of the construction of the Lebesgue-Stieltjes measures.

Let $f\colon \mathbb{R} \to \mathbb{R}$ be a (non-strictly) increasing left continuous function. Since f is monotone, the limits

$$f(-\infty) := \lim_{x\to-\infty} f(x) \qquad \text{and} \qquad f(\infty) := \lim_{x\to\infty} f(x)$$

exist in $\overline{\mathbb{R}}$. On a semi-algebra $\mathcal{S}(\mathbb{R})$ we define a function μ_f by

$$\begin{aligned} \mu_f(\emptyset) &= 0 \\ \mu_f([a,b)) &= f(b) - f(a) \qquad (a < b,\ a,b \in \mathbb{R}) \\ \mu_f((-\infty,b)) &= f(b) - f(-\infty) \\ \mu_f([a,\infty)) &= f(\infty) - f(a). \end{aligned}$$

2.7.2 Theorem *The function $\mu = \mu_f$ is a semi-measure on $\mathcal{S}(\mathbb{R})$.*

Proof. By definition we have $\mu(\emptyset) = 0$. Let $[a, b)$ be a finite union of pairwise disjoint intervals from $\mathcal{S}(\mathbb{R})$. This intervals are of the form $[a_j, b_j)$. We enumerate them as

$$a = a_1 < b_1 = a_2 < b_2 = a_2 < \ldots < b_n = b.$$

Then

$$\mu([a,b)) = f(b) - f(a) = \sum_{n=1}^{n}((f(b_j) - f(a_j)) = \sum_{j=1}^{n}\mu([a_j, b_j)).$$

If the interval is infinite, i.e., of the form $[a, \infty)$ or $(-\infty, b)$, we argue similarly as above.

Suppose an interval I from $\mathcal{S}(\mathbb{R})$ is a union of a sequence of pairwise disjoint intervals from $(\mathcal{S}, \mathbb{R})$. Suppose first that $I = [a, b)$ for some $a < b$ and $[a, b) = \bigcup_{j=1}^{\infty}[a_j, b_j)$. Fix $\epsilon > 0$. Since f is left continuous, there exist $c < b$ and $c_j < a_j$ for $j = 1, 2, \ldots$ such that

$$f(c) > f(b) - \frac{\epsilon}{2} \qquad \text{and} \qquad f(c_j) > f(a_j) - \frac{\epsilon}{2 \cdot 2^j}$$

for all $j = 1, 2, \ldots$. Then the compact interval $[a, c]$ is contained in the union of open intervals (c_j, b_j), so that a finite union actually covers $[a, c]$. The point a is contained in one of the (c_j, b_j). By reenumerating assume $a \in (c_1, b_1)$. If $b_1 > c$, then $[a, c] \subseteq (c_1, b_1)$. If $b_1 \leq c$, then after reenumerating $b_1 \in (c_2, b_2)$. If $b_2 > c$, then $[a, c] \subseteq (c_1, b_1) \cup (c_2, b_2)$. Otherwise, again after reenumerating $b_2 \in (c_3, b_3)$. After finitely many steps we find c_j such that $c_1 < a$, $c_j < b_{j-1} < b_j$ for $j = 2, \ldots, n$ and $c < b_n$. Then

$$\begin{aligned}\sum_{j=1}^{\infty}(f(b_j) - f(a_j)) \geq \sum_{j=1}^{n}(f(b_j) - f(a_j)) &\geq \sum_{j=1}^{n}(f(b_j) - f(c_j)) - \frac{\epsilon}{2} \\ &= -f(c_1) + \sum_{j=2}^{n-1}(f(b_j) - f(c_{j+1})) + f(b_n) - \frac{\epsilon}{2} \\ &\geq -f(a) + f(c) - \frac{\epsilon}{2} \geq f(b) - f(a) - \epsilon.\end{aligned}$$

The fourth inequality followed from monotonicity of f since $c_{j+1} < b_j$, $c_1 < a$ and $c < b_n$. Since $\epsilon > 0$ was arbitrary, we have $\sum_{j=1}^{\infty}\mu([a_j, b_j)) \geq f(b) - f(a) = \mu([a, b))$.

Now pick an interval of the form $[a, \infty)$ and assume $[a, \infty) = [a', \infty) \cup \bigcup_{j=1}^{\infty}[a_j, b_j)$. Then the desired inequality follows from the identity

$$\mu([a, \infty)) = \mu([a', \infty)) + \sum_{j=1}^{\infty}\mu([a_j, b_j)).$$

Suppose now $[a,\infty) = \bigcup_{j=1}^{\infty}[a_j, b_j)$ and pick n large enough. Then $[a,n) = \bigcup_{j=1}^{\infty}[a_j, \min\{n, b_j\})$ so that $f(n) - f(a) \le \sum_{j=1}^{\infty} f(b_j) - f(a_j)$. Now let $n \to \infty$ to get $f(\infty) - f(a) \le \sum_{j=1}^{\infty} f(b_j) - f(a_j)$. □

By Theorem 2.7.2 every increasing left-continuous function determines a semi-measure μ_f on the semi-algebra $\mathcal{S}(\mathbb{R})$. By Theorem 2.6.3 this semi-measure can be uniquely extended to the measure (we denote it again by) μ_f on the algebra $\mathcal{A}$ generated by $\mathcal{S}(\mathbb{R})$. By Theorem 2.5.2 the measure μ_f on $\mathcal{A}$ can be uniquely extended to the complete measure (we denote it again by) μ_f to the σ-algebra $\mathcal{A}_{\mu_f}$ by Caratheódory's theorem. This obtained measure μ_f is called the ***Lebesgue-Stieltjes measure*** associated with f. Since $\mathcal{S}(\mathbb{R}) \subseteq \mathcal{A}_{\mu_f}$, we first conclude $\mathcal{B}(\mathbb{R}) \subseteq \mathcal{A}_{\mu_f}$ and since $\mathcal{A}_{\mu_f}$ is complete we conclude $\mathcal{L}(\mathbb{R}) \subseteq \mathcal{A}_{\mu_f}$.

The most important case of a Lebesgue-Stieltjes measure is obtained by taking $f(x) = x$. The Lebesgue-Stieltjes measure associated to f is called the ***Lebesgue measure*** on $\mathbb{R}$ and is denoted by m.

2.7.3 Corollary *For each $A \in \mathcal{L}(\mathbb{R})$ and $x \in \mathbb{R}$ we have $m(x+A) = m(A)$ and $m(xA) = |x|m(A)$.*

Proof. Since the length on $\mathcal{S}(\mathbb{R})$ is invariant under translation, we have $m(x+A) = m(A)$ for all $x \in \mathbb{R}$ and $A \in \mathcal{S}(\mathbb{R})$. Hence, the semi-measure m_x defined as $m_x(A) := m(x+A)$ agrees with the semi-measure m on $\mathcal{S}(\mathbb{R})$. By uniqueness of the extensions from semi-algebra to the algebra, and finally to the σ-algebra by Carathéodory's theorem we have $m(x+A) = m(A)$ for all $x \in \mathbb{R}$ and $A \in \mathcal{L}(\mathbb{R})$.

To prove $m(xA) = |x|m(A)$, note first that for sets in $\mathcal{S}(\mathbb{R})$ the equality holds. Hence, the semi-measure m_x defined as $m_x(A) = m(xA)$ agrees on $\mathcal{S}(\mathbb{R})$ with the semi-measure m'_x defined as $m'_x(A) = |x|m(A)$. Now proceed as in the first part of the proof. □

Fix a Lebesgue-Stieltjes measure μ_f induced by some increasing left-continuous function f on $\mathbb{R}$. Let $\mathcal{A}_{\mu_f}$ be the domain of μ_f and let $\mathcal{A}$ be the algebra generated by $\mathcal{S}(\mathbb{R})$. Then for each $A \in \mathcal{A}_{\mu_f}$ we have

$$\mu_f(A) = \inf\left\{\sum_{n=1}^{\infty} \mu_f(A_n) : \ (A_n) \subseteq \mathcal{A} \text{ is a cover for } A\right\}.$$

The sets A_n are finite disjoint unions of members of $\mathcal{S}(\mathbb{R})$. Since each of the sets of the form $[a,\infty)$ and $(-\infty,b)$ can be written as a countable disjoint union of the sets of the form $[a, b)$, we have

$$\mu_f(A) = \inf\left\{\sum_{n=1}^{\infty} \mu_f([a_n, b_n)) : \ A \subseteq \bigcup_{n=1}^{\infty} [a_n, b_n)\right\}$$
$$= \inf\left\{\sum_{n=1}^{\infty} (f(b_n) - f(a_n)) : \ A \subseteq \bigcup_{n=1}^{\infty} [a_n, b_n)\right\}.$$

From now on when we say μ is a Lebesgue-Stieltjes measure we will implicitly think that $\mu = \mu_f$ for some increasing left-continuous function $f\colon \mathbb{R} \to \mathbb{R}$.

Half-open intervals can be replaced by open ones.

2.7.4 Proposition *Let μ be a Lebesgue-Stieltjes measure. For each set $A \in \mathcal{A}_\mu$ we have*

$$\mu(A) = \inf\left\{\sum_{n=1}^{\infty} (\mu((a_n, b_n)) : \ A \subseteq \bigcup_{n=1}^{\infty} (a_n, b_n)\right\}.$$

Proof. Denote the right hand-side as $\nu(A)$. Suppose $A \subseteq \bigcup_{n=1}^{\infty} (a_n, b_n)$. Each interval (a_n, b_n) is a countable disjoint union of the family of intervals $[a_{n,m}, b_{n,m})$ and so

$$\mu(A) \leq \sum_{n,m=1}^{\infty} \mu([a_{n,m}, b_{n,m})) = \sum_{n=1}^{\infty} \mu((a_n, b_n))$$

and so $\mu(A) \leq \nu(A)$. For the converse, pick $\epsilon > 0$ and find intervals $[a_n, b_n)$ such that $A \subseteq \bigcup_{n=1}^{\infty} [a_n, b_n)$ and $\sum_{n=1}^{\infty} \mu([a_n, b_n)) \leq \mu(A) + \epsilon$. Since f is left-continuous, for each $n \in \mathbb{N}$ there exists $\delta_n > 0$ such that

$$F(a_n) - F(a_n - \delta_n) < \frac{\epsilon}{2^n}.$$

Then $A \subseteq \bigcup_{n=1}^{\infty} (a_n - \delta_n, b_n)$ and

$$\sum_{n=1}^{\infty} \mu((a_n - \delta_n, b_n)) \leq \sum_{n=1}^{\infty} \mu([a_n, b_n)) + \epsilon \leq \mu(A) + 2\epsilon.$$

This proves $\nu(A) \leq \mu(A)$. □

As an application of Proposition 2.7.4 one can quickly show that for any Lebesgue-Stieltjes measure μ and any $A \in \mathcal{L}(\mathbb{R})$ we have

$$\mu(A) = \inf\{\mu(U) : \ A \subseteq U \text{ and } U \subseteq \mathbb{R} \text{ is open}\}. \tag{2.3}$$

As an application of the last equality one can show (with more work) that

$$\mu(A) = \sup\{\mu(K) : \ K \subseteq A \text{ and } K \subseteq \mathbb{R} \text{ is compact}\}. \tag{2.4}$$

Borel measures that enjoy (2.3) are called ***outer regular***. Borel measures that enjoy (2.4) are called ***outer regular***. Borel measures which are simultaneously inner and outer regular are called ***regular measures***. Hence, every Lebesgue-Stieltjes measure is regular.

3 Measurable Functions

3.1 Measurable Maps

Let $(X, \mathcal{A})$ and $(Y, \mathcal{B})$ be measurable spaces. A map $f\colon X \to Y$ is measurable if $f^{-1}(B) \in \mathcal{A}$ for each $B \in \mathcal{B}$. We say that f is $(\mathcal{A}, \mathcal{B})$-measurable.

3.1.1 Example Constant maps are always measurable. Indeed, suppose $f\colon X \to Y$ is a map that maps everything to a point y_0. If $B \in \mathcal{B}$ is any measurable set, then $f^{-1}(B)$ is either X or $\emptyset$ which are both contained in every σ-algebra on X.

3.1.2 Lemma *The composite of measurable maps is measurable.*

Proof. Suppose $f\colon (X, \mathcal{A}) \to (Y, \mathcal{B})$ and $g\colon (Y, \mathcal{B}) \to (Z, \mathcal{C})$ are measurable. Pick any $C \in \mathcal{C}$. Then

$$(g \circ f)^{-1}(C) = f^{-1}(g^{-1}(C))$$

implies $(g \circ f)^{-1}(C) \in \mathcal{A}$. □

3.1.3 Proposition *Let $(X, \mathcal{A})$ and $(Y, \mathcal{B})$ be measurable spaces and $f\colon X \to Y$ a map. Suppose $\mathcal{B} = \sigma(\mathcal{F})$. Then f is measurable iff $f^{-1}(F) \in \mathcal{A}$ for each $F \in \mathcal{F}$,*

Proof. Let $\mathcal{E} = \{B \subseteq Y : f^{-1}(B) \in \mathcal{A}\}$. We claim that $\mathcal{E}$ is a σ-algebra. Obviously $Y \in \mathcal{E}$. Also $f^{-1}(E^c) = f^{-1}(E)^c$ implies $E^c \in \mathcal{E}$ whenever $E \in \mathcal{E}$. If $(E_n) \subseteq \mathcal{E}$, then

$$f^{-1}\Big(\bigcup_{n=1}^{\infty} E_n\Big) = \bigcup_{n=1}^{\infty} f^{-1}(E_n) \in \mathcal{A}$$

implies $\bigcup_{n=1}^{\infty} E_n \in \mathcal{E}$, so that $\mathcal{E}$ is a σ-algebra. Since $\mathcal{F} \subseteq \mathcal{E}$, by definition $\mathcal{B} = \sigma(\mathcal{F}) \subseteq \mathcal{E}$. Hence, for each $B \in \mathcal{B}$ we have $f^{-1}(B) \in \mathcal{A}$.

The converse implication is clear. □

3.1.4 Corollary *Let $(X, \mathcal{A})$ be a measurable space and Y a topological space. A map $f\colon X \to Y$ is $(\mathcal{A}, \mathcal{B}(Y))$ measurable iff $f^{-1}(U) \in \mathcal{A}$ for each open subset $U \subseteq Y$.*

3.1.5 Corollary *Every continuous map $f\colon X \to Y$ between topological spaces is measurable with respect to Borel σ-algebras on X and Y.*

If X and Y are topological spaces, measurable maps with respect to Borel σ-algebras are called ***Borel maps***. The following proposition follows from the fact that the sets within the statements generate $\mathcal{B}(\mathbb{R})$.

3.1.6 Proposition *Let $(X, \mathcal{A})$ be a measurable space. For a given function $f\colon X \to \mathbb{R}$ the following statements are equivalent:*

(i) *f is $(\mathcal{A}, \mathcal{B}(\mathbb{R}))$-measurable.*

(ii) *$f^{-1}((-\infty, a)) \in \mathcal{A}$ for each $a \in \mathbb{R}$ or $f^{-1}((a, \infty)) \in \mathcal{A}$ for each $a \in \mathbb{R}$.*

(iii) *$f^{-1}((-\infty, a]) \in \mathcal{A}$ for each $a \in \mathbb{R}$ or $f^{-1}([a, \infty)) \in \mathcal{A}$ for each $a \in \mathbb{R}$.*

(iv) *$f^{-1}([a, b)) \in \mathcal{A}$ for all $a, b \in \mathbb{R}$ or $f^{-1}((a, b]) \in \mathcal{A}$ for all $a, b \in \mathbb{R}$.*

(v) *$f^{-1}([a, b]) \in \mathcal{A}$ for all $a, b \in \mathbb{R}$*

(vi) *$f^{-1}((a, b)) \in \mathcal{A}$ for all $a, b \in \mathbb{R}$*

Also the composite of two Borel maps is always a Borel map.

The Extended Real Line

The ***extended real line*** $\overline{\mathbb{R}} := [-\infty, \infty] = \mathbb{R} \cup \{-\infty, \infty\}$ is equipped with the topology such that

- on $\mathbb{R}$ it is the usual Eucliedean topology;
- fundamental system of neighborhoods of ∞ are half rays $(a, \infty]$, $(a \in \mathbb{R})$;
- fundamental system of neighborhoods of $-\infty$ are half rays $[-\infty, a)$, $(a \in \mathbb{R})$;

It is easy to see that $x_n \to \infty$ in $\overline{\mathbb{R}}$ iff $\lim\limits_{n\to\infty} x_n = \infty$ in the sense of Analysis 1 course.

By the definition of topology on $\overline{\mathbb{R}}$, we see that the usual Euclidean topology on $\mathbb{R}$ is precisely the relative topology on $\mathbb{R}$ induced by the topology of $\overline{\mathbb{R}}$. What is the connection between $\mathcal{B}(\mathbb{R})$ and $\mathcal{B}(\overline{\mathbb{R}})$? The answer to this question is provided by the following lemma.

3.1.7 Proposition *Let X be a topological space and Y its subspace with the relative topology. Then*

$$\mathcal{B}(Y) = \{A \cap Y : \ A \in \mathcal{B}(X)\}.$$

Proof. It is easy to see that the set

$$\mathcal{B} := \{A \cap Y : \ A \in \mathcal{B}(X)\}$$

is a σ-algebra on Y. Since $\mathcal{B}$ contains all open sets of Y, we conclude $\mathcal{B}(Y) \subseteq \mathcal{B}$.

To prove the converse statement, consider the inclusion $\iota\colon Y \hookrightarrow X$. Then ι is continuous and hence Borel measurable. If $A \in \mathcal{B}(X)$, then $A \cap Y = \iota^{-1}(A) \in \mathcal{B}(Y)$. This proves $\mathcal{B} \subseteq \mathcal{B}(Y)$ and the proof is finished. □

3.1.8 Theorem *Consider the extended real line $\overline{\mathbb{R}}$.*

(i) *Borel sets of $\overline{\mathbb{R}}$ are precisely the ones that are of the form $A \cup B$ where A is Borel in $\mathbb{R}$ and $B \subseteq \{-\infty, \infty\}$.*

(ii) *If $(X, \mathcal{A})$ is a measure space and $f \colon X \to \overline{\mathbb{R}}$ is a function, then f is measurable iff $f^{-1}([-\infty, a)) \in \mathcal{A}$ for each $a \in \mathbb{Q}$ iff $f^{-1}([-\infty, a]) \in \mathcal{A}$ for each $a \in \mathbb{R}$.*

Proof. (i) Denote $\mathcal{B} = \{A \cup B : \ A \in \mathcal{B}(\mathbb{R}) \text{ and } B \subseteq \{-\infty, \infty\}\}$. It is easy to see that $\mathcal{B}$ is σ-algebra on $\overline{\mathbb{R}}$ which contains all open sets of $\overline{\mathbb{R}}$. Hence, $\mathcal{B}(\overline{\mathbb{R}}) \subseteq \mathcal{B}$.

For the converse, take a set $E \in \mathcal{B}$. Since $\mathcal{B}(\mathbb{R}) \subseteq \mathcal{B}(\overline{\mathbb{R}})$ we have $E \cap \mathbb{R} \in \mathcal{B}(\overline{\mathbb{R}})$. Since singletons $\{\infty\}$ and $\{-\infty\}$ are closed in $\overline{\mathbb{R}}$, the set E is Borel in $\overline{\mathbb{R}}$.

(ii) is left for the reader. □

For the purpose of this course we define $0 \cdot \infty = \infty \cdot 0 = 0$.

3.1.9 Corollary *If $f \colon X \to \overline{\mathbb{R}}$ is measurable, then for each $\lambda \in \mathbb{R}$ the function $\lambda \cdot f$ is measurable.*

Proof. If $\lambda = 0$, then $\lambda f = 0$ is measurable. If $\lambda > 0$, then for each $a \in \mathbb{R}$ we have

$$\begin{aligned}(\lambda f)^{-1}([-\infty, a]) &= \{x \in \overline{\mathbb{R}} : \ (\lambda f)(x) \in [-\infty, a]\} \\ &= \{x \in \overline{\mathbb{R}} : \ f(x) \in [-\infty, \tfrac{a}{\lambda}]\} = f^{-1}([-\infty, \tfrac{a}{\lambda}])\end{aligned}$$

which is a measurable set by Theorem 3.1.8. Similarly one can tackle the problem when $\lambda < 0$. □

3.1.10 Example For a subset A of X let χ_A be the characteristic function of A. For $a \in \mathbb{R}$ we have

$$\chi_A^{-1}([-\infty, a]) = \begin{cases} X, & a \geq 1 \\ A^c, & 0 \leq a < 1 \\ \emptyset, & a < 0 \end{cases}$$

Hence, χ_A is measurable iff A is measurable.

3.2 Product σ-algebras

Let $(X_1, \mathcal{A}_1)$ and $(X_2, \mathcal{A}_2)$ be measurable spaces. A subset of the cartesian product $X = X_1 \times X_2$ of the form $A_1 \times A_2$ where $A_i \in \mathcal{A}_i$ is called a ***measurable rectangle***. The ***product σ-algebra*** $\mathcal{A}_1 \otimes \mathcal{A}_2$ on X is the σ-algebra on X generated by all measurable rectangles.

3.2.1 Lemma *Let $(X, \mathcal{A})$ and $(Y, \mathcal{B})$ be measurable spaces.*

(i) *The product σ-algebra $\mathcal{A} \otimes \mathcal{B}$ on $X \times Y$ is the coarsest σ-algebra with the property that the projection maps are measurable.*

(ii) *If $\mathcal{A}$ and B are generated by $\mathcal{E}$ and $\mathcal{F}$, respectively, then $\mathcal{A} \otimes \mathcal{B}$ is generated by*

$$\mathcal{G} := \{A \times Y : \ A \in \mathcal{E}\} \cup \{X \times B : \ B \in \mathcal{F}\}.$$

Proof. (i) Suppose $q_1 : X \times Y \to X$ and $q_2 : X \times Y \to Y$ are measurable with respect to some σ-algebra $\mathcal{C}$ on $X \times Y$. Then for any measurable set $A \in \mathcal{A}$ and $B \in \mathcal{B}$ the set

$$A \cap B = (A \times Y) \cap (X \times B) = q_1^{-1}(A) \cap q_2^{-1}(B)$$

is measurable. Hence, $\mathcal{C}$ contains all measurable rectangles which implies $\mathcal{A} \otimes \mathcal{B} \subseteq \mathcal{C}$.

(ii) Obviously we have $\sigma(\mathcal{G}) \subseteq \mathcal{A} \otimes \mathcal{B}$. On the other hand, the collection $\{A \subseteq X : \ q_1^{-1}(A) \in \sigma(\mathcal{G})\}$ is a σ-algebra which contains $\mathcal{E}$. Hence, $\mathcal{A} \subseteq \{A \subseteq X : \ q_1^{-1}(A) \in \sigma(\mathcal{G})\}$. This means that for any $A \in \mathcal{A}$ we have $A \times Y \in \sigma(\mathcal{G})$. Similarly, $X \times B \in \sigma(\mathcal{G})$ for each $B \in \mathcal{B}$. Finally, $\mathcal{A} \otimes \mathcal{B} \in \sigma(\mathcal{G})$. □

In the following result we consider the relationship between Borel σ-algebra on the product of topological spaces and the product of two Borel σ-algebras.

3.2.2 Proposition *Let X and Y be topological spaces. The space $X \times Y$ is equipped with the product topology. Then*

(i) $\mathcal{B}(X) \otimes \mathcal{B}(Y) \subseteq \mathcal{B}(X \times Y)$.

(ii) *If X and Y are second countable, then $\mathcal{B}(X) \otimes \mathcal{B}(Y) = \mathcal{B}(X \times Y)$.*

(iii) *If X and Y are separable metric spaces, then $\mathcal{B}(X) \otimes \mathcal{B}(Y) = \mathcal{B}(X \times Y)$.*

Proof. (i) Denote by $\mathcal{C}$ the σ-algebra generated by the family

$$\mathcal{F} := \{U \times Y : \ U \in \tau_X\} \cup \{X \times V : \ V \in \tau_Y\}.$$

By Lemma 3.2.1 we have $\mathcal{B}(X) \otimes \mathcal{B}(Y) = \mathcal{C}$. Since sets in $\mathcal{F}$ are open, (i) follows.

(ii) By (i) it suffices to prove $\mathcal{B}(X \times Y) \subseteq \mathcal{B}(X) \otimes \mathcal{B}(Y)$. By definition of Borel σ-algebra it suffices to prove $\mathcal{B}(X) \otimes \mathcal{B}(Y)$ contains all open sets of the product topology. The set of all open rectangles forms a basis for the product topology on $X \times Y$. Since X and Y are second countable, so it is $X \times Y$. Hence, each open set in $X \times Y$ is in $\mathcal{B}(X) \otimes \mathcal{B}(Y)$.

(iii) Separable metric spaces are second countable. □

3.2.3 Example The inclusion in Proposition 3.2.2 can be proper. Consider X to be a discrete topological space with cardinality greater than continuum.

Since $X \times X$ is Hausdorff, the diagonal $\triangle$ of $X \times X$ is closed and hence Borel in $X \times X$. However, a nontrivial proof which we omit shows that $\triangle \notin \mathcal{B}(X) \otimes \mathcal{B}(X)$.

If we have three measurable spaces $(X, \mathcal{A})$, $(Y, \mathcal{B})$ and $(Z, \mathcal{C})$, then one can first consider $(X \times Y, \mathcal{A} \otimes \mathcal{B})$ and then $((X \times Y) \times Z, (\mathcal{A} \otimes \mathcal{B}) \otimes \mathcal{C})$. It can be checked that $(\mathcal{A} \otimes \mathcal{B}) \otimes \mathcal{C}$ is generated by the family

$$\{(A \times B) \times C : \ A \in \mathcal{A}, \ B \in \mathcal{B}, \ C \in \mathcal{C}\}.$$

More generally, we have the following result.

3.2.4 Proposition *Let $(X_i, \mathcal{A}_i)$ be measurable spaces for $i = 1, \ldots, n$. Then the product σ-algebra is an associative operation; all iterated product σ-algebras are equal and we denote it by $\bigotimes_{i=1}^{n} \mathcal{A}_i$. It is generated by the set of all measurable rectangles*

$$\{A_1 \times \cdots \times A_n : \ A_i \in \mathcal{A}_i \text{ for } i = 1, \ldots, n\}.$$

By induction one can prove the following:

3.2.5 Corollary *For $n \in \mathbb{N}$ we have $\mathcal{B}(\mathbb{R}^n) = \bigotimes_{j=1}^{n} \mathcal{B}(\mathbb{R})$.*

Having in mind that $\mathbb{R}^2 \approx \mathbb{C}$ via the map $(x, y) \mapsto x + iy$ we can identify them as topological spaces. The preceding corollary yields the following useful fact.

3.2.6 Corollary *We have $\mathcal{B}(\mathbb{C}) = \mathcal{B}(\mathbb{R} \times \mathbb{R}) = \mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R})$.*

Now we turn our attention to measurable maps into the product space.

3.2.7 Proposition *Let $(Y_i, \mathcal{B}_i)$ $(i = 1, 2)$ be measurable spaces and let $\mathcal{B}_1 \otimes \mathcal{B}_2$ be the product σ-algebra on $Y_1 \times Y_2$. If $(X, \mathcal{A})$ is a measurable space and $f \colon X \to Y$ any map, then f is measurable iff the components $q_1 \circ f$ and $q_2 \circ f$ are measurable.*

Proof. If f is measurable, then also $q_1 \circ f$ and $q_2 \circ f$ are measurable since they are composites of measurable maps.

For the converse, assume $q_1 \circ f$ and $q_2 \circ f$ are measurable. Denote $\mathcal{C} = \{B \subseteq Y_1 \times Y_2 : \ f^{-1}(B) \subseteq \mathcal{A}\}$. Since $\mathcal{C}$ is a σ-algebra by Proposition 3.1.3, it suffices to prove that $\mathcal{C}$ contains all measurable rectangles and therefore the product σ-algebra. Suppose $B = B_1 \times B_2$ where $B_1 \in \mathcal{B}_1$ and $B_2 \in \mathcal{B}_2$. Then

$$\begin{aligned} f^{-1}(B) &= \{x \in A : \ f(x) \in B\} = \{x \in A : \ f_1(x) \in B_1 \text{ and } f_2(x) \in B_2\} \\ &= f_1^{-1}(B_1) \cap f_2^{-1}(B_2) \in \mathcal{A}. \end{aligned}$$

□

3.2.8 Corollary *Let $(X, \mathcal{A})$ be a measurable space and $Y \in \{\mathbb{R}, \mathbb{C}, [0, \infty]\}$. If $f, g\colon X \to Y$ are measurable, so are $f + g$ and $f \cdot g$.*

Proof. Define the map $F\colon X \to Y \times Y$ by $F(x) := (f(x), g(x))$. Since f, g are measurable, Proposition 3.2.7 implies F is measurable. Since the sum $s\colon (a, b) \mapsto a + b$ and the product $p\colon (a, b) \to a \cdot b$ are continuous and so Borel measurable, $f+g = s \circ F$ and $f \cdot g = p \circ F$ are measurable by Lemma 3.1.2. □

3.2.9 Corollary *Linear combinations of measurable functions with values in either $\mathbb{C}$ or $\mathbb{R}$ are again measurable.*

3.2.10 Proposition *Suppose Y_1 and Y_2 are second countable topological spaces and $(X, \mathcal{A})$ a measurable space. The map $f\colon X \to Y_1 \times Y_2$ is Borel measurable iff $f_i = p_i \circ f$ are Borel measurable.*

Proof. If f is Borel measurable, then $p_i \circ f$ as a composite of Borel measurable maps is again Borel measurable.

For the converse, assume f_1 and f_2 are Borel measurable. Then the map f is $(\mathcal{A}, \mathcal{B}(Y_1) \otimes \mathcal{B}(Y_2))$-measurable by Proposition 3.2.7. Now apply Proposition 3.2.2 to get $\mathcal{B}(Y_1 \times Y_2) = \mathcal{B}(Y_1) \otimes \mathcal{B}(Y_2)$. □

3.2.11 Corollary *Let $f\colon X \to \mathbb{C}$ be any map. If $\mathcal{A}$ is a σ-algebra on X, then f is measurable iff $\operatorname{Re} f\colon X \to \mathbb{R}$ and $\operatorname{Im} f\colon X \to \mathbb{R}$ are measurable.*

3.3 Sequence of Measurable Functions

Given a sequence $(a_n) \subseteq [-\infty, \infty]$ we can form a new sequence $\widetilde{a}_n := \sup_{k \geq n} a_k$. Then $(\widetilde{a}_n)$ is decreasing but not strictly decreasing. Hence, it has a limit in $[-\infty, \infty]$. This limit is called the ***upper limit*** or ***limes superior***. We denote it by

$$\limsup_n a_n := \inf_n \sup_{k \geq n} a_k = \lim_n \sup_{k \geq n} a_k.$$

Similarly we can define the ***lower limit*** or ***limes inferior*** as

$$\liminf_n a_n := \sup_n \inf_{k \geq n} a_k = \lim_n \inf_{k \geq n} a_k.$$

A sequence of functions $f_n\colon X \to [-\infty, \infty]$ converges ***pointwise*** to a function f if $\lim_{n \to \infty} f_n(x) = f(x)$ for each $x \in X$. More generally, for a sequence $f_n\colon X \to [-\infty, \infty]$ we can define functions

$$\begin{aligned}
(\sup_n f_n)(x) &:= \sup_n f_n(x) \\
(\inf_n f_n)(x) &:= \inf_n f_n(x) \\
(\limsup_n f_n)(x) &:= \inf_n \sup_{k \geq n} f_n(x) \\
(\liminf_n f_n)(x) &:= \sup_n \inf_{k \geq n} f_n(x).
\end{aligned}$$

The supremum of an uncountable family of measurable functions can be nonmeasurable.

3.3.1 Example Let X be any uncountable set and let $\mathcal{A}$ be the σ-algebra of all sets which are countable or their complement is countable. Let E be any set which is neither countable nor its complement is countable. Then χ_E is non-measurable. Note that

$$\chi_E = \sup\{\chi_{\{x\}} : \ x \in E\}$$

is a supremum of a family of measurable functions.

3.3.2 Lemma *For any sequence of measurable functions $f_n \colon X \to \overline{\mathbb{R}}$ the functions $\sup\limits_n f_n$ and $\inf\limits_n f_n$ are also measurable.*

Proof. Denote $g = \sup_n f_n$. Since for each $a \in [-\infty, \infty]$ we have

$$g^{-1}([-\infty, a]) = \bigcap_{n=1}^{\infty} f_n^{-1}[-\infty, a]),$$

the function g is measurable by Theorem 3.1.8. Since $\inf_n f_n = -\sup_n(-f_n)$, the first part of the proof yields that $\inf_n f_n$ is measurable. □

3.3.3 Corollary *Let $f_n \colon X \to \overline{\mathbb{R}}$ be a sequence of measurable functions.*

(i) *Then $\limsup\limits_n f_n$ and $\liminf\limits_n f_n$ are measurable.*

(ii) *If $f_n \to f$ pointwise, then f is measurable.*

Proof. (i) By Lemma 3.3.2 the function $g_n = \sup_{k \geq n} f_n$ is measurable. Again by Lemma 3.3.2 the function $\inf_n g_n$ is measurable. Hence, $\limsup_n f_n$ is measurable. Similar proof works for $\liminf_n f_n$.

(ii) If $f_n \to f$ pointwise, then $f = \limsup_n f_n = \liminf_n f_n$ is measurable by (i). □

3.4 Approximation by Step Functions

A function $f \colon X \to \mathbb{C}$ with a finite range is called a ***step function***. Let $\{a_1, \ldots, a_n\}$ be the range of the step function $f \colon X \to \mathbb{C}$. For each j let

$$A_j = \{x \in X : \ f(x) = a_j\}.$$

Then f can be written as a linear combination of characteristic functions of sets A_j:

$$f = \sum_{j=1}^{n} a_j \chi_{A_j}.$$

The space of all complex bounded measurable functions on a measurable space X is denoted by $B(X)$. Similarly, the space of all complex step functions on a measurable space X is denoted by $S(X)$. Obviously $S(X) \subseteq B(X)$. The space $B(X)$ has the natural norm, i.e., the supremum norm. If (f_n) is a Cauchy sequence in $B(X)$, then f_n converges pointwise to some function f. Since functions (f_n) are measurable, we conclude f is measurable and hence $B(X)$ is a Banach space.

3.4.1 Theorem *Let $f\colon X \to [0,\infty]$ be a measurable function. Then there exists an increasing sequence (s_n) of measurable step functions such that $s_n \to f$ pointwise. The convergence is uniform on each set where f is bounded.*

Proof. For $n \in \mathbb{N}$ and $k = 1, 2, \ldots, n \cdot 2^n$ define measurable sets

$$E_{n,k} := \left\{x \in X : \tfrac{k-1}{2^n} \leq f(x) < \tfrac{k}{2^n}\right\} = f^{-1}\left(\left[\tfrac{k-1}{2^n}, \tfrac{k}{2^n}\right)\right)$$

and

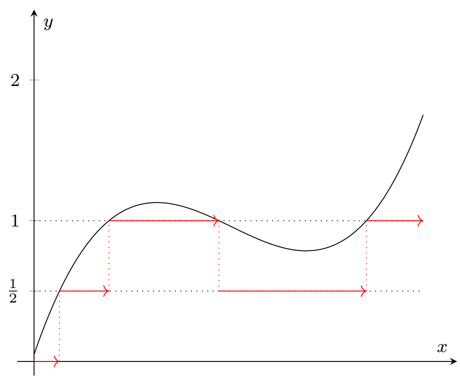
$$F_n = \{x \in X : n \leq f(x)\} = f^{-1}([n,\infty]).$$

For a fixed n we have that the sets $F_n, E_{n,1}, \ldots, E_{n,n2^n}$ are pairwise disjoint with union X.. Now define

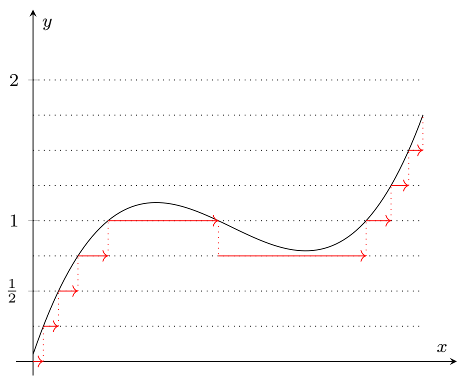
$$s_n = \sum_{k=1}^{n2^n} \frac{k-1}{2^n}\chi_{E_{n,k}} + n\chi_{F_n}.$$

It can be seen that $s_n \leq s_{n+1}$ and $0 \leq f - s_n \leq 2^{-n}$ for each n on the set $f^{-1}([0,n])$. If $f(x) < \infty$, then $s_n(x) \to f(x)$. If $f(x) = \infty$, then $s_n(x) = n$ and so $s_n(x) \to f(x)$.

If f is bounded on some set A, then there exists $n_0 \in \mathbb{N}$ such that $F_m \cap A = \emptyset$ for $m > n_0$. This means that $0 \leq f - s_m \leq 2^{-m}$ on A for all $m > n_0$, so that $s_n \to f$ uniformly on A. □



Approximation with s_1



Approximation with s_2

3.4.2 Corollary *Let $f\colon X \to \mathbb{C}$ be a measurable function. Then there exists a sequence (s_n) of measurable step functions such that $0 \leq |s_1| \leq |s_2| \leq \cdots \leq |f|$ and $s_n \to f$ pointwise. The convergence is uniform on each set where f is bounded.*

Proof. Write $f = u + iv$ and $u = u^+ - u^-$ and $v = v^+ - v^-$. By Theorem 3.4.1 there exist increasing sequences of measurable step functions (p_n^+), (p_n^-), (q_n^+) and (q_n^-) that converge pointwise to u^+, u^-, v^+ and v^-, respectively. Then $s_n = p_n^+ - p_n^- + i(q_n^+ - q_n^-)$ is measurable step function and the sequence (s_n) has the desired proeprties. □

3.5 Modes of Convergences

A sequence of measurable functions (f_n) on a space with measure $(X, \mathcal{A}, \mu)$ converges to a function f ***almost uniformly*** if for each $\epsilon > 0$ there exists a set $A \in \mathcal{A}$ with $\mu(A^c) < \epsilon$ and $f_n|_A \to f|_A$ uniformly.

We say that some property $\mathcal{P}$ holds ***almost everywhere*** if the set

$$\{x \in X : \ x \text{ does not have the property } \mathcal{P}\}$$

has zero measure. Therefore, a sequence of functions (f_n) is said to converge almost everywhere to f if

$$\{x \in X : \ f_n(x) \not\to f(x)\}$$

has measure zero.

3.5.1 Proposition *If a sequence (f_n) of measurable functions is almost uniform convergent to f, then it converges to f almost everywhere.*

Proof. Pick $m \in \mathbb{N}$ and find $A_m \in \mathcal{A}$ such that $\mu(A_m^c) < \frac{1}{m}$ and $f_n \to f$ uniformly on A_m. Then $f_n \to f$ also pointwise on A_m. Hence, $f_n \to f$ pointwise on $A := \bigcup_m A_m$. To finish the proof note that $\mu(A^c) = 0$. □

If the measure is finite, then almost everywhere convergence implies almost uniform convergence.

3.5.2 Jegorov's Theorem *Suppose $(X, \mathcal{A}, \mu)$ is a space with finite measure. If a sequence of measurable functions $f_n \colon X \to \mathbb{C}$ converges to f almost everywhere, then it converges to f also almost uniformly.*

Proof. Write $X = X' \cup N$ where $\mu(N) = 0$ and $f_n \to f$ on X'. Since f_n is measurable on X it is measurable also on X'. We will prove that $f_n \to f$ almost uniformly on X' and since $\mu(N) = 0$, it will follow that $f_n \to f$ almost uniformly on X. Hence, WLOG assume $N = \emptyset$ and so $f_n \to f$ pointwise. Also, by replacing f_n by $f_n - f$ we can assume that $f = 0$.

Define the sets

$$A_{k,m} := \{x \in X : \ |f_n(x)| \le \tfrac{1}{m} \text{ for all } n \ge k\}.$$

Since for each $x \in X$ we have $f_n(x) \to 0$, for a fixed m we have

$$\bigcup_{k=1}^{\infty} A_{k,m} = X.$$

Since for a fixed m the sequence $(A_{k,m})$ is increasing we have $\lim_{k\to\infty} \mu(A_{k,m}) = \mu(X)$. Fix $\epsilon > 0$ and find k_m such that

$$\mu(A_{k_m,m}) \geq \mu(X) - \frac{\epsilon}{2^m}.$$

Define $A := \bigcap_{m=1}^{\infty} A_{k_m,m}$ and note that

$$\mu(A^c) = \mu\left(\bigcup_{m=1}^{\infty} A^c_{k_m,m}\right) \leq \sum_{m=1}^{\infty} \mu(A^c_{k_m,m}) \leq \sum_{m=1}^{\infty} \frac{\epsilon}{2^m} = \epsilon.$$

We claim that $f_n \to f$ converges uniformly on A. Indeed, if $x \in A$, then $x \in A_{k_m,m}$ for some m, so that $|f_n(x)| \leq \frac{1}{m}$ for all $n \geq k_m$. So for a given $\epsilon > 0$ pick m large enough so that $\frac{1}{m} < \epsilon$ and hence $|f_n(x)| < \epsilon$ for all $n \geq k_m$ and $x \in A$. □

A sequence of measurable functions $f_n \colon X \to \mathbb{C}$ on a space with measure $(X, \mathcal{A}, \mu)$ ***converges in measure*** to a function f if for each $\epsilon > 0$ we have

$$\lim_{n\to\infty} \mu(\{x \in X : \ |f_n(x) - f(x)| \geq \epsilon\}) = 0.$$

3.5.3 Proposition *If a sequence (f_n) of measurable functions converges to f almost uniformly, then it also converges in measure.*

Proof. Again we assume $f = 0$. Pick $\epsilon, \delta > 0$ and denote

$$F_n = \{x \in X : \ |f_n(x)| \geq \epsilon\}.$$

Since $f_n \to 0$ almost uniformly there is $A \in \mathcal{A}$ with $\mu(A^c) < \delta$ and $f_n|_A \to 0$ uniformly. There exists n_0 such that for all $n \geq n_0$ and $x \in A$ we have $|f_n(x)| < \epsilon$. This yields $F_n \subseteq A^c$ and so $\mu(F_n) < \delta$ for all $n \geq n_0$. By definition of the limit we have $\mu(F_n) \to 0$. □

3.5.4 Corollary *Let $(X, \mathcal{A}, \mu)$ be a space with finite measure. If $f_n \to f$ almost everywhere, then $f_n \to f$ in measure.*

Proof. By Jegorov's Theorem 3.5.2 $f_n \to f$ almost uniformly so that by Proposition 3.5.3 $f_n \to f$ in measure. □

In Example 3.5.5(i) we see that Jegorov's theorem fails when $\mu(X) = \infty$.

3.5.5 Example

(i) The sequence $\chi_{(n,n+1)}$ on the real line $\mathbb{R}$ equipped with the Lebesgue measure converges to 0 everywhere yet not almost uniformly nor in measure.

(ii) Let us equip the interval $[0,1]$ with the Lebesgue measure. Define the sequence of step functions $f_n\colon [0,1] \to \mathbb{R}$ as follows: $f_1 \equiv 1$, $f_2 = \chi_{[0,\frac{1}{2}]}$, $f_3 = \chi_{[\frac{1}{2},1]}$, $f_4 = \chi_{[0,\frac{1}{4}]}$, $f_5 = \chi_{[\frac{1}{4},\frac{1}{2}]}$, ... In general, each function is of the form $\chi_{[\frac{k}{2^m},\frac{k+1}{2^m}]}$ for some $m \in \mathbb{N}$ and $k = 0, 1, \ldots, 2^m - 1$. Then $f_n \to 0$ in measure yet in each point $x \in [0,1]$ infinitely many functions f_n have attain the value zero and one. Hence $(f_n(x))$ does not converge for any $x \in [0,1]$.

4 Integral

4.1 Integration of Step Functions

Let $(X, \mathcal{A}, \mu)$ be a space with a positive measure μ. For a measurable step function $s = \sum_{j=1}^{n} c_j \chi_{E_j}$ we say that it is written ***canonically*** if $c_j \in [0, \infty)$ are pairwise different and $E_j \in \mathcal{A}$ are pairwise disjoint with union X. We define the ***integral*** of the function s on X as

$$\int_X s d\mu = \sum_{j=1}^{n} c_j \mu(E_j). \tag{4.1}$$

Here we adopt the convention $0 \cdot \infty = 0$. If the space X is clear from the context, we write $\int s d\mu$ instead of $\int_X s d\mu$. In the case $A \in \mathcal{A}$ we have $\int_X \chi_A d\mu = \mu(A)$. It is easy to see that (4.1) holds even for step functions which are not represented canonically, i.e., some of c_i are the same. If s is a nonnegative measurable step function and $A \in \mathcal{A}$, then the function $s\chi_A$ is also nonnegative, step and measurable, so we can define

$$\int_A s d\mu := \int_X s \chi_A d\mu.$$

4.1.1 Lemma *If $s \colon X \to [0, \infty)$ is a nonnegative measurable step function on a space with measure $(X, \mathcal{A}, \mu)$, then*

$$\nu(A) = \int_A s d\mu$$

defines a positive measure on $\mathcal{A}$.

Proof. Since $\chi_\emptyset = 0$, we have $\nu(\emptyset) = \int_\emptyset s d\mu = \int_X 0 d\mu = 0$. Choose now a sequence (A_n) of pairwise disjoint sets in $\mathcal{A}$ and denote their union by A. Let $s = \sum_{j=1}^{n} c_j \chi_{E_j}$ be the canonical representation of the step function s. Then $s\chi_A = \sum_{j=1}^{n} c_j \chi_{A \cap E_j}$ for any subset $A \subseteq X$. Hence

$$\nu(A) = \int_A s d\mu = \int_X s \chi_A d\mu = \sum_{j=1}^{n} c_j \mu(A \cap E_j).$$

Since μ is σ-additive, we have $\mu(A \cap E_j) = \sum_{k=1}^{\infty} \mu(A_k \cap E_j)$ from where it follows

$$\nu(A) = \sum_{j=1}^{n} c_j \mu(A \cap E_j) = \sum_{j=1}^{n} \sum_{k=1}^{\infty} c_j \mu(A_k \cap E_j) = \sum_{k=1}^{\infty} \sum_{j=1}^{n} c_j \mu(A_k \cap E_j)$$
$$= \sum_{k=1}^{\infty} \int_{A_k} s d\mu = \sum_{k=1}^{\infty} \nu(A_k).$$

□

4.1.2 Lemma *Let $s, t \colon X \to [0, \infty)$ be measurable step functions and $c \in [0, \infty)$. Then*

(i) $\int (s + t) d\mu = \int s d\mu + \int t d\mu$,

(ii) $\int cs d\mu = c \int s d\mu$.

Proof. (i) Let $s = \sum_{i=1}^{m} c_i \chi_{E_i}$ and $t = \sum_{j=1}^{n} d_j \chi_{F_j}$ be canonical representations for s and t. Then

$$s + t = \sum_{i=1}^{m} \sum_{j=1}^{n} (c_i + d_j) \chi_{E_i \cap F_j}$$

is a representation for $s + t$ which in general is not canonical. However, we still have

$$\int (s + t) d\mu = \sum_{i=1}^{m} \sum_{j=1}^{n} (c_i + d_j) \mu(E_i \cap F_j).$$

Since $\bigcup_{j=1}^{n} (E_i \cap F_j) = E_i$ and $\bigcup_{i=1}^{m} (E_i \cap F_j) = F_j$, we have

$$\int (s + t) d\mu = \sum_{i=1}^{m} \sum_{j=1}^{n} c_i \mu(E_i \cap F_j) + \sum_{i=1}^{m} \sum_{j=1}^{n} d_j \mu(E_i \cap F_j)$$
$$= \sum_{i=1}^{m} c_i \mu(E_i) + \sum_{j=1}^{n} d_j \mu(F_j) = \int s d\mu + \int t d\mu.$$

(ii) It is obvious and left for the reader. □

Integral is monotone:

4.1.3 Lemma *Let $0 \leq s \leq t < \infty$ be measurable step functions. Then*

$$\int s d\mu \leq \int t d\mu.$$

Proof. Write $t = (t - s) + s$ and note that $t - s$ is also a nonnegative measurable step function. Since the integral is additive on step functions, we have

$$\int t d\mu = \int (t - s) d\mu + \int s d\mu \geq \int s d\mu.$$

□

4.2 Integral of a Nonnegative Measurable Function

As usual, let $(X, \mathcal{A}, \mu)$ be a space with positive measure. For any measurable function $f\colon X \to [0, \infty]$ we denote by $\mathcal{S}_f$ the set of all measurable step functions $s\colon X \to [0, \infty)$ with $0 \le s \le f$. For a nonnegative measurable function $f\colon X \to [0, \infty]$ we define

$$\int_X f d\mu := \sup_{s \in \mathcal{S}_f} \int_X s d\mu.$$

Obviously for every $s \in \mathcal{S}_f$ we have $\int_X s d\mu \le \int_X f d\mu$. Also, since the integral is monotone on nonnegative step functions, the above definition agrees with the definition of the integral of a step function in this special case.

4.2.1 Example

(i) Let X be a nonempty set together with the power σ-algebra. Pick $x \in X$ and consider the Dirac measure δ_x. If f is any nonnegative function, then $\int_X f d\delta_x = f(x)$.

Indeed, if $0 \le s \le f$ is any step function, then we have $\int s d\delta_x = s(x) \le f(x)$. Hence $\int f d\delta_x \le f(x)$. On the other hand, consider $s = f(x)\chi_{\{x\}}$. Then $\int_X f d\delta_x \ge \int_X s d\delta_x = f(x)$.

(ii) Consider the power σ-algebra on $\mathbb{N}$. Let μ be the counting measure on $(\mathbb{N}, \mathbb{P}(\mathbb{N}))$. Let $f\colon \mathbb{N} \to [0, \infty]$ be any function. It is not hard to see that $\int_{\mathbb{N}} f d\mu = \sum_{n=1}^{\infty} f(n)$. The details are left for the reader.

Integral is monotone.

4.2.2 Lemma *Let $f, g\colon X \to [0, \infty]$ be measurable functions with $f \le g$. Then*

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

Proof. This follows from the definition of the integral and the fact that $\mathcal{S}_f \subseteq \mathcal{S}_g$. □

4.2.3 Lebesgue's Monotone Convergence Theorem *Let $f_n\colon X \to [0, \infty]$ be an increasing sequence of measurable functions. Then*

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

Proof. The sequence of integrals $\int f_n d\mu$ is increasing; hence it converges to some $a \in [0, \infty]$. Due to the monotonocity of the integral we have

$$a \le \int f d\mu.$$

For the proof of the converse inequality it suffices to prove that for each $c \in (0,1)$ and $s \in \mathcal{S}_f$ we have

$$\int cs d\mu = c \int s d\mu \leq a$$

since we can let $c \to 1$ and take the supremum of over all $s \in \mathcal{S}_f$ to finally get $\int f d\mu \leq a$. For $n \in \mathbb{N}$ define

$$A_n := \{x \in X :\ cs(x) \leq f_n(x)\}.$$

Then A_n is measurable and $A_n \subseteq A_{n+1}$ for $n \in \mathbb{N}$. The sequence (A_n) satisfies $\bigcup_{n=1}^{\infty} A_n = X$.

Define a measure ν by

$$\nu(B) = \int_B (cs) d\mu.$$

Since the integral is monotone, we have

$$\int (cs) d\mu = \nu(X) = \lim_{n\to\infty} \nu(A_n) = \lim_{n\to\infty} \int_{A_n} (cs) d\mu \leq \lim_{n\to\infty} \int_X f_n d\mu \leq a.$$

□

4.2.4 Theorem *For any sequence $f_j \colon X \to [0,\infty]$ of measurable functions we have*

$$\int_X \sum_{j=1}^{\infty} f_j d\mu = \sum_{j=1}^{\infty} \int_X f_j d\mu.$$

Proof. We first show

$$\int (f_1 + f_2) d\mu = \int f_1 d\mu + \int f_2 d\mu$$

for measurable functions $f_1, f_2 \colon X \to [0,\infty]$. By Theorem 3.4.1 there exist increasing sequences (s_n) and (t_n) of measurable step functions such that $s_n \uparrow f$ and $t_n \uparrow g$. Since $s_n + t_n$ is a also a measurable function and $s_n + t_n \to f + g$, by the Lebesgue Monotone Convergence Theorem 4.2.3 we have

$$\begin{aligned} \int (f+g) d\mu &= \lim_{n\to\infty} \int (s_n + t_n) d\mu = \lim_{n\to\infty} \int s_n d\mu + \lim_{n\to\infty} \int t_n d\mu \\ &= \int f d\mu + \int g d\mu. \end{aligned}$$

For each $n \in \mathbb{N}$ define $g_n = f_1 + \cdots + f_n$. Then (g_n) is an increasing sequence of nonnegative measurable functions with $g_n \to \sum_{j=1}^{\infty} f_j$. Again, by the Lebesgue Monotone Convergence Theorem 4.2.3 we have

$$\sum_{j=1}^{\infty} \int f_j d\mu = \lim_{n\to\infty} \sum_{j=1}^{n} \int f_j d\mu = \lim_{n\to\infty} \int g_n d\mu = \int \lim_{n\to\infty} g_n d\mu = \int \sum_{j=1}^{\infty} f_j d\mu.$$

□

4.2.5 Proposition *Suppose $f\colon X \to [0,\infty]$ is a measurable function and define*

$$\nu(A) = \int_A f d\mu := \int f\chi_A d\mu.$$

(i) *Then ν is a positive measure on $(X, \mathcal{A})$.*

(ii) *If $\mu(A) = 0$, then $\nu(A) = 0$.*

(iii) *If $g\colon X \to [0,\infty]$ is measurable, then*

$$\int g d\nu = \int g f d\mu. \tag{4.2}$$

Proof. (i) It is obvious that $\nu(\emptyset) = 0$. For a sequence (A_n) of pairwise disjoint sets in $\mathcal{A}$ we denote by A the union of all A_n. Then

$$\nu(A) = \int_A f d\mu = \int f\chi_A d\mu = \int \sum_{n=1}^{\infty} f\chi_{A_n} d\mu = \sum_{n=1}^{\infty} \int f\chi_{A_n} d\mu = \sum_{n=1}^{\infty} \nu(A_n).$$

(ii) Take $0 \le s \le f\chi_A$ and write $s = \lambda_1\chi_{A_1} + \cdots + \lambda_n\chi_{A_n}$. Since $s = 0$ on A^c, we have $s = s\chi_A$, so that $\int s d\mu = \sum_{k=1}^{n} \lambda_k \mu(A \cap A_k) = 0$.

(iii) If $g = \chi_A$ with A measurable, then $\int g d\nu = \int \chi_A d\nu = \nu(A) = \int f\chi_A d\mu$. Additivity of the integral yields that $\int g d\nu = \int g f d\mu$ for any nonnegative measurable step function. The general case follows from the Lebesgue Monotone Convergence Theorem 4.2.3. □

The formula (4.2) is usually written $d\nu = f d\mu$ for short. We proceed with yet again corollary of the Lebsegue Monotone Convergence Theorem 4.2.3.

4.2.6 Corollary *Let $f\colon X \to [0,\infty]$ be a measurable function and $c \ge 0$. Then*

$$\int c f d\mu = c \int f d\mu.$$

Proof. Let (s_n) be an increasing sequence of nonnegative measurable step functions with $s_n \to f$. Then

$$\int c f d\mu = \lim_{n\to\infty} \int c s_n d\mu = c \int \lim_{n\to\infty} s_n d\mu = c \int f d\mu. \qquad \square$$

4.2.7 Fatou Lemma *For a sequence $f_n\colon X \to [0,\infty]$ of measurable functions we have*

$$\int \liminf_n f_n d\mu \le \liminf_n \int f_n d\mu.$$

Proof. Define $g_n = \inf_{k\geq n} f_k$. Then $g_n\colon X\to[0,\infty]$ is measurable and $g_n \leq g_{n+1}$. By Lebsegue Monotone Convergence Theorem 4.2.3 we have

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

Since $\liminf_n f_n = \lim_{n\to\infty} g_n$ and $\int g_n d\mu \leq \int f_k d\mu$ for all $k\geq n$, the desired inequality follows. □

4.2.8 Proposition *Let $f\colon X\to[0,\infty]$ be a measurable function.*

(i) *If $\int f d\mu < \infty$ then $\mu(\{x\in X:\ f(x)=\infty\})=0$.*

(ii) *$\int_X f d\mu = 0$ iff $f=0$ almost everywhere.*

Proof. (i) Denote $A := \{x\in X:\ f(x)=\infty\}$. Then A is measurable and $0\leq \infty\cdot\chi_A\leq f$. From the assumption it follows $\infty\cdot\mu(A)<\infty$. This only happens if $\mu(A)=0$.

(ii) Suppose first $f=0$ on A and $\mu(A^c)=0$. Then

$$\int_X f d\mu = \int_A f d\mu + \int_{A^c} f d\mu = \int_{A^c} f d\mu = 0.$$

To prove the converse, for each $n\in\mathbb{N}$ define

$$A_n = \{x\in X:\ f(x)\geq \tfrac{1}{n}\}.$$

Then

$$\int_X f d\mu \geq \int_{A_n} f d\mu \geq \tfrac{1}{n}\mu(A_n)$$

implies $\mu(A_n)=0$. Since

$$\{x\in X:\ f(x)\neq 0\} = \bigcup_{n=1}^{\infty} A_n$$

we conclude $f=0$ almost everywhere. □

4.3 Integral of a Complex Measurable Function

Throughout this section $(X,\mathcal{A},\mu)$ is a space with measure. A measurable function $f\colon X\to\mathbb{C}$ is ***integrable*** whenever

$$\|f\|_1 := \int |f| d\mu < \infty.$$

The set of all integrable functions on $(X,\mathcal{A},\mu)$ is denoted by $L^1(X,\mathcal{A},\mu)$. Usually we will write $L^1(\mu)$ or even L^1 when $\mathcal{A}$ and μ are clear from the context. In the case of the Lebesgue measure m on $\mathbb{R}$ the functions in $L^1(m)$ are called ***Lebesgue integrable***.

Let $f\colon X \to \mathbb{C}$ be any function. We can write $f = \operatorname{Re} f + i \operatorname{Im} f$. We already know f is measurable iff $\operatorname{Re} f$ and $\operatorname{Im} f$ are measurable. Since $|\operatorname{Re} f|, |\operatorname{Im} f| \leq |f| \leq |\operatorname{Re} f| + |\operatorname{Im} f|$ we conclude $f \in L^1(\mu)$ iff $\operatorname{Re} f, \operatorname{Im} f \in L^1(\mu)$. If $f\colon X \to \mathbb{R}$ is a real function, then

$$f^+ = \frac{1}{2}(|f| + f) \qquad \text{and} \qquad f^- = \frac{1}{2}(|f| - f)$$

are called the ***negative*** and the ***positive part*** of f. It is obvious that f is measurable iff f^+ and f^- are measurable. It is easy to see that $f \in L^1(\mu)$ iff $f^+, f^- \in L^1(\mu)$.

If $f\colon X \to [-\infty, \infty]$ is a real function such that at least one of the integrals $\int_X f^+ d\mu$ or $\int_X f^- d\mu$ is finite, then we can define

$$\int_X f d\mu := \int_X f^+ d\mu - \int_X f^- d\mu.$$

If $f \in L^1(\mu)$ is complex, then we define

$$\int_X f d\mu := \int_X \operatorname{Re} f d\mu + i \int_X \operatorname{Im} f d\mu$$

if both integral exist.

4.3.1 Lemma *If we equip $L^1(\mu)$ with pointwise operations, then $L^1(\mu)$ becomes a vector space and $\|\cdot\|_1$ is a seminorm on $L^1(\mu)$.*

Proof. Pick $\alpha, \beta \in \mathbb{C}$ and $f, g \in L^1(\mu)$. Since $|\alpha f + \beta g| \leq |\alpha|\ |f| + |\beta|\ |g|$, monotonicity and additivity of the integral of a nonnegative function yields

$$\int_X |\alpha f + \beta g| d\mu \leq |\alpha| \int_X |f| d\mu + |\beta| \int_X |g| d\mu < \infty.$$

The reader can prove it by himself that $\|\cdot\|_1$ is a seminorm on $L^1(\mu)$. □

If $\|f\|_1 = 0$, then $f = 0$ almost everwhere, so that in general $\|\cdot\|_1$ is not a norm on $L^1(\mu)$. On $L^1(\mu)$ we introduce the relation $\sim$ by $f \sim g$ iff $f = g$ almost everywhere. For $[f] \in L^1(\mu)/_\sim$ we define $\|[f]\|_1 := \|f\|_1$. If $f \sim g$, then $\|f\|_1 = \|g\|_1$ and $\int f d\mu = \int g d\mu$. This means that $\|\cdot\|_1$ is well-defined on $L^1(\mu)/_\sim$.

4.3.2 Corollary *The space $(L^1(\mu)/_\sim, \|\cdot\|_1)$ is a normed space.*

From now on the quotient space $L^1(\mu)/_\sim$ is denoted by $L^1(\mu)$ and we will write $\|f\|_1$ instead of $\|[f]\|_1$. Whenever we will work with elements in $L^1(\mu)$ we will work with representatives of the respectful equivalence classes.

We proceed with the properties of the integral of L^1-functions.

4.3.3 Proposition *The following assertions hold for L^1-functions.*

(i) *Integral is a linear functional on $L^1(\mu)$:*

$$\int (\alpha f + \beta g) d\mu = \alpha \int f d\mu + \beta \int g d\mu$$

for $\alpha, \beta \in \mathbb{C}$ and $f, g \in L^1(\mu)$.

(ii) *For real L^1-functions $f, g\colon X \to \mathbb{R}$ with $f \leq g$ we have*

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

(iii) *For $f \in L^1(\mu)$ we have*

$$\left| \int f d\mu \right| \leq \int |f| d\mu.$$

Proof. (i) We first prove that for real functions f, g we have $\int (f+g) d\mu = \int f d\mu + \int g d\mu$. To prove this, we first write $h = f + g$. Then $h^+ + f^- + g^- = h^- + f^+ + g^+$ and since the integral is additive on nonnegative measurable functions we conclude

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

By the definition of the integral of a real function after re-arranging the terms we get

$$\int h d\mu = \int f d\mu + \int g d\mu.$$

Now pick $\alpha \in \mathbb{R}$ and a real function $f \in L^1(\mu)$. If $\alpha \geq 0$, we have $(\alpha f)^+ = \alpha f^+$ and $(\alpha f)^- = \alpha f^-$, so that by Corollary 4.2.6 we have

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

Other cases are left for the reader.

(ii) follows immediately from

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

(iii) If $\int f d\mu = 0$ there is nothing to prove. Suppose $\alpha := \int f d\mu \neq 0$. Let ω we a complex number with $\omega\alpha = |\alpha|$. Then

$$\begin{aligned} \left| \int f d\mu \right| &= |\alpha| = \omega\alpha = \omega \int f d\mu = \int (\omega f) d\mu = \int \mathrm{Re}(\omega f) d\mu \\ &\leq \int |\omega f| d\mu = \int |f| d\mu. \end{aligned}$$

□

4.3.4 Lebesgue's Dominated Convergence Theorem *Let (f_n) be a sequence of measurable complex functions on a measure space which converges to a measurable function f almost everywhere. If there exists a function $g \in L^1(\mu)$ with $|f_n(x)| \leq g(x)$ for almost every $x \in X$ then $f \in L^1(\mu)$ and*

$$\lim_{n\to\infty} \int |f_n - f| d\mu = 0 \qquad \textit{and} \qquad \lim_{n\to\infty} \int f_n d\mu = \int f d\mu.$$

Proof. Since the integral over sets of measure zero is always zero and since changing the function on a set of measure zero is not affecting the integral, by a similar argument as in the proof of Jegorov's Theorem 3.5.2 we may assume $f_n \to f$ pointwise and $|f_n| \leq g$ for all $n \in \mathbb{N}$ and $x \in X$.

Since $|f_n| \leq g$ and $|f| \leq g$ we gave $|f_n - f| \leq 2g$. Functions $2g - |f_n - f|$ are nonnegative and measurable and pointwise converge to $2g$. By Fatou lemma 4.2.7 we have

$$\begin{aligned}\int 2g d\mu &= \int \liminf_{n\to\infty}(2g - |f_n - f|)d\mu \leq \liminf_{n\to\infty} \int (2g - |f_n - f|)d\mu \\ &= \int 2g d\mu - \limsup_{n\to\infty} \int |f_n - f| d\mu.\end{aligned}$$

This implies

$$0 \leq \liminf_{n\to\infty} \int |f_n - f| d\mu \leq \limsup_{n\to\infty} \int |f_n - f| d\mu \leq 0,$$

from where it folows $\lim\limits_{n\to\infty} \int |f_n - f| d\mu = 0$. By Proposition 4.3.3(iii) we finally have

$$\left| \int f_n d\mu - \int f d\mu \right| \leq \int |f_n - f| d\mu \to 0. \qquad \square$$

We will conclude this section by proving that $L^1(\mu)$ is a Banach space. To prove this result we need the following corollary of the Lebesgue's Dominated Convergence Theorem 4.3.4 which is also of independent interest.

4.3.5 Proposition *Let (f_n) be a sequence in $L^1(\mu)$ such that*

$$\sum_{n=1}^{\infty} \|f_n\|_1 < \infty.$$

Then the series $\sum\limits_{n=1}^{\infty} f_n$ converges almost everywhere to a function $f \in L^1(\mu)$ and

$$\int f d\mu = \sum_{n=1}^{\infty} \int f_n d\mu.$$

Proof. Define $g(x) := \sum\limits_{n=1}^{\infty} |f_n(x)|$. Then $\int g d\mu < \infty$, so that $g(x) < \infty$ for almost every $x \in X$. Then the series $\sum_{n=1}^{\infty} f_n(x)$ is absolutely convergent almost everywhere. If $g_n := f_1 + \cdots + f_n$, then $|g_n| \leq g$, so that by Lebesgue's Dominated Convergence Theorem 4.3.4 we have

$$\sum_{n=1}^{\infty} \int f_n d\mu = \lim_{n\to\infty} \sum_{j=1}^{n} \int f_j d\mu = \lim_{n\to\infty} \int g_n d\mu = \int \lim_{n\to\infty} g_n d\mu = \int \sum_{n=1}^{\infty} f_n d\mu.$$

□

In literature, any theorem that resembles to the following one is called the ***"Riesz-Fischer Theorem"***.

4.3.6 Riesz-Fischer Theorem *The space $L^1(\mu)$ is a Banach space. Furthermore, if $f_n \to f$ in $L^1(\mu)$, then some subsequence f_{n_k} converges to f almost everywhere.*

Proof. Since we already know that $L^1(\mu)$ is a normed space, it suffices to prove that $L^1(\mu)$ is complete.

Let (f_n) be a Cauchy sequence in $L^1(\mu)$. We first prove that some subsequence of (f_n) converges in $L^1(\mu)$. Since (f_n) is Cauchy, by an easy induction one can prove that there is a strictly increasing sequence (n_k) of positive integers such that

$$\|f_n - f_m\|_1 < 2^{-k}$$

for all $n, m \geq n_k$. By Proposition 4.3.5 the sum

$$f := f_{n_1} + \sum_{j=1}^{\infty} (f_{n_{j+1}} - f_{n_j})$$

converges almost everywhere and $f \in L^1(\mu)$. Partial sums

$$f_{n_1} + \sum_{j=1}^{k} (f_{n_{j+1}} - f_{n_j}) = f_{n_{k+1}}$$

converge to f almost everywhere and since

$$|f_{n_{k+1}}| \leq |f_{n_1}| + \sum_{j=1}^{k} |f_{n_{j+1}} - f_{n_j}| \leq |f_{n_1}| + \sum_{j=1}^{\infty} |f_{n_{j+1}} - f_{n_j}| =: g \in L^1(\mu),$$

by the Lebesgue's Dominated Convergence Theorem 4.3.4 we conclude

$$\lim_{k\to\infty} \int |f_{n_{k+1}} - f| d\mu = 0.$$

Therefore, for each ϵ there exists $k_1 \in \mathbb{N}$ such that $\|f - f_{n_k}\| < \frac{\epsilon}{2}$ for all $k \geq k_1$. Due to the Cauchy property we can also find $k_2 \in N$ such that $\|f_n - f_m\|_1 < \frac{\epsilon}{2}$ for all $n, m \geq n_{k_2}$. Hence, if $k_0 = \max\{k_1, k_2\}$, then for all $n \geq n_{k_0}$ we have

$$\|f - f_n\|_1 \leq \|f - f_{n_{k_1}}\|_1 + \|f_{n_{k_0}} - f_n\|_1 < \epsilon.$$

This yields $f_n \to f$ in $L^1(\mu)$. □

4.4 Riemann vs. Lebesgue

Let $f \colon [a, b] \to \mathbb{R}$ be a function. For any partion P

$$P: \qquad a = x_0 < x_1 < \ldots < x_n = b$$

of $[a, b]$ define

$$m_j = \inf_{x_{j-1} \leq x \leq x_j} f(x) \qquad \text{and} \qquad M_j = \sup_{x_{j-1} \leq x \leq x_j} f(x).$$

The lower and the upper Darboux sum of f with respect to the partition P are defined as

$$S_P = \sum_{j=1}^{n} m_j \Delta_j x \qquad \text{and} \qquad Z_P = \sum_{j=1}^{n} M_j \Delta_j x$$

where $\Delta_j = x_j - x_{j-1}$. If $\sup_P S_P = \inf_P Z_P$, then we say that f is ***Riemann integrable*** on $[a, b]$ and we call this value $\sup_P S_P = \inf_P Z_P$ the ***Riemann integral*** of f on $[a, b]$. By a result of Analysis 1 we already know that every Riemann integrable function on $[a, b]$ is bounded.

4.4.1 Theorem *A Riemann integrable function $f \colon [a, b] \to \mathbb{R}$ is Lebesgue integrable and*

$$\int_a^b f(x)dx = \int_{[a,b]} f dm.$$

Proof. For a partition P define step functions

$$s_P = \sum_{j=1}^{m} m_j \chi_{[x_{j-1}, x_j)} + f(b)\chi_{\{b\}} \quad \text{and} \quad z_P = \sum_{j=1}^{m} M_j \chi_{[x_{j-1}, x_j)} + f(b)\chi_{\{b\}}.$$

Then

$$\int_{[a,b]} s_P dm = S_P \quad \text{and} \quad \int_{[a,b]} z_P dm = Z_P$$

and $s_P(x) \leq f(x) \leq z_P(x)$ for each $x \in [a, b]$. Denote by $\Delta_P := \max_{1 \leq j \leq n} \Delta_j x$ the width of the partition P. If f is Riemann integrable then there exists a sequence (P_k) of finer partitions (each partition is included in the next one) with $\lim\limits_{k\to\infty} \Delta_k \to 0$ and

$$\lim_{k\to\infty} \int_{[a,b]} s_{P_k} dm = \lim_{k\to\infty} S_{P_k} = \int_a^b f(x)dx = \lim_{k\to\infty} Z_{P_k} = \lim_{k\to\infty} \int_{[a,b]} z_{P_k} dm.$$

Since (s_{P_k}) is increasing it increases to a Borel function s pointwise. Similarly, since (z_{P_k}) is decreasing it decreases to a Borel function pointwise. Since $s_{P_k} \leq f \leq z_{P_k}$ we conclude $s \leq f \leq z$. By the Lebesgue Dominated Convergence Theorem 4.3.4 we conclude

$$\int_{[a,b]} sdm = \int_a^b f(x)dx = \int_{[a,b]} zdx.$$

Since $\int\limits_{[a,b]} (z-s)dm = 0$ and since $z \geq s$ we conclude $z = s = f$ almost everywhere with respect to the Lebesgue measure. Hence, f is Lebesgue measurable. Since $f = s$ almost everywhere, we have

$$\int_{[a,b]} fdm = \int_{[a,b]} sdm = \int_a^b f(x)dx.$$

□

4.4.2 Proposition *A bounded function $f\colon [a,b] \to \mathbb{R}$ is Riemann integrable iff it is continuous almost everywhere.*

4.5 Product Measure

Suppose we have two spaces with measure $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \lambda)$. The product σ-algebra $\mathcal{A} \otimes \mathcal{B}$ is the smallest σ-algebra on $X \times Y$ which contains all measurable rectangles. The goal of this section is to produce a measure on the measurable space $(X \times Y, \mathcal{A} \otimes \mathcal{B})$ such that the measure of a rectangle is a product of lengths of its sides. Let us denote by $\mathcal{S}$ the set

$$\{A \times B : \ A \in \mathcal{A}, B \in \mathcal{B}\}$$

of all measurable rectangles in $X \times Y$. We also define

$$\Theta(A \times B) := \mu(A) \times \lambda(B)$$

for any $A \times B \in \mathcal{S}$.

4.5.1 Proposition *The mapping $\Theta\colon \mathcal{S} \to [0, \infty]$ is a semi-measure on a semi-algebra $\mathcal{S}$.*

Proof. Obviously we have $\emptyset = \emptyset \times \emptyset \in \mathcal{S}$ and

$$(A \times B) \cap (C \times D) = (A \cap C) \times (B \cap D) \in \mathcal{S}.$$

Since

$$(A \times B)^c = (A^c \times B) \cup (X \times B^c)$$

the complement of $A \times B$ is a finite union of disjoint members of $\mathcal{S}$. Hence, $\mathcal{S}$ is a semi-algebra.

To prove that Θ is a semi-measure on $\mathcal{S}$, note first that we have $\Theta(\emptyset) = 0$. We will prove that for any sequence $(A_n \times B_n)$ of disjoint measurable rectangles whose union is a measurable rectangle $A \times B$ we have

$$\Theta(A \times B) = \sum_{n=1}^{\infty} \Theta(A_n \times B_n).$$

Since the family $(A_n \times B_n)_n$ consists of pairwise disjoint rectangles with union $A \times B$, for each $(x, y) \in X \times Y$ we have

$$\chi_A(x) \cdot \chi_B(y) = \chi_{A \times B}(x, y) = \sum_{n=1}^{\infty} \chi_{A_n \times B_n}(x, y) = \sum_{n=1}^{\infty} \chi_{A_n}(x)\chi_{B_n}(y).$$

Integration over y yields

$$\begin{aligned} \chi_A(x)\lambda(B) &= \int_Y \chi_A(x)\chi_B(y)d\lambda(y) = \int_Y \sum_{n=1}^{\infty} \chi_{A_n}(x)\chi_{B_n}(y)\, d\lambda(y) \\ &= \sum_{n=1}^{\infty} \chi_{A_n}(x) \int_Y \chi_{B_n}(y)d\lambda(y) = \sum_{n=1}^{\infty} \chi_{A_n}(x)\lambda(B_n). \end{aligned}$$

Integration over x yields

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

□

4.5.2 Corollary *The semi-measure $\Theta \colon \mathcal{S} \to [0, \infty]$ can be extended to the measure on the σ-algebra $\mathcal{A} \otimes \mathcal{B}$. If μ and λ are σ-finite, then the extension is unique.*

Proof. First we uniquely extend Θ to the measure on the algebra generated by $\mathcal{S}$. By Carathódory's theorem the measure on the algebra can be extended to the σ-algebra which contains all measurable rectangles and so $\mathcal{A} \otimes \mathcal{B}$.

Suppose μ and λ are σ-finite and pick pairwise disjoint sequences $(X_n) \subseteq \mathcal{A}$ and $(Y_n) \subseteq \mathcal{B}$ of sets with finite measure such that $X = \bigcup_n X_n$ and $Y = \bigcup_m Y_m$. Then

$$X \times Y = \bigcup_{n,m} (X_n \times Y_m)$$

and since $\Theta(X_n \times Y_m) = \mu(X_n) \times \lambda(Y_m) < \infty$, the extension is unique by Theorem 2.5.2. □

The measure obtained in the corollary is denoted by $\mu \times \lambda$ and it is called the ***product measure*** of measures μ and λ.

For any subset $E \subseteq X \times Y$ and for all $x \in X$ and $y \in Y$ we define the x-***section*** E_x and y-***section*** E^y of the set E as

$$E_x = \{y \in Y : (x, y) \in E\} \qquad \text{and} \qquad E^y = \{x \in x : (x, y) \in E\}.$$

Similarly, if $f \colon X \times Y \to Z$ is a function, we define its sections as

$$f_x(y) := f(x, y) \qquad \text{and} \qquad f^y(x) := f(x, y).$$

The sections f_x and f^y are functions on Y and X, respectively.

4.5.3 Example Let E be any subset of $X \times Y$. Then for all $x \in X$ and $y \in Y$ we have $(\chi_E)_x = \chi_{E_x}$ and $(\chi_E)^y = \chi_{E^y}$.

4.5.4 Proposition

(i) *If $E \in \mathcal{A} \otimes \mathcal{B}$, then $E_x \in \mathcal{B}$ and $E^y \in \mathcal{A}$ for all $x \in X$ and $y \in Y$.*

(ii) *If f is a $\mathcal{A} \otimes \mathcal{B}$-measurable function on $X \times Y$, then f_x and f^y are are measurable with respect to $\mathcal{B}$ and $\mathcal{A}$, respectively, for all $x \in X$ and $y \in Y$.*

Proof. (i) Let $\mathcal{C}$ be the family of all subsets of $X \times Y$ such that $E_x \in \mathcal{B}$ for each $x \in X$ and $E^y \in \mathcal{A}$ for each $y \in Y$. Since it is easy to see that $\emptyset_x = \emptyset$, $(E_x^c) = (E_x)^c$ and $\left(\bigcup_n E_n\right)_x = \bigcup_n (E_n)_x$, and that similar identites hold for y-sections, $\mathcal{C}$ is a σ-algebra. Also, in the case $A \times B$ and $x \in X$ we have $(A \times B)_x = B$ if $x \in A$ and $(A \times B)_x = \emptyset$ otherwise. This proves $\mathcal{C}$ also contains all measurable rectangles and so $\mathcal{A} \otimes \mathcal{B} \subseteq \mathcal{C}$.

(ii) Pick a measurable subset D from the target space Z and pick $x \in X$. Since f is $\mathcal{A} \otimes \mathcal{B}$-measurable, the set $f^{-1}(D)$ is contained in $\mathcal{A} \otimes \mathcal{B}$, so that by (i) we have $(f^{-1}(D))_x \in \mathcal{B}$. To finish the proof observe $f_x^{-1}(D) = (f^{-1}(D))_x$. □

A ***monotone class*** on a set X is a family $\mathcal{M}$ of subsets of X which satisfies

(i) the union of any increasing sequence $A_1 \subseteq A_2 \subseteq \cdots$ of sets in $\mathcal{M}$ is again $\mathcal{M}$;

(ii) the intersection of any decreasing sequence $A_1 \supseteq A_2 \supseteq \cdots$ of sets in $\mathcal{M}$ is again $\mathcal{M}$.

It is easy to see that for a given subset $\mathcal{S} \subseteq \mathbb{P}(X)$ there exists the smallest monotone class which contains $\mathcal{S}$. In fact, it equals to the intersection of all monotone classes that contain $\mathcal{S}$.

Obviously any σ-algebra is a monotone class. However, if $\mathcal{A}$ is an algebra that is also a monotone class, then it is a σ-algebra. Indeed, if (A_n) is a sequence in $\mathcal{A}$, then $B_n = \bigcup_{i=1}^n A_i$ is an increasing sequence in $\mathcal{A}$ and since $\mathcal{A}$ is a monotone class, $\bigcup_n A_n = \bigcup_n B_n \in \mathcal{A}$.

4.5.5 Monotone Class Lemma *If $\mathcal{A}$ is an algebra on X, then the monotone class $\mathcal{M}$ generated by $\mathcal{A}$ is σ-algebra.*

It is obvious that $\mathcal{M}$ is actually the σ-algebra generated by $\mathcal{A}$.

Proof. By the remark preceding the theorem it suffices to prove that $\mathcal{M}$ is an algebra.

For each $A \in \mathcal{M}$ define

$$\mathcal{M}(A) = \{B \in \mathcal{M} : \ A \setminus B \in \mathcal{M}, B \setminus A \in \mathcal{M} \text{ and } A \cap B \in \mathcal{M}\}.$$

By routine set-theoretic manipulation one can show that $\mathcal{M}(A)$ is a monotone class. By the symmetric definition for sets $A, B \in \mathcal{M}$ we have $A \in \mathcal{M}(B)$ iff $B \in \mathcal{M}(A)$. If $A \in \mathcal{A}$, then for each $B \in \mathcal{A}$ we have $B \in \mathcal{M}(A)$. This implies $\mathcal{A} \subseteq \mathcal{M}(A)$ for each $A \in \mathcal{A}$. Since $\mathcal{M}(A)$ is a monotone class, we have $\mathcal{M} \subseteq \mathcal{M}(A)$. This implies that each $B \in \mathcal{M}$ is contained in $\mathcal{M}(A)$ and so $A \in \mathcal{M}(B)$. This yields $\mathcal{A} \subseteq \mathcal{M}(B)$ and so $\mathcal{M} \subseteq \mathcal{M}(B)$. This proves $\mathcal{M}$ is closed under intersections and differences, i.e., if $A, B \in \mathcal{M}$, then $A \cap B$ and $A \setminus B$ are in $\mathcal{M}$. Since $X \in \mathcal{A} \subseteq \mathcal{M}$ we see that $\mathcal{M}$ is closed under taking complements. □

The following theorem can be considered as the "baby" version of Fubini's theorem.

4.5.6 Theorem *Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \lambda)$ be σ-finite measure spaces. For all $E \in \mathcal{A} \otimes \mathcal{B}$ the functions*

$$x \mapsto \lambda(E_x) \qquad \text{and} \qquad y \mapsto \mu(E^y)$$

are measurable and

$$(\mu \times \lambda)(E) = \int_X \lambda(E_x) d\mu(x) = \int_Y \mu(E^y) d\lambda(y).$$

Proof. Define the family $\mathcal{E}$ as the family of all subsets of $E \in \mathcal{A} \otimes \mathcal{B}$ for which the theorem holds. Suppose first μ and λ are finite measures.

To prove $\mathcal{E}$ contains all measurable rectangles, pick $E = A \times B$ with $A \in \mathcal{A}$ and $B \in \mathcal{B}$. If $y \in Y$, then $E^y = A$ if $y \in B$ and $E^y = \emptyset$ if $y \notin B$. This implies $\mu(E^y) = \mu(A)\chi_B(y)$. Similarly we see $\lambda(E_x) = \lambda(B)\chi_A(x)$. Hence, if $E = A \times B$ is a measurable rectangle, the functions $x \mapsto \lambda(E_x)$ and $y \mapsto \mu(E^y)$ are $\mathcal{A}$ and $\mathcal{B}$ measurable, respectively. By integration we see

$$\int_X \lambda(E_x) d\mu(x) = \mu(A) \cdot \lambda(B) = \int_Y \mu(E^y) d\lambda(y)$$

so that $E \in \mathcal{E}$. If C and D are disoint, then $\int_{C \cup D} = \int_C + \int_D$ implies that $\mathcal{E}$ also contains all finite unions of disjoint measurable rectangles. Hence, by Proposition 2.6.2 the algebra $\mathcal{E}_0$ generated by the semi-algebra of all measurable rectangles is contained in $\mathcal{E}$.

We claim $\mathcal{E}$ is a monotone class. If we prove this, then the monotone class $\mathcal{E}'$ generated by $\mathcal{E}_0$ equals $\mathcal{A} \otimes \mathcal{B}$ by the Monotone Class Lemma 4.5.5. Then the proof will be finished since $\mathcal{A} \otimes \mathcal{B} = \mathcal{E}' \subseteq \mathcal{E} \subseteq \mathcal{A} \otimes \mathcal{B}$.

Pick an increasing sequence (E_n) of sets in $\mathcal{E}$. Denote by E the union $\bigcup_n E_n$. For each $y \in Y$ the sequence (E_n^y) is increasing and $E^y = \bigcup_n E_n^y$. Since the functions $f_n(y) := \mu(E_n^y)$ are measurable and f_n converges pointwise to the function $f(y) := \mu(E^y)$, the function f is also measurable, so that by the Lebesgue's Monotone Convergence Theorem 4.2.3 we have

$$\int_Y \mu(E^y) d\lambda(y) = \lim_{n \to \infty} \int_Y \mu(E_n^y) d\lambda(y) = \lim_{n \to \infty} (\mu \times \lambda)(E_n) = (\mu \times \lambda)(E).$$

Similarly we can prove $g(x) := \lambda(E^x)$ is measurable and that

$$\int_X \lambda(E_x) d\mu(x) = (\mu \times \lambda)(E)$$

which proves $E \in \mathcal{E}$. To prove that $\mathcal{E}$ is closed under taking intersections of decreasing sequences, we use Lebesgue's Dominated Convergence Theorem 4.3.4. This is legal since measures μ and λ are finite.

Suppose X and Y are σ-finite measures. Then $X = \bigcup_n X_n$ and $Y = \bigcup_n Y_n$ for some increasing sequences of sets with finite measure (X_n) and (Y_n), respectively. Pick $E \in \mathcal{A} \otimes \mathcal{B}$. Then the set $E \cap (X_j \times Y_j)$ is contained in the σ-algebra $\mathcal{A}|_{X_j} \otimes \mathcal{B}|_{Y_j}$ so that the functions $x \mapsto \lambda(E_x \cap Y_j)$ and $y \mapsto \mu(E^y \cap Y_j)$ are measurable on X_j and Y_j, respectively, and we have

$$(\mu \times \lambda)(E \cap (X_j \times Y_j)) = \int_{X_j} \lambda(E_x \cap Y_j) d\mu(x) = \int_{Y_j} \mu(E^y \cap X_j) d\lambda(y).$$

Extend these functions by zero on Y_j^c and X_j^c, respectively so that we obtain measurable functions ϕ_j and ψ_j. Since ϕ_j and ψ_j converge to $x \mapsto \lambda(E_x)$ and $x \mapsto \mu(E^y)$, respectively the limit functions are $\mathcal{A}$ and $\mathcal{B}$-measurable, respectively. An application of the Lebesgue's Monotone Convergence Theorem 4.2.3 completes the proof. □

4.5.7 Tonelli-Fubini's theorem *Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \lambda)$ be σ-finite measure space.*

(i) *For any measurable function $f\colon X \times Y \to [0, \infty]$ the functions $g(x) := \int_Y f_x(y) d\lambda(y)$ and $h(y) := \int_X f^y(x) d\mu(x)$ are measurable with respect to $\mathcal{A}$ and $\mathcal{B}$, respectively, and we have*

$$\int\limits_{X\times Y} f(x, y) d(\mu \times \lambda)(x, y) = \int\limits_X \Big(\int\limits_Y f(x, y) d\lambda(y) \Big) d\mu(x)$$
$$= \int\limits_Y \Big(\int\limits_X f(x, y) d\mu(x) \Big) \lambda(y).$$

(ii) *For any $L^1(\mu \times \lambda)$ we have $f_x \in L^1(\lambda)$ for almost every $x \in X$ and $f^y \in L^1(\mu)$ for almost every $y \in Y$. The functions g and h from (i) are in $L^1(\mu)$ and $L^1(\lambda)$, respectively and the conclusion from (i) holds. The same conclusion holds also if we replace $f \in L^1(\mu \times \lambda)$ by the assumption that f is measurable and at least one of the repeated integrals*

$$\int\limits_X \Big(\int\limits_Y |f(x, y)| d\lambda(y) \Big) d\mu(x) \qquad \text{and} \qquad \int\limits_Y \Big(\int\limits_X |f(x, y)| d\mu(x) \Big) d\lambda(y)$$

is finite.

Assertions in (i) and (ii) are called Tonelli's and Fubini's theorem, respectively. Tonelli's and Fubini's theorem state that under appropriate conditions double integral equals to any of the repeated ones.

Proof. (i) If $f = \chi_E$ for some $E \in \mathcal{A} \otimes \mathcal{B}$, then $f_x = \chi_{E_x}$ and $f^y = \chi_{E^y}$. Also $g(x) = \int_Y f_x(y) d\lambda(y) = \lambda(E_x)$. Similarly $h(y) = \mu(E^y)$. In this case Tonelli's theorem reduces to Theorem 4.5.6. By linearity (i) also holds for nonnegative measurable step functions.

If $f\colon X \times Y \to [0, \infty]$ is a measurable function, find an increasing sequence (s_n) of measurable step functions which converges pointwise to f. Then $g_n(x) := \int_Y (s_n)_x(y) d\lambda(y)$ increases to g and so g is measurable. By the Lebesgue's Monotone Convergence Theorem 4.2.3 we have

$$\int\limits_{X\times Y} f(x, y) d(\mu \times \lambda)(x, y) = \lim_{n\to\infty} \int\limits_{X\times Y} s_n(x, y) d(\mu \times \lambda)(x, y)$$
$$= \lim_{n\to\infty} \int\limits_X g_n(x) d\mu(x) = \int\limits_X g(x) d\mu(x)$$
$$= \int\limits_X \Big(\int\limits_Y f(x, y) d\lambda(y) \Big) d\mu(x).$$

The reader can similarly prove the remaining part of (i).

(ii) In the general case, by writing $f = \operatorname{Re} f + i \operatorname{Im} f$ it suffices to consider only the case when f is real. If f is real we can write $f = f^+ - f^-$ so that we only need to consider the case when f is nonnegative.

If $f \in L^1(\mu \times \lambda)$ is nonnegative, then (i) implies $\int_X g(x) d\mu(x) < \infty$, so that $g(x) < \infty$ for almost every $x \in X$. This means $f_x \in L^1(\lambda)$ for almost every $x \in X$. Similarly $f^y \in L^1(\mu)$ for almost every $y \in Y$. $\square$

5 Complex Measures

5.1 Variation of a Complex Measure

A ***complex measure*** on a measurable space $(X, \mathcal{A})$ is a countably additive map $\lambda\colon \mathcal{A} \to \mathbb{C}$. Countable additivity means

$$\lambda\Big(\bigcup_{n=1}^{\infty} A_n\Big) = \sum_{n=1}^{\infty} \lambda(A_n)$$

for an sequence of pairwise disjoint sets in $\mathcal{A}$.

Observe, that if we take $A_n = \emptyset$ for each $n \in \mathbb{N}$, on the account that λ maps into $\mathbb{C}$ we have $\lambda(\emptyset) = 0$. Also, since for any permutation π of $\mathbb{N}$ we have

$$\sum_{n=1}^{\infty} \lambda(A_n) = \lambda\Big(\bigcup_{n=1}^{\infty} A_n\Big) = \lambda\Big(\bigcup_{n=1}^{\infty} A_{\pi(n)}\Big) = \sum_{n=1}^{\infty} \lambda(A_{\pi(n)}),$$

the series $\sum\limits_{n=1}^{\infty} \lambda(A_n)$ is absolutely convergent.

5.1.1 Example Let $(X, \mathcal{A})$ be a measurable space with with a positive measure μ. For any function $f \in L_1(\mu)$ we define

$$(f \cdot \mu)(A) := \int_A f d\mu \qquad (A \in \mathcal{A}).$$

Then $f \cdot \mu$ defines a complex measure on $\mathcal{A}$. Usually we will denote this measure by $f\mu$ instead of $f \cdot \mu$. Sometimes we will also write $f d\mu$.

The following proof is very similar to the proof in the case of positive measures. The proof is omitted.

5.1.2 Proposition *A finitely additive function $\lambda\colon \mathcal{A} \to \mathbb{C}$ is a complex measure iff λ satisfies one of the following:*

(i) *For any increasing sequence (A_n) of sets in $\mathcal{A}$ we have $\lambda\Big(\bigcup\limits_{n=1}^{\infty} A_n\Big) = \lim\limits_{n\to\infty} \lambda(A_n)$.*

(ii) *For any decreasing sequence (A_n) of sets in $\mathcal{A}$ we have $\lambda\Big(\bigcap\limits_{n=1}^{\infty} A_n\Big) = \lim\limits_{n\to\infty} \lambda(A_n)$.*

If (A_n) is a sequence of pairwise disjoint sets in $\mathcal{A}$ and $A = \bigcup_n A_n$, then (A_n) is called a ***countable partition*** of A. If λ is a complex measure on $\mathcal{A}$, we define its ***total variation*** $|\lambda|$ as

$$|\lambda|(A) = \sup \sum_{n=1}^{\infty} |\lambda(A_n)|$$

where the supremum runs over all countable partitions of the set A. Since $\{A, \emptyset, \emptyset, \ldots\}$ is a countable partition for A we have $|\lambda(A)| \leq |\lambda|(A)$.

Given a complex measure λ on a σ-algebra $\mathcal{A}$, one can define

$$(\operatorname{Re}\lambda)(A) := \operatorname{Re}\lambda(A) \qquad \text{and} \qquad (\operatorname{Im}\lambda)(A) := \operatorname{Im}\lambda(A).$$

Then $\operatorname{Re}\lambda$ and $\operatorname{Im}\lambda$ are complex measures with values in $\mathbb{R}$. Complex measures with values in $\mathbb{R}$ are called ***real measures*** . Real measures $\operatorname{Re}\lambda$ and $\operatorname{Im}\lambda$ are called the ***real part*** and the ***imaginary part*** of λ. A complex measure λ always satisfies $|\lambda| \leq |\operatorname{Re}\lambda| + |\operatorname{Im}\lambda|$. To see this, pick $B \in \mathcal{A}$. Then

$$|\lambda(B)| = |\operatorname{Re}(B) + i \operatorname{Im}\lambda(B)| \leq |\operatorname{Re}(B)| + |\operatorname{Im}(B)|.$$

If (A_n) is countable partition for A, then

$$\sum_{n=1}^{\infty} |\lambda(A_n)| \leq \sum_{n=1}^{\infty} |\operatorname{Re}\lambda(A_n)| + \sum_{n=1}^{\infty} |\operatorname{Im}\lambda(A_n)| \leq |\operatorname{Re}\lambda|(A) + |\operatorname{Im}\lambda|(A).$$

This immediately yields $|\lambda|(A)| \leq |\operatorname{Re}\lambda|(A) + |\operatorname{Im}\lambda|(A)$.

5.1.3 Theorem *For any complex measure λ the total variation $|\lambda|$ is a finite positive measure.*

Proof. We first prove that $|\lambda|$ is a positive measure. It is obvious that $|\lambda|(\emptyset) = 0$. Let A be the disjoint union of a sequence (A_n) in $\mathcal{A}$. Pick $\epsilon > 0$. For each $n \in \mathbb{N}$ pick a nonnegative real number $a_n \leq |\lambda|(A_n)$ and a countable partition $(A_{n,j})$ for A_n such that

$$\sum_{j=1}^{\infty} |\lambda(A_{n,j})| \geq a_n - \frac{\epsilon}{2^n}.$$

Then the family $(A_{n,j})_{n,j}$ is a countable partition for A, from where it follows

$$|\lambda|(A) \geq \sum_{n=1}^{\infty} \sum_{j=1}^{\infty} |\lambda(A_{n,j}) \geq \sum_{n=1}^{\infty} a_n - \epsilon.$$

Letting $a_n \to |\lambda|(A_n)$ for each $n \in \mathbb{N}$ and then $\epsilon \to 0$ we obtain

$$|\lambda|(A) \geq \sum_{n=1}^{\infty} |\lambda|(A_n).$$

To prove the converse inequality, pick a countable partition (B_j) for A. Then $(B_j \cap A_n)$ is a countable partition for A_n. Since

$$\sum_{j=1}^{\infty} |\lambda(B_j)| = \sum_{j=1}^{\infty} \left| \sum_{n=1}^{\infty} \lambda(B_j \cap A_n) \right| \leq \sum_{n=1}^{\infty} \sum_{j=1}^{\infty} |\lambda(A_n \cap B_j)| \leq \sum_{n=1}^{\infty} |\lambda|(A_n).$$

By the definition of the total variation we conclude

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

To prove that $|\lambda$ is finite, from the decompositon $\lambda = \operatorname{Re}\lambda + i \operatorname{Im}\lambda$ we conclude that we only need to consider the case when λ is a real measure. Assume $|\lambda|(X) = \infty$. Then there exists countable partitions (X_n) of X with arbitrary large sums of the series $\sum_{n=1}^{\infty} |\lambda(X_n)|$. Let (X_n) be any such countable partition of X. Define

$$S = \{n \in \mathbb{N} : \ \lambda(X_n) \geq 0\} \qquad \text{and} \qquad T := \{n \in \mathbb{N} : \ \lambda(X_n) < 0\}.$$

Since

$$\sum_{n=1}^{\infty} |\lambda(X_n)| = \sum_{n \in S}^{\infty} \lambda(X_n) - \sum_{n \in T}^{\infty} \lambda(X_n),$$

at least one of $\sum\limits_{n \in S}^{\infty} \lambda(X_n)$ and $\sum\limits_{n \in T}^{\infty} \lambda(X_n)$ is large by absolute value. This means that the measure of at least one of the sets $\bigcup\limits_{n \in S} X_n$ and $\bigcup\limits_{n \in T} X_n$ is large. This means that in $\mathcal{A}$ there exists a set Y_1 such that $|\lambda(Y_1)|$ is arbitrarily large in absolute value.

Define $A_1 = X$ and find $Y_1 \in \mathcal{A}$ such that

$$|\lambda(Y_1)| \geq |\lambda(A_1)| + 1.$$

Since $|\lambda|(A_1 \setminus Y_1) + |\lambda|(Y_1) = |\lambda|(A_1) = \infty$, the measure $|\lambda|$ of at least one of the sets Y_1 and $A_1 \setminus Y_1$ is ∞. Denote this set by A_2. Then $|\lambda|(A_2) = \infty$. If $A_2 = Y_1$, then obviously $|\lambda(A_2)| \geq 1$. If $A_2 = A_1 \setminus Y_1$, then $|\lambda(A_2 \setminus Y_1)| \geq ||\lambda(Y_1)| - |\lambda(A_1)|| \geq 1$. Hence, in any case we have $|\lambda(A_2)| \geq 1$.

Since $|\lambda|(A_2) = \infty$, there exists a measurable set $Y_2 \in \mathcal{A}$ with $Y_2 \subseteq A_2$ such that

$$|\lambda(Y_2)| \geq |\lambda(A_2)| + 2.$$

Similarly as above we can see that one of the sets Y_2 and $A_2 \setminus Y_2$ which we denote by A_3 satisfies

$$|\lambda|(A_3)| = \infty \qquad \text{and} \qquad |\lambda(A_3)| \geq 2.$$

Proceeding by induction we can construct a decreasing sequence (A_n) in $\mathcal{A}$ with $\lambda(A_n) \geq n-1$ for each $n \in \mathbb{N}$. Since λ is a real measure, the set $A := \bigcap_{n=1}^{\infty} A_n$ satisfies

$$|\lambda(A)| = \left|\lambda\Big(\bigcap_{n=1}^{\infty} A_n\Big)\right| = \lim_{n\to\infty} |\lambda(A_n)| = \infty.$$

This contradicts the fact that $\lambda(A) \in \mathbb{R}$. □

The set of all complex measures on $(X, \mathcal{A})$ is denoted by $M(X)$. On $M(X)$ we can define operations

$$(\mu + \lambda)(A) := \mu(A) + \lambda(A) \qquad \text{and} \qquad (\alpha\lambda)(A) := \alpha\lambda(A).$$

It is easy to see that $M(X)$ is a vector space. For $\lambda \in M(X)$ we define $\|\lambda\| := |\lambda|(X)$.

5.1.4 Theorem *The space of all complex measures $M(X)$ equipped with $\|\cdot\|$ is a Banach space.*

Proof. It suffices to prove that every absolutely convergent series in $M(X)$ converges in $M(X)$. Let (λ_n) be a absolutely convergent series in $M(X)$. Then

$$\sum_{n=1}^{\infty} |\lambda_n|(X) < \infty.$$

Hence, the series

$$\lambda(E) = \sum_{n=1}^{\infty} \lambda_n(E)$$

converges in $\mathbb{C}$ for each $E \in \mathcal{A}$. The reader should check that λ is a complex measure and that λ is the limit of the sequence of partial sums of the series $\sum_{n=1}^{\infty} \lambda_n$ in $M(X)$. □

5.2 Absolute Continuity and Mutual Singularity

Let μ be a positive measure and suppose λ is either a complex measure or a positive measure on a σ-algebra $\mathcal{A}$. The measure λ is ***absolutely continuous*** with respect to μ if $\mu(A) = 0$ implies $\lambda(A) = 0$. If λ is absolutely continuous with respect to μ, we denote it by $\lambda \ll \mu$.

5.2.1 Lemma *$\lambda \ll \mu$ iff $|\lambda| \ll \mu$.*

Proof. ($\Rightarrow$) Suppose $\mu(A) = 0$. Pick any countable partition (A_n) for A. Since $\mu(A_n) = 0$, we have $\lambda(A_n) = 0$ and so $\sum_{n=1}^{\infty} |\lambda(A_n)| = 0$. Taking the supremum over all countable partitions we conclude $|\lambda|(A) = 0$.

($\Leftarrow$) Since $|\lambda(A)| \leq |\lambda|(A)$ for each $A \in \mathcal{A}$, we conclude that $|\lambda|(A) = 0$ implies $\lambda(A) = 0$. □

The following theorem justifies the term "absolute continuiuty".

5.2.2 Theorem *Let λ and μ be a complex and a positive measure, respectively, on a σ-algebra $\mathcal{A}$. Then $\lambda \ll \mu$ iff for each $\epsilon > 0$ there exists $\delta > 0$ such that for each $A \in \mathcal{A}$ we have $|\lambda(A)| < \epsilon$ whenever $\mu(A) < \delta$.*

The condition in the theorem is called the ***absolute continuity condition***.

Proof. ($\Rightarrow$) Suppose $\mu(A) = 0$. Pick $\epsilon > 0$. Then we find $\delta > 0$ such that $\mu(A) = 0 < \delta$ implies $|\lambda(A)| < \epsilon$. Since $\epsilon > 0$ is arbitrary, we have $\lambda(A) = 0$.

($\Leftarrow$) Suppose (AC) is not satisfied. Then there exists $\epsilon > 0$ such that for all $\delta > 0$ there exist a set $A \in \mathcal{A}$ such that $\mu(A) < \delta$ but $|\lambda(A)| \geq \epsilon$.

For each $n \in \mathbb{N}$ find $A_n \in \mathcal{A}$ such that $\mu(A_n) < \frac{1}{2^n}$ and $|\lambda(A_n)| \geq \epsilon$. For each $n \in \mathbb{N}$ define the set $B_n = \bigcup_{k \geq n} A_k$ and note that $\mu(B_n) \leq \frac{1}{2^{n-1}}$. Define $A := \bigcap_{n=1}^{\infty} B_n$. Since the sequence (B_n) is decreasing, we have $\mu(A) = 0$. On the other hand,

$$|\lambda|(A) = \lim_{n \to \infty} |\lambda|(B_n) \geq |\lambda|(A_n) \geq |\lambda(A_n)| \geq \epsilon$$

implies $|\lambda| \not\ll \mu$. By Lemma 5.2.1 we conclude $\lambda \not\ll \mu$. □

A positive or a complex measure λ on a σ-algebra $\mathcal{A}$ is said to be ***concentrated*** on the set $A \in \mathcal{A}$ whenever $\lambda(B) = \lambda(B \cap A)$ for each $B \in \mathcal{A}$. Measures λ and μ are ***mutually singular*** whenever they are concentrated on disjoint sets. Mutual singularity of μ and λ is denoted by $\mu \perp \lambda$.

5.2.3 Proposition *Let λ and μ be two complex measures.*

(i) *If λ is positive, then λ is concentrated on A iff $\lambda(A^c) = 0$.*

(ii) *λ is concentrated on A iff $|\lambda|$ is concentrated on A.*

(iii) *$\lambda \perp \mu$ iff there exists $A \in \mathcal{A}$ with $|\lambda|(A) = 0$ and $|\mu|(A^c) = 0$.*

5.3 Positive and Negative part of a Real Measure

Let λ be a real measure on $(X, \mathcal{A})$. A set $P \in \mathcal{A}$ is said to be ***λ-positive*** whenever $\lambda(B) \geq 0$ for each measurable subset $B \subseteq P$. Similarly we can define ***λ-negative*** sets. A set $N \in \mathcal{A}$ is ***λ-null*** if $\lambda(B) = 0$ for every measurable set $B \subseteq N$. It is obvious that a set is λ-null iff it is λ-positive and λ-negative. Also, a set N is λ-null iff $|\lambda|(N) = 0$.

5.3.1 Lemma *If λ is a real measure on $(X, \mathcal{A})$, every measurable set A contains a λ-positive set P with $\lambda(P) \geq \lambda(A)$.*

Proof. We first prove that for each $\epsilon > 0$ there exists a measurable subset $P_\epsilon \subseteq A$ with $\lambda(P_\epsilon) \geq \lambda(A)$ and $\lambda(B) \geq -\epsilon$ for each measurable subset $B \subseteq P_\epsilon$. If this is not true, then there exists $\epsilon > 0$ such that for each measurable subset $C \subseteq A$ with $\lambda(C) \geq \lambda(A)$ we would have $\lambda(B) < -\epsilon$ for some measurable subset $B \subseteq C$.

First find $B_1 \subseteq A$ such that $\lambda(B_1) < -\epsilon$. Since $A \setminus B_1$ satisfies $\lambda(A \setminus B_1) = \lambda(A) - \lambda(B_1) > \lambda(A)$, there exists $B_2 \subseteq A \setminus B_1$ such that $\lambda(B_2) < -\epsilon$. By induction we find a sequence (B_n) of pairwise disjoint measurable subsets such that $B_n \subseteq A \setminus (B_1 \cup \cdots \cup B_{n-1})$ and $\lambda(B_n) < -\epsilon$. Then the set $B := \bigcup_n B_n$ does not have a real measure since $\lambda(B_n) \not\to 0$.

For each $k \in \mathbb{N}$ we can find a sequence (P_k) of measurable sets such that

$$A \supseteq P_1 \supseteq P_2 \supseteq \cdots$$

and $\lambda(P_k) \geq \lambda(A)$ and $\lambda(B) \geq -\frac{1}{k}$ whenever B is a measurable subset of P_k. Define $P := \bigcap_{k=1}^{\infty} P_k$. Then $\lambda(P) \geq \lambda(A)$ and if $B \subseteq P$, then $\lambda(B) \geq -\frac{1}{k}$ for each k implies $\lambda(B) \geq 0$. Hence, P is λ-positive set. $\square$

5.3.2 The Hahn Decomposition Theorem *For each real measure λ on $(X, \mathcal{A})$ there exist a λ-positive set P and a λ-negative set N such that $P \cup N = X$ and $P \cap N = \emptyset$. If $\widetilde{P}$ and $\widetilde{N}$ is another pair, then $\widetilde{P} \triangle P$ and $\widetilde{N} \triangle N$ are λ-null sets.*

Any pair (P, N) which satisfies the Hahn Decomposition theorem is called the Hahn decomposition of a real measure λ.

Proof. Let us define

$$s := \sup\{\lambda(A) : \ A \text{ is a } \lambda\text{-positive set}\}.$$

Let (P_k) be an increasing sequence of λ-positive sets with $\lambda(P_k) \uparrow s$. Define $P := \bigcup_k P_k$. If $A \subseteq P$ is measurable, then

$$\lambda(A) = \lim_k \lambda(P_k \cap A) \geq 0,$$

so that P is λ-positive. It is obvious that $\lambda(P) = \lim_k \lambda(P_k) = s$.

We claim that the set $N := X \setminus P$ is a λ-negative set. If not, there exists a measurable set $A \subseteq N$ with $\lambda(A) > 0$. By Lemma 5.3.1 there exists a λ-positive set $B \subseteq A$ with $\lambda(B) > 0$. Then $P \cup B$ is a λ-positive set with $\lambda(P \cup B) > s$. This contradiction yields that N is λ-negative.

Suppose $\widetilde{P}$ and $\widetilde{N}$ are another pair of sets that satisfy the conclusion of the theorem. Then $P \cap \widetilde{N}$ and $\widetilde{P} \cap N$ are λ-null sets from where it immediately follows

$$\lambda(P \setminus \widetilde{P}) + \lambda(\widetilde{P} \setminus P) = \lambda(P \cap \widetilde{N}) + \lambda(\widetilde{P} \cap N) = 0.$$

Similarly we can prove $\lambda(N \triangle \widetilde{N}) = 0$. □

5.3.3 The Jordan Decomposition Theorem *For a real measure λ on $(X, \mathcal{A})$ there exist unique positive measure λ^+ and λ^- on $\mathcal{A}$ such that $\lambda = \lambda^+ - \lambda^-$ and $\lambda^+ \perp \lambda^-$*

Measures λ^+ and λ^- are called the ***positive part*** and the ***negative part*** of a real measure λ, respectively.

Proof. Let (P, N) be a Hahn decomposition of λ. Defne $\lambda^+(A) := \lambda(A \cap P)$ and $\lambda^-(A) := -\lambda(A \cap N)$. Then λ^+ and λ^- are mutually singular positive measures which satisfy $\lambda = \lambda^+ - \lambda^-$.

Suppose μ_1 and μ_2 are positive measures on $\mathcal{A}$ with $\lambda = \mu_1 - \mu_2$ and $\mu_1 \perp \mu_2$. Suppose μ_1 and μ_2 are concentrated on E and F, respectively. By replacing E with $E \cup (E \cup F)^c$ we may assume $E \cup F = X$. Since (E, F) is another Hahn's decomposition of λ, we conclude that the set $E \triangle P$ is a λ-null set. Since for each $A \in \mathcal{A}$ we have

$$\begin{aligned}
\mu_1(A) = \mu_1(A \cap E) &= \mu_1(A \cap E) - \mu_2(A \cap E) = \lambda(A \cap E) \\
&= \lambda(A \cap E \cap P) + \lambda(A \cap E \cap P^c) = \lambda(A \cap E \cap P) \\
&= \lambda(A \cap E) = \lambda(A \cap E \cap P) + \lambda(A \cap E^c \cap P) = \lambda(A \cap P) \\
&= \lambda^+(A),
\end{aligned}$$

we conclude $\mu_1 = \lambda^+$ and so $\mu_2 = \lambda^-$. □

5.4 Lebesgue-Radon-Nikodým Theorem

If $f \in L^1(\mu)$, then the complex measure $f\mu$ is absolutely continuous with respect to μ. In this section we prove the converse statement (see 5.4.3). Before we proceed to 5.4.3 we need the following result.

5.4.1 Lemma *Suppose μ and λ are finite measures. Then either $\mu \perp \lambda$ or there exists $\epsilon > 0$ and $E \in \mathcal{A}$ such that $\mu(E) > 0$ and $\lambda \geq \epsilon\mu$ of E.*

The condition in the preceding lemma yields that E is a positive set for the real measure $\lambda - \epsilon\mu$.

Proof. Let $X = P_n \cup N_n$ be a Hahn decomposition for $\lambda - \frac{1}{n}\mu$ and define $P := \bigcup_n P_n$ and $N := \bigcap_n N_n$. Then N is a negtive set for $\lambda - \frac{1}{n}\mu$ for all n, so that $0 \leq \lambda(N) \leq \frac{1}{n}\mu(N) \to 0$. Hence, $\lambda(N) = 0$ and λ is concentrated on $N^c = P$. If $\mu(P) = 0$, then μ is concentrated on $P^c = N$ and so $\mu \perp \lambda$. Otherwise, if $\mu(P) > 0$, then $\mu(P_n) > 0$ for some $n \in \mathbb{N}$ and so P_n is a positive set for $\lambda - \frac{1}{n}\mu$. □

5.4.2 Proposition *For $f \in L^1(\mu)$ we have $\int\limits_A f d\mu = 0$ for each $A \in \mathcal{A}$ iff $f = 0$ almost everywhere.*

5.4.3 Lebesgue-Radon-Nikodým Theorem *Let λ be a complex measure and let μ be a σ-finite positive measure on a measurable space $(X, \mathcal{A})$. Then there exist unique complex measures λ_a and λ_s such that*

$$\lambda_a \ll \mu, \quad \lambda_s \perp \mu \quad \textit{and} \quad \lambda = \lambda_a + \lambda_s$$

and a unique $f \in L^1(\mu)$ such that

$$\lambda_a(A) = \int_A f d\mu \qquad (A \in \mathcal{A}).$$

The identity $\lambda = \lambda_a + \lambda_s$ is called the ***Lebesgue decomposition*** of λ with respect to μ. The function f is called the ***Radon-Nikodým derivative*** of λ with respect to μ. The Radon-Nikodým derivative of λ_a with respect to μ is denoted by $\frac{d\lambda_a}{d\mu}$. If $\lambda \ll \mu$, then Theorem 5.4.3 is called ***Radon-Nikodým theorem***.

Proof. The uniqueness part: Suppose there exist another pair of complex measures λ'_a and λ'_s such that $\lambda = \lambda'_a + \lambda'_s$, $\lambda'_a \ll \mu$ and $\lambda'_s \perp \mu$. Since $\lambda_a + \lambda_s = \lambda = \lambda'_a + \lambda'_s$ we conclude $\lambda_a - \lambda'_a = \lambda'_s - \lambda_s$ from where it follows $\lambda_a - \lambda'_a \ll \mu$ and $\lambda_a - \lambda'_a = \lambda'_s - \lambda_s \perp \mu$. Hence $\lambda'_a - \lambda_a \perp \lambda'_a - \lambda_a$ so that $\lambda'_a = \lambda_a$. This yields also $\lambda'_s = \lambda_s$.

Suppose there exists another function $\widetilde{f} \in L^1(\mu)$ such that $\lambda_a = f\mu = \widetilde{f}\mu$. This means that for any $A \in \mathcal{A}$ we have

$$\int_A f d\mu = \int_A \widetilde{f} d\mu.$$

Then $g := f - \widetilde{f}$ satisfies $\int_A g d\mu = 0$ so that $g \equiv 0$ almost everywhere on X by Proposition 5.4.2. This means $f \equiv \widetilde{f}$ almost everywhere on X so that $f = \widetilde{f}$ as elements of $L^1(\mu)$.

The existence part: Given any complex measure ν, it is easy to see that $\nu \ll \mu$ iff $\operatorname{Re}\nu, \operatorname{Im}\nu \ll \mu$ and that $\nu \perp \lambda$ iff $\operatorname{Re}\nu, \operatorname{Im}\nu \perp \lambda$. Hence, by decomposing λ into its real and imaginary part we may assume λ is real. Given any real measure ν, it is easy to see that $\nu \ll \mu$ iff $\nu^+, \nu^- \ll \mu$ and that $\nu \perp \mu$ iff $\nu^+, \nu^- \perp \mu$. Hence, by decomposing λ into its positive and negative part it suffices to consider only the case when λ is a positive measure.

Suppose first that μ is a finite measure. Let us define the set

$$\mathcal{D} := \{g\colon X \to [0,\infty] : \ g \text{ is measurable and } \int_A g d\mu \leq \lambda(A) \text{ for each } A \in \mathcal{A}\}.$$

We will show that $\sup D$ is the required function.

If $g_1, g_2 \in \mathcal{D}$, then $g := \max\{g_1, g_2\} \in \mathcal{D}$. To see this, define $E = \{x \in X : g_1(x) \geq g_2(x)\}$. Then for $A \in A$ we have

$$\int_A g d\mu = \int_{A\cap E} g_1 d\mu + \int_{A\cap E^c} g_2 d\mu \leq \lambda(A\cap E) + \lambda(A \cap E^c) = \lambda(A).$$

Define

$$s := \sup\left\{\int_X g d\mu : \ g \in \mathcal{D}\right\}$$

and note $s \leq \lambda(X)$. Then there exists an increasing sequence (g_n) in $\mathcal{D}$ such that $\int_X g_n d\mu \to s$. By Lebesgue's Monotone Convergence Theorem 4.2.3 we have $f \in \mathcal{D}$ and $\int_X f d\mu = s$. Define $\lambda_a := f\mu$. Then $\lambda_a \ll \mu$. Since $f \in \mathcal{D}$, the measure $\lambda_s := \lambda - f\mu$ is positive. We claim $\lambda_s \perp \mu$.

Assume $\lambda_s \not\perp \mu$. By Lemma 5.4.1 there exists $\epsilon > 0$ and $E \in \mathcal{A}$ with $\mu(E) > 0$ such that the $\lambda_s \geq \epsilon\mu$ on E. Since $\epsilon\chi_E\mu \leq \lambda_s = \lambda - f\mu$, we conclude $(f + \epsilon\chi_E)\mu \leq \lambda$. Hence, $h := f + \epsilon\chi_E \in \mathcal{D}$, so that

$$s \geq \int_X h d\mu = \int_X f d\mu + \epsilon\mu(E) \geq s.$$

This is impossible since $\mu(E) > 0$.

For the general case, assume μ is σ-finite. There exists a sequence of pairwise disjoint sets of finite measure (A_n) with $X = \bigcup_n A_n$. Define $\mu_n(A) := \mu(A_n \cap A)$ and $\lambda_n(A) := \lambda(A_n \cap A)$. Then μ_n and λ_n are positive measures with μ_n finite for each $n \in \mathbb{N}$. There exist measures $\lambda_{n,a}$ and $\lambda_{n,s}$ such that $\lambda_{n,a} \ll \mu_n$, $\lambda_{n,s} \perp \mu_n$ and $\lambda_{n,a} + \lambda_{n,s} = \lambda_n$. The required measures are

$$\lambda_a := \sum_{n=1}^{\infty} \lambda_{n,a} \qquad \text{and} \qquad \lambda_s = \sum_{n=1}^{\infty} \lambda_{n,s}.$$

□

6 L^p-spaces

6.1 Convex Functions and Inequalities

A function $\varphi\colon (a,b) \to \mathbb{R}$ is convex whenever for all $x, y \in (a,b) \in (a,b)$ and any $\lambda \in [0,1]$ we have

$$\varphi((1-\lambda)x + \lambda y) \leq (1-\lambda)\varphi(x) + \lambda\varphi(y).$$

Here $-\infty \leq a < b \leq \infty$. If $a < s < t < u < b$, then by writing $t = (1-\lambda)s + \lambda u$ for appropriate $\lambda \in (0,1)$ and rewriting convexity condition one gets

$$\frac{\varphi(t) - \varphi(s)}{t-s} \leq \frac{\varphi(u) - \varphi(t)}{u-t}.$$

6.1.1 Jensen's Inequality *Let μ be a probability measure on $(X, \mathcal{A})$. If $\varphi\colon (a,b) \to \mathbb{R}$ is a convex function and if the range of $f \in L^1(\mu)$ is contained in (a,b), then*

$$\varphi\left(\int f d\mu\right) \leq \int (\varphi \circ f) d\mu.$$

Proof. Write $t = \int f d\mu$. Since the function g defined as $g(x) = f(x) - a$ strictly positive, its integral is strictly positive so that $t > a$. Similarly $t < b$. Define

$$\beta := \sup_{s \in (a,t)} \frac{\varphi(t) - \varphi(s)}{t-s}.$$

Since for each $s \in (a,t)$ and each $u \in (t,b)$ we have

$$\frac{\varphi(t) - \varphi(s)}{t-s} \leq \beta \leq \frac{\varphi(u) - \varphi(t)}{u-t},$$

we get

$$\varphi(s) \geq \varphi(t) + \beta(s-t)$$

for all $s \in (a,b)$. If $x \in X$, then $f(x) \in (a,b)$, so that

$$\varphi(f(x)) \geq \varphi(t) + \beta(f(x) - t).$$

Integrating the last inequality and using the fact that $\mu(X) = 1$ gives

$$\int (\varphi \circ f) d\mu \geq \varphi(t) + \beta\left(\int f d\mu - t\right) = \varphi(t).$$

□

Let $X = \{p_1, \dots, p_n\}$ be a finite set equipped by the power σ-algebra and the probability measure. Denote $\alpha_j = \mu(\{p_j\})$ and $x_j = f(p_j)$ for $j = 1, \dots, n$. In this case Jensen's Inequality 6.1.1 yields

$$\varphi\left(\sum_{j=1}^{n} \alpha_j x_j\right) \le \sum_{j=1}^{n} \alpha_j \varphi(x_j).$$

Now chose our convex function to be the exponential function, i.e., $\varphi(x) = e^x$. Denote $y_j = e^{x_j}$, so that

$$y_1^{\alpha_1} \cdot \ldots \cdot y_n^{\alpha_n} \le \alpha_1 y_1 + \cdots + \alpha_n y_n \tag{6.1}$$

holds for any positive numbers $y_1, \dots, y_n$. Take $\alpha_1 = \cdots = \alpha_n = \frac{1}{n}$ to get the classical ***Arithmetic-Geometric Mean Inequality***

$$\sqrt[n]{y_1 \cdot \ldots \cdot y_n} \le \frac{y_1 + \ldots + y_n}{n}.$$

Finally, in (6.1) take $n = 2$, $y_1 = x^p$, $y_2 = y^q$, $\alpha_1 = \frac{1}{p}$ and $\alpha_2 = \frac{1}{q}$. Then we get the ***Young Inequality***

$$xy \le \frac{x^p}{p} + \frac{y^q}{q}.$$

The following lemma can be proved by standard methods of calculus.

6.1.2 Lemma *In Young's inequality the equality occurs iff* $x^p = y^q$.

6.2 L^p-spaces

Let $(X, \mathcal{A}, \mu)$ be a space with a positive measure. For each $p \in [1, \infty)$ let $L^p(\mu)$ be the set of all measurable functions $f \colon X \to \mathbb{C}$ such that $|f|^p \in L^1(\mu)$. For $f \in L^p(\mu)$ we define

$$\|f\|_p := \left(\int |f|^p d\mu\right)^{\frac{1}{p}}.$$

If $f \in L^p(\mu)$, then $\lambda f \in L^p(\mu)$ for any $\lambda \in \mathbb{F}$ and $\|\lambda f\|_p = |\lambda|\,\|f\|_p$. Suppose that f and g are in $L^p(\mu)$. Since the function $t \mapsto t^p$ is convex on $[0, \infty)$ for $p \in (1, \infty)$, for $x_1, x_2 \ge 0$ we have

$$\left(\frac{x_1 + x_2}{2}\right)^p \le \frac{x_1^p}{2} + \frac{x_2^p}{2}.$$

This yields that

$$\|f + g\|_p^p = \int_X |f + g|^p \, d\mu \le 2^{p-1} \int_X \left(|f|^p + |g|^p\right) \, d\mu < \infty,$$

so that $f+g \in \mathcal{L}^p(\mu)$ as well. This proves that $\mathcal{L}^p(\mu)$ is a vector space. The quotient vector space $L^p(\mu) := \mathcal{L}^p(\mu)/_\sim$ where $\sim$ is the equivalence relation which identifies almost everywhere equal functions from $\mathcal{L}^p(\mu)$ is again a vector space. Our goal is to prove that $\|\cdot\|_p$ is a complete norm on $L^p(\mu)$, i.e., $(L^p(\mu), \|\cdot\|_p)$ is a Banach space. Once we prove that $\|\cdot\|_p$ is a norm on $L^p(\mu)$, completeness will follow from similar arguments as in the proof of completeness of L^1-spaces. The only difficult thing that remains to be proved is the triangle inequality for $\|\cdot\|_p$.

If a pair of numbers $p, q \in (1,\infty)$ satisfies $\frac{1}{p}+\frac{1}{q}=1$, then p and q are called ***conjugated exponents***.

6.2.1 Hölder's inequality *Let $p, q \in (1,\infty)$ be conjugated exponents. Then*

$$\left|\int fg d\mu\right| \leq \left(\int |f|^p d\mu\right)^{\frac{1}{p}} \left(\int |g|^q d\mu\right)^{\frac{1}{q}}$$

for all $f \in L^p(\mu)$ and $g \in L^q(\mu)$.

Hölder's inequality yields that whenever $f \in L^p(\mu)$ and $g \in L^q(\mu)$, then $fg \in L^1(\mu)$ and $\|fg\|_1 \leq \|f\|_p \|g\|_q$. When $p=q=2$ we get the Cauchy-Schwarz inequality for the Hilbert space $L^2(\mu)$.

Proof. Assume first f and g are nonnegative. Denote

$$A = \left(\int f^p d\mu\right)^{\frac{1}{p}} \qquad \text{and} \qquad B = \left(\int g^q d\mu\right)^{\frac{1}{q}}.$$

If $A=0$ or $B=0$, then $f=0$ or $g=0$ almost everywhere and the inequality holds. Assume $A, B > 0$. If $A=\infty$ or $B=\infty$ then again the inequality holds. Assume now $A, B \in (0,\infty)$. Define functions $F=\frac{f}{A}$ and $G=\frac{g}{B}$. By Young's inequality we have

$$0 \leq FG \leq \frac{F^p}{p}+\frac{G^q}{q}.$$

Since $\int F^p d\mu = \int G^q d\mu = 1$, by integration we obtain

$$\int (FG) d\mu \leq \frac{1}{p}+\frac{1}{q}=1.$$

This yields $\int fg d\mu \leq AB$.

In general, $|f| \in L^p(\mu)$ and $|g| \in L^q(\mu)$ yields $|fg| \in L^1(\mu)$. Now apply Proposition 4.3.3. □

Suppose $f, g \neq 0$ almost everywhere. An examination of the proof of Hölder inequality yields that we have $\|fg\|_1 = \|f\|_p\|g\|_q$ iff the equality

$$FG = \frac{F^p}{p}+\frac{G^q}{q}$$

holds almost everywhere iff $F^p = G^q$ holds almost everywhere. Finally, $\|fg\|_1 = \|f\|_p\|g\|_q$ iff $f^p = \alpha g^q$ holds almost everywhere where $\alpha = \frac{\|f\|_p^p}{\|g\|_q^q}$. If $f = 0$ or $g = 0$ almost everywhere, then the equality obviously holds.

6.2.2 Lemma *Let $p, q \in (1, \infty)$ be conjugate exponents. For each $f \in L^p(\mu)$ we have*

$$\|f\|_p = \sup\left\{\left|\int fgd\mu\right| : \ g \in L^q(\mu), \ \|g\|_q = 1\right\}.$$

Proof. If $f = 0$ almost everywhere, then the equality in the lemma is obvious. Denote by s the supremum from the lemma. By Hölder's inequality we immediately conclude $s \leq \|f\|_p$. To prove the converse inequality we will actually find $g \in L^q(\mu)$ with $\|g\|_q = 1$ such that $\int fgd\mu = \|f\|_p$. Define a function g on X as

$$g(x) = \begin{cases} \frac{|f(x)|^{p-1}}{\|f\|_p^{p-1}} \cdot \frac{\overline{f(x)}}{|f(x)|} & : \ f(x) \neq 0 \\ 0 & : \ f(x) = 0 \end{cases}.$$

It is easy to see that $\|g\|_q = 1$ and $\int fgd\mu = \|f\|_p$. □

6.2.3 Minkowski's inequality *If $p \geq 1$ and $f, g \in L^p(\mu)$, then*

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p.$$

Proof. If $p = 1$, then this is the triangle inequality for L^1-functions. Assume $p > 1$. Take $h \in L^q(\mu)$ with $\|h\|_q \leq 1$. By Lemma 6.2.2 we have $\left|\int fhd\mu\right| \leq \|f\|_p$ and $\left|\int ghd\mu\right| \leq \|g\|_p$ from where it follows

$$\left|\int (f+g)hd\mu\right| \leq \|f\|_p + \|g\|_p.$$

Apply Lemma 6.2.2 once more. □

Now we introduce the class of essentially bounded functions. A function f is ***essentially bounded*** whenever there exists $M \geq 0$ such that $|f| \leq M$ for almost every $x \in X$. We define

$$\|f\|_\infty := \inf\{M \geq 0 : \ |f(x)| \leq M \text{ for almost every } x \in X\}.$$

It is easy to see that $|f| \leq \|f\|_\infty$ almost everywhere. The number $\|f\|_\infty$ is called the ***essential supremum*** of $|f|$ and is sometimes written as

$$\|f\|_\infty := \operatorname{ess\,sup}_{x \in X} |f(x)|.$$

The space of all essentially bounded functions is denoted by $L^\infty(\mu)$.

6.2.4 Lemma *If $f \in L^\infty(\mu)$, then $|f(x)| \leq \|f\|_\infty$ for almost every $x \in X$.*

Proof. The set

$$A_n := \{x \in X : \ |f(x)| > \|f\|_\infty + \tfrac{1}{n}\}$$

has zero measure. Define $A := \bigcap_n A_n^c$. Then $\mu(A^c) = 0$ and if $x \in A$ we have $|f(x)| \leq \|f\|_\infty + \frac{1}{n}$ for each $n \in \mathbb{N}$. This yields $|f(x)| \leq \|f\|_\infty$ for almost every $x \in X$. □

Hölder's inequality can be extended to functions $f \in L^1(\mu)$ and $g \in L^\infty(\mu)$:

$$\|fg\|_1 \leq \|g\|_\infty \|f\|_1.$$

6.2.5 Proposition *Let (f_n) be a sequence in $L^\infty(\mu)$.*

(i) *If (f_n) is Cauchy in $L^\infty(\mu)$, then there exists $A \in \mathcal{A}$ with $\mu(A^c) = 0$ and (f_n) is uniformly Cauchy on A.*

(ii) *If $f_n \to f$ in $L^\infty(\mu)$, then there exists $A \in \mathcal{A}$ with $\mu(A^c) = 0$ and $f_n \to f$ uniformly on A.*

Proof. (i) For $m, n \in \mathbb{N}$ define

$$A_{n,m} = \{x \in X : \ |f_n(x) - f_m(x)| > \|f_n - f_m\|_\infty\}.$$

Then $\mu(A_{n,m}) = 0$ for all $m, n \in \mathbb{N}$. Define $B := \bigcup_{m,n=1}^\infty A_{m,n}$. Then $\mu(B) = 0$ and $|f_n - f_m| \leq \|f_n - f_m\|_\infty$ on $A := B^c$.

The proof of (ii) is similar to the proof of (i) and is left for the reader. □

Again we introduce the relation $\sim$ on $L^\infty(\mu)$ such that $f \sim g$ iff $f = g$ almost everywhere. This is an equivalence relation. The quotient space $L^\infty(\mu)/_\sim$ is again denoted by $L^\infty(\mu)$. On $L^\infty(\mu)$ we define $\|f\|_\infty := \operatorname{ess\,sup}_{x\in X} |f(x)|$. The reader should check that $\|f\|_\infty$ is well defined, i.e., if $f \sim g$, then $\|f\|_\infty = \|g\|_\infty$.

6.2.6 Theorem *For any $p \in [1, \infty]$ the space $L^p(\mu)$ is a Banach space.*

Proof. The case $1 < p < \infty$ is very similar to $p = 1$ and so we omit it.

Assume $p = \infty$. It is obvious that $\|f\|_\infty \geq 0$ and since $|f(x)| \leq \|f\|_\infty$ for almost every $x \in X$ we conclude $f = 0$ almost everywhere, i.e., $f = 0$ in $L^\infty(\mu)$. Since $|f(x) + g(x)| \leq \|f\|_\infty + \|g\|_\infty$ for almost every $x \in X$, we have $\|f + g\|_\infty \leq \|f\|_\infty + \|g\|_\infty$.

Pick any $\lambda \in \mathbb{C}$. If $\lambda = 0$, then $\|\lambda f\|_\infty = |\lambda|\|f\|_\infty = 0$. If $\lambda \neq 0$, then $|\lambda f(x)| \leq |\lambda|\|f\|_\infty$ for almost every $x \in X$ implies $\|\lambda f\|_\infty \leq |\lambda|\|f\|_\infty$. The converse inequality follows from $\|f\|_\infty \leq \frac{1}{|\lambda|}\|\lambda f\|_\infty$.

Now we prove completeness of $L^\infty(\mu)$. Let (f_n) be a Cauchy sequence in $L^\infty(\mu)$. Then there exists $A \in \mathcal{A}$ with $\mu(X \setminus A) = 0$ such that (f_n) is uniformly Cauchy on A. Since by redifining functions f_n on $X \setminus A$ by zero we do not affect the corresponding equivalence classes we may assume from the start that (f_n) is uniformly Cauchy on X. This (f_n) converges uniformly to a measurable function f. Uniform limit of bounded functions is bounded. □

By Corollary 3.4.2 we already know that step functions are dense in $L^\infty(\mu)$. However, the same is true in L^p-spaces.

Let s be a step function. Then s is of the form $s = \alpha_1\chi_{A_i} + \cdots + \alpha_n\chi_{A_n}$. Since $\int |s|^p d\mu = |\alpha_1|^p\mu(A_1) + \cdots + |\alpha_n|^p\mu(A_n)$, we conclude $s \in L^p(\mu)$ iff the set $\{x \in X : s(x) \neq 0\}$ has finite measure.

6.2.7 Corollary *For $1 \leq p \leq \infty$ the set of all step functions in $L^p(\mu)$ is dense in $L^p(\mu)$.*

Proof. Suppose $p < \infty$. If $f \in L^p(\mu)$ is nonnegative, then there is an increasing sequence (s_n) of step functions such that $s_n \to f$ pointwise. Since $0 \leq (f - s_n)^p \leq f^p$, we can apply Lebesgue's Dominated Convergence Theorem 4.3.4 to see that $s_n \to f$ in L^p. The general case follows easily from the approximation of complex functions with step functions. ☐

Let μ be the counting measure on measurable space $(\mathbb{N}, \mathbb{P}(\mathbb{N}))$. If $1 \leq p < \infty$, then $f \in L^p(\mu)$ iff

$$\sum_{n=1}^{\infty} |f(n)|^p < \infty.$$

If $f \in L^p(\mu)$, then $\|f\|_p = (\sum_{n=1}^{\infty} |f(n)|^p)^{\frac{1}{p}} < \infty$. In this special case we get the sequence space ℓ^p.

When $p = \infty$, we have $f \in L^\infty(\mu)$ iff $\sup_{n\in\mathbb{N}} |f(n)| < \infty$. If $f \in L^\infty(\mu)$, we get $\|f\|_\infty = \sup_{n\in\mathbb{N}} |f(n)|$. In this special case we get the sequence space ℓ^∞.

6.2.8 Corollary *For $1 \leq p \leq \infty$ the space ℓ^p is a Banach space.*

6.3 Duals of L^p-spaces

A linear functional $f\colon V \to \mathbb{F}$ on a normed space V is is said to be ***bounded*** if

$$\sup\{|f(x)| : \|x\| \leq 1\} < \infty.$$

The set of all bounded linear functionals on V is denoted by V^* and it is called the ***dual space*** of V. Given a bounded linear functional $f\colon V \to \mathbb{F}$, the number

$$\|f\| := \sup\{|f(x)| : \|x\| \leq 1\}$$

is called the ***operator norm*** of the functional f. It is easy to see that it satisfies

$$\|f\| = \sup\{|f(x)| : \|x\| = 1\} = \sup_{x\neq 0} \frac{|f(x)|}{\|x\|}.$$

6.3.1 Proposition *Let V be a normed space and let $f\colon V \to \mathbb{F}$ be a linear functional.*

(i) *V^* equipped with the operator norm is a Banach space.*

(ii) *f is bounded iff f is (uniformly) continuous.*

Suppose now that $p \in [1, \infty]$ and let q be the conjugated exponent of p. For $f \in L^p(\mu)$ we define the linear functional ρ_f on $L^q(\mu)$ by

$$\rho_f(g) = \int fg d\mu.$$

By Lemma 6.2.2 we have that $\|\rho_f\| = \|f\|_p$ whenever $p \in (1, \infty)$.

6.3.2 Proposition *Let p, q are conjugate exponents. Then for each $f \in L^p(\mu)$ the following statements hold.*

(i) *If $p \in [1, \infty)$, then $\|\rho_f\| = \|f\|_p$.*

(ii) *If $p = \infty$, then $\|\rho_f\| \leq \|f\|_\infty$.*

(iii) *If $p = \infty$ and μ is σ-finite, then $\|\rho_f\| = \|f\|_\infty$.*

Proof. (i) For $1 < p < \infty$ the conclusion follows from Lemma 6.2.2. Suppose that $p = 1$ and take any function $g \in L^1(\mu)$. Since

$$\left| \int_X fg \, d\mu \right| \leq \int_X |fg| \, d\mu \leq \|g\|_\infty \int_X |f| \, d\mu = \|g\|_\infty \, \|f\|_1,$$

we have that $\|\rho_f\| \leq \|f\|_1$. We also need to prove the converse inequality.

If $f = 0$, then $\rho_f = 0$, so that we have the equality $\|\rho_f\| = \|f\|_1$. Assume now that $f \neq 0$ in $L^1(\mu)$ and take

$$g(x) = \begin{cases} \frac{\overline{f(x)}}{|f(x)|} & : \ f(x) \neq 0 \\ 0 & : \ f(x) = 0 \end{cases}.$$

It should be obvious that $\|g\|_\infty = 1$ and that $\int_X fg \, d\mu = \|f\|_1$. This implies that $\|\rho_f\| \geq \|f\|_1$, so that $\|\rho_f\| = \|f\|_1$. (ii) Check the proof of (i).

(iii) Assume now $p = \infty$ ($q = 1$) and μ is σ-finite. To prove that $\|\rho_f\| = \|f\|_\infty$, pick $\epsilon > 0$ and define

$$F = \{x \in X : \ |f(x)| \geq \|f\|_\infty - \epsilon\}.$$

Then $\mu(F) > 0$. Since μ is σ-finite, there exists $A \subseteq F$ with $0 < \mu(A) < \infty$. Define

$$g(x) = \begin{cases} \mu(A)^{-1} \chi_A(x) \frac{\overline{f(x)}}{|f(x)|} & : \ f(x) \neq 0 \\ 0 & : \ f(x) = 0 \end{cases}.$$

Then $\|g\|_1 = 1$ and so

$$\|\rho_f\| \geq |\rho_f(g)| = \mu(A)^{-1} \int_A |f| d\mu \geq \mu(A)^{-1} \int_A (\|f\|_\infty - \epsilon) d\mu = \|f\|_\infty - \epsilon.$$

Since $\epsilon > 0$ was arbitrary, we obtain that $\|\rho_f\| \geq \|f\|_\infty$. An application of (ii) yields (iii). □

This proves that the mapping $f \mapsto \rho_f$ is an isometry from $L^q(\mu)$ into $L^p(\mu)^*$ We will prove that this mapping is surjective whenever $p \in [1, \infty)$. Suppose $f \in L^p(\mu)$ and $g \in L^q(\mu)$ where $p, q \in (1, \infty)$ are conjugated exponents, i.e., $\frac{1}{p} + \frac{1}{q} = 1$. Hölder's inequality can be read as

$$\|fg\|_1 \leq \|f\|_p \|g\|_q.$$

Also, if $f, g \in L^p(\mu)$, then Minkowski's inequality can be read as $\|f + g\|_p \leq \|f\|_p + \|g\|_q$.

6.3.3 Theorem *Let μ be a σ-finite measure and $p \in [1, \infty)$. Then for each functional $\rho \in L^p(\mu)^*$ there exists a unique function $g \in L^q(\mu)$ such that*

$$\rho(f) = \int_X fg d\mu$$

and $\|\rho\| = \|g\|_q$.

Proof. Assume first μ is a finite measure. Then for each $A \in \mathcal{A}$ we have $\chi_A \in L^p(\mu)$. Define

$$\lambda(A) = \rho(\chi_A).$$

We claim that λ is a complex measure. To see this, take a sequence (A_n) of pairwise disjoint measurable sets and denote their union by A. Then $\chi_A = \sum\limits_{j=1}^{\infty} \chi_{A_j}$ the series $\sum_{j=1}^{\infty} \chi_{A_j}$ converges in $L^p(\mu)$ to χ_A since

$$\|\chi_A - \sum_{j=1}^{n} \chi_{A_j}\|_p^p = \sum_{j=n+1}^{\infty} \int_X \chi_{A_j}^p d\mu = \sum_{j=n+1}^{\infty} \mu(A_j) \to 0.$$

Since ρ is continuous, we have

$$\lambda(A) = \rho(\chi_A) = \sum_{j=1}^{\infty} \rho(\chi_{A_j}) = \sum_{j=1}^{\infty} \lambda(A_j).$$

If $\mu(A) = 0$, then $\chi_A = 0$ in $L^p(\mu)$, so that $\lambda(A) = \rho(\chi_A) = 0$. Hence $\lambda \ll \mu$. By the Radon-Nikodým Theorem 5.4.3 there exists a function $g \in L^1(\mu)$ such that

$$\rho(\chi_A) = \lambda(A) = \int_A g d\mu = \int_X g\chi_A d\mu$$

for each measurable subset A of X. We claim $g \in L^q(\mu)$. For each $n \in \mathbb{N}$ define $X_n = \{x \in X : |g(x)| \leq n\}$. Since μ is finite, the function $g_n := g\chi_{X_n} \in L^q(\mu)$ for every n. The function g_n defines a bounded functional ρ_n on $L^p(\mu)$. If $A \subseteq X_n$ is a measurable subset, then

$$\rho(\chi_A) = \rho(\chi_A \chi_{X_n}) = \int_A g d\mu = \rho_n(\chi_A).$$

This implies that bounded functionals ρ and ρ_n agree on the set of all step functions in $L^p(\mu|_{X_n})$ and since step functions are dense in L^p-spaces, we conclude that ρ and ρ_n agree on $L^p(\mu|_{X_n}) \equiv L^p(\mu)\chi_{X_n}$. This yields

$$\|\rho_n|_{L^p(\mu|_{X_n})}\| = \|g_n|_{X_n}\|_q = \|\rho|_{L^p(\mu|_{X_n})}\|_q \leq \|\rho\|.$$

This means

$$\int_X |g|^q \chi_{X_n}\, d\mu = \|g_n\|_q^q \leq \|\rho\|^q$$

so that by Lebesgue's Monotone Convergence Theorem 4.2.3 we conclude $\|g\|_q \leq \|\rho\|$. This proves that $g \in L^q(\mu)$ and that the f defines a bounded linear functional ρ_g which satisfies $\rho_g(\chi_A) = \rho(\chi_A)$ for all measurable subsets A of X. Then ρ_g and ρ agree on all step functions and since they are dense in $L^p(\mu)$, we have $\rho = \rho_g$.

Assume now that μ is σ-finite. Then there exists an increasing sequence (X_n) of measurable subsets of finite measure whose union is X. Since $L^p(\mu|_{X_n})$ and $L^p(X)\chi_{X_n}$ are isometrically isomorphic as Banach spaces, we will consider $L^p(\mu|_{X_n})$ as a subspace of $L^p(\mu)$. By the case above for each $n \in \mathbb{N}$ there exists $g_n \in L^q(\mu|_{X_n})$ such that

$$\rho(f) = \int_{X_n} f g_n d\mu$$

for each function $f \in L^p(\mu|_{X_n})$ and $\int_{X_n} |g_n|^q d\mu = \|g_n\|_q^q \leq \|\rho\|$. Since the functions g_n are uniquely defined almost everywhere, we have $g_{n+1}|_{X_n} = g_n$ almost everywhere. There exists a function $g\colon X \to \mathbb{C}$ such that $g|_{X_n} = g_n$ almost everywhere. By using Lebesgue Monotone Convergence Theorem once again we get

$$\int_X |g|^q d\mu = \lim_{n\to\infty} \int_X |g_n|^q d\mu \leq \|\rho\|^q.$$

Take any $f \in L^p(\mu)$. Then for each $n \in \mathbb{N}$ we have $\rho(f\chi_{X_n}) = \int_{X_n} f g_n d\mu$. Since $f\chi_{X_n} \to f$ in $L^p(\mu)$ and ρ is continuous, the left-hand side converges to $\rho(f)$. Since the sequence $\chi_{X_n} f g_n \to fg$ is dominated by $|fg|$ and since it converges almost everywhere to fg, by the Lebesgue Dominated Convergence Theorem the right-hand side converges to $\int_X fg d\mu$. □

7 Measures on Locally Compact Spaces

Throughout this chapter X is assumed to be a locally compact Hausdorff space.

7.1 Continuous Functions with Compact Support

Let X be a topological space. A set A which contains a point $x \in X$ is called a ***neighborhood*** of x whenever there exists an open set U such that $x \in U \subseteq A$. Let X and Y be topological spaces. A mapping $f\colon X \to Y$ is said to be ***continuous*** at a point $x \in X$ whenever for each neighborhood U of a point $f(x) \in Y$ there is a neighborhood V of a point $x \in X$ such that $f(V) \subseteq U$. If f is continuous at each point $x \in X$, then we say f is continuous on X. Continuity is equivalent to the following:

- $f^{-1}(U)$ is an open set in X for every open set U in Y.

If f is continuous at every point $x \in X$, then f is ***continuous*** on X.

A topological space X is called ***Hausdorff*** if for each pair of distinct points x, y in X there exist disjoint neighborhoods U and V for x and y, respectively. We say that X is compact whenever every open cover of X contains a finite subcover. A topological space X is ***locally compact*** whenever every point $x \in X$ has a compact neighborhood. If X is Hausdorff, then any point has a fundamental system consisting of compact neighborhoods.

7.1.1 Lemma *Suppose U is open and $K \subseteq U$ is compact. Then there exists an open set V with the compact closure $\overline{V}$ such that $K \subseteq V \subseteq \overline{V} \subseteq U$.*

Given a function f on X its ***support*** $\operatorname{supp} f$ is denoted as the closure of the set $\{x \in X : f(x) \neq 0\}$. The set of all continuous functions with compact support is denoted by $C_c(X)$. Pick any continuous function $f\colon X \to [0,1]$. Given a compact set K we denote $K \preceq f$ whenever $f \equiv 1$ on K. Given an open set U we denote $f \prec U$ whenever $\operatorname{supp} f \subseteq U$. The notation $K \preceq f \prec U$ means $K \preceq f$ and $f \prec U$.

7.1.2 Theorem *Let K be a compact subset of a locally compact Hausdorff space and $U \supseteq K$ an open set. Then any continuous function $f\colon K \to [0,1]$ can be extended to a function $F \in C_c(X)$ such that* $\operatorname{supp} F \subseteq U$.

Proof. First find a relatively compact open set V such that $K \subseteq V \subseteq \overline{V} \subseteq U$. Let X^+ be the one-point compactification of X. Since K is compact in X, it is also compact in X^+ and hence closed. Since V is open in X, it is also open in X^+, so that $X^+ \setminus V$ is closed in X^+. The function $g\colon K \cup (X^+ \setminus V) \to [0,1]$ defined as $g|_K \equiv f$ and $f|_{X^+\setminus V}$ is continuous. Since X^+ is normal (follows from the fact that X^+ is compact and Hausdorff) by Tietze's extension theorem there exists a continuous function $G\colon X^+ \to [0,1]$ such that $G|_K \equiv f$ and $G|_{X^+\setminus V} \equiv 0$. Denote by F the restriction of G to X. Then $F\colon X \to [0,1]$ is continuous, $F \equiv f$ on K and $\{x \in X : F(x) \neq 0\} \subseteq V$. Since $\operatorname{supp} F \subseteq \overline{V} \subseteq U$, we conclude $F \in C_c(X)$ ☐

7.1.3 Urysohn's lemma *Let X be a locally compact Hausdorff space and $K \subseteq U$ a compact and an open subset, respectively. Then there exists $f \in C_c(X)$ such that $K \preceq f \prec U$.*

Proof. Take $f \equiv 1$ on K and apply Theorem 7.1.2 to obtain a function $F\colon X \to [0,1]$ with compact support such that $F|_K \equiv 1$ and $\operatorname{supp} F \subseteq U$. ☐

7.2 Positive Linear Functionals on $C_c(X)$

A linear functional φ on a vector space $C_c(X)$ is said to be ***positive*** whenever $\varphi(f) \geq 0$ for every nonnegative function $f \in C_c(X)$. It is easy to see that φ is nonnegative iff $\varphi(f) \geq \varphi(g)$ whenever $f \geq g$. Recall that a measure μ on X is Borel if it is defined on Borel subsets of X.

7.2.1 Example Let μ be a Borel measure on X which is finite on all compact subsets. Then for any $f \in C_c(X)$, the function f is zero outside $\operatorname{supp} f$, and since $\operatorname{supp} f$ is compact, f is bounded. This yields that

$$\int_X |f| d\mu = \int_{\operatorname{supp} f} |f| d\mu < \infty.$$

Hence, $f \in L^1(\mu)$ and so the mapping

$$f \mapsto \int_X f d\mu$$

is a well-defined positive linear functional on $C_c(X)$.

The Riesz representation theorem will say that every positive functional on $C_c(X)$ is actually defined as the integral over some sufficiently nice measure. That the measure is sufficiently nice means that it needs to satisfy some additional restrictions.

A positive Borel measure μ is ***outer regular*** on a Borel set A whenever

$$\mu(A) = \inf\{\mu(U) : U \text{ is open and } A \subseteq U\}.$$

Similarly we say that a positive Borel measure μ is ***inner regular*** on a Borel set A whenever

$$\mu(A) = \sup\{\mu(K) : \ K \text{ is compact and } K \subseteq A\}.$$

Radon measure is a Borel measure on X which is finite on all compact sets, inner regular on all open sets and outer regular on all Borel sets.

7.2.2 Lemma *Radon measure is determined by its values on compact sets.*

Proof. Take a Borel set A in X. Since μ is outer regular, μ is uniquely determined by its values on open sets. Since μ is inner regular on all open sets, μ is uniquely determined by its values on compact sets. □

7.2.3 Lemma *Let μ be a Radon measure on X and let φ be the positive functional on $C_c(X)$ induced by μ. Then*

$$\mu(U) = \sup\{\varphi(f) : \ f \prec U\}. \tag{7.1}$$

Furthermore, μ is uniquely determined by φ.

Proof. Denote by s the supremum from lemma. Suppose $f \prec U$. Then $f \leq \chi_U$, so that

$$\varphi(f) = \int_X f d\mu \leq \int_X \chi_U d\mu = \mu(U).$$

This proves $s \leq \mu(U)$. For the converse, pick any compact set $K \subseteq U$ and find a function $f \in C_c(X)$ with $K \preceq f \prec U$. Then $\chi_K \leq f \leq \chi_U$ as above shows $\mu(K) \leq \varphi(f) \leq \mu(U)$. Since μ is Radon, we have $s = \mu(U)$. Outer regularity of μ implies that values of φ uniquely determine measures of Borel sets. □

7.2.4 Lemma *Suppose φ is a positive functional on $C_c(X)$ and suppose μ is an outer regular Borel measure which satisfies* (7.1) *for all open sets. Then*

$$\mu(K) = \inf\{\varphi(f) : \ K \preceq f\} \tag{7.2}$$

holds for any compact set K. In particular $\mu(K) < \infty$ for any compact set K.

Proof. Pick $\epsilon > 0$ and $f \in C_c(X)$ such that $K \preceq f$. Define $U := \{x \in X : f(x) > 1 - \epsilon\}$. Then U is open and if $g \prec U$ then we have $g \leq \frac{f}{1-\epsilon}$. This yields $\varphi(g) \leq \frac{\varphi(f)}{1-\epsilon}$. By the assumption we conclude $\mu(U) \leq \frac{\varphi(f)}{1-\epsilon}$. If we let $\epsilon \to 0$ we obtain $\mu(U) \leq \varphi(f)$. Since $K \subseteq U$ we have $\mu(K) \leq \varphi(f)$ and $\mu(K) < \infty$.

To prove $\mu(K) = \inf\{\varphi(f) : \ K \preceq f\}$, pick $\epsilon > 0$ and find an open set U which contains K such that $\mu(K) > \mu(U) - \epsilon$. Now find $f \in C_c(X)$ with $K \preceq f \prec U$. Then

$$\mu(K) > \mu(U) - \epsilon \geq \varphi(f) - \epsilon$$

gives the desired conclusion. □

7.2.5 Riesz representation theorem *For each positive functional φ on $C_c(X)$ there exists a unique Radon measure μ on X such that*

$$\varphi(f) = \int_X f d\mu$$

for each $f \in C_c(X)$. The measure μ satisfies (7.1) *and* (7.2).

Proof. Suppose there exist two Radon measures μ_1 and μ_2 for which the assertions of the Riesz representation theorem hold. Then (7.1) assures that μ_1 and μ_2 agree on all open subsets of X. Since both μ_1 and μ_2 are outer regular, μ_1 and μ_2 agree on each Borel subset. This proves the uniqueness part.

The proof of the existence is hard and long. We will present only the main idea. We define the measure μ first on open sets. If U is open in X we define

$$\mu(U) = \sup\{\varphi(f) : \ f \preceq U\}.$$

If Y is an arbitrary Borel set in X, we define

$$\mu^*(Y) = \inf\{\mu(U) : \ Y \subseteq U \text{ is open}\}.$$

Obviously μ^* agrees with μ on open sets.

Step 1: μ^* *is an outer measure.* If one defines

$$\nu(Y) = \inf\{\sum_{j=1}^{\infty} \mu(U_j) : \ (U_j) \text{ is a countable open cover for } Y\},$$

then ν is an outer measure on X by Proposition 2.4.1. The next thing that one needs to prove is that whenever (U_j) is any countable open cover for an open set U, then

$$\mu(U) \leq \sum_{j=1}^{\infty} \mu(U_j).$$

This would then imply $\mu^* = \nu$ and so μ^* is an outer measure.

Step 2: *Every open set is μ^*-measurable.* If we prove this, then the σ-algebra $\mathcal{A}_{\mu^*}$ of all μ^*-measurable sets contains the Borel σ-algebra $\mathcal{B}(X)$ and $\mu^*|_{\mathcal{B}(X)}$ is a Borel measure which agrees on open sets with μ. This measure is outer regular on Borel sets and finite on compact sets by Lemma 7.2.4. By (7.1) it follows that $\mu^*|_{\mathcal{A}_{\mu^*}}$ is inner regular on open sets. To see this, pick an open set U which contains $\operatorname{supp} f$. Then $f \prec U$ implies $\varphi(f) \leq \mu(U)$, so that by definition of μ^* it follows $\varphi(f) \leq \mu^*(\operatorname{supp} f)$.

Step 3: $\varphi(f) = \int f d\mu$ *for every $f \in C_c(X)$.* □

7.3 Regularity of Radon Measures

A Borel measure μ is called ***regular*** whenever it is simultaneously inner and outer regular.

7.3.1 Proposition *A Radon measure is inner regular on every set on which it is σ-finite.*

Proof. We first prove that a Radon measure μ is inner regular on each subset $Y \subseteq X$ with finite measure. Pick $\epsilon > 0$. Since μ is outer regular, there exists an open set $U \supseteq Y$ such that $\mu(U) < \mu(Y) + \epsilon$. Since μ is inner regular on open sets there exists a compact set $K \subseteq U$ such that $\mu(U) - \epsilon < \mu(K)$. Since $\mu(U \setminus Y) = \mu(U) - \mu(Y) < \epsilon$, there exists an open se $V \supseteq U \setminus Y$ such that $\mu(V) < \epsilon$. The set $H := K \setminus V$ is a compact subset of Y which satisfies

$$\mu(H) = \mu(K) - \mu(K \cap V) > \mu(U) - \epsilon - \mu(V) > \mu(Y) - 2\epsilon.$$

If Y is σ-finite and $\mu(Y) = \infty$, then there exists an increasing sequence of sets Y_j of finite measure whose union in Y. By the first case, for each $j \in \mathbb{N}$ there exists a compact subset $K_j \subseteq Y_j$ such that $\mu(K_j) \geq \mu(Y_j) - 1$. Since $\mu(Y_j) \to \infty$ as $j \to \infty$, we have $\mu(K_j) \to \infty$ as $j \to \infty$. □

A topological space X is called ***σ-compact*** if it can be written as a union of countably many compact sets.

7.3.2 Corollary *Let μ be a Radon measure on X.*

(i) *If μ is σ-finite, then μ is regular.*

(ii) *If X is σ-compact, then μ is σ-finite and regular.*

Proof. Since every Radon measure is finite on compact sets, every σ-compact topological space is σ-finite. □

In particular, every Radon measure on a compact topological space is regular.

7.4 Approximation with Continuous Functions

In the Introductory course to Functional analysis, the completion of the normed space $L^2_{\text{cont}}[a,b]$ was denoted by $L^2[a,b]$. In this section we prove that the L^2-space from the Introductory course to Functional analysis is actually our space $L^2(m)$ where m denotes the restriction of the Lebesgue measure m to $[a,b]$.

7.4.1 Proposition *For any Radon measure μ on X the set $C_c(X)$ is dense in $L^p(\mu)$ for each $p \in [1, \infty)$.*

Proof. By Corollary 6.2.7 step functions of the space $L^p(\mu)$ are dense in $L^p(\mu)$. Therefore, the proof is finished once we show that every step function in $L^p(\mu)$ can be approximated by functions in $C_c(X)$. Since step functions are linear combinations of characteristic functions, it suffices to consider only the case of a function χ_A where $\mu(A) < \infty$. Pick $\epsilon > 0$. Since μ is regular on A, there exist an open and a compact set U and K, respectively, such that $K \subseteq A \subseteq U$ and $\mu(A \setminus K) < \epsilon$ and $\mu(U \setminus A) < \epsilon$. By Urysohn's lemma there exists $f \in C_c(X)$ such that $K \preceq f \prec U$. Then

$$\|\chi_A - f\|_p^p = \int_{U \setminus K} |\chi_A - f|^p d\mu \leq \mu(U \setminus K) < 2\epsilon. \qquad \square$$

7.4.2 Corollary *The completion of $L^2_{cont}[a,b]$ is the space $L^2(m)$.*

We finish this section with Luzin's approximation theorem and its consequence.

7.4.3 Luzin's theorem *Let μ be a Radon measure on X and let $f\colon X \to \mathbb{C}$ be a measurable function which is nonzero on a set of finite measure. Then for each $\epsilon > 0$ there exists $g \in C_c(X)$ such that*

$$\mu(\{x \in X : \; g(x) \neq f(x)\}) < \epsilon$$

and

$$\sup_{x \in X} |g(x)| \leq \sup_{x \in X} |f(x)|.$$

Proof. Assume first $f \in L^1(\mu)$. By Proposition 7.4.1 there exists a sequence (f_n) in $C_c(X)$ which converges to f in $L^1(\mu)$. Applying 4.3.6 we may assume that $f_n \to f$ almost everwhere. Since the set $A := \{x \in X : \; f(x) \neq 0\}$ has a finite measure, by Jegorov's theorem there exists $B \subseteq A$ such that $\mu(A \setminus B) < \frac{\epsilon}{3}$ and that $f_n|_B \to f|_B$ uniformly. Find a compact set $K \subseteq B$ and an open set $U \supseteq A$ such that $\mu(B \setminus K) < \frac{\epsilon}{3}$ and $\mu(U \setminus A) < \frac{\epsilon}{3}$. Since $f_n|_B \to f|_B$ uniformly, the function $f|_B$ and so $f|_K$ is continuous. By the variant of the Tietze Extension Theorem 7.1.2 for locally compact spaces there exists $h \in C_c(X)$ such that $\operatorname{supp} h \subseteq U$ and $h|_K \equiv f|_K$. Since $\{f \neq h\} \subseteq U \setminus K$, we have

$$\mu(\{f \neq h\}) \leq \mu(U \setminus K) \leq \mu(U \setminus A) + \mu(A \setminus B) + \mu(B \setminus K) < \epsilon.$$

We will adapt the function h such that $\sup_{x \in X} |h(x)| \leq \sup_{x \in X} |f(x)|$. If $\sup_{x \in X} |f(x)|$ then there is nothing to prove. Assume $c := \sup_{x \in X} |f(x)| < \infty$ and define the function $\phi\colon \mathbb{C} \to \mathbb{C}$ by

$$\phi(z) := \begin{cases} z & : \; |z| \leq c \\ c\frac{z}{|z|} & : \; |z| > c \end{cases}.$$

We claim that $g := \phi \circ h$ is the required function. The function ϕ is continuous, so that $g \in C(X)$ Since $g(x) \neq 0$ iff $h(x) \neq 0$, we have $\operatorname{supp} g = \operatorname{supp} h$, so that $g \in C_c(X)$. If $g(x) = f(x)$, then $|g(x)| \leq c$, so that $g(x) = \phi(h(x)) = h(x) = f(x)$. Therefore, g and f agree do not agree on some subset of $U \setminus K$ whose measure is less than ϵ.

Consider now the case $f \notin L^1(\mu)$ and define the set $D_n = \{x \in X : |f(x)| > n\}$ for each $n \in \mathbb{N}$. Since (D_n) is a decreasing sequence, $\bigcap_n D_n = \emptyset$, $D_n \subseteq A$ and $\mu(A) < \infty$, we have $\mu(D_n) \to 0$. Consider the function $g_n := f\chi_{D_n^c}$. Then f and g_n agree everywhere except on D_n. Find n_0 such that $\mu(D_n) < \frac{\epsilon}{2}$, Since g_n is in $L^1(\mu)$ there exists $g \in C_c(X)$ such that $\mu(\{g \neq g_n\}) < \frac{\epsilon}{2}$ and $\sup_{x \in X} |g(x)| \leq \sup_{x \in X} |g_n(x)|$. Then $\mu(\{f \neq g\}) < \epsilon$ and $\sup_{x \in X} |g(x)| \leq \sup_{x \in X} |g_n(x)| \leq \sup_{x \in X} |f(x)|$. □