Real analysis.

Lecturer Aleksandr Rotkevich, spring 2024.

1 Integrability and continuous functions

1.1 σ-algebras and measurable functions.

Definition 1.1. *Let X be any set.*

1. *A σ-algebra on X is a family $\mathcal{A}$ of subsets of X such that the following conditions hold:*

   (a) $\varnothing \in \mathcal{A}$, $X \in \mathcal{A}$.

   (b) If $E \in \mathcal{A}$, then the complement set $X \backslash E = \{x \in X \mid x \notin E\}$ is also in $\mathcal{A}$.

   (c) If (E_n) is any countable family of subsets $Y_n \in \mathcal{A}$, then

   $$\bigcup_{n \geqslant 1} E_n \in \mathcal{A} \text{ and } \bigcap_{n \geqslant 1} E_n \in \mathcal{A}.$$

   A set $Y \in \mathcal{A}$ is said to be measurable with respect to $\mathcal{A}$. The pair $(X, \mathcal{A})$ is called a measurable space.

2. *Let $(X, \mathcal{A})$ and $(X', \mathcal{A}')$ be measurable spaces. A map $f : X \to X'$ is measurable with respect to $\mathcal{A}$ and $\mathcal{A}'$ (sometimes we write $f : (X, \mathcal{A}) \to (X, \mathcal{A})$ to indicate $\sigma-$algebras) if, for all $E \in \mathcal{A}'$, the inverse image*

   $$f^{-1}(E) = \{x \in X \mid f(x) \in E\}$$

   is in $\mathcal{A}$.

Lemma 1.2. *1. Let $(X, \mathcal{A})$ be a measurable space. The identify map*

$$id : (X, \mathcal{A}) \to (X, \mathcal{A})$$

sending X to itself is measurable.

2. Any constant map is measurable.

3. Let $f : X \to X'$ and $X'g :\to X''$ be measurable maps. The composition

$$g \circ f : X \to X''$$

is also measurable.

Proof. **1.** Notice that $id^{-1}(E) = E$ for every $E \in \mathcal{A}$.

2. Let $f : X \to X'$. If $f = c \in X'$ is constant on X then for every $E \subset X'$

$$f^{-1}(E) = \begin{cases} \varnothing, & c \notin E; \\ X, & c \in E. \end{cases}$$

In both cases $f^{-1}(E)$ is measurable.

3. Let $E \in \mathcal{A}''$. Then $g^{-1}(E) \in \mathcal{A}'$ and, hence, $(g \circ f)^{-1}(E) = f^{-1}(g^{-1}(E)) \in \mathcal{A}$ and $g \circ f$ is measurable. $\square$

Examples 1. For any X, one can take $\mathcal{A} = \mathcal{A}_{\min} = \{X, \varnothing\}$; this is the smallest possible σ-algebra on X.

2. Similarly, the largest possible σ-algebra on X is the set $\mathcal{A}_{\max} = 2^X$ of all subsets of X. Of course, any map $(X, \mathcal{A}_{\max}) \to (X', \mathcal{A}')$ is measurable. Although we will see that $\mathcal{A}_{\max}$ is not suitable for defining integration when X is a "big" set, however, it is the most usual σ-algebra used when X is either finite or countable.

3. Let $\mathcal{A}'$ be a σ-algebra on a set X', and let $f : X \to X'$ be any map; then defining

$$\mathcal{A} = f^{-1}(\mathcal{A}') = \left\{ f^{-1}(Y) \mid Y \in \mathcal{A}' \right\}$$

we obtain a σ-algebra on X, called the inverse image of $\mathcal{A}'$. Indeed, the formulas

$$f^{-1}(\varnothing) = \varnothing, \quad f^{-1}(X' \backslash Y') = X \backslash f^{-1}(Y'),$$

$$f^{-1}\left(\bigcup_{i \in I} Y_i \right) = \bigcup_{i \in I} f^{-1}(Y_i), \quad f^{-1}\left(\bigcap_{i \in I} Y_i \right) = \bigcap_{i \in I} f^{-1}(Y_i)$$

(valid for any index set I) show that $\mathcal{A}$ is a σ-algebra. (On the other hand, the direct image $f(\mathcal{A})$ is not a σ-algebra in general, since $f(X)$ might not be all of X', and this would prevent X' to lie in $f(\mathcal{A})$). This inverse image is such that

$$f : \left(X, f^{-1}(\mathcal{A}) \right) \to (X', \mathcal{A}')$$

becomes measurable.

The following special case is often used. Let $(X, \mathcal{A})$ be a measurable space and let $X' \subset X$ be any fixed subset of X (not necessarily in $\mathcal{A}$). Then we can define a σ-algebra on X' by putting

$$\mathcal{A}' = \{ Y \cap X' \mid Y \in \mathcal{A} \};$$

it is simply the inverse image σ-algebra $i^{-1}(\mathcal{A})$ associated with the inclusion map $i : X' \hookrightarrow X$. Note that if $X' \in \mathcal{A}$, the following simpler description

$$\mathcal{A}' = \{ Y \in \mathcal{A} \mid Y \subset X' \}$$

(that is not valid if $X' \notin \mathcal{A}$).

4. Let $(\mathcal{A}_i)_{i\in I}$ be σ-algebras on a fixed set X, with I an arbitrary index set. Then the intersection

$$\bigcap_i \mathcal{A}_i = \{Y \subset X \mid Y \in \mathcal{A}_i \text{ for all } i \in I\}$$

is still a σ-algebra on X.

5. Let $(X, \mathcal{A})$ be a measurable space, and $Y \subset X$ an arbitrary subset of X. Then $Y \in \mathcal{A}$ if and only if the characteristic function

$$\chi_Y : (X, \mathcal{A}) \to (\{0,1\}, \mathcal{A}_{\max})$$

defined by

$$\chi_Y(x) = \begin{cases} 1 & \text{if } x \in Y \\ 0 & \text{otherwise} \end{cases}$$

is measurable. This is clear, since $\chi_Y^{-1}(\{0\}) = X \backslash Y$ and $\chi_Y^{-1}(\{1\}) = Y$.

Definition 1.3. 1. *Let X be a set and f any family of subsets of X. The $\sigma-$ algebra generated by $\mathcal{M}$, denoted $\sigma(\mathcal{M})$, is the smallest σ-algebra containing $\mathcal{M}$, i.e., it is given by*

$$\sigma(\mathcal{M}) = \{Y \subset X \mid Y \in \mathcal{A} \textit{ for any } \sigma\textit{-algebra } \mathcal{A} \textit{ with } \mathcal{M} \subset \mathcal{A}\}$$

(in other words, it is the intersection of all σ-algebras containing $\mathcal{M}$).

2. *Let $(X, \mathcal{T})$ be a topological space. The Borel σ-algebra on X, denoted $\mathcal{B}_X$, is the σ-algebra generated by the collection $\mathcal{T}$ of open sets in X.*

3. *Let $(X, \mathcal{A})$ and $(X', \mathcal{A}')$ be measurable spaces. The product σ-algebra on $X \times X'$ is the σ-algebra denoted $\mathcal{A} \otimes \mathcal{A}'$ which is generated by all the sets of the type $Y \times Y'$ where $Y \in \mathcal{A}$ and $Y' \in \mathcal{A}'$.*

Remark 1.4. ***1.*** *If $(X, \mathcal{T})$ is a topological space, we can prove that $\mathcal{B}$ is generated either by the closed sets or the open sets (since the closed sets are the complements of the open ones, and conversely). If $X = \mathbb{R}$ with its usual topology (which is the most important case), the Borel σ-algebra contains all intervals, whether closed, open, half-closed, half-infinite, etc. Moreover, the Borel σ-algebra on $\mathbb{R}$ is in fact generated by the much smaller collection of open intervals (a, b), or by the intervals $(-\infty, a)$ where $a, b \in \mathbb{R}$.*

By convention, when X is a topological space, we consider the Borel σ-algebra on X when speaking of measurability issues, unless otherwise specified. This applies in particular to functions $(X, \mathcal{A}) \to \mathbb{R}$. To say that such a function is measurable means with respect to the Borel σ-algebra.

2. *If $(X, \mathcal{A})$ and $(X', \mathcal{A}')$ are measurable spaces, one may check also that the product σ-algebra $\mathcal{A} \otimes \mathcal{A}'$ defined above is the smallest σ-algebra on $X \times X'$ such that the projection maps*

$$p_1 : X \times X' \to X \text{ and } p_2 : X \times X' \to X'$$

are both measurable.

Indeed, for a given σ-algebra $\mathcal{N}$ on $X \times X'$, the projections are measurable if and only iff $p_2^{-1}(Y) = X \times Y' \in \mathcal{N}$ and $p_1^{-1}(Y) = Y \times X' \in \mathcal{N}$ for any $Y \in \mathcal{A}$, $Y' \in \mathcal{A}'$. Since

$$Y \times Y' = (Y \times X') \cap (X \times Y'),$$

we see that these two types of sets generate the product σ-algebra.

3. *The Borel σ-algebra on $\mathbb{C} = \mathbb{R}^2$ is the same as the product σ-algebra $\mathcal{B} \otimes \mathcal{B}$, where $\mathcal{B}$ denotes the Borel σ-algebra on $\mathbb{R}$; this is due to the fact that any open set in $\mathbb{C}$ is a countable union of sets of the type $I_1 \times I_2$ where $I_i \subset \mathbb{R}$ is a semiopen interval. Moreover, one can check that the restriction to the real line $\mathbb{R}$ of the Borel σ-algebra on $\mathbb{C}$ is simply the Borel σ-algebra on $\mathbb{R}$ (since the inclusion map $\mathbb{R} \hookrightarrow \mathbb{C}$ is continuous).*

Similarly, Borel $\sigma-$algebra on $\mathbb{R}^n$ is the same as the product σ-algebra $\mathcal{B} \otimes \mathcal{B} \otimes \ldots \otimes \mathcal{B}$.

4. *If we state that $(X, \mathcal{A}) \to \mathbb{R}$ is a measurable function with no further indication, it is implied that the target $\mathbb{R}$ is given with the Borel σ-algebra. Similarly for target space $\mathbb{C}$.*

5. *We will also consider functions with values in the extended real line*

$$f : X \to \overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}.$$

Measurability is then always considered with respect to the σ-algebra on $\overline{\mathbb{R}}$ generated by $\mathcal{B}_{\mathbb{R}}$ and the singletons $\{-\infty\}$ and $\{+\infty\}$. In other words, f is measurable if and only if

$$f^{-1}(U), \quad f^{-1}(+\infty) = \{x \mid f(x) = +\infty\}, \quad f^{-1}(-\infty) = \{x \mid f(x) = -\infty\}$$

are in $\mathcal{A}$, where U is arbitrary open subset of X.

Lemma 1.5. *1. Let $(X, \mathcal{A})$ and $(X', \mathcal{A}')$ be measurable spaces such that $\mathcal{A}' = \sigma(\mathcal{M})$ is generated by a collection of subsets $\mathcal{M}$. A map $f : X \to X'$ is measurable if and only if $f^{-1}(\mathcal{M}) \subset \mathcal{A}$.*

2. In particular, for any measurable space $(X, \mathcal{A})$, a function f :

$X \to \mathbb{R}$ is measurable if and only if Lebesgue set

$$f^{-1}((-\infty, a)) = \{x \in X \mid f(x) < a\}$$

is measurable for all $a \in \mathbb{R}$.

3. *Let $(X, \mathcal{A})$, $(X', \mathcal{A}')$ and $(X'', \mathcal{A}'')$ be measurable spaces. A map $f : X'' \to X \times X'$ is measurable for the product σ-algebra on $X \times X'$ if and only if $p_1 \circ f : X'' \to X$ and $p_2 \circ f : X'' \to X'$ are measurable.*

Proof. Part 2 is a special case of the first.

This first point follows from the formula

$$f^{-1}(\sigma(\mathcal{M})) = \sigma\left(f^{-1}(\mathcal{M})\right), \tag{1}$$

since we deduce then that, using the assumption $f^{-1}(\mathcal{M}) \subset \mathcal{A}$, that

$$f^{-1}(\mathcal{A}') = f^{-1}(\sigma(\mathcal{M})) = \sigma\left(f^{-1}(\mathcal{M})\right) \subset \sigma(\mathcal{A}) = \mathcal{A},$$

which is exactly what it means for f to be measurable.

To prove (1) notice that $f^{-1}(\sigma(\mathcal{M}))$ is a $\sigma-$algebra and that $f^{-1}(\mathcal{M}) \subset f^{-1}(\sigma(\mathcal{M}))$. Then

$$\sigma(f^{-1}(\mathcal{M})) \subset \sigma(f^{-1}(\sigma(\mathcal{M}))) = f^{-1}(\sigma(\mathcal{M})).$$

Conversely, notice that

$$\mathcal{A}'' = \left\{Y \mid f^{-1}(Y) \in \sigma\left(f^{-1}(\mathcal{M})\right)\right\}$$

is a σ-algebra on X' , which contains $\mathcal{M}$, and therefore also $\sigma(\mathcal{M})$. Consequently, we get

$$f^{-1}(\sigma(\mathcal{M})) \subset f^{-1}(\mathcal{A}'') \subset \sigma\left(f^{-1}(\mathcal{M})\right),$$

as desired.

As for the 3rd assertion, compositions $p_1 \circ f$ and $p_2 \circ f$ are measurable if f is. Conversely, to check that f is measurable, it suffices by 1 to check that $f^{-1}(Y \times Y') \in \mathcal{A}''$ for any $Y \in \mathcal{A}, Y' \in \mathcal{A}'$. But since $f(x) = (p_1 \circ f(x), p_2 \circ f(x))$, we get

$$f^{-1}(Y \times Y') = (p_1 \circ f)^{-1}(Y) \cap (p_2 \circ f)^{-1}(Y'),$$

which proves the result. □

Corollary 1.5.1.

- *Let $f : X \to X'$ be a continuous map between topological spaces. Then f is measurable with respect to the Borel σ-algebras on X and X'.*

- *Let $(X, \mathcal{A})$ be a measurable space, and let $f, g : X \to \mathbb{C}$ be measurable maps. Then $f \pm g$ and fg are measurable, and if $g(x) \neq 0$ for $x \in X$, the inverse $1/g$ is measurable. In particular, the set of complex-valued measurable functions on X is a vector space with operations given by addition of functions and multiplication by constants.*

- *A function $f : X \to \mathbb{C}$ is measurable for the Borel σ-algebra on $\mathbb{C}$ if and only if $\mathrm{Re}(f)$ are $\mathrm{Im}(f)$ are measurable as functions $X \to \mathbb{R}$.*

Proof. Assertion 1 is immediate from Lemma, since continuity means that $f^{-1}(U)$ is open for any open set in X'.

To prove second assertion, we write, for instance,

$$f + g = p \circ (f \times g),$$

where

$$p : \mathbb{C} \times \mathbb{C} \to \mathbb{C}$$

is the addition map, and $f \times g : x \mapsto (f(x), g(x))$.

According to previous Lemma this last map $f \times g$ is measurable, and according to first assertion, the map p is measurable (because it is continuous). By measurability of composition, $f + g$ is also measurable. Similar arguments apply to $f - g$, fg and $1/g$.

Finally, 3rd assertion is a special case of the 2nd. $\square$

Theorem 1.6. *Let $(X, \mathcal{A})$ be a measurable space. Let $(f_n), n \geqslant 1$, be a sequence of measurable real-valued functions. Then*

1. *the functions defined by*

$$\limsup f_n(x), \quad \liminf f_n, \quad \sup_n f_n(x), \quad \inf_n f_n$$

   are measurable.

2. *In particular, if $f_n(x) \to f(x)$ for any $x \in X$, i.e. (f_n) converges pointwise to a limiting function f, then this function f is measurable.*

Proof. **1.** Let $g = \sup\limits_n f_n$, $h = \inf\limits_n f_n$. Measurability of Lebesgue sets of functions g and h follows from identities

$$E(g > a) = \bigcup_n E(f_n > a), \quad E(h < a) = \bigcup_n E(f_n < a).$$

We will prove the first identity. The condition $x \in E(g > a)$ means that $\sup f_n(x) > a$, or, equivalently, $f_n(x) > a$ for some n, that is $x \in \bigcup\limits_n E(f_n > a)$.

2. We will prove the measurability of the upper limit. Since $g_n = \sup_{k \geqslant n} f_k$ is decreasing, then by the definition of upper limit

$$\overline{\lim_{n \to \infty}} f_n = \lim_{n \to \infty} g_n = \inf_{n \in \mathbb{N}} g_n.$$

Applying the first assertion of the theorem twice we see that functions g_n and $\overline{\lim_{n \to \infty}} f_n$ are measurable. □

1.2 Measure on a $\sigma-$algebra.

Definition 1.7. *1. Let $(X, \mathcal{A})$ be a measurable space. A measure μ on $(X, \mathcal{A})$ is a map*

$$\mu : \mathcal{A} \to [0, +\infty]$$

such that $\mu(\varnothing) = 0$ and

$$\mu\left(\bigcup_n E_n\right) = \sum_n \mu\left(E_n\right)$$

for any countable family of pairwise disjoints measurable sets $E_n \in \mathcal{A}$. The triple $(X, \mathcal{A}, \mu)$ is called a measured space.

2. A measure μ is said to be finite if $\mu(X) < +\infty$, and is said to be σ-finite if one can write X as a countable union of subsets with finite measure: there exist $X_n \in \mathcal{A}$ such that

$$X = \bigcup_{n \geqslant 1} X_n, \text{ and } \mu\left(X_n\right) < +\infty \text{ for all } n.$$

3. A probability measure is a measure μ such that $\mu(X) = 1$.

4. *Measure on a Borel $\sigma-$algebra is called a Borel measure.*

Theorem 1.8 (Properties of a measure.)**.** *Let μ be a measure on $(X, \mathcal{A})$.*

(1) ***Monotonicity.*** *For $E, F \in \mathcal{A}$, with $E \subset F$, we have $\mu(E) \leqslant \mu(F)$, and more precisely*

$$\mu(F) = \mu(E) + \mu(F \backslash E).$$

(2) *For $E, F \in \mathcal{A}$, we have*

$$\mu(E \cup F) + \mu(E \cap F) = \mu(E) + \mu(F).$$

(3) ***Lower continuity.*** *If $E_1 \subset \cdots \subset E_n \subset \cdots$ is an increasing sequence of measurable sets, then*

$$\mu\left(\bigcup_n E_n\right) = \lim_{n \to +\infty} \mu\left(E_n\right) = \sup_{n \geqslant 1} \mu\left(E_n\right).$$

(4) ***Upper continuity.*** *If $E_1 \supset \cdots \supset E_n \supset \cdots$ is a decreasing sequence of measurable sets, and if furthermore $\mu\left(E_1\right) < +\infty$, then we have*

$$\mu\left(\bigcap_n E_n\right) = \lim_{n \to +\infty} \mu\left(E_n\right) = \inf_{n \geqslant 1} \mu\left(E_n\right).$$

(5) *For any countable family (E_n) of measurable sets, we have*

$$\mu\left(\bigcup E_n\right) \leqslant \sum \mu\left(E_n\right)$$

Proof. **(1)** Since $E, F \in \mathcal{A}$ then $F \backslash E \in \mathcal{A}$ by definition of $\sigma-$algebra. Then $F = E \cup (F \backslash E)$ and

$$\mu(F) = \mu(E) + \mu(F \backslash E).$$

(2) Exercise
(3) If $\mu A_k = +\infty$ for some k then by monotonicity of a measure $\mu A_n = +\infty$ for every $n > k$ and $\mu A = +\infty$. Consequently, in this case the assertion is satisfied.

Assume that $\mu A_k < +\infty$ for every k. We can express A as the union of mutually disjoint sets

$$A = A_1 \cup \bigcup_{k=1}^{\infty} \left(A_{k+1} \backslash A_k\right).$$

By the first assertion and finiteness of μA_k

$$\mu\left(A_{k+1} \backslash A_k\right) = \mu A_{k+1} - \mu(A_k).$$

Using countable additivity of a measure μ we see that

$$\mu A = \mu A_1 + \sum_{k=1}^{\infty} \mu\left(A_{k+1} \backslash A_k\right) = \mu A_1 + \lim_{n \to \infty} \sum_{k=1}^{n-1} \mu\left(A_{k+1} \backslash A_k\right) =$$

$$\mu A_1 + \lim_{n \to \infty} \sum_{k=1}^{n-1} \left(\mu A_{k+1} - \mu A_k\right) = \lim_{n \to \infty} \mu A_n \quad (2)$$

(4) Notice that, by monotonicity of a measure, $\mu A_k < +\infty$ for every k. Express A_1 as the union

$$A_1 = A \cup \bigcup_{k=1}^{\infty} \left(A_k \backslash A_{k+1}\right),$$

which is mutually disjoint by monotonicity of A_k. Then since $\mu A_{k+1} < +\infty$ we see that

$$\mu\left(A_k \backslash A_{k+1}\right) = \mu A_k - \mu A_{k+1}.$$

Using countable additivity of μ we obtain

$$\begin{aligned}\mu A_1 &= \mu A + \sum_{k=1}^{\infty} \mu\left(A_k \backslash A_{k+1}\right) = \mu A + \sum_{k=1}^{\infty}\left(\mu A_k - \mu A_{k+1}\right) = \\ &= \mu A + \lim_{n\to\infty} \sum_{k=1}^{n-1}\left(\mu A_k - \mu A_{k+1}\right) = \mu A + \lim_{n\to\infty}\left(\mu A_1 - \mu A_n\right) = \\ &= \mu A + \mu A_1 - \lim_{n\to\infty} \mu A_n. \quad (3)\end{aligned}$$

Hence, $\lim_{n\to\infty} \mu A_n = \mu A$ since $\mu A_1 < \infty$. □

Definition 1.9. *Let $(X, \mathcal{A}, \mu)$ be a measured space. A subset $Y \subset X$ is said to be μ-negligible if it is a subset of a set of measure 0, i.e. there exists $E \in \mathcal{A}$ such that*

$$Y \subset E, \ \text{and } \mu(E) = 0.$$

For instance, any set in $\mathcal{A}$ of measure 0 is negligible.

If $\mathcal{P}(x)$ is a mathematical property parametrized by $x \in X$, then one says that $\mathcal{P}$ is true μ-almost everywhere if

$$\{x \in X \mid \mathcal{P}(x) \text{ is not true }\}$$

is μ-negligible.

Definition 1.10. *If all μ-negligible sets are measurable, the measured space is said to be complete.*

Remark 1.11. *The following procedure can be used to construct a natural complete "extension" of a measured space* $(X, \mathcal{A}, \mu)$.

Let $\mathcal{A}_0$ *denote the collection of* μ*-negligible sets. Then define*

$$\mathcal{A}' = \{E \subset X \mid E = E_1 \cup E_0 \text{ with } E_0 \in \mathcal{A}_0 \text{ and } E_1 \in \mathcal{A}\},$$

the collection of sets which, "differ" from a measurable set only by a μ*-negligible set. Define then* $\mu'(E) = \mu(E_1)$ *if* $E = E_0 \cup E_1 \in \mathcal{A}'$ *with* E_0 *negligible and* E_1 *measurable.*

Then the triple $(X, \mathcal{A}', \mu')$ *is a complete measured space; the* σ*-algebra* $\mathcal{A}'$ *contains* $\mathcal{A}$*, and* $\mu' = \mu$ *on* $\mathcal{A}$.

We now describe some important ways to operate on measures, and to construct new measures from old ones.

Lemma 1.12. *Let* $(X, \mathcal{A}, \mu)$ *be a measured space.*

1. *For any finite collection* $\mu_1, \ldots, \mu_n$ *of measures on* $(X, \mathcal{A})$*, and any choice of real numbers* $\alpha_i \in [0, +\infty[$*, the measure* $\mu = \sum \alpha_i \mu_i$ *is defined by*

$$\mu(Y) = \sum_{1 \leqslant i \leqslant n} \alpha_i \mu_i(Y)$$

   for $Y \in \mathcal{A}$*; it is a measure on* $(X, \mathcal{A})$.

2. *Let* $f : (X, \mathcal{A}) \to (X', \mathcal{A}')$ *be a measurable map. Let*

$$f_*(\mu)(Y) = \mu\left(f^{-1}(Y)\right) \text{ for } Y \in \mathcal{A}'.$$

   Then $f_*(\mu)$ *is a measure one* X'*, called the image measure of* μ *under* f*. It is also sometimes denoted* $f(\mu)$.

If we have another measurable map $g : (X', \mathcal{A}') \to (X'', \mathcal{A}'')$, then we have

$$(g \circ f)_*(\mu) = g_* (f_*(\mu)) . \tag{4}$$

3. *For any measurable subset $Y \subset X$, the restriction of μ to the σ-algebra of measurable subsets of Y is a measure on Y for this σ-algebra.*

Proof. **1.** First, notice that

$$\mu(\varnothing) = \sum_{1 \leqslant i \leqslant n} \alpha_i \mu_i(\varnothing) = 0.$$

Now we will check countable additivity. Assume that sets E_k are disjoint. Then

$$\mu\left(\bigcup_{k=1}^{\infty} E_k\right) = \sum_{i=1}^{n} \alpha_i \mu_i \left(\bigcup_{k=1}^{\infty} E_k\right) = \\ \sum_{i=1}^{n} \alpha_i \sum_{k=1}^{\infty} \mu_i (E_k) = \sum_{k=1}^{\infty} \sum_{i=1}^{n} \alpha_i \mu_i (E_k) = \sum_{k=1}^{\infty} \mu(E_k)$$

2. First, we see that $f^{-1}(\varnothing) = \varnothing$ and $f_*(\mu)(\varnothing) = \mu(\varnothing) = 0$. To check countable additivity we use the following property of preimages

$$f^{-1}\left(\bigcup_{k=1}^{\infty} E_k\right) = \bigcup_{k=1}^{\infty} f^{-1}(E_k).$$

Also if sets E_k are measurable and disjoint then preimages are $f^{-1}(E_k)$

are measurable and disjoint. Hence

$$f_*(\mu)\left(\bigcup_{k=1}^{\infty} E_k\right) = \mu\left(f^{-1}\left(\bigcup_{k=1}^{\infty} E_k\right)\right) =$$
$$\mu\left(\bigcup_{k=1}^{\infty} f^{-1}(E_k)\right) = \sum_{k=1}^{\infty} \mu\left(f^{-1}(E_k)\right) = \sum_{k=1}^{\infty} f_*(\mu)(E_k).$$

The formula (4) is due to the simple fact that

$$(g \circ f)^{-1}(E'') = f^{-1}\left(g^{-1}(E'')\right)$$

for every $E'' \subset X''$.

3. Let

$$\mathcal{A}_Y = \{E \in \mathcal{A} : E \subset Y\}$$

and

$$\mu|_Y(E) = \mu(E), \quad E \in \mathcal{A}_Y.$$

Then $\mathcal{A}_Y$ is a $\sigma-$algebra of subsets of Y. If $E_k \in \mathcal{A}$ are disjoint such that $E_k \subset Y$ then $E = \bigcup_{k=1}^{\infty} E_k \in A$ and $E \subset Y$. Moreover,

$$\mu|_Y(E) = \mu(E) = \sum_{k=1}^{\infty} \mu(E_k) = \sum_{k=1}^{\infty} \mu|_Y(E_k)$$

and $\mu|_Y$ is a measure on $\mathcal{A}_Y$. $\square$

1.3 The Lebesgue measure

The following result was obtained in the course Mathematical analysis (3).

Theorem 1.13. *There exists a complete $\sigma-$algebra $\mathcal{A}_n$ of subsets of $\mathbb{R}^n$ and unique measure μ such that*

$$\mu\big([a_1, b_1) \times \ldots \times [a_n, b_n)\big) = (b_1 - a_1) \ldots (b_n - a_n)$$

for any cell in $\mathbb{R}^n$. This $\sigma-$algebra and a measure is called the Lebesgue measure.

Remark 1.14. *This theorem is obtained from the Caratheodory theorem. The restriction of a Lebesgue measure to Borel $\sigma-$algebra is also called Lebesgue measure.*

1.4 Properties of Lebesgue $\sigma-$algebra and Lebesgue measure

1. Lebesgue measure is σ-finite.
2. Lebesgue $\sigma-$algebra contains Borel $\sigma-$algebra. In particular, open, closed sets, sets of type F_σ and G_δ are measurable (see MA(3)).
3. Any at most countable subset of $\mathbb{R}^n$ is Lebesgue measurable and has measure 0.
4. Let $E \in \mathcal{A}_n$. Then there exists a set H of type F_σ and a set K of type G_δ such that

$$H \subset E \subset K, \quad \mu(K \backslash H) = 0.$$

5. A set E is Lebesgue-measurable if it can be expressed as the union of increasing sequence of compact sets and a set of measure 0 , that is

$$E = \bigcup_{k=1}^{\infty} F_k \cup e,$$

where F_k are compact sets, $F_k \subset F_{k+1}$, $\mu e = 0$.

6. Let $A : \mathbb{R}^n \to \mathbb{R}^n$ be linear transformation and $E \subset \mathbb{R}^n$ be measurable. Then $A(E)$ is measurable and

$$\mu(A(E)) = |\det A|\, \mu(A).$$

In particular, the Lebesgue measure is invariant with respect to a shift.

1.5 Borel measures and regularity properties.

Recall the following definitions

- A Borel σ-algebra of subsets of topological space X is a $\sigma-$algebra generated by topology (family of open subsets) of X.

- A Borel measure μ on a topological space X is a measure defined on a Borel $\sigma-$algebra.

Definition 1.15. *A Borel measure μ on a topological space X is said to be regular if, for any Borel set $Y \subset X$, we have*

- $\mu(Y) = \inf\{\mu(U) \mid U \text{ is an open set containing } Y\}$;

- $\mu(Y) = \sup\{\mu(K) \mid K \text{ is a compact subset contained in } Y\}$.

This property, when it holds, gives a link between general Borel sets and the more "regular" open or compact sets.

In the course Mathematical Analysis (3) we have shown, in particular the following:

Theorem 1.16. *The Lebesgue measure on $\mathcal{B}_{\mathbb{R}^n}$ is regular.*

The following criterion shows that regularity can be achieved under quite general conditions:

Theorem 1.17. *Let X be a locally compact topological space, in which any open set is a countable union of compact subsets. Then any Borel measure μ such that $\mu(K) < +\infty$ for all compact sets $K \subset X$ is regular; in particular, any finite measure, including probability measures, is regular.*

w/o proof

2 Integration with respect to a measure

In this section we assume that $(X, \mathcal{A}, \mu)$ is a measured space, $E \in \mathcal{A}$ and that functions that we consider have values in $\overline{\mathbb{R}}$.

We recall the decomposition for real-valued function $f : E \to \mathbb{R}$

$$f = f_+ - f_-, \quad f_+(x) = \max(0, f(x)), \quad f_-(x) = \max(0, -f(x)),$$

and we observe that $|f| = f_+ + f_-$.

Also we remind results on approximation of measurable function by simple and step functions.

Theorem 2.1. *Let (X, A) be measurable space, $E \in \mathcal{A}$ and $f : E \to \mathbb{R}$ be measurable function.*

1. *If $f \geq 0$ then there exists a sequence φ_n of simple functions such that for every $x \in E$*

$$\varphi_n(x) \leq \varphi_{n+1}, \text{ and } f(x) = \lim_{n\to\infty} \varphi_n(x) = \sup_n \varphi_n(x).$$

2. *There exists a sequence φ_n of step functions such that for every $x \in E$*

$$|\psi_n(x)| \leq |f(x)|, \text{ and } f(x) = \lim_{n\to\infty} \psi_n(x).$$

3. *If f is bounded then convergence in previous assertions is uniform.*

Proof. **1.** Let $n \in \mathbb{N}$ and denote

$$E_{in} = \begin{cases} E\left(\frac{i}{2^n} \leqslant f < \frac{i+1}{2^n}\right), & i \in [0 : n2^n - 1], \\ E(f \geqslant n), & i = n2^n. \end{cases}$$

Sets E_{in} are measurable, disjoint and $\bigcup\limits_{i=0}^{n2^n} E_{in} = E$. Let

$$\varphi_n = \sum_{i=0}^{n2^n} \frac{i}{2^n} \chi_{E_{in}}.$$

In other words, $\varphi_n = \frac{i}{2^n}$ on E_{in} and $\varphi_n = 0$ on E. Functions φ_n are simple.

Let $x \in E$. Assume that $[+\infty] = +\infty$ and check that

$$\varphi_n(x) = \min\left\{\frac{[2^n f(x)]}{2^n}, n\right\}.$$

If $x \in E_{in}$ for $i = n2^n$ then $f(x) \geqslant n$. Hence, $\frac{[2^n f(x)]}{2^n} \geqslant n = \varphi_n(x)$. If $x \in E_{in}$ for $i \in [0 : n2^n - 1]$ then $i \leqslant 2^n f(x) < i + 1$. Hence, $n > \frac{[2^n f(x)]}{2^n} = \frac{i}{2^n} = \varphi(x)$.

Now, we will prove that $\varphi_n(x) \leqslant \varphi_{n+1}(x)$. It is enough to check that

$$\frac{[2^n f(x)]}{2^n} \leqslant \frac{\left[2^{n+1} f(x)\right]}{2^{n+1}}.$$

Let $A = 2^n f(x)$. Then, this estimate follows from general estimate for the integer part of real number:

$$[2A] \geq [2[A]] = 2[A].$$

It remains to prove that $\varphi_n(x) \underset{n\to\infty}{\longrightarrow} f(x)$. If $f(x) = +\infty$ then $\varphi_n(x) = n$ for every $n \in \mathbb{N}$ and $\varphi_n(x) \to +\infty$. Assume that $f(x) \in [0, +\infty)$. Then for every $n > f(x)$ there exists such $i \in [0 : n2^n - 1]$ that $\frac{i}{2^n} \leqslant f(x) < \frac{i+1}{2^n}$, that is $x \in E_{\text{in}}$. Consequently, $\varphi_n(x) = \frac{i}{2^n}$ and

$$0 \leqslant f(x) - \varphi_n(x) < \frac{1}{2^n}.$$

This proves (pointwise) convergence and finalizes the proof.

2. By the previous theorem there exist the increasing sequences of functions φ_n^* and φ_n^{**} pointwise convergent to f_+ and f_-. Functions $\psi_n = \varphi_n^* - \varphi_n^{**}$ are step functions. Moreover, $\psi_n \to f_+ - f_- = f$.

It remains to check the estimate. If $f(x) \geqslant 0$ then $f_-(x) = 0$ and $\varphi_n^{**}(x) = 0$. Consequently,

$$|\psi_n(x)| = \varphi_n^*(x) \leqslant f_+(x) = f(x).$$

If $f(x) < 0$ then $f_+(x) = 0$ and $\varphi_n^*(x) = 0$, Consequently,

$$|\psi_n(x)| = \varphi_n^{**}(x) \leqslant f_-(x) = -f(x).$$

3. To prove uniform convergence notice that if $0 \leq f \leq M$ then $\varphi_n(x) = \frac{[2^n f(x)]}{2^n}$ for $n > M$ and

$$|f(x) - \varphi_n(x)| < \frac{1}{2^n}$$

for every $x \in E$. For the case of function that obtains negative and positive function consider decomposition $f = f_+ - f_-$ and notice that boundedness of f implies boundedness of $f_\pm$. $\square$

Definition 2.2. *Integral $\int_E f d\mu$ of function $f \in S(E)$ with respect to a measure μ on E is defined in the following way.*

1. *Let f be simple:*

$$f = \sum_{k=1}^{N} c_k \chi_{A_k}, \quad A_k \in \mathcal{A}, A_k \text{ are disjoint, } c_k \in [0, +\infty).$$

Then we let

$$\int_E f d\mu = \sum_{k=1}^{N} c_k \mu \left(A_k \cap E\right).$$

2. *let $f \geqslant 0$. Then we let*

$$\int_E f d\mu = \sup_{\substack{\varphi \text{ simple} \\ \varphi \leqslant f \text{ on } E}} \int_E \varphi d\mu.$$

3. *Let f be arbitrary measurable function. Then we let*

$$\int_E f d\mu = \int_E f_+ d\mu - \int_E f_- d\mu$$

if at least one of the integrals $\int_E f_\pm d\mu$ is finite. If both integrals are infinite then symbol $\int_E f d\mu$ has no value.

4. *Let $f : E \to \mathbb{C}$ be complex-valued measurable function. Then*

$$\int_E f d\mu = \int_E \mathrm{Re}(f) d\mu + i \int_E \mathrm{Im}(f) d\mu \in \mathbb{C}.$$

if both integrals exist in sense of the integral of real-valued function.

We use notation $\int\limits_E f(x) d\mu(x)$ to indicate the variable of integration. If the measure is fixed we write $\int\limits_E f$.

Definition 2.3. *Integral with respect to Lebesgue measure is called Lebesgue integral.*

Now we will shortly remind you the key properties of the integral.

Theorem 2.4 (Monotonicity of the integral)**.** *Assume that $f, g : E \to \overline{\mathbb{R}}$, $f \leqslant g$ on E and integrals $\int_E f d\mu$, $\int_E g d\mu$ exist. Then*

$$\int_E f d\mu \leqslant \int_E g d\mu$$

Proof. **Case 1.** Assume that functions f and g are simple,

$$f = \sum_{k=1}^{N} c_k \chi_{A_k}, \quad g = \sum_{i=1}^{M} d_i \chi_{B_i},$$

where $c_k, d_i \in [0, +\infty), A_k, B_i \in \mathcal{A}$, A_k are mutually disjoint B_i are mutually disjoint, $\bigcup\limits_{k=1}^{N} A_k = \bigcup\limits_{i=1}^{M} B_i = X$. Let $D_{ki} = A_k \cap B_i$. Then D_{ki} are mutually disjoint, $\bigcup\limits_{i=1}^{M} D_{ki} = A_k$, $\bigcup\limits_{k=1}^{N} D_{ki} = B_i$. By definition of the integral and additivity of a measure we see that

$$\int_E f = \sum_{k=1}^{N} c_k \mu \left(A_k \cap E\right) = \sum_{k=1}^{N} c_k \mu \bigcup_{i=1}^{M} \left(D_{ki} \cap E\right) =$$
$$\sum_{k=1}^{N} \sum_{i=1}^{M} c_k \mu \left(D_{ki} \cap E\right),$$

and, analogously,

$$\int_E g = \sum_{i=1}^{M} \sum_{k=1}^{N} d_i \mu \left(D_{ki} \cap E\right).$$

In this sum we can take into account only such terms for which $\mu\left(D_{ki}\cap E\right) > 0$. If $\mu\left(D_{ki}\cap E\right) > 0$ then there exists $x\in D_{ki}\cap E$. Then $f(x) = c_k, g(x) = d_i$ and, consequently, $c_k \leqslant d_i$. Thus, $\int_E f \leqslant \int_E g$.

Case 2. Let $f, g \geqslant 0$. The condition $f \leqslant g$ implies the inclusion

$$\{\varphi : \varphi \text{ is simple}, \varphi \leqslant f \text{ on } E\} \subset \{\varphi : \varphi \text{ is simple}, \varphi \leqslant g \text{ on} E\}.$$

Consequently the supremum of the integrals $\int_E \varphi$ over the set in the left-hand side of the inclusion is less or equal then over the set in the right-hand side.

Case 3. Let f and g be arbitrary. The inequality $f \leqslant g$ is equivalent to the system of inequalities $f_+ \leqslant g_+$, $f_- \geqslant g_-$. By the previous case

$$\int_E f_+ \leqslant \int_E g_+, \quad \int_E f_- \geqslant \int_E g_-.$$

Substracting the second inequality from the first one we obtain the assertion the theorem. □

Corollary 2.4.1 (Monotonicity of the integral by set.)**.** *Let $f \in S(E)$, $f \geqslant 0$, $E_1 \subset E$, $E_1 \in \mathcal{A}$. Then $\int_{E_1} f d\mu \leqslant \int_E f d\mu$.*

Proof. Indeed, by Lemma

$$\int_{E_1} f = \int_E f\chi_{E_1} \leqslant \int_E f$$

□

Theorem 2.5 (B.Levy monotone convergence theorem)**.** *Let $f_n \in S(E)$, $f_n \geqslant 0$, $f_n \leqslant f_{n+1}$, $f = \lim\limits_{n\to\infty} f_n$. Then*

$$\int_E f_n d\mu \underset{n\to\infty}{\longrightarrow} \int_E f d\mu$$

Proof. By Theorem on the limit of measurable functions $f \in S(E)$. The monotonicity of the integral implies that the sequence $I_n = \int_E f_n$ is increasing. Consequently it has finite or infinite limit $\alpha \in [0, +\infty]$. Since $f_n \leqslant f$ then $\int_E f_n \leqslant \int_E f$ for every n and $\alpha \leqslant \int_E f$.

It remains to prove the inverse inequality. Let φ be simple function, $\varphi = \sum\limits_{k=1}^{N} c_k \chi_{A_k}$, $\varphi \leqslant f$ on E. Let $q \in (0,1)$ and denote $E_n = E\left(f_n \geqslant q\varphi\right)$. Then $E_n \subset E_{n+1}$ since the sequence $\{f_n\}$ is increasing. We will prove that $\bigcup\limits_{n=1}^{\infty} E_n = E$. the inclusion of the left-hand side into the right-hand side is obvious. Let $x \in E$. If $\varphi(x) = 0$ then $x \in E_n$ for every n. If $\varphi(x) > 0$ then $f(x) \geqslant \varphi(x) > q\varphi(x)$. By the definition of the limit starting from dome number $f_n(x) > q\varphi(x)$, that is $x \in E_n$.

By monotonicity of the integral we see that

$$\int_E f_n \geqslant \int_{E_n} f_n \geqslant \int_{E_n} q\varphi = \sum_{k=1}^{N} q \cdot c_k \mu\left(A_k \cap E_n\right).$$

Properties of E_n imply the following relations

$$\left(A_k \cap E_n\right) \subset \left(A_k \cap E_{n+1}\right), \quad \bigcup_{n=1}^{\infty} \left(A_k \cap E_n\right) = A_k \cap E$$

and, by continuity of the measure

$$\mu\left(A_k \cap E_n\right) \underset{n\to\infty}{\longrightarrow} \mu\left(A_k \cap E\right).$$

Consequently letting n to ∞ and q to 1 we see that

$$\alpha \geqslant \sum_{k=1}^{N} q \cdot c_k \mu\left(A_k \cap E\right) = q \int_E \varphi, \quad \alpha \geqslant \int_E \varphi.$$

Taking supremum over φ we finalize the proof of the theorem, $\alpha \geqslant \int_E f$. $\square$

Theorem 2.6 (Linearity of the integral)**.** ***(1)*** *Assume that* $\alpha \in \mathbb{R}$, $f : E \to \overline{\mathbb{R}}$ *and the integral* $\int_E f d\mu$ *exists. Then the integral* $\int_E \alpha f d\mu$ *exists and*

$$\int_E \alpha f d\mu = \alpha \int_E f d\mu.$$

(2) *Assume that there exist integrals* $\int_E f d\mu$, $\int_E g d\mu$ *and their sum. Then the integral* $\int_E (f+g) d\mu$ *exists and*

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

Proof. **(1)** If $\alpha = 0$ the assertion is trivial. Consider other cases.
Case 1. Let $\alpha > 0$.
Case 1.1. Assume that f is simple. Then the function αf is also simple and

$$\int_E \alpha f = \sum_k (\alpha c_k)\, \mu\,(A_k \cap E) = \alpha \sum_k c_k \mu\,(A_k \cap E) = \alpha \int_E f.$$

Case 1.2. Let $f \geqslant 0$. Consider the increasing sequence of simple functions $\{\varphi_n\}$ that converges to f. Then $\{\alpha\varphi_n\}$ is the increasing sequence of simple functions that converges to αf. By the previous case

$$\int_E \alpha\varphi_n = \alpha \int_E \varphi_n$$

Then, considering the limit, by Levy theorem we obtain the assertion of the theorem.

Case 1.3. Let f then

$$\int_E (\alpha f)_\pm = \int_E \alpha f_\pm = \alpha \int_E f_\pm$$

and this finalizes the proof of the case $\alpha > 0$.

Case 2. Let $\alpha < 0$. Since $\alpha = (-1) \cdot |\alpha|$ it is enough to consider the case $\alpha = -1$. Notice that

$$(-f)_+ = \max\{-f, 0\} = f_-, \quad (-f)_- = \max\{-(-f), 0\} = f_+,$$

and by the definition of the integral

$$\int_E (-f) = \int_E (-f)_+ - \int_E (-f)_- = \int_E f_- - \int_E f_+ = -\int_E f.$$

(2) Case 1. Let f and g be simple. Without loss of generality we may assume that they are equal to 0 outside of E :

$$f = \sum_{k=1}^{N} c_k \chi_{A_k}, \quad g = \sum_{i=1}^{M} d_i \chi_{B_i},$$

where $c_k, d_i \in [0, +\infty)$, $A_k, B_i \in \mathcal{A}$, A_k are mutually disjoint, B_i are mutually disjoint, $\bigcup_{k=1}^{N} A_k = \bigcup_{i=1}^{M} B_i = E$. Denote $D_{ki} = A_k \cap B_i$. Then D_{ki} are mutually disjoint, $\bigcup_{i=1}^{M} D_{ki} = A_k$, $\bigcup_{k=1}^{N} D_{ki} = B_i$. Then by the definition of the integral and by the additivity of the measure

$$\int_E (f+g) = \sum_{k,i} (c_k + d_i)\, \mu D_{ki} = \sum_k c_k \sum_i \mu D_{ki} + \sum_i d_i \sum_k \mu D_{ki} =$$

$$= \sum_k c_k \mu A_k + \sum_i d_i \mu B_i = \int_E f + \int_E g.$$

Case 2. Let $f, g \geqslant 0$. Consider increasing sequences of simple functions $\{\varphi_n\}$ and $\{\psi_n\}$ that converge to f and g respectively. Then $\{\varphi_n + \psi_n\}$ is the increasing sequence of simple functions that converges to $f + g$. By the previous case

$$\int_E (\varphi_n + \psi_n) = \int_E \varphi_n + \int_E \psi_n$$

and the result follows from Levy theorem.

Case 3. Let f and g be of arbitrary sign and $h = f + g$. The function h is defined on the set $E_1 = E \setminus (e_1 \cup e_2)$, where

$$e_1 = E(f = +\infty) \cap E(g = -\infty), \quad e_2 = E(f = -\infty) \cap E(g = +\infty).$$

We will check that $\mu e_1 = \mu e_2 = 0$, that will mean that h is defined a.e. on E. If $\mu e_1 > 0$ then $\int_E f_+ = +\infty$ since for every $R \in (0, +\infty)$

$$\int_E f_+ \geqslant \int_{e_1} f_+ = \int_{e_1} f \geqslant R \cdot \mu e_1.$$

Analogously, $\int_E g_- = +\infty$. Consequently, $\int_E f = +\infty, \int_E g = -\infty$ and the sum of thes integrals is not defined, which contradicts the assumption of the theorem. The identity $\mu e_2 = 0$ can be checked in the same way.

The integral on E can be substituted by the integral on E_1. In the following we consider functions f, g and h on E_1. It is clear that

$$h_+ - h_- = h = f + g = f_+ - f_- + g_+ - g_- = (f_+ + g_+) - (f_- + g_-).$$

Let's prove that (some values may be infinite !)

$$h_+ + f_- + g_- = h_- + f_+ + g_+.$$

From the definition of the positive and negative parts follows that

$$h_+ = \max\{f+g,0\} \leqslant f_+ + g_+, \quad h_- = \max\{-f-g,0\} \leqslant f_- + g_-. \tag{5}$$

If $f_- + g_- < +\infty$ then $h_- < +\infty$ and (5). If $f_- + g_- = +\infty$ then $h_- = +\infty$ and the identity (5) has the form $+\infty = +\infty$. By the previous reasoning

$$\int_{E_1} h_+ + \int_{E_1} f_- + \int_{E_1} g_- = \int_{E_1} h_- + \int_{E_1} f_+ + \int_{E_1} g_+$$

We'll prove that in this identity we also can move terms from one part to another, that is

$$\int_{E_1} h_+ - \int_{E_1} h_- = \int_{E_1} f_+ + \int_{E_1} g_+ - \left(\int_{E_1} f_- + \int_{E_1} g_-\right).$$

In particular, integrals $\int_{E_1} h_+, \int_{E_1} h_-$ can not be infinite simultaneously. This will finalize the prove since this identity implies that

$$\int_{E_1} h = \int_{E_1} f + \int_{E_1} g.$$

If $\int_{E_1} f_- + \int_{E_1} g_- < +\infty$ then, since $\int_{E_1} h_- < +\infty$, we can move terms from one part to another. If $\int_{E_1} f_- + \int_{E_1} g_- = +\infty$, that is one of the integrals is infinite, then $\int_{E_1} f_+ + \int_{E_1} g_+ < +\infty$, otherwise one of the integral $\int_{E_1} f, \int_{E_1} g$ or their sum doesn't exist. Thus

$$\int_{E_1} h_- = +\infty, \quad \int_{E_1} h_+ < +\infty,$$

and the equality has the following form: $-\infty = -\infty$. $\square$

Corollary 2.6.1 (Levy theorem for series)**.** *Series with nonnegative terms can be integrated term-by-term. If* $f_k \geqslant 0$, $f_k \in S(E)$ *then*

$$\int_E \sum_{k=1}^{\infty} f_k d\mu = \sum_{k=1}^{\infty} \int_E f_k d\mu.$$

Proof. By Theorem 2.6 we have

$$\int_E \sum_{k=1}^{n} f_k = \sum_{k=1}^{n} \int_E f_k, \ n \in \mathbb{N}.$$

The sequence $F_n = \sum\limits_{k=1}^{n} f_k$ is increasingly convergent to $F = \sum\limits_{k=1}^{\infty} f_k$. Then, by the definition of the sum and Levy theorem the left-hand side converges to $\int_E \sum\limits_{k=1}^{\infty} f_k$ and the right-hand side to $\sum\limits_{k=1}^{\infty} \int_E f_k$. □

2.1 Chebyshev inequality

Lemma 2.7 (Chebyshev inequality)**.** *Let* $f : X \to \mathbb{R}$ *be measurable function,* $t \in (0, +\infty)$*. Then*

$$\mu E(|f| \geqslant t) \leqslant \frac{1}{t} \int_E |f| d\mu.$$

Proof. First, the set $E_t = E(|f| \geqslant t)$ is measurable and

$$\int_E |f| d\mu \geq \int_{E_t} |f| d\mu \geq \int_{E_t} t d\mu = t\mu(E_t).$$

□

Corollary 2.7.1. *1. Assume that $\int_E |f|\, d\mu$ is finite then f is finite a.e. on E.*

2. Assume that $\int_E |f|\, d\mu = 0$. Then $f = 0$ a.e.

Proof. **1.** Notice that

$$\mu E(|f| = +\infty) \leq \mu E(|f| \geqslant t) \leq \frac{1}{t} \int_E |f|\, d\mu$$

for every $t > 0$. Letting $t \to \infty$ we conclude $\mu E(|f| = +\infty) = 0$.
2. Notice that

$$E(f \neq 0) = \bigcup_{n=1}^{\infty} E\left(|f| \geqslant \frac{1}{n}\right)$$

and that

$$E\left(|f| \geqslant \frac{1}{n}\right) \leq n \int_E |f|\, d\mu = 0.$$

Hence, $\mu E(f \neq 0) = 0$ and $f = 0$ a.e. □

Remark 2.8. *Notice that if $f = 0$ a.e. on $E \in \mathcal{A}$ then $\int_E f = 0$. In particular if $\mu E = 0$ then $\int_E f = 0$ for every function f that is measurable on E.*

Proof. Assume firs that $f \geq 0$. Let

$$\varphi(x) = \sum_{k=1}^{N} c_k \chi_{A_k}(x), \quad A_k \cap A_j = \varnothing,\ k \neq j, \quad X = \bigcup_{k=1}^{\infty} A_k, \quad c_k \in [0, \infty]$$

be a simple function such that $\varphi \leq f$. Then $\varphi = 0$ a.e. on E and $\mu(A_k \cap E) = 0$ if $c_k > 0$. Consequently,

$$\int_E \varphi = \sum_{k=1}^{N} c_k \mu(A_k \cap E) = 0.$$

Finally,

$$\int_E f = \sup\left\{\int_E \varphi : \varphi \text{ is simple and } \varphi \le f\right\} = \sup 0 = 0.$$

$\square$

2.2 Integral as a measure. General formula for the change of the variable.

Theorem 2.9. *Let $(X, \mathcal{A}, \mu)$ be measured space and f be a nonnegative measurable function. Define*

$$\mu_f(E) = \int_E f d\mu$$

for $E \in \mathcal{A}$.

Then μ_f is a measure on $(X, \mathcal{A})$, such that any μ-negligible set is μ_f negligible. Moreover, we can write

$$\int_E g d\mu_f = \int_E g f d\mu \tag{6}$$

for any $E \in \mathcal{M}$ and any measurable $g \geqslant 0$.

Proof. First, by monotonicity of the integral μ_f is nonnegative. Also $\mu_f(\varnothing) = 0$. It remains to check that μ_f is countable additive. Consider a sequence $\{E_n\}$ of pairwise disjoint measurable sets and let $E = \bigcap\limits_{n=1}^{\infty} E_n$. Then

$$\chi_E(x) = \sum_{n=1}^{\infty} \chi_{E_n}(x)$$

and

$$\mu_f(E) = \int_E f\chi_E d\mu = \sum_{n=1}^{\infty} \int_X \chi_{E_n}(x) = \sum_{n=1}^{\infty} \mu_f(E_n).$$

Assume now that $E_1 \subset E$ such that $\mu(E) = 0$. Then $\mu_f(E) = 0$ Remark 2.8. Thus every negligible set with respect to measure μ is also negligible for μ_f.

The formula (6), By definition, is valid if g is a characteristic function of a measurable set. By linearity of the integral, it is valid also for any simple function. Let s_n be a sequence of simple functions monotonicaly convergent to g. Applying the monotone convergence theorem (B. Levi) we see that

$$\int_E g d\mu_f = \lim_{n\to+\infty} \int_E s_n d\mu_f = \lim_{n\to+\infty} \int_E \varphi_n f d\mu = \int_Y g f d\mu.$$

$\square$

Theorem 2.10 (General formula for the change of the variable)**.** *Assume that $\varphi : X \to X'$ is a measurable map. Then for any $g \geqslant 0$ measurable on X' and any $E \in \mathcal{A}'$, we have the change of variable formula*

$$\int_{\varphi^{-1}(E)} (g\circ\varphi) d\mu = \int_{\varphi^{-1}(E)} g(\varphi(x)) d\mu(x) = \int_E g d\varphi_*(\mu), \qquad (7)$$

where $\varphi_*(\mu)$ is the image of measure μ with respect to the map φ.

Proof. Finally, we prove (7) by checking it for more and more general functions g. First of all, for $g = \chi_F$, the characteristic function of a set $F \in \mathcal{A}'$, we have

$$g\circ\varphi = \chi_{\varphi^{-1}(F)}$$

(since the left-hand side takes values 0 and 1 , and is equal to 1 if and only if $\varphi(x) \in F$), and the formula becomes the definition

$$(\varphi_*(\mu))(E \cap F) = \mu\left(\varphi^{-1}(E \cap F)\right).$$

Next, observe that if (7) holds for two functions, it holds for their linear combination, by additivity of the integral on both sides. This means that the formula also holds for all step functions. And then, finally, if $g \geqslant 0$ is the pointwise non-decreasing limit of a sequence s_n of step functions, we have

$$(s_n \circ \varphi)(x) = s_n(\varphi(x)) \leqslant s_{n+1}(\varphi(x)) \to (g \circ \varphi)(x),$$

and

$$(s_n \circ \varphi)(x)\chi_E(x) \to (g \circ \varphi)(x)\chi_E(x).$$

Consequently, the monotone convergence theorem shows that

$$\int_E s_n(\varphi(x))d\mu(x) \to \int_E g(\varphi(x))d\mu(x),$$

which concludes the proof. □

EXAMPLE 1. Let $(X, \mathcal{A}, \mu)$ be measured space. Consider a measurable function $g : X \to \mathbb{R}$ and denote by $\nu = g(\mu)$ the image of μ. This is a Borel measure on $\mathbb{R}$ and by (7), applied to $|g|$ we see that

$$\int_X |g(x)|\, d\mu(x) = \int_{\mathbb{R}} |x| d\nu(x)$$

2.3 Integrable functions.

Definition 2.11. *Let $(X, \mathcal{A}, \mu)$ be a measured space and $E \in \mathcal{A}$.*

- *A measurable function $f : E \to \mathbb{R}$ is said to be integrable on E with respect to the measure μ if the non-negative function $|f| = f_+ + f_-$ has finite integral*

$$\int_E |f| d\mu < +\infty.$$

- *A complex-valued measurable function $f : E \to \mathbb{C}$ is said to be integrable on E if $|f| = \sqrt{\mathrm{Re}(f)^2 + \mathrm{Im}(f)^2} \geqslant 0$ has finite integral*

$$\int_E |f| d\mu < +\infty.$$

- *We denote by $L(E, \mu)$, or sometimes simply $L(\mu)$ or $L(E)$, the set of all μ-integrable complex-valued functions defined on E.*

Definition 2.12. *The function that is integrable with respect to Lebesgue measure is called Lebesgue integrable.*

We first observe that these definitions lead to well-defined real (or complex) numbers for the values of the integrals. For instance, since

$$0 \leqslant f^{\pm} \leqslant |f|$$

if f is μ-integrable. If f is complex-valued, we have

$$|\mathrm{Re}(f)| \leqslant |f|\,, \text{ and } \ |\mathrm{Im}(f)| \leqslant |f|,$$

which implies that the real and imaginary parts of f are themselves μ-integrable.

One may also remark immediately that

$$\int_X \bar{f} d\mu = \overline{\int_X f d\mu}$$

if f (equivalently, $\bar{f}$) is integrable.

Remark. Notice here that this definition, which is extremely general (without assumption on the structure of X or boundedness of f, etc), has the nature of "absolutely convergent" integral; this may seem like a restriction, but it is essential for the resulting process to behave reasonably. In particular, note that if f is μ-integrable on X, then it follows that f is μ-integrable on Y for any measurable subset $Y \subset X$, because

$$\int_Y |f| d\mu = \int_X |f| \chi_Y d\mu,$$

and $|f|\chi_Y \leqslant |f|$. This innocuous property would fail for most definitions of a nonabsolutely convergent integral.

Lemma 2.13. ***1.*** *The set $L(X,\mu)$ is a $\mathbb{C}-$vector space. Moreover, the map*

$$\|f\|_1 = \int_X |f| d\mu$$

is a semi-norm on $L(X,\mu)$, i.e., we have

$$\|af\|_1 = |a| \, \|f\|_1$$

for $a \in \mathbb{C}$ and $f \in L(X,\mu)$, and

$$\|f+g\|_1 \leqslant \|f\|_1 + \|g\|_1,$$

for $f, g \in L(X,\mu)$. Moreover, $\|f\|_1 = 0$ if and only if f is zero μ-almost everywhere.

2. *The map*

$$f \mapsto \int_X f d\mu$$

is a linear map, it is non-negative in the sense that $f \geqslant 0$ *implies* $\int f d\mu \geqslant 0$*, and it satisfies*

$$\left|\int_X f d\mu\right| \leqslant \int_X |f|\, d\mu = \|f\|_1 .$$

3. *For any measurable map* $\varphi : (X, \mathcal{A}) \to (X', \mathcal{A}')$*, and any measurable function* g *on* X'*, we have*

$$\int_{\varphi^{-1}(E)} g(\varphi(x)) d\mu(x) = \int_E g(y) d\varphi_*(\mu)(y)$$

in the sense that if either of the two integrals written are defined, i.e., the corresponding function is integrable with respect to the relevant measure, then the other is also integrable, and their respective integrals are equal.

Proof. Notice first that

$$|\alpha f + \beta g| \leq |\alpha|\,|f| + |\beta|\,|g|$$

immediately shows (using the additivity and monotony of integrals of non-negative functions) that

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

for any $f, g \in L^1(\mu)$ and $\alpha, \beta \in \mathbb{C}$. This proves both $\alpha f + \beta g \in L^1(\mu)$ and the triangle inequality (take $\alpha = \beta = 1$). Moreover, if $g = 0$, we have an identity $|\alpha f| = |\alpha||f|$, and hence also $\|\alpha f\|_1 = |\alpha|\|f\|_1$. $\square$

EXAMPLES. We start with the simplest examples of measures.

1. Let X be an arbitrary set and μ the counting measure on X. Of course, we expect that

$$\int_X f(x)d\mu = \sum_{x\in X} f(x),$$

for a function f on X, but since there is no assumption on an ordering of the summation set, or that it be countable, we see that even this simple case offers something new.

One can then check that a function is integrable with respect to this measure if it is "absolutely summable", in the following sense: given a family $(f(x))_{x\in X}$, it is said to be absolutely summable with sum $S \in \mathbb{C}$ if and only if, for any $\varepsilon > 0$, there exists a finite subset $X_0 \subset X$ for which

$$\left|\sum_{x\in X_1} f(x) - S\right| < \varepsilon,$$

for any finite subset X_1 with $X_0 \subset X_1 \subset X$. In particular, if $X = \mathrm{N}$ with the counting measure, we are considering series

$$\sum_{n\geqslant 1} a_n,$$

of complex numbers, and a sequence (a_n) is integrable if and only if the series converges absolutely. In particular, $a_n = (-1)^n/(n+1)$ does not define an integrable function.

Then, by Tonelli theorem, provided each $a_{i,j}$ is $\geqslant 0$, we can exchange two series:

$$\sum_{i\geqslant 1}\sum_{j\geqslant 1} a_{i,j} = \sum_{j\geqslant 1}\sum_{i\geqslant 1} a_{i,j}.$$

This is not an obvious fact, and it is quite nice to recover this as part of the general theory.

2. Let $\mu = \delta_{x_0}$ be the Dirac measure at $x_0 \in X$, i.e.

$$\delta(A) = \begin{cases} 1, & x_0 \in A; \\ 0, & x_0 \notin A; \end{cases}$$

then any function $f : X \to \mathbb{C}$ is μ-integrable and we have

$$\int_X f(x) d\delta_{x_0}(x) = f(x_0).$$

More generally, let $x_1, \ldots, x_n \in X$ be finitely many points in X; one can construct the probability measure

$$\delta = \frac{1}{n} \sum_{1 \leqslant i \leqslant n} \delta_{x_n}$$

such that

$$\int_X f(x) d\delta(x) = \frac{1}{n} \sum_{1 \leqslant i \leqslant n} f(x_i),$$

which is some kind of "sample sum" which is very useful in applications (since any "integral" which is really numerically computed is in fact a finite sum of a similar type).

Lemma 2.14 (Fatou's lemma.)**.** *Let (f_n) be a sequence of non-negative measurable functions $f_n : X \to [0, +\infty]$. We then have*

$$\int_X \left(\varliminf_{n \to +\infty} f_n \right) d\mu \leqslant \varliminf_{n \to +\infty} \int_X f_n d\mu,$$

and in particular, if $f_n(x) \to f(x)$ for all x, we have

$$\int_X f(x) d\mu(x) \leqslant \varliminf_{n \to +\infty} \int_X f_n(x) d\mu(x).$$

Proof. **1.** Denote $g = \underline{\lim} f_n$, $g_n = \inf_{k \geqslant n} f_k$. Then $g_n \leqslant g_{n+1}$, $g = \lim g_n$ and $g_n \leqslant f_n$. Moreover, $g_n \in S(E)$ by the theorem **??** on the limit of measurable functions. By monotonicity of the integral

$$\int_E g_n \leqslant \int_E f_n.$$

Consequently, by Levy theorem

$$\int_E g = \lim \int_E g_n = \underline{\lim} \int_E g_n \leqslant \underline{\lim} \int_E f_n.$$

2. This assertion follows from the first one since $f = \underline{\lim} f_n$ a.e. on E. □

Theorem 2.15 (Lebesgue dominated convergence theorem.)**.** *Let $(X, \mathcal{A}, \mu)$ be a measured space, $E \subset X$ and (f_n) be a sequence of complex-valued μ-integrable on E functions. Assume that, for almost every $x \in E$, we have*

$$f_n(x) \to f(x) \in \mathbb{C}$$

as $n \to +\infty$, so f is a complex-valued function on E. Then f is measurable; moreover, if there exists $\Phi(x) \in L^1(E, \mu)$ such that

$$|f_n(x)| \leqslant \Phi(x) \text{ for all } n \geqslant 1 \text{ for a.e. } x \in E,$$

then the limit function f is μ-integrable, and it satisfies

$$\int_E f(x) d\mu(x) = \lim_{n \to +\infty} \int_E f_n(x) d\mu(x).$$

In addition, we have in fact

$$\int_E |f_n - f| \, d\mu \to 0 \text{ as } n \to +\infty, \tag{8}$$

or in other words, f_n converges to f in $L^1(E, \mu)$ for the distance given by $d(f, g) = \|f - g\|_1$.

Proof. The following assertions hold simultaneously a.e.

1. $|f_n(x)| \leqslant \Phi(x)$ for every $n \in \mathbb{N}$;

2. $f_n(x) \to f(x)$.

Considering the limit in the first one we see that $|f(x)| \leqslant \Phi(x)$ for a.e. $x \in E$. Consequently, $f_n, f \in L(E, \mu)$ and by the Chebyshev inequality functions f_n, f, Φ are finite a.e. Noticing that $\Phi + f_n \geqslant 0$ a.e. and applying Fatu's theorem we obtain

$$\int_E \Phi + \int_E f = \int_E (\Phi + f) \leqslant \underline{\lim} \int_E (\Phi + f_n) = \\ = \underline{\lim} \left(\int_E \Phi + \int_E f_n \right) = \int_E \Phi + \underline{\lim} \int_E f_n.$$

Hence

$$\int_E f \leqslant \underline{\lim} \int_E f_n \tag{9}$$

Analogously,

$$\int_E \Phi - \int_E f = \int_E (\Phi - f) \leqslant \underline{\lim} \int_E (\Phi - f_n) = \\ = \int_E \Phi + \underline{\lim} \left(- \int_E f_n \right) = \int_E \Phi - \overline{\lim} \int_E f_n.$$

Consequently,

$$\overline{\lim} \int_E f_n \leqslant \int_E f. \tag{10}$$

Considering estimates (9) and (10) we see that

$$\int_E f \leqslant \underline{\lim} \int_E f_n \leqslant \overline{\lim} \int_E f_n \leqslant \int_E f.$$

This implies the existence of the limit $\lim \int_E f_n$ and that this limit is equal to the integral $\int_E f$.

Finally, to get (8) we notice that

$$|f_n(x) - f(x)| \leq 2g(x) \quad \text{and} \quad |f_n(x) - f(x)|$$

and applying the previous result to the sequence $|f_n(x) - f(x)|$ we see that

$$\int_E |f_n(x) - f(x)|\, d\mu \to 0,\ n \to +\infty.$$

$\square$

Example 1. Here is a first general example: assume that μ is a finite measure (for instance, a probability measure). Then the constant functions are in $L^1(\mu)$, and the domination condition holds, for instance, for any sequence of functions (f_n) which are uniformly bounded (over n) on X.

Example 2. Let $X_n \in \mathcal{A}$ be an increasing sequence of measurable sets such that

$$X = \bigcup_{n \geqslant 1} X_n$$

and let $Y_n = X \backslash X_n$ be the sequence of complementary sets. For any $f \in L(\mu)$, we have

$$\lim_{n \to +\infty} \int_{X_n} f(x) d\mu(x) = \int_X f(x) d\mu(x),$$
$$\lim_{n \to +\infty} \int_{Y_n} f(x) d\mu(x) = 0.$$

Proof. We apply the dominated convergence theorem to the sequence

$$f_n(x) = f(x) \chi_{X_n}(x).$$

Since $X_n \subset X_{n+1}$, it follows that, for any fixed $x \in X$, we have $f_n(x) = f(x)$ for all n large enough (but depending on x!). Thus $f_n(x) \to f(x)$ pointwise. Moreover, the domination condition is satisfied since

$$|f_n(x)| = |f(x)|\chi_{X_n}(x) \leqslant |f(x)|$$

for all x and all n and since, by assumption, $g = |f|$ is μ-integrable. Consequently, we obtain

$$\int_{X_n} f(x)d\mu(x) = \int_X f_n(x)d\mu(x) \to \int_X f(x)dx.$$

For the second statement, we just need to remark that

$$\int_{X_n} f(x)d\mu(x) + \int_{Y_n} f(x)d\mu(x) = \int_X f(x)d\mu(x)$$

for $n \geqslant 1$. □

For instance, consider $f \in L(X, \mu)$, and define

$$X_n = \{x \in X || f(x) | \leqslant n\}$$

for $n \geqslant 1$. We have $X_n \subset X_{n+1}$, of course, and

$$\bigcup_{n \geqslant 1} X_n = X.$$

We therefore conclude that

$$\int_X f(x)d\mu(x) = \lim_{n \to +\infty} \int_{X_n} f(x)d\mu(x) = \lim_{n \to +\infty} \int_{\{|f| \leqslant n\}} f(x)d\mu(x).$$

This is often quite useful because it shows that one may often prove general properties of integrals by restricting first to situations where the function that is integrated is bounded.

3 L^p-spaces.

3.1 Lebesgue space L^1

One of the most important use of integration theory is the construction of function spaces with excellent analytic properties, in particular completeness (i.e., convergence of Cauchy sequences). Indeed, we have the following property, which we will prove below: given a measure space $(X, \mathcal{A}, \mu)$, and a sequence (f_n) of integrable functions on X such that, for any $\varepsilon > 0$, we have

$$\|f_n - f_m\|_1 < \varepsilon$$

for all n, m large enough (a Cauchy sequence), there exists a limit function f, unique except that it may be changed arbitrarily on any μ-negligible set, such that

$$\lim_{n\to+\infty} \|f_n - f\| = 0$$

(but f is not usually a pointwise limit of the (f_n)).

Before considering this, we need to be involved in a bit of abstract manipulations in order to have structures which coincide exactly with the expected topological notions. In particular, note that the function $\|f\|_1$, as defined, does not have the property that $\|f\|_1 = 0$ implies that $f = 0$. The solution is to consider systematically functions defined "up to a perturbation defined on a set of measure zero".

Definition 3.1. *Let $(X, \mathcal{M}, \mu)$ be a measured space. The space $L^1(X, \mu) = L^1(\mu)$ is defined to be the quotient vector space*

$$\{f : X \to \mathbb{C} \mid f \textit{ is integrable } \}/N$$

where

$$N = \{f \mid f \text{ is measurable and } f \text{ is zero } \mu\text{-almost everywhere } \} = \left\{ f : \int_X |f| \, d\mu = 0 \right\}. \quad (11)$$

This is a normed vector space for the norm

$$\|f\|_1 = \int_X |f| d\mu$$

in particular $\|f\|_1 = 0$ *if and only if* $f = 0$ *in* $L^1(X, \mu)$*.*

From now on, the notation $L^1(\mu)$ and $L^1(X, \mu)$ refer exclusively to this definition

This means that an element of $L^1(X, \mu)$ can not properly be thought of as a function defined everywhere on X, the equivalence relation does not induce much difficulty. One usually works with "actual" functions f, and one remembers to say that $f = g$ if it is shown that f and g coincide except on a set of measure zero. Most importantly, in order to define a map

$$L^1(X, \mu) \xrightarrow{\phi} M,$$

for any set M, one usually defines a map on actual (integrable) functions, before checking that the value of $\phi(f)$ is the same as that of $\phi(g)$ if $f = g$ almost everywhere. This is, for instance, exactly why the norm is well-defined on $L^1(\mu)$.

On the other hand, fix a point x_0 and consider the map

$$\phi : f \mapsto f(x_0)$$

defined on all integrable functions on X.

If $\mu(\{x_0\}) = 0$, any two functions which differ only at x_0 are equal almost everywhere, and therefore the value of $\phi(f)$ is not well-defined on the equivalence classes which are the elements of $L^1(X,\mu)$. Hence, under this condition, one can not speak sensibly of the value of a function in $L^1(\mu)$ at x_0.

It will be quickly seen that manipulating elements of $L^1(\mu)$ is quite easy, and it is frequent to abuse notation by stating that they are "integrable functions on X". Also, one often has a function f which is constructed in a way which makes sense almost everywhere, say outside E_0, and is integrable on $X\backslash E_0$; then defining $f(x) = 0$ if $x \in E_0$, one obtains a well-defined element of $L^1(\mu)$.

Here is the first main result concerning the space L^1, which will lead quickly to its completeness property.

Theorem 3.2. *Let $(X, \mathcal{A}, \mu)$ be a measured space, and let (f_n) be a function of elements $f_n \in L^1(\mu)$. Assume that the series $\sum f_n$ is normally convergent in $L^1(\mu)$, i.e., that*

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

Then the series

$$\sum_{n=1}^{\infty} f_n(x)$$

converges almost everywhere in X, and if $f(x)$ denotes its limit, well-defined almost everywhere, we have $f \in L^1(\mu)$ and

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

Moreover the convergence is also valid in the norm $\|\cdot\|_1$*, i.e., we have* $\left\|\sum_{k=1}^{n} f_k - f\right\| \to 0$ *as* $n \to +\infty$.

Proof. Define

$$h(x) = \sum_{n=1}^{\infty} |f_n(x)| \in [0, +\infty]$$

for $x \in X$. Notice that if the values of f_n are modified on a set Y_n of measure zero, then h is changed, at worse, on the set

$$Y = \bigcup_{n=1}^{\infty} Y_n$$

which is still of measure zero. This indicates that h is well-defined almost everywhere, and it is measurable. It is also non-negative, and hence its integral is well-defined in $[0, +\infty]$. By the monotone convergence theorem and assumption, we have

$$\int_X h d\mu = \sum_{n=1}^{\infty} \int_X |f_n| \, d\mu = \sum_{n \geqslant 1} \|f\|_1 < +\infty,$$

and, in particular, h is finite almost everywhere.

For any x such that $h(x) < \infty$, the numeric series $\sum f_n(x)$ converges absolutely, and hence has a sum $f(x)$ such that $|f(x)| \leqslant h(x)$. We can extend f to X by defining, for instance, $f(x) = 0$ on the negligible set (say, Z) of those x where $h(x) = +\infty$.

In any case, we have $|f| \leqslant h$, and therefore the above inequality implies that $f \in L^1(\mu)$. We can now apply the dominated convergence theorem to the sequence of partial sums

$$u_N(x) = \sum_{n=1}^{N} f_n(x), \quad x \notin Z, \quad u_N(x) = 0, \quad x \in Z,$$

which satisfy $u_N(x) \to f(x)$ for all x and $|u_N(x)| \leqslant h(x)\chi_Z(x)$ for all $N \in \mathbb{N}$ and $x \in X$. We conclude that

$$\int_X f d\mu = \lim_{N\to+\infty} \int_X u_N d\mu = \lim_{N\to+\infty} \sum_{n=1}^{N} \int_X f_n d\mu = \sum_{n=1}^{\infty} \int_X f_n d\mu.$$

□

Theorem 3.3 (Completeness of the space L^1.). *Let $(X, \mathcal{A}, \mu)$ be a measured space.*

1. *Let (f_n) be a Cauchy sequence in $L^1(\mu)$ then it is convergent, there exists a unique $f \in L^1(\mu)$ such that $f_n \to f$ in $L^1(\mu)$, i.e., such that*

$$\lim_{n\to+\infty} \|f_n - f\|_1 = 0.$$

In other words, the normed vector space $L^1(\mu)$ is complete (and is, consequently, Banach space).

2. *Moreover, there exists a subsequence (f_{n_k}) of (f_n) such that*

$$\lim_{k\to+\infty} f_{n_k}(x) = f(x)$$

for μ-almost all x.

Proof. We recall the Cauchy condition: for any $\varepsilon > 0$, there exists $N(\varepsilon) \geqslant 1$ such that

$$\|f_n - f_m\|_1 < \varepsilon$$

for all $n \geqslant N$ and $m \geqslant N$.

Consider successively $\varepsilon = \varepsilon_k = 2^{-k}$ for $k \geqslant 1$. Then, using induction on k, we see that there exists a strictly increasing sequence (n_k) such that

$$\left\| f_{n_{k+1}} - f_{n_k} \right\|_1 < 2^{-k}$$

for $k \geqslant 1$. We claim that $\left(f_{n_k}\right)$ converges almost everywhere. Indeed, the series

$$\sum_{k=1}^{\infty} \left(f_{n_{k+1}} - f_{n_k}\right)$$

is a series with terms $g_k = f_{n_{k+1}} - f_{n_k} \in L^1(\mu)$ such that

$$\sum_{k \geqslant 1} \|g_k\|_1 \leqslant \sum_{k=1}^{\infty} 2^{-k} < +\infty.$$

Then, by Theorem 3.2 this series converges almost everywhere, and in $L^1(\mu)$, to a function $g \in L^1(\mu)$. But since the partial sums are given by the telescoping sums

$$\sum_{k=1}^{K} g_k = \left(f_{n_{K+1}} - f_{n_K}\right) + \left(f_{n_K} - f_{n_{K-1}}\right) + \cdots + \left(f_{n_2} - f_{n_1}\right) = f_{n_{K+1}} - f_{n_1},$$

this means that the subsequence $\left(f_{n_k}\right)$ converges almost everywhere, with limit $f = g + f_{n_1}$.

Moreover, since the convergence is valid in L^1, we have

$$\lim_{k \to +\infty} \left\| f - f_{n_k} \right\|_1 = 0$$

which means that f is a limit point of the Cauchy sequence (f_n). However, it a Cauchy sequence with a limit point, in a metric space, converges necessarily to this limit point. Here is the argument for completeness: for all n and k, we have

$$\|f - f_n\|_1 \leqslant \left\|f - f_{n_k}\right\|_1 + \left\|f_{n_k} - f_n\right\|_1,$$

and we know that $\left\|f_n - f_{n_k}\right\|_1 < \varepsilon$ if n and n_k are both $\geqslant N(\varepsilon)$. Moreover, $\left\|f - f_{n_k}\right\|_1 < \varepsilon$ for $k > K(\varepsilon)$.

Now fix some $k > K(\varepsilon)$ such that, in addition, we have $n_k > N(\varepsilon)$ (this is possible since n_k is strictly increasing). Then, for all $n > N(\varepsilon)$, we obtain $\|f - f_n\|_1 < 2\varepsilon$, and this proves the convergence of (f_n) towards f in $L^1(\mu)$.

□

3.2 Lebesgue spaces L^p with $1 \leq p < \infty$.

Definition 3.4. *Let $(X, \mathcal{A}, \mu)$ be a measure space and let $p \in [1, +\infty)$ be a real number. The $L^p(X, \mu) = L^p(\mu)$ is defined to be the quotient vector space*

$$\{f : X \to \mathbb{C} : |f|^p \text{ is integrable } \} / N$$

where N is defined as in (11). This space is equiped with the norm

$$\|f\|_p = \left(\int_X |f|^p \, d\mu\right)^{1/p}.$$

The first thing to do is to check that the definition actually makes sense, i.e., that the set of functions f where $|f|^p$ is integrable is a vector space, and then that $\|\cdot\|_p$ defines a norm on $L^p(\mu)$, which is not entirely obvious (in contrast with the case $p = 1$). For this, we need some inequalities which are also quite useful for other purposes.

First, we recall that a function $\varphi : I \to [0, +\infty)$ defined on an interval I is said to be convex if it is such that

$$\varphi(t_1 x_1 + t_2 x_2) \leqslant t_1 \varphi(x_1) + t_2 \varphi(x_2) \tag{12}$$

for any $x_1, x_2 \in I$ and any real numbers, $t_1, t_2 \geqslant 0$ such that $t_1 + t_2 = 1$. If φ is twice-differentiable with continuous derivatives then φ is convex on an open interval if and only if $\varphi'' \geqslant 0$.

By induction, we can see that

$$\varphi(t_1 x_1 + t_2 x_2 + \ldots + t_n x_n) \leqslant t_1 \varphi(x_1) + t_2 \varphi(x_2) + \ldots t_n \varphi(x_n) \tag{13}$$

for any $x_1, x_2, \ldots, x_n \in I$ and any real numbers $t_1, t_2, \ldots, t_n \geqslant 0$ such that $t_1 + t_2 + \ldots + t_n = 1$.

Lemma 3.5 (Jensen, Hölder and Minkowski inequalities)**.** *Let $(X, \mathcal{A}, \mu)$ be a fixed measure space.*

1. *Assume μ is a probability measure. Then, for any function*

$$\varphi : [0, +\infty) \to [0, +\infty)$$

which is non-decreasing, continuous and convex, and any measurable function

$$f : X \to [0, +\infty)$$

we have Jensen's inequality:

$$\varphi\left(\int_X f(x) d\mu(x)\right) \leqslant \int_X \varphi(f(x)) d\mu(x) \tag{14}$$

with the convention $\varphi(+\infty) = +\infty$.

2. *Let $p > 1$ be a real number and let $q > 1$ be the "dual" real number such that $p^{-1} + q^{-1} = 1$. Then for any measurable functions*

$$f, g : X \to [0, +\infty]$$

we have Hölder's inequality:

$$\int_X fg d\mu \leqslant \left(\int_X f^p d\mu \right)^{1/p} \left(\int_X g^q d\mu \right)^{1/q} = \|f\|_p \|g\|_q. \qquad (15)$$

3. *Let $p > 1$ be any real number. Then for any measurable functions*

$$f, g : X \to [0, +\infty]$$

we have Minkowski's inequality:

$$\|f + g\|_p = \left(\int_X (f+g)^p d\mu \right)^{1/p} \leqslant$$
$$\left(\int_X f^p d\mu \right)^{1/p} + \left(\int_X g^p d\mu \right)^{1/p} = \|f\|_p + \|g\|_p \quad (16)$$

Proof. **1.** First, assume that f is not constant, otherwise, the inequality is obvious. **Case 1.** Consider first the case when f is a simple function

$$f = \sum_{k=1}^{N} c_k \chi_{A_k}$$

for a disjoint family A_k such that $\bigcup\limits_{k=1}^{N} A_k = X$.

Then since $\sum_{k=1}^{N} \mu(A_k) = 1$ we see that

$$\varphi\left(\int_X f(x)d\mu(x)\right) = \varphi\left(\sum_{k=1}^{N} c_k\mu(A_k)\right) \leq \sum_{k=1}^{N} \varphi\left(c_k\right)\mu(A_k) =$$
$$\leqslant \int_X \varphi(f(x))d\mu(x)$$

Next, for an arbitrary $f \geqslant 0$, we select a sequence of step functions (s_n) which is non-decreasing and converges pointwise to f. By the monotone convergence theorem, we have

$$\int s_n d\mu \to \int f d\mu$$

and similarly, since φ is itself non-decreasing, we have

$$\varphi\left(s_n(x)\right) \to \varphi(f(x))$$

and $\varphi \circ s_n \leqslant \varphi \circ s_{n+1}$, so that

$$\int \varphi\left(s_n(x)\right) d\mu \to \int \varphi(f(x))d\mu$$

by the monotone convergence theorem again.

Finally, the continuity of φ and Jensen's inequality for step functions implies that

$$\varphi\left(\int_X f d\mu\right) = \lim_{n\to+\infty} \varphi\left(\int_X s_n d\mu\right) \leqslant \lim_{n\to+\infty} \int_X \varphi\left(s_n(x)\right) d\mu = \int_X \varphi(f(x))d\mu.$$

2. Recall that if $1/p + 1/q = 1$ then by Young's inequality

$$xy \leqslant \frac{x^p}{q} + \frac{y^q}{q}, \text{ for } x, y \geqslant 0.$$

Assuming this, we continue as follows to prove (15). It is enough to consider the cases when $|f|^p$ and $|g|^q$ are integrable (otherwise, the result is trivial). Replacing f with $f/\|f\|_p$ and g with $g/\|g\|_q$, homogeneity shows that it is enough to prove the inequality when $\|f\|_p = \|g\|_q = 1$. But in that case, we use the pointwise Young inequality

$$f(x)g(x) \leqslant \frac{f(x)^p}{p} + \frac{g(x)^q}{q}.$$

We integrate to obtain

$$\int_X f(x)g(x)d\mu \leqslant \frac{\|f\|_p}{p} + \frac{\|f\|_q}{q} = 1.$$

3. The inequality is obvious if either $\|f\|_p = +\infty$ or $\|g\|_p = +\infty$. In the remaining case, we first note that

$$(f+g)^p \leqslant (2\max(f,g))^p \leqslant 2^p\left(f^p + g^p\right)$$

so that we also get $\|f+g\|_p < +\infty$. Now, we consider the auxiliary function

$$h = (f+g)^{p-1},$$

which has the property that $(f+g)^p = fh + gh$, so applying Hölder's inequality (15), we derive

$$\int_X (f+g)^p d\mu = \int_X fhd\mu + \int_X ghd\mu \leqslant \|f\|_p \|h\|_q + \|g\|_p \|h\|_q .$$

But

$$\|h\|_q = \left(\int_X (f+g)^{q(p-1)}d\mu\right)^{1/q} = \|f+g\|_p^{p/q} \text{ since } q(p-1) = p$$

Moreover, $1 - 1/q = 1/p$, hence (considering separately the case where h vanishes almost everywhere) the last inequality gives

$$\|f+g\|_p = \left(\int_X |f+g|^p \, d\mu\right)^{p(1-1/q)} \leqslant \|f\|_p + \|g\|_p$$

after dividing by $\|h\|_q = \|f+g\|_p^{p/q} \in (0, +\infty)$. □

Remark 3.6. *Consider the space* $(\mathbb{R}, \mathcal{B}, \mu)$. *Consider the function* f *defined by*

$$f(x) = \frac{\sin(x)}{x} \text{ if } x \neq 0 \text{ and } f(0) = 1$$

is in $L^2(\mathbb{R})$ *(because its square decays faster than* $1/x^2$ *for* $|x| \geqslant 1$ *and is bounded for* $|x| \leqslant 1$ *), although it is not in* $L^1(\mathbb{R})$. *In particular, note that a function in* $L^p, p \neq 1$, *may well have the property that it is not integrable. This is one of the main ways to go around restriction to absolute convergence in Lebesgue's definition: many properties valid in* L^1 *are still valid in* L^p, *and can be applied to functions which are not integrable.*

Theorem 3.7 (Completeness of L^p-spaces)**.** *Let* $(X, \mathcal{A}, \mu)$ *be a fixed measure space.*

1. *Let* $p \in (1, +\infty)$ *be a real number,* (f_n) *a sequence of elements* $f_n \in L^p(\mu)$. *If*

$$\sum_{n \geqslant 1} \|f_n\|_p < +\infty,$$

the series

$$\sum_{n \geqslant 1} f_n$$

converges almost everywhere, and in L^p-norm, to a function $g \in L^p(\mu)$.

2. *For any $p \geqslant 1$, the space $L^p(\mu)$ is a Banach space for the norm $\|\cdot\|_p$; more precisely, for any Cauchy sequence (f_n) in $L^p(\mu)$, there exists $f \in L^p(\mu)$ such that $f_n \to f$ in $L^p(\mu)$, and in addition there exists a subsequence $\left(f_{n_k}\right), k \geqslant 1$, such that*

$$\lim_{k \to +\infty} f_{n_k}(x) = f(x) \text{ for almost all } x \in X.$$

3. *In particular, for $p = 2$, $L^2(\mu)$ is a Hilbert space with respect to the inner product*

$$\langle f, g \rangle = \int_X f(x)\overline{g(x)} d\mu.$$

Proof. Assuming the first assertion is proved, the same argument used to prove Lemma 3.3 can be copied line by line, replacing every occurrence of $\| \cdot \|_1$ with $\| \cdot \|_p$, proving second assertion.

In order obtain the proof of the firs assertion, we also follow the proof of Lemma 3.3. Let

$$h(x) = \sum |f_n(x)| \in [0, +\infty]$$

for $x \in X$, and observe that (by continuity of $y \mapsto y^p$ on $[0, +\infty]$) we also have

$$h(x)^p = \lim_{N \to +\infty} \left(\sum_{n=1}^{N} |f_n(x)| \right)^p.$$

This expresses h^p as a non-decreasing limit of non-negative functions, and therefore by the monotone convergence theorem, and the triangle inequality, we derive

$$\left(\int_X h(x)^p d\mu(x)\right)^{1/p} = \lim_{N\to+\infty}\left(\int_X \left(\sum_{n=1}^{N}|f_n(x)|\right)^p d\mu(x)\right)^{1/p} = \lim_{N\to+\infty}\left\|\sum_{n=1}^{N}|f_n|\right\|_p \leqslant \sum_{n=1}^{\infty}\|f_n\|_p < +\infty,$$

by assumption. It follows that h^p, and also h, is finite almost everywhere. This implies (as was the case for $p = 1$) that the series

$$f(x) = \sum_{n\geqslant 1} f_n(x)$$

converges absolutely almost everywhere. Since $|f(x)| \leqslant h(x)$, we have $f \in L^p(\mu)$, and since

$$\left\|f - \sum_{n=1}^{N} f_n\right\|_p = \left\|\sum_{n=N}^{\infty} f_n\right\|_p \leqslant \sum_{n=N}^{\infty}\|f_n\|_p \to 0,$$

the convergence is also valid in L^p. □

3.3 Essentially bounded function. Space L^∞.

Definition 3.8. *A measurable function*

$$f : X \to \mathrm{C}$$

is said to be essentially bounded by $M \geqslant 0$ if

$$\mu(\{x||f(x)\,|> M\}) = 0. \tag{17}$$

Lemma 3.9. *Let f be a measurable function on X and let*

$$\|f\|_\infty = \inf\{M : \ f \textit{ is essentially bounded by } M\} \in [0, +\infty] \tag{18}$$

Then f is essentially bounded by $\|f\|_\infty$, or in other words the infimum is attained. Moreover, the quotient vector space

$$L^\infty(\mu) = \{f \mid \|f\|_\infty < +\infty\} /N,$$

where N is the subspace (11) of measurable functions vanishing almost everywhere, is a normed vector space with norm $\|\cdot\|_\infty$.

If $f \in L^1(\mu)$ and $g \in L^\infty(\mu)$, we have $fg \in L^1(\mu)$ and

$$\int_X |fg| d\mu \leqslant \|f\|_1 \|g\|_\infty. \tag{19}$$

The last inequality should be thought of as the analogue of Hölder's inequality for the case $p = 1, q = +\infty$.

Note that the obvious inequality $\|f\|_\infty \leqslant \sup\{f(x)\}$ is not, in general, an equality, even if the supremum is attained. For instance, if we consider $([0, 1], \mathcal{B}, \lambda)$, and take $f(x) = x\chi_{\mathbb{Q}}(x)$, we find that $\|f\|_\infty = 0$, although the function f has maximum equal to 1 on $[0, 1]$.

The definition of the L^∞ norm is most commonly used as follows: we have

$$|f(x)| \leqslant M$$

μ-almost everywhere, if and only if

$$M \geqslant \|f\|_\infty$$

Proof of Lemma 3.9. We must check that $M = \|f\|_\infty$ satisfies the condition (17) when $M < +\infty$ (the case $M = +\infty$ is obvious). To prove this, we note that there exists by definition a sequence (M_n) of real numbers such that

$$M_{n+1} \leqslant M_n, \quad M_n \to M,$$

and M_n satisfies the condition (17). Then

$$\mu(\{x || f(x) |> M\}) = \mu\left(\bigcup_n \{x || f(x) |> M_n\}\right) = 0,$$

as the measure of a countable union of sets of measure zero.

It is also immediate that $\|f\|_\infty = 0$ is equivalent with f being zero μ-almost everywhere, since the definition becomes

$$\mu(\{x \mid f(x) \neq 0\}) = \mu(\{x || f(x) |> 0\}) = 0.$$

Since all the other axioms defining a normed vector space can be checked very easily, there only remains to prove (19). However, this is clear by monotony by integrating the upper bound

$$|fg| \leqslant \|g\|_\infty |f|$$

which, by the above, is valid except on a set of measure zero. □

Lemma 3.10 (Completeness of L^∞)**.** *Let $(X, \mathcal{A}, \mu)$ be a measure space.*

1. *Assume that (f_n) is a sequence of measurable functions on X with $f_n \in L^\infty(\mu)$. If*

$$\sum \|f_n\|_\infty < +\infty$$

the series

$$\sum_{n=1}^{\infty} f_n$$

converges μ-almost everywhere, and in L^∞, to function $g \in L^\infty(\mu)$.

2. *The space $L^\infty(\mu)$ is a Banach space. More precisely, for any Cauchy sequence (f_n) in $L^\infty(\mu)$, there exists $f \in L^\infty(\mu)$ such that $f_n \to f$ in $L^\infty(\mu)$, and the convergence also holds almost everywhere.*

Proof. **1.** The method is the same as the one used before. Let

$$h(x) = \sum_{n \geqslant 1} |f_n(x)|\,, \quad g(x) = \sum_{n \geqslant 1} f_n(x)$$

so that the assumption quickly implies that both series converge almost everywhere. Since

$$|g(x)| \leqslant h(x) \leqslant \sum_{n=1}^{\infty} \|f_n\|_\infty$$

almost everywhere, it follows that $g \in L^\infty(\mu)$. Finally

$$\sum f_n = g$$

with convergence in the space L^∞, since

$$\left\| g - \sum_{n=1}^{N} f_n \right\|_\infty = \left\| \sum_{n=N+1}^{\infty} f_n \right\|_\infty \leqslant \sum_{n=N+1}^{\infty} \|f_n\|_\infty \to 0.$$

2. Let (f_n) be a Cauchy sequence in $L^\infty(\mu)$. Then

$$|f_n(x) - f_m(x)| \leqslant \|f_n - f_m\|_\infty$$

for any fixed n and m, and for almost all $x \in X$. To see this consider $A_{n,m}$ to be the exceptional set (of measure zero) such that the inequality holds outside $A_{n,m}$. Let A be the union over n, m of the $A_{n,m}$. Since this is a countable union, we still have $\mu(A) = 0$.

Now, for all $x \notin A$, the sequence $(f_n(x))$ is a Cauchy sequence in C, and hence it converges, say to an element $f(x) \in$ C. The resulting function f (extended to be zero on A, for instance) is of course measurable. Now we need to check that f is in L^∞.

For this, note that (by the triangle inequality) the sequence $(\|f_n\|_\infty)$ of the norms of f_n is itself a Cauchy sequence in $\mathbb{R}$. Let $M \geqslant 0$ be its limit, and let B_n be the set of x for which $|f_n(x)| > \|f_n\|_\infty$, which has measure zero. Again, the union B of all B_n has measure zero, and so has $A \cup B$.

Now, for $x \notin A \cup B$, we have

$$|f_n(x)| \leqslant \|f_n\|_\infty$$

for all $n \geqslant 1$, and

$$f_n(x) \to f(x).$$

For such x, it follows by letting $n \to +\infty$ that $|f(x)| \leqslant M$, and consequently we have shown that $|f(x)| \leqslant M$ almost everywhere. This gives $\|f\|_\infty \leqslant M < +\infty$.

Finally, we show that (f_n) converges to f in L^∞ (convergence almost everywhere is already established). Fix $\varepsilon > 0$, and then let N be such that

$$\|f_n - f_m\|_\infty < \varepsilon$$

when $n, m \geqslant N$. We obtain $|f_n(x) - f_m(x)| < \varepsilon$ for all $x \notin A \cup B$ and all $n, m \geqslant N$. Taking any $m \geqslant N$, and letting $n \to +\infty$, we get

$$|f(x) - f_m(x)| < \varepsilon \text{ for almost all } x \in X,$$

and this means that $\|f - f_m\|_\infty < \varepsilon$ when $m > N$. Consequently, we have shown that $f_n \to f$ in $L^\infty(\mu)$. $\square$

3.4 Comparison of Lebesgue spaces

Remark 3.11. *Usually, there is no obvious relation between the various spaces $L^p(\mu)$, $p \geqslant 1$. For instance, consider $X = \mathrm{R}$ with the Lebesgue measure. The function $f(x) = \inf(1, 1/|x|)$ is in L^2, but not in L^1, whereas $g(x) = x^{-1/2}\chi_{[0,1]}$ is in L^1, but not in L^2. Other examples can be given for any choices of p_1 and p_2, we never have $L^{p_1}(\mathbb{R}) \subset L^{p_2}(\mathbb{R})$*

Theorem 3.12 (Comparison of L^p spaces for finite measure spaces.)**.** *Let $(X, \mathcal{A}, \mu)$ be a measured space with $\mu(X) < +\infty$. For any p_1, p_2 with*

$$1 \leqslant p_2 \leqslant p_1 \leqslant +\infty,$$

there is a continuous inclusion map

$$L^{p_1}(\mu) \hookrightarrow L^{p_2}(\mu),$$

and indeed

$$\|f\|_{p_2} \leqslant \mu(X)^{1/p_2 - 1/p_1} \|f\|_{p_1}.$$

Proof. We need only prove the inequality as stated for $f \geqslant 0$. We use Hölder's inequality for this purpose: for any $p \in [1, +\infty]$, with dual exponent q, we have

$$\int_X f^{p_2} d\mu = \int_X f^{p_2} \cdot 1 d\mu \leqslant \left(\int_X f^{pp_2} d\mu\right)^{1/p} \left(\int_X d\mu(x)\right)^{1/q}$$

and if we pick $p = p_1/p_2 \geqslant 1$, with $q^{-1} = p_2/p_1 - 1$, we obtain

$$\int_X f^{p_2} d\mu \leqslant \mu(X)^{p_2/p_1 - 1} \|f\|_{p_1}^{p_2}$$

and taking the $1/p_2$-th power, this gives

$$\|f\|^{p_2} \leqslant \mu(X)^{1/p_1 - 1/p_2} \|f\|_{p_1},$$

as claimed. $\square$

4 Integration with respect to Lebesgue measure

In this section, we consider the special case of the measure space $(\mathbb{R}^n, \mathcal{A}_n, \mu)$, where μ is the Lebesgue measure on a Lebesgue $\sigma-$algebra $\mathcal{A}_n$. We will see basic examples of μ-integrable functions (and of functions which are not), and will explain the relation between this integral and the Riemann integral in the case of sufficiently smooth functions.

EXAMPLE 1 (Integrable functions). In order to prove that a given measurable function f is integrable (with respect to the Lebesgue measure on a subset $X \subset \mathbb{R}^n$, or for a more general measure space), the most common technique is to find a "simple" comparison function g which is known to be integrable and for which it is known that

$$|f(x)| \leqslant g(x), \quad x \in X.$$

Frequently, one uses more than one comparison function: for instance, one finds disjoint subsets X_1, X_2 such that $X = X_1 \cup X_2$, and functions

g_1, g_2 integrable on X_1 and X_2 respectively, such that

$$|f(x)| \leqslant \begin{cases} g_1(x) & \text{if } x \in X_1, \\ g_2(x) & \text{if } x \in X_2. \end{cases}$$

With some care, this can be applied with infinitely many subsets. For instance, consider

$$X = \big[1, +\infty\big), \quad f(x) = x^{-\nu}$$

where $\nu \geqslant 0$. Then f is μ-integrable on X if (in fact, only if) $\nu > 1$. Indeed, note that

$$0 \leqslant f(x) \leqslant n^{-\nu}, \quad x \in [n, n+1), \quad n \geqslant 1,$$

and therefore, using the monotone convergence theorem, we get

$$\int_X x^{-\nu} d\mu(x) \leqslant \sum_{n \geqslant 1} n^{-\nu} < +\infty$$

if $\nu > 1$ (the converse is proved in the next example). Note that the range of integration here, the unbounded interval $[0, +\infty)$, is treated just like any other; there is no distinction in principle between bounded and unbounded intervals in Lebesgue's theory, like there was in Riemann's integral. However,

$$\int_{[0,+\infty)} f(x) d\mu(x) = \lim_{n \to +\infty} \int_{[0,n]} f(x) d\mu(x)$$

for any μ-integrable function on X (Exercise: explain this).

Similarly, Lebesgue's integral deals uniformly with unbounded functions (either on bounded or unbounded intervals). Consider $X = [0,1]$ and

$$f(x) = \begin{cases} x^{-\nu} & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$$

for some fixed $\nu \geqslant 0$. We now note that

$$0 \leqslant f(x) \leqslant \left(\frac{1}{n+1}\right)^{-\nu} = (n+1)^{\nu}, \quad x \in \big((n+1)^{-1}, n^{-1}\big], \quad n \geqslant 1,$$

and summing over $n \geqslant 1$, keeping in mind the size

$$\frac{1}{n} - \frac{1}{n+1} = \frac{1}{n(n+1)} \leqslant \frac{2}{(n+1)^2}$$

of each interval in the subdivision, we obtain

$$\int_{[0,1]} x^{-\nu} d\mu(x) \leqslant 2 \sum_{n \geqslant 1} (n+1)^{\nu-2}$$

which is finite for $\nu - 2 < -1$, i.e., for $\nu < 1$. One can also deal easily with functions with singularities located at arbitrary points of $\mathbb{R}$, without requiring to place them at the extremities of various intervals.

Another very common comparison function, that may be applied to any set with finite measure, is a constant function: for $X = [a,b]$ for instance, any measurable bounded function is μ-integrable on X, because the constant function 1 has integral $\leqslant (b-a)$.

EXAMPLE 2. Non-integrable functions. Except for the requirement of measurability, which is very rarely an issue in practice, the condition of integrability of a function f on a measurable subset $X \subset \mathbb{R}$

with respect to μ, is analogue to the condition of absolute convergence for series

$$\sum_{n\geqslant 1} a_n, \quad a_n \in \mathbb{C}.$$

As such, corresponding to the existence of series like

$$\sum_{n\geqslant 1} \frac{(-1)^{n+1}}{n}$$

which are convergent but not absolutely so, there are many examples of functions f which are not integrable although some specific approximating limits may exist. For instance, consider

$$f(x) = \begin{cases} \frac{\sin(x)}{x} & \text{if } x > 0, \\ 1 & \text{if } x = 0, \end{cases}$$

defined for $x \in [0, +\infty)$. It is well-known that

$$\lim_{B\to+\infty} \int_{[0,B]} f(x)d\mu(x) = \frac{\pi}{2}, \tag{20}$$

which means that the integral of f on $[0, +\infty)$ exists in the sense of "improper" Riemann integrals (see also the next example; f is obviously integrable on each interval $[0, B]$ because it is bounded by 1 there). However, f is not in $L^1(\mathbb{R}, \mu)$. Indeed, note that

$$f(x) \geqslant \frac{\sqrt{2}}{2x} \geqslant \frac{\sqrt{2}}{2(2n+1)\pi}$$

whenever $x \in I_n \left[\frac{\pi}{2} + 2n\pi - \frac{\pi}{4}, \frac{\pi}{2} + 2n\pi + \frac{\pi}{4}\right]$, $n \geqslant 1$, so that

$$\int_{[0,+\infty)} f(x)d\mu(x) \geqslant \frac{\sqrt{2}}{2\pi} \sum_{n=1}^{+\infty} \frac{1}{2n+1} = +\infty$$

which proves the result.

If one managed to change the definition of integral further so that f becomes integrable on $[0, +\infty)$ with integral $\pi/2$, we would have to draw the undesirable conclusion that the restriction of an integrable function to some subset of its domain may fail to be integrable. Avoiding this is one of the main reasons for using a definition related to absolute convergence. This is not to say that formulas like (20) have no place in analysis: simply, they have to be studied separately, and should not be thought of as being truly analogous to a statement of integrability.

EXAMPLE 3. Comparison with the Riemann integral. We have already seen examples of μ-integrable step-functions which are not Riemann-integrable. However, we now want to discuss the relation between the two notions in the case of sufficiently regular functions. For clarity, we denote here

$$(L)\int_{[a,b]} f(x)d\mu(x)$$

the Riemann integral, and use

$$(R)\int_{a}^{b} f(x)dx$$

for the Lebesgue integral.

Theorem 4.1. *Let $I = [a, b]$ be a compact interval, and let*

$$f : I \to \mathbb{R}$$

be a measurable function which is Riemann-integrable. Then f is μ-integrable, and we have the equality

$$(R)\int_{a}^{b} f(x)dx = (L)\int_{[a,b]} f(x)d\mu(x). \tag{21}$$

This already allows us to compute many Lebesgue integrals by using known Riemann integral identities.

Proof. The proof of this fact is very simple: for any partition

$$a = x_0 < x_1 < \cdots < x_n = b$$

of the interval of integration, let

$$S_+(f) = \sum_{i=0}^{n-1} (x_i - x_{i-1}) \sup_{x_i \leqslant x \leqslant x_{i+1}} f(x) = \sum_{i=0}^{n-1} M_i \Delta x_i,$$
$$S_-(f) = \sum_{i=0}^{n-1} (x_{i+1} - x_i) \min_{x_i \leqslant x \leqslant x_{i+1}} f(x) = \sum_{i=0}^{n-1} m_i \Delta x_i,$$

be the upper and lower Darboux's sums (see MA(2)) for the integral of f. We can immediately see that

$$S_-(f) \leqslant \underbrace{\sum_{i=0}^{n-1} (L) \int_{[y_i, y_{i+1}]} f(x) d\mu(x)}_{=(L) \int_{[a,b]} f(x) d\mu(x)} \leqslant S_+(f),$$

and hence equality (21) is an immediate consequence of the definition of the Riemann integral. □

This applies, in particular, to any function f which is either continuous or piecewise continuous.

One can also study general Riemann-integrable functions, although this requires a bit of care with measurability. The following result indicates clearly the restriction that Riemann's condition imposes:

Theorem 4.2. *Let $f : [a,b] \to \mathbb{C}$ be any function. Then f is Riemann-integrable if and only if f is bounded, and the set of points where f is not continuous is negligible with respect to the Lebesgue measure.*

Theorem 4.3. *Let $I = [a, +\infty)$ and let $f : I \to \mathbb{C}$ be such that the improper Riemann integral*

$$\int_a^{+\infty} f(x)dx$$

converges absolutely. Then $f \in L^1(I, \mu)$ and

$$\int_a^{+\infty} f(x)dx = \int_I f(x)d\mu(x)$$

Proof. Let

$$X_n = [a, a+n)$$

for $n \geqslant 1$ and $f_n = f\chi_{X_n}$. Then

$$|f_n| \leqslant |f_{n+1}| \to |f|$$

and by Levy monotone convergence theorem

$$\int_I |f(x)|d\mu(x) = \lim_n \int_I |f_n|\, d\mu(x) = \lim_n \int_a^{a+n} |f(x)|dx = \int_a^{+\infty} |f(x)|dx < +\infty,$$

where we used the monotone convergence theorem, and the earlier comparison (21) together with the assumption on the Riemann-integrability of f. It follows that f is Lebesgue-integrable, and, hence, we get

$$\int_I f(x)d\mu(x) = \int_a^{+\infty} f(x)dx$$

using the definition of the Riemann integral again. □

Remark 4.4. *A similar argument applies to an unbounded function defined on an interval $[a, b)$ such that the improper Riemann integral*

$$\int_a^b f(x)dx$$

converges absolutely.

4.1 The Multiple Lebesgue Integral

In this section, we consider a few properties of the integral with respect to the Lebesgue measure on a multi-dimensional space. As in the previous section, the integral with respect to the Lebesgue measure is called the Lebesgue integral and is denoted by $\int_E f(x)dx$ by analogy with the one-dimensional case. The Lebesgue measure itself is usually denoted by μ, without indicating the dimension.

Note that the integrals with respect to the planar, three-dimensional, and m-dimensional Lebesgue measures are usually called the double, triple, and m-multiple integrals, respectively, and are often conveniently denoted by the symbols $\iint$, $\iiint$, and $\int \cdots \int$.

The theorem below deals with a power of the norm, which in many cases serves as a reference function with which one compares other functions when studying their integrability.

Theorem 4.5. *Let B be a ball in $\mathbb{R}^m$ of radius r centered at a point $a \in \mathbb{R}^m$. Given $q > 0$, set*

$$f(x) = \frac{1}{\|x - a\|^q}$$

for $x \in \mathbb{R}^m \backslash \{a\}$. *Then:*

1. $f \in L^1(B)$ *if and only if* $q < m$;

2. $f \in L^1(\mathbb{R}^m \backslash B)$ *if and only if* $q > m$.

Proof. First recall that the volume (m-dimensional Lebesgue measure) of an m dimensional ball of radius R is equal to $\alpha_m R^m$, where α_m is the volume of a ball of unit radius. Hence the volume of the spherical layer

$$E(R) = \left\{ x \in \mathbb{R}^m \,\middle|\, \frac{R}{2} \leqslant \|x - a\| < R \right\}$$

is equal to

$$\mu(E(R)) = \alpha_m R^m - \alpha_m \left(\frac{R}{2}\right)^m = \alpha_m \left(2^m - 1\right) \left(\frac{R}{2}\right)^m = \beta_m R^m,$$

where $\beta_m = \alpha_m \left(1 - 2^{-m}\right)$.

Now consider the partition of the ball B into the spherical layers

$$B = \{a\} \bigcup \bigcup_{k \geqslant 1} E_k, \quad E_k = E\left(\frac{r}{2^k}\right).$$

Then

$$\mu\left(E_k\right) = \beta_m \left(\frac{r}{2^k}\right)^m \qquad \text{for all } k \in \mathbb{N}.$$

Furthermore,

$$\left(\frac{2^{k-1}}{r}\right)^q \leqslant f(x) \leqslant \left(\frac{2^k}{r}\right)^q \text{ for } x \in E_k.$$

Integrating this inequality, we see that

$$\left(\frac{2^{k-1}}{r}\right)^q \cdot \beta_m \left(\frac{r}{2^k}\right)^m \leqslant \int_{E_k} f(x) dx \leqslant \left(\frac{2^k}{r}\right)^q \cdot \beta_m \left(\frac{r}{2^k}\right)^m,$$

i.e.,

$$A2^{k(q-m)} \leqslant \int_{E_k} f(x)dx \leqslant B2^{k(q-m)},$$

where A and B are positive constants that do not depend on k. The obtained two-sided bound implies that convergence of a series

$$\sum_{k=1}^{\infty} \int_{E_k} f(x)dx$$

is equivalent to convergence of a series

$$\sum_{k=1}^{\infty} 2^{k(q-m)}$$

But the second series has a finite sum only for $q < m$, while the sum of the first series, by the countable additivity of the integral, is equal to $\int_B f(x)dx$; the first claim of the theorem follows.

The proof of the second claim is entirely similar (one should consider the spherical layers $E\left(2^k r\right)$). $\square$

Corollary 4.5.1. *Let B be a ball in $\mathbb{R}^m$ of radius r centered at a point $a \in \mathbb{R}^m$. Given $q > 0$, set*

$$f(x) = \frac{1}{\|x-a\|^q}$$

for $x \in \mathbb{R}^m \backslash \{a\}$. Then:

1. *$f \in L^p(B)$ if and only if $q < m/p$;*

2. *$f \in L^p(\mathbb{R}^m \backslash B)$ if and only if $q > m/p$.*

Proof. It is enough to notice that $f \in L^p(X)$ if and only if $|f|^p \in L^1(X)$. $\square$

Remark 4.6. *Another approach to test integrability of the function is the application of Tonelli theorem. For example, let $E \subset \mathbb{R}^2$ be measurable and $E(x) = \{y \in \mathbb{R} : (x, y) \in E\}$ be a section of a set E. Then*

$$\iint_E |f(x,y)|\, dxdy = \iint_{\mathbb{R}} dx \int_{E(x)} |f(x,y)|\, dy.$$

Since function $|f|$ is nonnegative this, in particular means that the LHS is finite if and only if the RHS is finite.

Also we can perform a change of the variable. Suppose

$$\Phi : G \to V$$

is diffeomorphism of domains $G, V \subset \mathbb{R}^n$ (Φ is C^1-smooth, invertible and $\Phi^{-1} : V \to G$ is also C^1-smooth), $E \subset G$ is measurable. Then

$$\iint_{\Phi(E)} |f|\, d\mu_n = \iint_E |f \circ \Phi|\, |\det \Phi'|\, d\mu_n.$$

In particular LHS is finite if and only if the RHS is finite.

The function $J = \det \Phi'$ is called Jacobian of the map Φ.

Example 4.1. *Find conditions on $p, q \in \mathbb{R}$ under which*

$$I = \iint_{|x|+|y|\geqslant 1} \frac{dxdy}{|x|^p + |y|^q} < \infty.$$

Solution.

Notice also that f is continuous, even with respect to every variable, and bounded on compacts that do not contain 0. *Hence, instead of a set* $\{(x,y)\in\mathbb{R}^2 : |x|+|y|\geq 1\}$ *we can test integrability on a set*

$$\{(x,y)\in\mathbb{R}^2 : x^p+y^q\geq 1,\ x,y>0\}.$$

Assume that $1<p,q\leq 2$. *We consider the change of variables*

$$x=u^{2/p},\quad x=v^{2/q}$$

with Jacobian equal to

$$\frac{4}{pq}u^{2/p-1}v^{2/q-1}.$$

Then the finitenss of the integral is equivalent to finitness of

$$\frac{4}{pq}\iint\limits_{\substack{u^2+v^2\geq 1\\ u,v>0}}\frac{u^{2/p-1}v^{2/q-1}dudv}{u^2+v^2}=\left[\begin{array}{l}u=r\cos t,\\ v=r\sin t\end{array},J=r\right]=$$

$$\frac{4}{pq}\iint\limits_{\substack{r\geq 1\\ 0<t<\pi/2}}r^{2/p+2/q-3}\cos^{2/p-1}t\sin^{2/q-1}tdrdt=$$

$$\frac{4}{pq}\int_1^{+\infty}r^{2/p+2/q-3}dr\int_0^{\pi/2}\cos^{2/p-1}t\sin^{2/q-1}tdt$$

The integral

$$\int_0^{\pi/2}\cos^{2/p-1}t\sin^{2/q-1}tdt$$

is finite if and only if $p,q>0$ *and the integral*

$$\int_1^{+\infty}r^{2/p+2/q-3}dr$$

is finite if and only if $\frac{1}{p} + \frac{1}{q} < 1$.

Consequently, the integral I *is finite only in case when* $p, q > 0$ *and satisfy the following estimate*

$$\frac{1}{p} + \frac{1}{q} < 1.$$

Example 4.2. *Assume that* φ *is a continuous function such that*

$$0 < m < \varphi(x, y) < M.$$

Find condition on p *under which*

$$I = \iint\limits_{x^2+y^2\leq 1} \frac{\varphi(x, y)}{(1 - x^2 - y^2)^p} dxdy$$

is finite.

Solution. *First notice that*

$$m\tilde{I} = m \iint\limits_{x^2+y^2\leq 1} \frac{dxdy}{(1 - x^2 - y^2)^p} dxdy \leq I \leq M \iint\limits_{x^2+y^2\leq 1} \frac{dxdy}{(1 - x^2 - y^2)^p} dxdy = M\tilde{I}$$

and the finiteness of the integral I *is equivalent to finiteness of the integral* $\tilde{I}$.

Consider the polar change of variables

$$x = r\cos t, \quad y = r\sin t, \quad 0 < r < R,\ 0 \leq t < 2\pi$$

with Jacobian $J = r$ *to see that*

$$\tilde{I} = \int_0^{2\pi} dt \int_0^1 \frac{rdr}{(1 - r^2)^p} = 2\pi \int_0^1 \frac{rdr}{(1 - r^2)^p} = [s = 1 - r^2] = 2\pi \int_0^1 \frac{ds}{s^p}.$$

Consequently, $\tilde{I}$ *is finite if and only if* $p < 1$.

5 Integration and continuous functions

5.1 Approximation of measurable functions by continuous functions

Definition 5.1. *Let X be a topological space, and*

$$f : X \to \mathbb{C}$$

a continuous function on X. The support of f, denoted $\mathrm{supp}(f)$, *is the closed subset*

$$\mathrm{supp}(f) = \overline{\{x \mid f(x) \neq 0\}} \subset X.$$

If $\mathrm{supp}(X) \subset X$ *is compact, the function f is said to be compactly supported. We denote by $C_c(X)$ the* C*-vector space of compactly supported continuous functions on X.*

Of course, if $x \notin \mathrm{supp}(f)$, we have $f(x) = 0$. However, the converse does not hold; for example, if $f(x) = x$ for $x \in \mathbb{R}$, we have

$$\mathrm{supp}(f) = \mathbb{R}, \quad f(0) = 0.$$

If X is itself compact, we have $C_c(X) = C(X)$, but otherwise the spaces are distinct (for example, a non-zero constant function has compact support only if X is compact). To check that $C_c(X)$ is a vector space notice that

$$\mathrm{supp}(\alpha f + \beta g) \subset \mathrm{supp}(f) \cup \mathrm{supp}(g).$$

Note also an important immediate property of compactly-supported functions: they are bounded on X. Indeed, we have

$$|f(x)| \leqslant \sup_{x \in \mathrm{supp}(f)} |f(x)|$$

for all x, and of course f is bounded on the compact $\mathrm{supp}(f)$. We denote

$$\|f\|_\infty = \sup\{|f(x)| \mid x \in X\},$$

for $f \in C_c(X)$, which is a norm on $C_c(X)$.

Definition 5.2. *A topological space is locally compact if every point has a neighbourhood which closure is compact.*

Theorem 5.3. *Let X be locally compact topological space.*

(1) *For any compact set $K \subset X$, and any open neighbourhood V of K, i.e., with $K \subset V \subset X$, there exists $f \in C_c(X)$ such that*

$$\chi_K \leqslant f \preccurlyeq \chi_V$$

where the notation

$$f \preccurlyeq \chi_V$$

means that

$$\begin{cases} f \leqslant \chi_V; \\ \mathrm{supp}(f) \subset V. \end{cases}$$

(2) ***Urysohn's lemma.*** *Let K_1 and K_2 be disjoint compact subsets of X. There exists $f \in C_c(X)$ such that $0 \leqslant f \leqslant 1$ and*

$$f(x) = \begin{cases} 0 & \textit{if } x \in K_1 \\ 1 & \textit{if } x \in K_2. \end{cases}$$

(3) *Let $K \subset X$ be a compact subset and let $V_1, \ldots, V_n$ be open sets such that*

$$K \subset V_1 \cup V_2 \cup \cdots \cup V_n.$$

For any function $g \in C_c(X)$, there exist functions $g_i \in C_c(X)$ such that $\operatorname{supp}(g_i) \subset V_i$ *for all i and*

$$\sum_{i=1}^{n} g_i(x) = g(x) \text{ for all } x \in K.$$

In addition, if $g \geqslant 0$, one can select g_i so that $g_i \geqslant 0$.

Proof. These results are standard facts of topology. **We will provide the proof for the case when X is a metric space.**

(1) Let $W \subset V$ be a relatively compact open neighborhood of K, and let $F = X \backslash W$ be the (closed) complement of W. We can then define

$$f(x) = \frac{d(x, F)}{d(x, F) + d(x, K)},$$

and check that it satisfies the required properties.

Part **(2)** can be deduced from (1) by applying the latter to $K = K_2$ and V any open neighborhood of K_2 which is disjoint of K_1.

For **(3)**, one shows first how to construct functions $f_i \in C_c(X)$ such that

$$0 \leqslant f_i \preccurlyeq \chi_{V_i}$$

and

$$1 = f_1(x) + \cdots + f_n(x)$$

for $x \in K$. The general statement follows by taking $g_i = g f_i$ we still have $g_i \leqslant V_i$ of course, and also $g_i \geqslant 0$ if $g \geqslant 0$. $\square$

Definition 5.4. *A topological space $(X, \mathcal{T})$ is locally compact if every point has neighborhood V such that $\overline{V}$ is compact.*

Definition 5.5. *Let X be a locally compact topological space. A Radon measure on X is a regular Borel measure μ that is finite on compact sets, i.e. for every compact set $K \subset X$ we have $\mu K < +\infty$.*

Recall that space $L^p(x, \mu)$ is a space of classes of equivalent functions f such that

$$\|f\|_p = \left(\int_X |f(x)|^p \, d\mu(x) \right)^{1/p}, \quad 1 \leq p < \infty;$$

$$\|f\|_\infty = esssup\,|f| = \inf\{M > 0 : |f| \leq M \text{ a.e.}\}.$$

Definition 5.6. *We say that a set of functions $V \subset L^p(X, \mu)$ is dense in $L^p(X, \mu)$ if for any $f \in L^p(X, \mu)$ and any $\varepsilon > 0$ there exists a function $g \in V$ such that*

$$\|f - g\|_p < \varepsilon.$$

Lemma 5.7. *Let μ be a Radon measure on a locally compact topological space X and $1 \leq p < \infty$. Then any set $E \subset X$ with finite measure and any $\varepsilon > 0$ there exists a $f \in C_c(X)$ such that*

$$\|f - \chi_E\|_p < \varepsilon.$$

Proof. Let $\varepsilon > 0$. Then, by regularity of measure μ there exists a compact set $K \subset E$ and an open set $G \supset E$ such that

$$\mu(G \backslash K) < \varepsilon^{1/p}.$$

Furthermore, by Uhrysohn's lemma, there exists a function f continuous on X such that

$$\chi_K \leq f \leq \chi_U$$

and supp $(f) \subset U$. In particular, f is compactly supported and

$$\|f - \chi_E\|_p^p = \int_X |f - \chi_E|^p \, d\mu = \int_{U \setminus K} |f - \chi_E|^p \, d\mu \leq \mu(U \setminus K) < \varepsilon^p$$

since $|f - \chi_E| \leq 1$ on X and f coincides with χ_E on U and on K. $\square$

Theorem 5.8 (Density of continuous functions in L^p.)**.** *Let μ be a Radon measure on a locally compact topological space X and $1 \leq p < \infty$. then the following assertions are satisfied.*

(1) The space of step functions is dense in $L^p(x, \mu)$, i.e.

(2) The (inclusion of) space $C_c(X)$ is dense in $L^p(X, \mu)$.

Proof. **(1)** Suppose now that $f \in Ł^p(X, \mu)$ is measurable and nonnegative. Then there exists a sequence ψ_n of step functions such that

$$|\psi_n(x)| \leq |f(x)|, \quad f(x) = \lim_{n \to +\infty} s_n(x), \quad x \in X.$$

We will show that this convergence is also true in L^p. To see this apply the Lebesgue theorem on dominated convergence to the sequence $g_n = |f - \psi_n|^p$. then $g_n \to 0$ a.e. and $g_n \leq 2^p(|f|^p + |f|^p) \leq 2^p |f|^p$. since $\int_X |f|^p \, d\mu < +\infty$ it follows that

$$\|f - \psi_n\|_p^p = \int_X |g_n| \, d\mu \to 0, \; n \to +\infty,$$

which proves the assertion.

(2) By Lemma 5.7 the statement of the theorem is true for every characteristic function χ_E of a measurable set E. Now, let $f \in L^p(X, \mu)$, $\varepsilon > 0$ and $\psi = \sum_{k=1}^{n} c_k \chi_{E_k}$ be such step function that $\|f - \psi\|_p < \varepsilon/2$.

For every $k = 1 \dots n$ with $c_k \neq 0$ we can choose a function g_k such that $\left\|\chi_{E_k} - g_k\right\|_p < \varepsilon / |2c_k|$. Let

$$g = \sum_{k=1}^{n} c_k \chi_{E_k}.$$

Then, by triangle inequality, we have

$$\|g - \psi\|_p \leq \sum_{k=1}^{n} |c_k| \left\|\chi_{E_k} - g_k\right\|_p < \varepsilon/2.$$

Finally,

$$\|g - f\|_p \leq \|g - \psi\|_p + \|\psi - f\|_p < \varepsilon.$$

□

Theorem 5.9 (Continuity of translation operators). *Let $m \in \mathbb{N}$ and $1 \leq p < +\infty$. Then for any $f \in L^p(\mathbb{R}^m)$ we have*

$$\lim_{h \to 0} \|f(\cdot + h) - f(\cdot)\|_p^p = \lim_{h \to 0} \int_{\mathbb{R}^m} |f(x+h) - f(x)|^p \, dx = 0.$$

Here $L^p(\mathbb{R}^m)$ is L^p-space with respect to the Lebesgue measure on $\mathbb{R}^m$.

Proof. Notice first that for any function $f \in L^p(\mathbb{R}^m)$ any fixed $h \in \mathbb{R}^m$ the function

$$g(x) = f(x + h)$$

is well defined in $L^p(\mathbb{R}^m)$ and $\|g\|_p = \|f\|_p$.

First, consider the case when $f \in C_c(\mathbb{R}^m)$. Let $h_n \in \mathbb{R}^m$ be a sequence such that $h_n \to 0$ and

$$\psi_n(x) = |f(x + h_n) - f(x)|^p.$$

Then $\psi_n(x) \to 0$ for every $x \in \mathbb{R}^n$ and

$$0 \leq \psi_n(x) \leq 2^p \left\| f \right\|_\infty^p \chi_Q(x),$$

where Q is a compact set that contains $h + supp\, f$ for every h with $|h| < 1$.

Consequently, by dominated convergence theorem

$$\int_{\mathbb{R}^m} \psi_n(x) dx = \int_{\mathbb{R}^m} \left| f(x + h_n) - f(x) \right|^p dx \to 0$$

and by Heine definition

$$\lim_{h \to 0} \lim \int_{\mathbb{R}^m} \left| f(x + h) - f(x) \right|^p dx = 0.$$

To consider the general case let $f \in L^p(\mathbb{R}^m)$ and $\varepsilon > 0$. Then, applying the approximation theorem, there exists $g \in C_c(\mathbb{R}^m)$ such that

$$\left\| f - g \right\|_p < \varepsilon.$$

Then

$$\begin{aligned} &\left| f(x + h) - f(x) \right|^p \leq \\ &\quad \big(\left| f(x + h) - g(x + h) \right| + \left| g(x + h) - g(x) \right| + \left| g(x) - f(x) \right| \big)^p \leq \\ &\qquad 3^p \big(\left| f(x + h) - g(x + h) \right|^p + \left| g(x + h) - g(x) \right|^p + \\ &\qquad\qquad \left| g(x) - f(x) \right|^p \big) \end{aligned}$$

so that

$$\int_{\mathbb{R}^m} |f(x+h) - f(x)|^p dx \leqslant 3^p \left(2 \left\| f - g \right\|_p^p + \int_{\mathbb{R}^m} \left| g(x + h) - g(x) \right|^p dx \right).$$

Since g is continuous then by previous case we can choose $\delta > 0$ such that for every $h \in \mathbb{R}^m$ with $|h| < \delta$ we have

$$\int_{\mathbb{R}^m} |g(x+h) - g(x)|^p \, dx < \varepsilon^p$$

and

$$\int_{\mathbb{R}^m} |f(x+h) - f(x)|^p dx < 3^{p+1} \varepsilon^p$$

□

5.2 The Riesz representation theorem

Lemma 5.10. *Let X be a topological space and μ a Borel measure which is finite on compact sets, i.e., such that*

$$\mu(K) < +\infty,$$

for any compact subset $K \subset X$. Then the map

$$\Lambda : C_c(X) \to \mathbf{C},$$

where

$$\Lambda(f) := \int_X f d\mu(x),$$

is well-defined, it is linear, and moreover it is positive: for any $f \in C_c(X)$ which is non-negative on X, we have $\Lambda(f) \geqslant 0$.

Proof. The only point that needs proof is that Λ is well-defined. However, if $f \in C_c(X)$, so is $|f|$, and since $|f|$ is bounded by the remark before the statement, and is zero outside the compact set $K = \operatorname{supp}(f)$, we have

$$\int_X |f(x)| d\mu(x) = \int_K |f(x)| d\mu(x) \leqslant \mu(K) \|f\|_\infty < +\infty$$

since we assumed that μ is finite on compact sets. $\square$

This proposition is extremely simple. In fact, for reasonable topological spaces (such as compact spaces, or $\mathbb{C}$) there is a converse: any linear map $C_c(X) \to \mathbb{C}$ which has the positivity property (that $\Lambda(f) \geqslant 0$ if $f \geqslant 0$) is obtained by integrating f against a fixed Borel measure μ. This powerful result is a very good way to construct measures. For instance, it may be applied to the Riemann integral, which is a well-defined map

$$\Lambda : C_c(\mathbb{R}) \to \mathbf{C},$$

and the resulting measure is the Lebesgue measure on $\mathbb{R}^1$.

Remark 5.11. *Although it is tempting to state that*

$$C_c(X) \subset L^1(\mu),$$

under the situation of the proposition, one must be aware that this is really an absure of notation, since $L^1(\mu)$ is the space of equivalence classes of functions, up to functions which are zero almost everywhere. Indeed, it the quotient map

$$C_c(X) \to L^1(\mu)$$

may be not injective. A simple example is the Borel measure

$$\mu = \big(1 - \chi_{[-1,1]}\big)\, d\lambda$$

on $\mathbb{R}$, where λ is the Lebesgue measure. Of course, λ is finite on compact sets, and since

$$\mu((-1/2, 1/2)) = 0$$

we see that two continuous functions f_1 and f_2 which differ only on the interval $(-1/2, 1/2)$ actually define the same element in $L^1(\mu)$.

Theorem 5.12 (Riesz Theorem)**.** *Let X be locally compact topological space and let*

$$\Lambda : C_c(X) \to \mathbf{C}$$

be a linear map such that $\Lambda(f) \geqslant 0$ if $f \geqslant 0$.

1. *There exists a σ-algebra $\mathcal{M} \supset \mathcal{B}_X$, and a complete measure μ on $\mathcal{M}$, such that μ is finite on compact sets, and*

$$\Lambda(f) = \int_X f d\mu(x) \text{ for all } f \in C_c(X).$$

2. *In fact, there exists such a unique measure μ for which the following additional properties hold:*

   (a) *For all $E \in \mathcal{M}$, we have*

$$\mu(E) = \inf\{\mu(U) \mid U \supset E \text{ is an open set containing } E\} \tag{22}$$

   (b) *For all $E \in \mathcal{M}$, if E is either open or has finite measure, we have*

$$\mu(E) = \sup\{\mu(K) \mid K \subset E \text{ is compact }\}. \tag{23}$$

3. *If X has the additional property that any open set in X is a countable union of compact sets, in which case X is called σ-compact, then the measure μ is unique as a measure on $(X, \mathcal{B}_X)$, i.e., without requiring (22) and (23).*

5.3 Construction of Lebesgue measure on $\mathbb{R}$ by Riesz' theorem

Let $X = \mathbb{R}$ and let Λ be the linear map

$$f \mapsto \int_{-\infty}^{\infty} f(x)dx = \int_a^b f(x)dx,$$

where the integral is a Riemann integral, and $[a, b] \subset \mathbb{R}$ is any interval such that $\operatorname{supp}(f) \subset [a, b]$. Applying Riesz's theorem, we obtain a Borel measure μ such that

$$\Lambda(f) = \int_{-\infty}^{+\infty} f(x)dx = \int_{\mathbb{R}} f(x)d\mu(x)$$

for any $f \in C_c(X)$. We claim that this measure is the Lebesgue measure.

Theorem 5.13. *The measure μ satisfies*

$$\mu([a, b]) = b - a$$

for any real numbers $a \leqslant b$. This measure is called the Lebesgue measure.

Proof. The case $a = b$ is obvious, and so we assume that $a < b$. Then we construct the following sequences $\{f_n\}, \{g_n\}$ of continuous functions with compact support (well-defined in fact for $n > 2(b - a))^{-1}$:

$$f_n(x) = \begin{cases} 1 & \text{if } a \leqslant x \leqslant b \\ 0 & \text{if } x \leqslant a - 1/n \text{ or } x \geqslant b + 1/n \\ nx - (na - 1) & \text{if } a - 1/n \leqslant x \leqslant a \\ -nx + (nb + 1) & \text{if } b \leqslant x \leqslant b + 1/n \end{cases}$$

and

$$g_n(x) = \begin{cases} 1 & \text{if } a + 1/n \leqslant x \leqslant b - 1/n \\ 0 & \text{if } a \leqslant x \text{ or } x \geqslant b \\ nx - na & \text{if } a \leqslant x \leqslant a + 1/n \\ -nx + nb & \text{if } b - 1/n \leqslant x \leqslant b \end{cases}$$

(a graph of these functions will convey much more information than these dry formulas). The definition implies immediately that

$$g_n \leqslant \chi_{[a,b]} \leqslant f_n$$

for $n > (2(b-a))^{-1}$, and after integrating with respect to μ, we derive the inequalities

$$\Lambda(g_n) = \int g_n d\mu \leqslant \mu([a,b]) \leqslant \int f_n d\mu = \Lambda(f_n)$$

for all n, using on the right and left the fact that integration of continuous functions is the same as applying Λ. In fact, the Riemann integrals of f_n and g_n can be computed very easily, and we derive

$$\Lambda(f_n) = (b-a) + \frac{1}{n} \text{ and } \Lambda(g_n) = (b-a) - \frac{1}{n},$$

so that $\mu([a,b]) = b - a$ follows after letting n go to infinity. □

Since it is easy to check that $\mathbb{R}$ is σ-compact (see below where the case of $\mathbb{R}^d$ is explained), we derive in fact that the Lebesgue measure is the unique measure on $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ which extends the length of intervals.