Chapter 6

Measure and integration theory

Remark 6.1. 1. Idea of the Riemann Integral: Approximate f by "stair functions", i.e. decompose the domain into intervals (rectangles, cubes, ...) and use

$$g(x) = \sum_{i=1}^{n} \alpha_i \chi_{[a_i, a_{i+1}]}(x)$$

where for $A \subset \mathbb{R}$ the *characteristic function* of A is defined:

$$\chi_A(x) = \begin{cases} 1 & x \in A \\ 0 & x \notin A \end{cases}$$

The integral of a stair function is:

$$\int g(x)dx = \sum_{i=1}^{n} \alpha_i (a_{i+1} - a_i)$$

2. Idea of the Lebesgue integral: Decompose the range of the function into intervals $[\alpha_i, \alpha_{i+1})$ and approximate by "simple functions"

$$g(x) = \sum_{i=1}^{n} \alpha_i \chi_{A_i}(x)$$

e.g. $A_i = f^{-1}([\alpha_i, \alpha_{i+1}))$ (not interval in general).
The integral of a simple function is given by:

$$\int g(x)\, dx = \sum_{i=1}^{n} \alpha \lambda(A_i)$$

where $\lambda(A_i)$ is the "length" of A_i (area, volume, measure).

Example 6.2. $f(x) = \chi_{\mathbb{Q} \cap [0,1]}(x)$ is not Rieman integrable, but it is Lebesgue integrable:

$$\int_0^1 f(x)dx = 1 \cdot \lambda(\mathbb{Q} \cap [0,1]) + 0 \cdot \lambda([0,1] \backslash \mathbb{Q}) = 0$$

Remark 6.3. Two advantages of the Lebesgue integral:

1. There are more integrable functions, meaning spaces of Lebesgue integrable functions are complete.
2. The Lebesgue integral can be defined on all spaces where one can define a measure λ (not only on $\mathbb{R}$ or $\mathbb{R}^n$).

Basics of measure theory

In 1924 Banach and Tarski managed to prove that there exists no volume map $\mathrm{vol} : \mathcal{P}(\mathbb{R}^3) \to [0, \infty)$ such that

1. $\mathrm{vol}(\emptyset) = 0$, $\mathrm{vol}([0,1]^3) = 1$
2. $X_1, ..., X_k \in \mathcal{P}(\mathbb{R}^3)$ pairwise disjoint, then
$$\mathrm{vol}\Big(\bigcup_{i=1}^{k} X_i) = \sum_{i=1}^{k} \mathrm{vol}(X_i)$$
3. Invariant under transformations. Let $v \in \mathbb{R}^3$, $A \in O(3)$, $X \in \mathbb{R}^3$, then
$$\mathrm{vol}(\{Ax + v \ : \ x \in X\}) =: \mathrm{vol}(A \cdot X + v) = \mathrm{vol}(X)$$

To circumvent this problem σ-algebras and measure theory was created.

Definition 6.4 (σ-algebra)**.**
A family $\mathcal{A} \subset \mathcal{P}(X)$ of subsets of a set X is called *σ-algebra*, if

1. $\emptyset \in \mathcal{A}$
2. $A \in \mathcal{A} \quad \Rightarrow \quad A^C \in \mathcal{A}$
3. $A_k \in A$ for $k \in \mathbb{N} \quad \Rightarrow \quad \bigcup_{k=1}^{\infty} A_k \in \mathcal{A}$

The elements of $\mathcal{A}$ are called the $\mathcal{A}$-*measurable sets*.

Proposition 6.5. *Let $\mathcal{A}$ be a σ-algebra on X. Then*

1. $X \in \mathcal{A}$
2. $A_k \in \mathcal{A}$ *for* $k \in \mathbb{N} \quad \Rightarrow \quad \bigcap_{k=1}^{\infty} A_k \in \mathcal{A}$
3. $A, B \in \mathcal{A} \quad \Rightarrow \quad A \backslash B \in \mathcal{A}$

Example 6.6. 1. $\mathcal{P}(X)$ and $\{\emptyset, X\}$ are σ-algebras on X.

2. If $A_j, j \in I$ are, σ-algebras on X, so is $\bigcap_{j\in I} A_j$.

Definition 6.7 (Generating system)**.**
Let $\mathcal{F} \subset \mathcal{P}(X)$. Then the σ-algebra generated by $\mathcal{F}$ is:

$$\mathcal{A}_{\mathcal{F}} = \bigcap_{\substack{\mathcal{B} \text{ is } \sigma\text{-alg.} \\ \mathcal{F} \subset \mathcal{B}}} \mathcal{B}$$

Any $\mathcal{F} \subset \mathcal{P}(X)$ that generates $\mathcal{A}$ is called *generating system* for $\mathcal{A}$.

Definition 6.8 (Borel σ-algebra)**.**
Let $(X, \mathcal{T})$ be a topological space. Then the σ-algebra

$$\mathcal{A}_{\mathcal{T}} = \mathcal{B}$$

generated by the topology is called the *Borel σ-algebra* on X.

Definition 6.9 (Measure)**.**
Let $\mathcal{A} \subset \mathcal{P}(X)$ be a σ-algebra. A map $\mu : \mathcal{A} \to [0, \infty]$ is called a *measure*, if

1. $\mu(\emptyset) = 0$
2. For pairwise disjoint sets $A_k \in \mathcal{A}$, $k \in \mathbb{N}$,

$$\mu\Big(\bigcup_{k=1}^{\infty} A_k\Big) = \sum_{k=1}^{\infty} \mu(A_k) \qquad (\sigma\text{-additivity})$$

We further call μ

1. a *finite* measure, $\mu(X) < \infty$,
2. a *σ-finite* measure, if there exists a decomposition $X = \bigcup_{k=1}^{\infty} A_k$ such that $\mu(A_k) < \infty \ \forall k$.

The pair $(X, \mathcal{A})$ is called a *measurable space*, the triple $(X, \mathcal{A}, \mu)$ is called a *measure space*.

Example 6.10. Let X be a set and $x_0 \in X$. Then

$$v : \mathcal{P}(X) \to [0, \infty], \quad A \mapsto v(A) = \begin{cases} |A| & \text{if } A \text{ is finite} \\ \infty & \text{otherwise} \end{cases} \qquad \text{"counting measure"}$$

and

$$\delta_{x_0} : \mathcal{P}(X) \to [0, \infty], \quad A \mapsto \delta_{x_0}(A) = \begin{cases} 1 & \text{if } x_0 \in A \\ 0 & \text{otherwise} \end{cases} \qquad \text{"Dirac measure at } x_0\text{''}$$

are measures.

Proposition 6.11. *Let μ be a measure on $(X, \mathcal{A})$ and $A, B \in \mathcal{A}$. Then*

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

and if $A \subset B$

$$\mu(B) = \mu(A) + \mu(B \backslash A) \quad \Rightarrow \mu(A) \leq \mu(B). \qquad \textit{monotony}$$

For $A_j \in \mathcal{A}$, $j \in \mathbb{N}$,

$$\mu\Big(\bigcup_{j=1}^{\infty} A_j\Big) \leq \sum_{j=1}^{\infty} \mu(A_j) \qquad \textit{sub-additivity}$$

and if $A_j \subset A_{j+1}$ $\forall j$, then

$$\lim_{j \to \infty} \mu(A_j) = \mu\Big(\bigcup_{j=1}^{\infty} A_j\Big)$$

Definition 6.12 (Measurable function and the push-forward of a measure)**.**
Let $(X, \mathcal{A})$ and $(Y, \mathcal{C})$ be measure spaces. A map $f : X \to Y$ is called *$\mathcal{A}$-$\mathcal{C}$-measurable*, if

$$C \in \mathcal{C} \quad \Rightarrow \quad f^{-1}(C) \in \mathcal{A}.$$

If μ is a measure on $(X, \mathcal{A})$ then

$$f^*\mu : \mathcal{C} \to [0, \infty], \quad C \mapsto f^*\mu(C) = \mu\left(f^{-1}(C)\right)$$

is called its *push-forward* under f.

Remark 6.13 (Terminology from probability theory). A measure space $(X, \mathcal{A}, \mu)$ with $\mu(X) = 1$ is called a *probability space*. Then the elements $A \in \mathcal{A}$ are called *events* and $\mu(A)$ the probability of the event. Measurable functions $f : X \to Y$, $(Y, \mathcal{C})$ a measurable space, are called *random variables* and the probability measure $f^*\mu$ is called the *distribution* of f.

Theorem 6.14 (Lebesgue measure)**.** *There is a unique measure λ on $(\mathbb{R}^n, \mathcal{B})$ that is translation invariant (i.e. $\lambda(A + x) = \lambda(A)$, $\forall A \in \mathcal{B} \forall x \in \mathbb{R}^n$) and normalised to $\lambda((0,1)^n) = 1$. It is called the Lebesgue-Borel measure and its completion is called the Lebesgue measure.*

Basics of integration theory

Definition 6.15 (Simple function)**.**
A function $g : X \to [-\infty, \infty]$ is called simple, if $g(X) = \{\alpha_1, \ldots, \alpha_k\}$ is finite, i.e.

$$g(x) = \sum_{j=1}^{k} \alpha_j \chi_{A_j}(x) \qquad \text{with } A_j \cap A_i = \emptyset \text{ for } i \neq j$$

Definition 6.16 (Integral of non-negative measurable functions)**.**
Let $(X, \mathcal{A}, \mu)$ be a measure space and $g : X \to [0, \infty]$ a simple and measurable, then

$$\int\limits_X g \, d\mu \doteq \sum_{j=1}^{k} \alpha_j \mu(A_j) \, .$$

For a measurable function $f : X \to [0, \infty]$

$$\int\limits_X f \, d\mu = \sup \left\{ \int\limits_X g \, d\mu \;\middle|\; g : X \to [0, \infty] \text{ simple, measurable and } g \leq f \right\} .$$

Definition 6.17 (Integral of measurable functions)**.**
A measurable function $f : X \to [-\infty, \infty]$ is *integrable*, if for $f_+ = \max\{f, 0\}$ and $f_- = \max\{-f, 0\}$ it holds that

$$\int f_+ \, d\mu < \infty \qquad\qquad \int f_- \, d\mu < \infty.$$

Then

$$\int f \, d\mu \doteq \int f_+ \, d\mu - \int f_- \, d\mu \, .$$

Proposition 6.18. *Let $f, g : X \to \mathbb{R}$ be integrable and $\alpha \in \mathbb{R}$. Then*

1. $\int \alpha f \, d\mu = \alpha \int f \, d\mu$
2. $\int (f + g) \, d\mu = \int f \, d\mu + \int g \, d\mu$
3. $\left| \int f \, d\mu \right| \leq \int |f| \, d\mu$
4. $f \leq g \quad \Rightarrow \int f \, d\mu \leq \int g \, d\mu$

Theorem 6.19 (Beppo-Levi, Monotone convergence)**.** *Let $f_n : X \to [0, \infty]$ measurable and $f_n \leq f_{n+1}$ for all $n \in \mathbb{N}$. Let $f := \lim\limits_{n \to \infty} f_n$ (pointwise), then*

$$\lim_{n \to \infty} \int f_n \, d\mu = \int f \, d\mu$$

Corollary 6.20 (Fatou's lemma)**.** *Let $f_n : X \to [0, \infty]$ be measurable. Then*

$$\int \liminf_{n \to \infty} f_n \, d\mu \leq \liminf_{n \to \infty} \int f_n \, d\mu$$

Definition 6.21 (Almost everywhere)**.**
We say that a local property holds *almost everywhere* with respect to a measure μ on X, if it holds for all $x \in A \subset X$ and

$$\mu(X \backslash A) = 0,$$

i.e. if it fails to hold a in a *null set* only.

Example 6.22. 1. Almost every real number is irrational with respect to Lebesgue's measure on $\mathbb{R}$.

2. Let $f : X \to [0, \infty]$ be measurable. Then

$$\int_X f \, d\mu = 0 \quad \Leftrightarrow \quad f = 0 \text{ almost everywhere}$$

3. Changing an integrable function f on a null set does not change $\int f \, d\mu$.

4. For integrable functions we do not include $\pm\infty$ into the range anymore.

Remark 6.23. 1. Every Rieman integrable function $f : [a, b] \to \mathbb{R}$ is also Lebesgue integrable and the integrals coincide.

2. A function $f : X \to \mathbb{C}$ is integrable, if $|f|$ is integrable and

$$\int f \, d\mu = \int Re \, f \, d\mu + i \int Im \, f \, d\mu$$

3. Analogously for $f : X \to W$ (W-finite dimensional).

4. For $f : X \to W$, W a Banach space, the generalisation is called the Bochner integral.

Definition 6.24 (L^p-spaces)**.**
Let $(X, \mathcal{A}, \mu)$ be a measure space and $1 \le p < \infty$. We define the vector space

$$\mathcal{L}^p(X, \mu) \doteq \{f : X \to \mathbb{R} \mid f \text{ is measurable and } |f|^p \text{ is integrable}\}$$

as well as

$$\|f\|_{L^p} \doteq \left(\int_X |f|^p \, d\mu \right)^{\frac{1}{p}} < \infty \, .$$

Moreover, $L^p(X, \mu) = \mathcal{L}^p(X, \mu)/\sim$ with respect to the equivalence relation

$$f \sim g \quad \Leftrightarrow \quad f = g \text{ almost everywhere} \, .$$

Definition 6.25 (L^∞ and the essential supremum)**.**
Let $(X, \mathcal{A}, \mu)$ be a measure space. For measurable $f : X \to \mathbb{R}$ we define

$$\|f\|_{L^\infty} \doteq \inf \left\{ 0 \le \lambda \le \infty \mid \mu \left(|f|^{-1}((\lambda, \infty]) \right) = 0 \right\} =: \operatorname{ess\,sup} |f| .$$

Using this definition one can define

$$\mathcal{L}^\infty(X, \mu) = \{f : X \to \mathbb{R} \mid f \text{ measurable and } \|f\|_{L^\infty} < \infty\}$$

and

$$L^\infty(X, \mu) = \mathcal{L}^\infty(X, \mu)/\sim \, .$$

Remark 6.26. It is almost immediate to generalize L^p spaces to complex-valued (or vector-valued) measurable functions.

Theorem 6.27 (Completeness of L^p-spaces)**.** *Let* $(X, \mathcal{A}, \mu)$ *be a measure space and* $1 \leq p \leq \infty$*. Then* $(L^p(X, \mu), \|\cdot\|_{L_p})$ *is a Banach space.*

Theorem 6.28 (Hölder inequality)**.** *Let* $f, g : X \to \mathbb{R}$ *be measurable and* $1 \leq p, q \leq \infty$ *such that* $\frac{1}{p} + \frac{1}{q} = 1$ *(conjugated exponents) where* $\frac{1}{\infty} = 0$*. Then*

$$\|fg\|_{L^1} \leq \|f\|_{L^p} \cdot \|g\|_{L^q}$$

Corollary 6.29. *Let* μ *be a finite measure on* X*. Then,*

$$L^p(X, \mu) \subset L^q(X, \mu)$$

for all $1 \leq q \leq p \leq \infty$*.*

Remark 6.30. For $p = q = 2$ this is the *Cauchy-Schwarz inequality* on the Hilbert space L^2. Hence for $f, g \in L^2 \Rightarrow \overline{f}g \in L^1$, since

$$\underbrace{\left| \int \overline{f} g \, d\mu \right|}_{=|\langle f,g|f,g \rangle_{L^2}|} \leq \int |\overline{f} g| \, d\mu \leq \|f\|_{L^2} \cdot \|g\|_{L^2} \, .$$

Theorem 6.31 (Minkowski inequality)**.** *Let* $f, g : X \to \mathbb{R}$ *be measurable and* $1 \leq p \leq \infty$*. Then*

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p \, .$$

Definition 6.32 (Convergence of measurable functions)**.**
Let (X, μ) be a measure space. A sequence (f_n) of real-valued measurable functions

1. converges to f *pointwise* if

$$\forall \varepsilon > 0 \, \forall x \in X \, \exists \, N_{\varepsilon,x} \in \mathbb{N} \quad \text{s.t.} \quad \forall n \geq N_{\varepsilon,x} \; : \; |f(x) - f_n(x)| < \varepsilon \, .$$

2. converges to f *uniformly* if

$$\forall \varepsilon > 0 \, \exists \, N_\varepsilon \in \mathbb{N} \quad \text{s.t.} \quad \forall x \in X \, \forall n \geq N_\varepsilon \; : \; |f(x) - f_n(x)| < \varepsilon \, .$$

3. converges to f *almost everywhere* if there is a set $Y \subset X$ with $\mu(Y) = 0$ such that $f_n \to f$ pointwise in $X \setminus Y$.

4. converges to f *almost uniformly* if there is a set $Y \subset X$ with $\mu(Y) = 0$ such that $f_n \to f$ uniformly in $X \setminus Y$.

5. converges to f in *in* L^p*-norm* if

$$\forall \varepsilon > 0 \, \exists \, N_\varepsilon \in \mathbb{N} \quad \text{s.t.} \quad \forall n \geq N_\varepsilon \; : \; \|f(x) - f_n(x)\|_{L^p} < \varepsilon \, .$$

6. converges to f *in measure* if

$$\forall \varepsilon > 0 \quad : \quad \lim_{n\to\infty} \mu(\{x \in X \ : \ |f(x) - f_n(x)| \geq \varepsilon\}) = 0\,.$$

Since we identify functions that coincide almost everywhere, we only care about the notions 3-6 of convergence. By definition it is clear that almost uniform convergence implies almost everywhere convergence. It is also not hard to show that L^p convergence implies convergence in measure for any $1 \leq p < \infty$. For the rest of the section let (X, μ) be a measure space and let (f_n) be a sequence of real-valued measurable functions.

Theorem 6.33 (Uniform - L^p)**.** *Let μ be a finite measure and suppose that (f_n) is a sequence in $L^p(X, \mu)$ which converges uniformly to a measurable function f. Then, $f \in L^p(X, \mu)$ and*

$$f_n \to f \quad \text{in } L^p\,.$$

Theorem 6.34 (Almost everywhere - L^p)**.** *Suppose that (f_n) is a sequence in $L^p(X, \mu)$ which converges almost everywhere to a measurable function f. If there exists a function $g \in L^p(X, \mu)$ such that*

$$|f_n(x)| \leq |g(x)| \quad \forall n \in \mathbb{N} \quad \text{and} \quad \text{a.e. } x \in X$$

then $f \in L^p(X, \mu)$ and

$$f_n \to f \quad \text{in } L^p\,.$$

Corollary 6.35. *Let μ be a finite measure on X and suppose that (f_n) is a sequence in $L^p(X, \mu)$ which converges almost everywhere to a measurable function f. Suppose that there is a real number $M \geq 0$ such that*

$$|f(x)| \leq M \quad \forall x \in X, n \in \mathbb{N}.$$

Then, $f \in L^p(X, \mu)$ and

$$f_n \to f \quad \text{in } L^p\,.$$

Theorem 6.36 (Measure - L^p)**.** *Suppose that (f_n) is a sequence in $L^p(X, \mu)$ which converges in measure to a measurable function f. If there exists a function $g \in L^p(X, \mu)$ such that*

$$|f_n(x)| \leq |g(x)| \quad \forall n \in \mathbb{N} \quad \text{and} \quad \text{a.e. } x \in X$$

then $f \in L^p(X, \mu)$ and

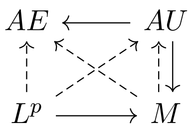
$$f_n \to f \quad \text{in } L^p\,.$$

Theorem 6.37 (Almost uniform - Measure)**.** *If $f_n \to f$ almost uniformly, then it also converges in measure. Conversely, if $f_n \to f$ in measure, then there exists a subsequence that it converges in almost uniformly to the same limit.*

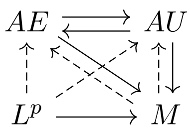
Theorem 6.38 (Almost everywhere - almost uniformly - measure convergence)**.** *Let μ be a finite measure and suppose that $f_n \to f$ almost everywhere. Then, $f_n \to f$ almost uniformly and in measure.*

We summarize all these convergence types and their interplay in the following diagrams. Solid arrow means that the convergence in the tail implies convergence in the nose. Dashed arrow means that convergence in the tail implies subconvergence (convergence through a subsequence) in the nose. The absence of arrow means a counter example can be found. Whenever L^p convergence is involved, it is understood that the functions are in L^p.

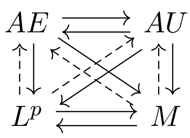
General measure space



Finite measure space



Dominated convergence



Exercises

1. (Proposition 6.5) Let $\mathcal{A}$ be a σ-algebra on X. Show that

1. $X \in \mathcal{A}$
2. $A_k \in \mathcal{A}$ for $k \in \mathbb{N} \quad \Longrightarrow \quad \bigcap_{k\in\mathbb{N}} A_k \in \mathcal{A}$
3. $A, B \in \mathcal{A} \quad \Longrightarrow \quad A \setminus B \in \mathcal{A}$

2. (Proposition 6.11) Let μ be a measure on $(X, \mathcal{A})$ and $A, B \in \mathcal{A}$. Show that

(i) if $A \subset B$, then $\mu(A) \leq \mu(B)$.

(ii) For $A_j \in \mathcal{A},\ j \in \mathbb{N}$,

$$\mu\Big(\bigcup_{j=1}^{\infty} A_j\Big) \leq \sum_{j=1}^{\infty} \mu(A_j)\,.$$

3. (Corollary 6.29) Let μ be a finite measure on $(X, \mathcal{A})$. Show that

$$L^p(X, \mu) \subset L^q(X, \mu)$$

for all $1 \leq q \leq p \leq \infty$.

4. Let (X, μ) be a measure space and let (f_n) be a sequence in $L^p(X, \mu)$.

(a) Show that if $f_n \to f$ in L^p, then $f_n \to f$ in measure.

(b) **(Theorem 6.33)** Show that if μ is a finite measure and $f_n \to f$ uniformly, then $f \in L^p(X, \mu)$ and $f_n \to f$ in L^p.

5. Find an example of

(a) a sequence in $L^p([0, 1])$ that converges pointwise but not in L^p to a function in $L^p([0, 1])$.

(b) a sequence in $L^p([0, 1])$ that converges in L^p but not almost everywhere.