Ergodic Theory and Transfer Operators
— Handout 1 —

Summer 2015

Measure theory and Lebesgue integration

1 Measures and measure spaces

1.1 Basic definitions and properties

We collect the most basic definitions in measure theory, followed by some results which will be useful in the lectures.

Definition 1 (Algebra and σ-algebra)**:** Consider a collection $\mathcal{A}$ of subsets of a set X, and the following properties:

(a) When $A \in \mathcal{A}$ then $A^c := X \setminus A \in \mathcal{A}$.

(b) When $A, B \in \mathcal{A}$ then $A \cup B \in \mathcal{A}$.

(b′) Given a finite or infinite sequence $\{A_k\}$ of subsets of X, $A_k \in \mathcal{A}$, then also $\bigcup_k A_k \in \mathcal{A}$.

If $\mathcal{A}$ satisfies (a) and (b), it is called an *algebra* of subsets of X; if it satisfies (a) and (b′), it is called a *σ-algebra* .

It follows from the definition that a σ-algebra is an algebra, and for an algebra $\mathcal{A}$ holds

- $\emptyset, X \in \mathcal{A}$;
- $A, B \in \mathcal{A} \Rightarrow A \cap B \in \mathcal{A}$;
- $A, B \in \mathcal{A} \Rightarrow A \setminus B \in \mathcal{A}$;
- if $\mathcal{A}$ is a σ-algebra , then $\{A_k\} \subset \mathcal{A} \Rightarrow \bigcap_k A_k \in \mathcal{A}$.

Definition 2 (Measure)**:** A function $\mu : \mathcal{A} \to [0, \infty]$ on a σ-algebra $\mathcal{A}$ is a *measure* if

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

(b) $\mu(A) \geq 0$ for all $A \in \mathcal{A}$; and

(c) $\mu\left(\bigcup_k A_k\right) = \sum_k \mu(A_k)$ if $\{A_k\}$ is a finite or infinite sequence of pairwise disjoint sets from $\mathcal{A}$, that is, $A_i \cap A_j = \emptyset$ for $i \neq j$. This property of μ is called *σ-additivity* (or *countable additivity*).

If, in addition, $\mu(X) = 1$, then μ is called a *probability measure*.

Definition 3:

(a) If $\mathcal{A}$ is a σ-algebra of subsets of X and μ is a measure on $\mathcal{A}$, then the triple $(X, \mathcal{A}, \mu)$ is called a *measure space*. The subsets of X contained in $\mathcal{A}$ are called *measurable*.

(b) If $\mu(X) < \infty$ (resp. $\mu(X) = 1$) then the measure space is called *finite* (resp. *probabilistic* or *normalized*).

(c) If there is sequence $\{A_k\} \subset \mathcal{A}$ satisfying $X = \bigcup_k A_k$ and $\mu(A_k) < \infty$ for all k, then the measure space $(X, \mathcal{A}, \mu)$ is called *σ-finite*.

A set $N \in \mathcal{A}$ with $\mu(N) = 0$ is called a *null set*. If a certain property involving the points of a measure space holds true except for a null set, we say the property holds *almost everywhere* (we write a.e., which, depending on the context, sometimes means "almost every"). We also use the word *essential* to indicate that a property holds a.e. (e.g. "essential bijection").

Theorem 4 (Hahn–Kolmogorov extension theorem)**:** Let X be a set, $\mathcal{A}_0$ an algebra of subsets of X, and $\mu_0 : \mathcal{A}_0 \to [0, \infty]$ a σ-additive function. If $\mathcal{A}$ is the σ-algebra generated[1] by $\mathcal{A}_0$, there exists a measure $\mu : \mathcal{A} \to [0, \infty]$ such that $\mu\big|_{\mathcal{A}_0} = \mu_0$. If μ_0 is σ-finite, the extension is unique.

Definition 5 (Cylinder)**:** Let $\mathcal{A}_k$ be a σ-algebra for $k \in \mathbb{N}$. Let $k_1 < k_2 < \ldots < k_r$ be integers and $A_{k_i} \in \mathcal{A}_{k_i}$, $i = 1, \ldots, r$. A *cylinder set* (also called *rectangle*) is set of the form

$$[A_{k_1}, \ldots, A_{k_r}] = \left\{ \{x_j\}_{j \in \mathbb{N}} \,\middle|\, x_{k_i} \in A_{k_i},\ 1 \le i \le r \right\}.$$

Definition 6: Let $(X_i, \mathcal{A}_i, \mu_i)$, $i \in \mathbb{N}$, be normalized measure spaces. The *product measure space* $(X, \mathcal{A}, \mu) = \prod_{i \in \mathbb{Z}} (X_i, \mathcal{A}_i, \mu_i)$ is defined by

$$X = \prod_{i \in \mathbb{N}} X_i \quad \text{and} \quad \mu\left([A_{k_1}, \ldots, A_{k_r}]\right) = \prod_{j=1}^{r} \mu_{k_j}(A_{k_j}).$$

An analogous definition holds if we replace $\mathbb{N}$ by $\mathbb{Z}$, i.e. if X consists of bi-infinite sequences.

One can see that finite unions of cylinders form an algebra in of subsets of X. By Theorem 4 it can be uniquely extended to a measure on $\mathcal{A}$, the smallest σ-algebra containing all cylinders.
It is often necessary to approximate measurable sets by sets of some sub-class (e.g. an algebra) of the given σ-algebra :

Theorem 7: Let $(X, \mathcal{A}, \mu)$ be a probability space, and let $\mathcal{A}_0$ be an algebra of subsets of X generating $\mathcal{A}$. Then, for each $\varepsilon > 0$ and each $A \in \mathcal{A}$ there is some $A_0 \in \mathcal{A}_0$ such that $\mu(A \Delta A_0) < \varepsilon$. Here, $E \Delta F := (E \setminus F) \cup (F \setminus E)$ denotes the *symmetric difference* of E and F.

1.2 The monotone class theorem

Definition 8: As sequence of sets $\{A_k\}$ is called *increasing* (resp. *decreasing*) if $A_k \subseteq A_{k+1}$ (resp. $A_k \supseteq A_{k+1}$) for all k.
The notation $A_k \uparrow A$ (resp. $A_k \downarrow A$) means that $\{A_k\}$ is an increasing (resp. decreasing) sequence of sets with $\bigcup_k A_k = A$ (resp. $\bigcap_k A_k = A$).

Definition 9 (Monotone class)**:** Let X be a set. A collection $\mathcal{M}$ of subsets of X is a *monotone class* if whenever $A_k \in \mathcal{M}$ and $A_k \uparrow A$, then $A \in \mathcal{M}$.

Theorem 10 (Monotone Class Theorem)**:** A monotone class which contains an algebra, also contains the σ-algebra generated by this algebra.

2 Lebesgue integration

Definition 11 (Borel σ-algebra / measure)**:** Let X be a topological space. The smallest σ-algebra containing all open subsets of X is called the *Borel σ-algebra* . If $\mathcal{A}$ is the Borel σ-algebra , then a measure μ on $\mathcal{A}$ is a *Borel measure* if the measure of any compact set is finite.

Definition 12 (Measurable function)**:** Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ be measure spaces, and $f : X \to Y$ a function. We call f *measurable*, if $f^{-1}(B) \in \mathcal{A}$ whenever $B \in \mathcal{B}$.

Lebesgue integration is concerned with integrals of measurable functions where $Y = \mathbb{R}$ (or $\mathbb{C}$) and $\mathcal{B}$ is the Borel σ-algebra on $\mathbb{R}$. For a detailed construction of the Lebesgue integral we refer to any textbook on measure theory.

[1] The σ-algebra generated by a collection $\mathcal{A}_0$ of subsets of X, also denoted by $\sigma(\mathcal{A}_0)$, is the smalles σ-algebra containing $\mathcal{A}_0$, i.e.

$$\sigma(\mathcal{A}_0) = \bigcap_{\mathcal{A} \text{ is a } \sigma\text{-algebra with } \mathcal{A}_0 \subseteq \mathcal{A}} \mathcal{A}.$$

Analogously we can define the algebra of subsets of X generated by some collection of subsets of X.

We merely note, that a bounded measurable function f can be approximated to arbitrary accuracy by *simple functions* of the form $f_n = \sum_{i=1}^n \lambda_i \chi_{A_i}$, where $\lambda_i \in \mathbb{R}$, and the A_i are disjoint measurable sets. Here, χ_A denotes the *characteristic function* of A, i.e. the function with $\chi_A(x) = 1$ for $x \in A$ and $\chi_A(x) = 0$ for $x \notin A$. The (Lebesgue) integral of simple functions is given by

$$\int f_n \, d\mu := \sum_{i=1}^n \lambda_i \mu(A_i),$$

and f is called (Lebesgue) integrable if for any convergent approximations of it by simple functions f_n the limit $\lim_{n\to\infty} \int f_n d\mu$ exists and is unique.

Definition 13 (L^p space)**:** For $p \in (0, \infty)$ the space $L^p_\mu(X)$ (sometimes also denoted as $L^p(X, \mu)$) consists of the equivalence classes[2] of measurable functions $f : X \to \mathbb{C}$ such that $\int |f|^p \, d\mu < \infty$. For $p \geq 1$, the L^p norm is defined by $\|f\|_p = (\int |f|^p d\mu)^{1/p}$. The space $L^\infty_\mu(X)$ consists of equivalence classes of essentially bounded functions.

If μ is finite, then $L^\infty_\mu(X) \subset L^p_\mu(X)$ for every $p > 0$. Here is a connection between L^p functions and continuous functions.

Theorem 14: If X is a topological space and μ is a Borel measure on X, then the space $C_0(X, \mathbb{C})$ of continuous, complex-valued, compactly supported functions on X is dense in $L^p_\mu(X)$ for all $p > 0$.

Hölder's inequality gives another connection between functions in L^p spaces: if $p \in [1, \infty]$ and q are such that $1/p + 1/q = 1$ (with the convention $1/\infty = 0$), $f \in L^p_\mu(X)$, and $g \in L^q_\mu(X)$, then one has

$$\int |fg| \, d\mu = \|fg\|_1 \leq \|f\|_p \|g\|_q.$$

For $p = 2$ the norm $\|\cdot\|_2$ comes from the inner product

$$\langle f, g \rangle = \int fg \, d\mu,$$

therefore $L^2_\mu(X)$ is a Hilbert space.
For sequences of functions we have the following results concerning interchangeability of integration and limits.

Theorem 15 (Fatou's lemma)**:** For a sequence $\{f_n\}$ of non-negative measurable functions, define $f : X \to [0, \infty]$ as the a.e. pointwise limit

$$f(x) = \liminf_{n\to\infty} f_n(x).$$

Then, f is measurable and

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

Theorem 16 (Lebesgue dominated convergence theorem)**:** Let $f : X \to [-\infty, \infty]$, $g : X \to [0, \infty]$ be measurable functions, and $f_n : X \to [-\infty, \infty]$ be measurable functions such that $|f_n(x)| \leq g(x)$ and $f_n(x) \to f(x)$ as $n \to \infty$ a.e. If g is integrable, then so are f and the f_n, furthermore

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

Note that non-negative integrable functions define finite measures: Let $f : X \to [0, \infty]$ be integrable, then $\mu_f : \mathcal{A} \to [0, \infty]$, defined via

$$\mu_f(A) := \int_A f \, d\mu = \int f \chi_A \, d\mu$$

is a measure. Here is the converse result:

[2] Two measurable functions are *equivalent* if they coincide up to a set of measure zero.

Theorem 17 (Radon–Nikodym)**:** Let $(X, \mathcal{A}, \mu)$ be a finite measure space, and $\nu : \mathcal{A} \to [0, \infty)$ a second measure with the property[3] $\nu(A) = 0$ whenever $\mu(A) = 0$. Then there exists a non-negative integrable function $f : X \to [0, \infty]$ such that

$$\nu(A) = \int_A f \, d\mu \, .$$

Theorem 18 (Fubini)**:** Let $(X, \mathcal{A}, \mu)$ be the product of the measure spaces $(X_i, \mathcal{A}_i, \mu_i)$, $i = 1, 2$, and let a μ-integrable function $f : X \to \mathbb{R}$ be given. Then, for a.e. $x_1 \in X_1$ the function $x_2 \mapsto f(x_1, x_2)$ is μ_2-integrable. Furthermore, the function

$$x_1 \mapsto \int_{X_2} f(x_1, x_2) \, d\mu_2(x_2)$$

is μ_1-integrable, and

$$\int_{X_1} \left(\int_{X_2} f(x_1, x_2) \, d\mu_2(x_2) \right) d\mu_1(x_1) = \iint_X f(x_1, x_2) \, d\mu(x_1, x_2) \, .$$

[3]We say ν is *absolutely continuous* with respect to μ, in shorthand $\nu \ll \mu$, iff $(\mu(A) = 0) \Rightarrow (\nu(A) = 0)$.