# Teoria Ergódica Diferenciável

**lecture 14: Mixing, weak mixing, ergodicity**

Instituto Nacional de Matemática Pura e Aplicada

Misha Verbitsky, November 1, 2017

## Hilbert spaces (reminder)

**DEFINITION:** **Hilbert space** is a complete, infinite-dimensional Hermitian space which is second countable (that is, has a countable dense set).

**DEFINITION:** **Orthonormal basis** in a Hilbert space $H$ is a set of pairwise orthogonal vectors $\{x_\alpha\}$ which satisfy $|x_\alpha| = 1$, and such that $H$ is the closure of the subspace generated by the set $\{x_\alpha\}$.

**THEOREM:** **Any Hilbert space has a basis, and all such bases are countable.**

**Proof:** A basis is found using Zorn lemma. If it's not countable, open balls with centers in $x_\alpha$ and radius $\varepsilon < 2^{-1/2}$ don't intersect, which means that the second countability axiom is not satisfied. ■

**THEOREM:** **All Hilbert spaces are isometric**.

**Proof:** Each Hilbert space has a countable orthonormal basis. ■

## Real Hilbert spaces (reminder)

**DEFINITION:** A Euclidean space is a vector space over $\mathbb{R}$ equipped with a positive definite scalar product $g$.

**DEFINITION:** **Real Hilbert space** is a complete, infinite-dimensional Euclidean space which is second countable (that is, has a countable dense set).

**DEFINITION:** **Orthonormal basis** in a Hilbert space $H$ is a set of pairwise orthogonal vectors $\{x_\alpha\}$ which satisfy $|x_\alpha| = 1$, and such that $H$ is the closure of the subspace generated by the set $\{x_\alpha\}$.

**THEOREM:** **Any real Hilbert space has a basis, and all such bases are countable.**

**Proof:** A basis is found using Zorn lemma. If it's not countable, open balls with centers in $x_\alpha$ and radius $\varepsilon < 2^{-1/2}$ don't intersect, which means that the second countability axiom is not satisfied. ■

**THEOREM:** **All real Hilbert spaces are isometric**.

**Proof:** Each Hilbert space has a countable orthonormal basis. ■

## Koopman operators

**DEFINITION:** Let $(M,\mu)$ be a space with finite measure, and $T:\ M \longrightarrow M$ a measurable map preserving measure. The triple $(M,\mu,T)$ is called **dynamical system**. The map $T$ defines a isometric embedding $T^*:$ Ł$^2(M,\mu) \longrightarrow L^2(M,\mu)$ on the space of square-integrable functions, called **the Koopman operator**.

**DEFINITION:** Two dynamical systems $(M,\mu,T)$ and $(M_1,\mu_1,T_1)$ are **spectral equivalent** if there exists an invertible map $\varphi:\ L^2(M,\mu) \longrightarrow L^2(M_1,\mu_1)$ such that the following diagram is commutative

$$\begin{array}{ccc} L^2(M,\mu) & \xrightarrow{\varphi} & L^2(M_1,\mu_1) \\ {\scriptstyle T^*}\downarrow & & \downarrow{\scriptstyle T_1^*} \\ L^2(M,\mu) & \xrightarrow{\varphi} & L^2(M_1,\mu_1) \end{array}$$

(this is the same as to say that the equivalence $\varphi$ exchanges the Koopman operators $T^*$ and $T_1^*$).

**DEFINITION:** A property $A$ of dynamical system is called **spectral invariant** if for each two spectral invariant systems $(M,\mu,T)$ and $(M_1,\mu_1,T_1)$, the property $A$ holds for $(M,\mu,T) \Leftrightarrow$ it holds for $(M_1,\mu_1,T_1)$.

**REMARK:** We shall see today that **ergodicity is a spectral invariant.**

## Adjoint maps (reminder)

**EXERCISE:** Let $(H, g)$ be a Hilbert space. Show that **the map $x \longrightarrow g(x, \cdot)$ defines an isomorphism $H \longrightarrow H^*$.**

**DEFINITION:** Let $A : H \longrightarrow H$ be a continuous linear endomorphism of a Hilbert space $(H, g)$. Then $\lambda \longrightarrow \lambda(A(\cdot))$ map $A^* : H^* \longrightarrow H^*$. Identifying $H$ and $H^*$ as above, we interpret $A^*$ as an endomorphism of $H$. It is called **adjoint endomorphism** (**Hermitian adjoint** in Hermitian Hilbert spaces).

**REMARK:** The map $A^*$ satisfies $g(x, A(y)) = g(A^*(x), y)$. This relation is often taken as a definition of the adjoint map.

**DEFINITION:** An operator $U : H \longrightarrow H$ is **orthogonal** if $g(x, y) = g(U(x), U(y))$ for all $x, y \in H$.

**CLAIM:** An operator $U$ **is orthogonal if and only if $U^*(U(x)) = x$ for all $x$.**

**Proof:** Indeed, orthogonality is equivalent to $g(x, y) = g(U^*U(x), y)$, which is equivalent to $U^*U = \mathrm{Id}$ because the form $g(z, \cdot)$ is non-zero for non-zero $z$. ∎

## Orthogonal maps and direct sum decompositions (reminder)

**LEMMA:** Let $U: H \longrightarrow H$ be an invertible orthogonal map. Denote by $H^U$ the kernel of $1-U$, that is, the space of $U$-invariant vectors, and let $H_1$ be the closure of the image of $1-U$. **Then $H = H^U \oplus H_1$ is an orthogonal direct sum decomposition.**

**Proof:** Let $x \in H^U$. Then

$$(U^* - 1)(x) = (U^* - 1)U(x) = (1-U)x = 0.$$

This gives $g(x, (U-1)y) = g((U^*-1)x, y) = 0$, hence $x \perp H_1$. Conversely, any vector $x$ which is orthogonal to $H_1$ satisfies $0 = g(x, (U-1)y) = g((U^*-1)x, y)$, giving

$$0 = (U^* - 1)(x) = (U^* - 1)U(x) = (1-U)x.$$

■

## Von Neumann erodic theorem (reminder)

**Corollary 1:** Let $U: H \longrightarrow H$ be an invertible orthogonal map, and $U_n := \frac{1}{n}\sum_{i=0}^{n-1} U^i(x)$. **Then $\lim_n U_n(x) = P(x)$, for all $x \in H$ where $P$ is orthogonal projection to $H^U$.**

**Proof:** By the previous lemma, it suffices to show that $\lim_n U_n = 0$ on $H_1$. However, the vectors of form $x = (1-U)(y)$ are dense in $H_1$, and for such $x$ we have $U_n(x) = U_n(1-U)(y) = \frac{1-U^n}{n}(y)$, and it converges to 0 because $\|U^n\| = 1$. ∎

**THEOREM:** Let $(M, \mu)$ be a measure space and $T: M \longrightarrow M$ a map preserving the measure. Consider the space $L^2(M)$ of functions $f: M \longrightarrow \mathbb{R}$ with $f^2$ integrable, and let $T^*: L^2(M) \longrightarrow L^2(M)$ map $f$ to $T^*f$. **Then the series $T_n(f) := \frac{1}{n}\sum_{i=0}^{n-1}(T^*)^i(f)$ converges in $L^2(M)$ to a $T^*$-invariant function.**

**Proof:** Corollary 1 implies that $T_n(f)$ converges to $P(f)$. ∎

## Ergodic measures and Cesàro sums

**From now on, all measure spaces we consider are tacitly assumed to have finite measure.**

**REMARK: a. e.** means "almost everywhere", that is, outside of a measure 0 set.

**THEOREM:** Let $(M,\mu)$ be a space with (finite) measure, and $T: M \longrightarrow M$ a measurable map. Then $T$ is ergodic if and only if for any bounded function $f: M \longrightarrow \mathbb{R}$, the function $\lim_n \frac{1}{n}\sum_{i=0}^{n-1}(T^*)^i f$ is constant a. e.

**Proof:** By von Neumann ergodic theorem, $\lim_n \frac{1}{n}\sum_{i=0}^{n-1}(T^*)^i f$ is $T$-invariant, hence constant a.e. whenever $T$ is $\mu$-ergodic. Conversely, if $T$ is not $\mu$-ergodic, there exists a bounded, measurable $T$-invariant function $f$ which is not constant a.e., and $\lim_n \frac{1}{n}\sum_{i=0}^{n-1}(T^*)^i f = f$ is not constant. ■

**REMARK:** Ergodicity would follow if $\lim_n \frac{1}{n}\sum_{i=0}^{n-1}(T^*)^i f = \text{const}$ for all $f = \chi_A$, where $A \subset M$ is a measurable subset, and $\chi_A$ its characteristic function.

**COROLLARY:** A dynamical system $(M,\mu,T)$ **is ergodic if and only if the eigenspace of the corresponding Koopman operator $T^*: L^2(M,\mu) \longrightarrow L^2(M,\mu)$ with eigenvalue 1 is 1-dimensional.**

## Convergence in density

**DEFINITION:** The **(asymptotic) density** of a subset $J \subset \mathbb{Z}^{\geqslant 1}$ is the limit $\lim_N \frac{|J\cap[1,N]|}{N}$. A subset $J \subset \mathbb{Z}^{\geqslant 1}$ has **density 1** if $\lim_N \frac{|J\cap[1,N]|}{N} = 1$.

**DEFINITION:** A sequence $\{a_i\}$ of real numbers **converges to $a$ in density** if there exists a subset $J \subset \mathbb{Z}^{\geqslant 1}$ of density 1 such that $\lim_{i\in J} a_i = a$. The convergence in density is denoted by $\operatorname{Dlim}_i a_i = a$.

**PROPOSITION: (Koopman-von Neumann, 1932)** Let $\{a_i\}$ be a sequence of bounded non-negative numbers, $a_i \in [0, C]$. **Then convergence to 0 in density is equivalent to the convergence of Cesàro sums:**

$$\operatorname{Dlim}_i a_i = 0 \Leftrightarrow \lim_N \frac{1}{N}\sum_{i=1}^N a_i = 0$$

**Proof:** See the next slide

## Convergence in density (2)

**Proof. Step 1:** If $\operatorname{Dlim} a_i = 0$, then for $J \subset \mathbb{Z}^{\geqslant 1}$ of density 1, one has $\lim_{i\in J} a_i = 0$, which gives

$$\lim_N \frac{1}{N}\sum_{i=1}^N a_i = \lim_N \frac{1}{N}\sum_{i\in[1,N]\cap J} a_i + \lim_N \frac{1}{N}\sum_{i\in[1,N]\backslash J} a_i$$

The first of the limits on RHS converges to 0 because $\lim_{i\in J} a_i = 0$, and the second limit is bounded by $\lim_N C\frac{|[1,N]\backslash|}{N}$, converging to 0. **The same argument proves that density convergence always implies Cesàro convergence for bounded sequences.**

**Step 2:** Conversely, if $\lim_N \frac{1}{N}\sum_{i=1}^N a_i = 0$, let $L_k$ be the set of all $n$ such that $a_n \geqslant \frac{1}{k}$. Clearly, $L_1 \subset L_2 \subset ....$ The density of each $L_k$ is 0 because $\frac{|L_k\cap[1,N]|}{N} \leqslant k\frac{1}{N}\sum_{i=1}^N a_i$, and the later term converges to 0. Define a sequence $n_k \leqslant n_{k+1} \leqslant ...$ in such a way that $\frac{|L_k\cap[1,n]|}{n} < \frac{1}{k}$ for all $n \geqslant n_k$, and let $L := \bigcup_k (L_k \cap [n_k,\infty[)$. Denote by $J$ the set $\mathbb{Z}^{\geqslant 1}\backslash L$. **Then $\lim_{i\in J} a_i = 0$, because on each interval $[n_k, n_{k+1}]$, for all $i \notin L$ one has $i \notin L_k$, giving $a_i \leqslant 1/k$.**

**Step 3:** It remains to show that $L$ has density 0. Let $m \in [n_{k-1}, n_k]$. Then $\frac{|L\cap[0,m]|}{m} \leqslant \frac{|L_k\cap[0,m]|}{m} \leqslant \frac{1}{k}$, hence $\lim_m \frac{|L\cap[0,m]|}{m} = 0$. ■

## Mixing, weak mixing, ergodicity

**DEFINITION:** Let $(M,\mu,T)$ be a dynamic system, with $\mu$ a probability measure. We say that

(i) $T$ **is ergodic** if $\lim_n \frac{1}{n}\sum_{i=0}^{n-1}\int_M (T^*)^i(\chi_A)\chi_B\mu = \mu(A)\mu(B)$, for all measurable sets $A, B \subset M$.

(ii) $T$ **is weak mixing** if $\underset{i\to\infty}{\operatorname{Dlim}}(T^*)^i(\chi_A)\chi_B = \mu(A)\mu(B)$.

(iii) $T$ **is mixing**, or **strongly mixing** if $\lim\limits_{i\to\infty}\int (T^*)^i(\chi_A)\chi_B = \mu(A)\mu(B)$.

**REMARK:** **The first condition is equivalent to the usual definition of ergodicity by the previous remark.** Indeed, from (usual) ergodicity it follows that $\lim_n \frac{1}{n}\sum_{i=0}^{n-1}(T^*)^i(\chi_A) = \mu(A)$, which gives $\lim_n \frac{1}{n}\sum_{i=0}^{n-1}(T^*)^i(\chi_A)\chi_B = \mu(A)\chi(B)$ and the integral of this function is precisely $\mu(A)\mu(B)$. Conversely, if $\lim_n \int (T^*)^i(\chi_A)\chi_B$ depends only on the measure of $B$, the function $\lim_n \int (T^*)^i(\chi_A)$ is constant, hence $T$ is ergodic in the usual sense.

**REMARK:** Clearly, (iii) $\Rightarrow$ (ii) $\Rightarrow$ (i) (the last implication follows because the density convergence implies the Cesàro convergence). ■

## Mixing, weak mixing, ergodicity: spectral invariance

Notice that the space generated by $\chi_A$ is $C^0$-dense in the space of all measurable functions. Therefore, in the definition of mixing/weak mixing/ergodicity we may replace $\chi_A$, $\chi_B$ by arbitrary $L^2$-integrable functions. Denote by $\langle\cdot,\cdot\rangle$ the scalar product on the Hilbert space $L^2(M,\mu)$.

**DEFINITION:** Let $(M,\mu,T)$ be a dynamic system. We say that

(i) $T$ **is ergodic** if $\lim_n \frac{1}{n}\sum_{i=0}^{n-1}\langle T^i(f),g\rangle\langle 1,1\rangle = \langle f,1\rangle\langle g,1\rangle$ for all $f,g\in L^2(M,\mu)$.

(ii) $T$ **is weak mixing** if $\operatorname{Dlim}_{n\to\infty}\langle T^n(f),g\rangle\langle 1,1\rangle = \langle f,1\rangle\langle g,1\rangle$.

(iii) $T$ **is mixing**, or **strongly mixing**, if $\lim_{n\to\infty}\langle T^n(f),g\rangle\langle 1,1\rangle = \langle f,1\rangle\langle g,1\rangle$.

**REMARK:** Notice that **these three notions are spectral invariants.** Indeed, the weakest of them already implies ergodicity, that is, the eigenspace of eigenvalue 1 of $T$ is 1-dimensional. This implies that $T$ determines the constant function in $L^2(M,\mu)$. However, the conditions (i)-(iii) are expressed in terms of 1, $T$ and the scalar product, hence spectral invariant.

## Mixing: probabilistic interpretation

**DEFINITION: Probability space** is the set $M$, elements of which are called **outcomes**, equipped with a $\sigma$-algebra of subsets, called **events**, and a probability measure $\mu$. In this interpretation, the measure of an event $U \subset M$ is its probability. A **random variable** is a measurable map $f: M \longrightarrow \mathbb{R}$. Its **expected value** is $E(f) := \int_M f\mu$. The **correlation** of random variables $f, g$ is $C(f,g) := E(fg) - E(f)E(g)$.

**REMARK: Mixing** means precisely that $\lim_i(C(T^i f, g)) = 0$, that is, for any two random variables $f$ and $g$, the correlation of $T^i f$ and $g$ converges to 0.

**REMARK:** Let $A \subset M$ be an event. **Conditional expectation** of the random variable $f$ is $E_A(f) := \frac{\int_A f\mu}{\mu(A)}$. This is an expectation of $f$ under the condition that the event $A$ happened. The conditional expectation $E_A(\chi_B) := \frac{\mu(A\cap B)}{\mu(A)}$ is probability that $B$ happens under the condition that $A$ happened.

Correlation between events $A, B \subset M$ is the measure of their independence:

$$E_A(\chi_B) = E(\chi_B) \Leftrightarrow \frac{\mu(A\cap B)}{\mu(A)} = \mu(B) \Leftrightarrow \mu(A\cap B) = \mu(A)\mu(B).$$

if correlation is 0, **this means that the probability of $A$ is entirely unaffected by $B$.**

## Mixing: coin tossing

**DEFINITION:** Let $P$ be a finite set, $P^{\mathbb{Z}}$ the product of $\mathbb{Z}$ copies of $P$, $\Sigma \subset \mathbb{Z}$ a finite subset, and $\pi_\Sigma : P^{\mathbb{Z}} \longrightarrow P^{|\Sigma|}$ projection to the corresponding components. **Cylindrical sets** are sets $C_R := \pi_\Sigma^{-1}(R)$, where $R \subset P^{|\Sigma|}$ is any subset.

**REMARK:** For Bernoulli space, **a complement to an cylindrical set is again an open set,** and the cylindrical sets **form a Boolean algebra.**

**DEFINITION: Bernoulli measure** on $P^{\mathbb{Z}}$ is $\mu$ such that $\mu(C_R) := \frac{|R|}{|P|^{|\Sigma|}}$.

**REMARK:** We consider $P^{\mathbb{Z}}$ as the set of outcomes of infinite sets of coin tossing. The corresponding events are observations of some of the tosses, and its measure is the probability of an event.

**THEOREM: (Lebesque approximation theorem)**
For each Lebesgue measurable set $S \subset P^{\mathbb{Z}}$ and $\varepsilon > 0$, there exists a cylindrical subset $C_R = \pi_\Sigma^{-1}(R)$ **such that** $\mu(C_R \triangle X) < \varepsilon$**.**

**Proof:** The $\sigma$-algebra of Lebesgue measurable sets is by definition a completion of the Boolean algebra of cylindrical sets. ■

## Bernoulli shifts are mixing

**DEFINITION: Bernoulli shift** maps a sequence $a_{-n}, a_{-n+1}, ..., a_0, a_1, ...$ to the sequence $b_{-n}, b_{-n+1}, ..., b_0, b_1, ...,\ b_i = a_{i-1}$.

**CLAIM: The corresponding $\mathbb{Z}$-action is (strongly) mixing on the Bernoulli space.**

**Proof. Step 1:** Since the set of characteristic functions of cylindrical sets is dense, it suffices to prove the mixing for $A$, $B$ cylindrical, $A = C_R = \pi_\Sigma^{-1}(R)$, $B = C_{R'} = \pi_{\Sigma'}^{-1}(R')$.

**Step 2:** Let $C_R = \pi_\Sigma^{-1}(R)$ and $C_{R'} = \pi_{\Sigma'}^{-1}(R')$ be two open sets, where $\Sigma \subset \mathbb{Z}$ and $\Sigma' \subset \mathbb{Z}$ don't intersect. Then $\mu(C_R \cap C_{R'}) = \mu(C_R)\mu(C_{R'})$: the corresponding correlations vanish. Indeed,

$$\mu(C_R \cap C_{R'}) = \frac{|R||R'|}{|P|^{|\Sigma|+|\Sigma'|}}.$$

**This is intiutively clear, because different coin tossings are independent.**

**Step 3:** For sufficiently big power $T^N$ of the Bernoulli shift, the sets $\Sigma \subset \mathbb{Z}$ and $\Phi(\Sigma')$ don't intersect, which gives $C(T^N(C_R), C_{R'}) = 0$. ■