

II.1 THE MEAN ERGODIC THEOREM2.1 VON NEUMANN'S MEAN ERGODIC THEOREM

Thm: Let (X, \mathcal{B}, μ, T) be a ppt. If $f \in L^2(X, \mu)$, then

$$\frac{1}{N} \sum_{k=0}^{N-1} f \circ T^k \xrightarrow{L^2} \bar{f} \quad \text{as } N \rightarrow \infty,$$

where \bar{f} is invariant. If μ is ergodic, \bar{f} is a.e. constant.

Proof: Note that $f \mapsto f \circ T$ is an L^2 -isometry, i.e. $\|f\|_2 = \|f \circ T\|_2$

Step 1: Assume $f = g - g \circ T$ for some $g \in L^2$.

Then

$$\begin{aligned} \left\| \frac{1}{N} \sum_{k=0}^{N-1} f \circ T^k \right\|_2 &= \left\| \frac{1}{N} (g - g \circ T^N) \right\|_2 \leq \frac{1}{N} (\|g\|_2 + \|g \circ T^N\|_2) \\ &= \frac{2}{N} \|g\|_2 \rightarrow 0 \quad \text{as } N \rightarrow \infty \end{aligned}$$

\Rightarrow claim holds for $f \in \mathcal{C} := \{f = g - g \circ T \mid g \in L^2\}$
with $\bar{f} = 0$.

Step 2: Claim holds for $f \in \bar{\mathcal{C}}$ (closure of \mathcal{C})

\curvearrowright exercise.

Importance: we will write $L^2 = \bar{\mathcal{C}} \oplus \bar{\mathcal{C}}^\perp$, and for this we need a closed subspace (check why!)

Step 3: $f \in \bar{\mathcal{C}}^\perp \Rightarrow f$ is invariant

$$\begin{aligned} \|f - f \circ T\|_2^2 &= \|f\|_2^2 - 2\langle f, f \circ T \rangle + \|f \circ T\|_2^2 \\ &= 2\|f\|_2^2 - 2\langle f, f \circ T - f + f \rangle \\ &= 2\langle f, f - f \circ T \rangle = 0 \quad \text{by assumption} \end{aligned}$$

hence, for every $f \in L^2$ we have $f_1 \in \bar{\mathcal{E}}$ and f_2 invariant
 s.t. $f = f_1 + f_2 \rightarrow$

$$\frac{1}{N} \sum_{k=0}^{N-1} f \circ T^k \xrightarrow{N \rightarrow \infty} \begin{matrix} \bar{f}_1 & + & \bar{f}_2 & = & f_2 \\ \parallel & & \parallel & & \\ 0 & & f_2 & , & \text{since } f_2 = f_2 \circ T \text{ a.e.} \end{matrix}$$

The final claim ("if μ ergodic...") follows from Thm 1.11 \square

2.2 CONSEQUENCES OF THE MET

- Proof shows: $\bar{f} =$ " L^2 -projection of f on the space of invariant functions "
- Cor.: A p.p.t. (X, \mathcal{B}, μ, T) is ergodic if and only if for all $A, B \in \mathcal{B}$:

$$\frac{1}{N} \sum_{k=0}^{N-1} \mu(A \cap T^{-k}B) \rightarrow \mu(A)\mu(B) \text{ as } N \rightarrow \infty$$

"Ergodicity = mixing on average"

Proof: " \Rightarrow " Let $f_m = \frac{1}{m} \sum_{k=0}^{m-1} \chi_B \circ T^k$ and $g = \chi_A$

If $f_m \xrightarrow{L^2} \bar{f}$, also $\langle f_m, g \rangle \xrightarrow{m \rightarrow \infty} \langle \bar{f}, g \rangle$.

From the MET: $\bar{f} = \text{const} = \int \chi_B d\mu = \mu(B)$, and the claim follows

" \Leftarrow " Take $A=B$ and assume it is an invariant set. We get $\mu(A) = \mu(A)^2$, hence ergodicity. \square

II.2 THE BIRKHOFF ERGODIC THEOREM

2.3 THE MAXIMAL ERGODIC THEOREM

Def: If $a_1, \dots, a_m \in \mathbb{R}$ and $1 \leq m \leq m$, then we say that a_k is an m -leader if $a_k + \dots + a_{k+p-1} \geq 0$ for some $1 \leq p \leq m$.

Lemma: For every n , $1 \leq n \leq m$, the sum of all n -leaders is ≥ 0 .

Proof: No n -leader \Rightarrow claim holds. Let a_k be the first n -leader, and $p \geq 1$ the smallest number s.t. $a_k + \dots + a_{k+p-1} \geq 0$. Then, for all j , $k \leq j \leq k+p-1$, we have $a_j + \dots + a_{k+p-1} \geq 0$ (otherwise $a_k + \dots + a_{j-1} \geq 0$, and p is not minimal) $\rightarrow a_j$ is also an n -leader \Rightarrow the sum of all n -leaders from a_1 to a_{k+p-1} is $a_k + \dots + a_{k+p-1} \geq 0$. Apply the same argument to the rest of the sequence, a_{k+p}, \dots, a_m , to get the claim. \blacksquare

Thm: Let (X, \mathcal{B}, μ, T) be a mpt, and $f \in L^1(X, \mu)$. Define

$$A = \left\{ x \in X \mid \sup_{n \geq 0} \sum_{k=0}^n f(T^k x) \geq 0 \right\} = \left\{ x \mid \sum_{k=0}^n f(T^k x) \geq 0 \text{ for some } n \right\}$$

Then

$$\int_A f d\mu \geq 0$$

Proof: Let $A_n := \left\{ x \in X \mid \sum_{i=0}^n f(T^i x) \geq 0 \text{ for some } 0 \leq \ell \leq n \right\}$. Then

$A_n \subseteq A_{n+1}$, and $A_n \uparrow A$. By the dominated conv. thm., it suffices to show that $\int_{A_n} f d\mu \geq 0 \ \forall n$.

Fix $m \in \mathbb{N}$. Let $s_m(x)$ denote the sum of n -leaders in the sequence $f(x), f(Tx), \dots, f(T^{m+m+1}x)$ (1)

Let $B_\ell \subset X$ be the set of points for which $f(T^\ell x)$ is an n -leader of (1). By the lemma:

$$0 \leq \int s_m(x) d\mu(x) = \sum_{\ell=0}^{m+m-1} \int_{B_\ell} f(T^\ell x) d\mu(x) \quad (2)$$

Since $f(T^\ell x) = f(T^{\ell-1}Tx)$, $x \in B_\ell \Leftrightarrow Tx \in B_{\ell-1}$, but only for $\ell \leq m-1$ (check this!)

$\Rightarrow B_\ell = T^{-1}(B_{\ell-1}) = \dots = T^{-\ell}(B_0)$ for $\ell \leq m-1$.

$$\Rightarrow \int_{B_k} f(T^k x) d\mu = \int_{T^{-k}(B_0)} f(T^k x) d\mu = \int_{B_0} f d\mu$$

↑
 μ is invariant

\Rightarrow In (2) the first m terms are equal, thus

$$0 \leq m \cdot \int_{B_0} f d\mu + \sum_{k=m}^{m+n-1} \int_{B_k} f(T^k x) d\mu \leq m \int_{A_m} f d\mu + m \int |f| d\mu$$

$\leq \int_X |f| d\mu$

Note $B_0 = A_m$

$$m \text{ arbitrary} \Rightarrow \int_{A_m} f d\mu \geq -\frac{m}{m} \int |f| d\mu \xrightarrow{m \rightarrow \infty} 0$$

this yields the claim. ▣

2.4 THE BIRKHOFF ERGODIC THEOREM

Thm: Let (X, \mathcal{B}, μ, T) be a ppt and $f \in L^1(X, \mu)$. Then

(a) The limit

$$\bar{f}(x) = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} f(T^k x)$$

exists for a.e. $x \in X$; $\quad \quad \quad =: f_n$

(b) If $f \in L^p(X, \mu)$, $1 \leq p < \infty$, then $\bar{f} \in L^p(X, \mu)$ and $f_n \xrightarrow{L^p} \bar{f}$;

(c) $\bar{f} = \bar{f} \circ T$ (a.e.);

$$(d) \int f d\mu = \int \bar{f} d\mu$$

Proof: (a): For $a, b \in \mathbb{R}$, $a < b$, consider

$$X(a, b) = \left\{ x \in X \mid \liminf_{n \rightarrow \infty} f_n(x) < a < b < \limsup_{n \rightarrow \infty} f_n(x) \right\}$$

Note: this set is measurable and T -invariant, and T preserves $\mu|_{X(a, b)}$. For every $x \in X(a, b)$ we have

$$\sup_{n \geq 0} \sum_{k=0}^n [f(T^k x) - b] \geq 0$$

\Rightarrow Apply Thm 2.3 (max. erg. thm.) to $T: X(a,b) \rightarrow X(a,b)$ 23
and $f-b$, yields

$$\left. \begin{array}{l} \int_{X(a,b)} f-b \, d\mu \geq 0 \\ \text{Similarly for } a-f: \\ \int_{X(a,b)} a-f \, d\mu \geq 0 \end{array} \right\} \Rightarrow \int_{X(a,b)} a-b \, d\mu \geq 0$$

$$\Rightarrow \mu(X(a,b)) = 0 \quad a < b$$

a, b arbitrary \Rightarrow f_n converges a.e. Let $\bar{f}(x) = \liminf_{n \rightarrow \infty} f_n(x)$

(b): \bar{f} is measurable and $f_n \rightarrow \bar{f}$ a.e., and

$$\int |\bar{f}| \, d\mu = \int \liminf_n |f_n| \, d\mu \stackrel{\text{a.e. conv.}}{=} \int \liminf_n |f_n| \, d\mu$$

Factor

$$\leq \liminf_n \int |f_n| \, d\mu \stackrel{\mu \text{ invariant}}{=} \liminf_n \frac{1}{n} \sum_{k=0}^{n-1} \int |f \circ T^k| \, d\mu$$

$$= \liminf_n \frac{1}{n} \sum_{k=0}^{n-1} \int |f| \, d\mu = \int |f| \, d\mu$$

$\Rightarrow \| \bar{f} \|_p \leq \| f \|_p$ *

Thus $\bar{f} \in L^1$. Analogous argumentation shows $\bar{f} \in L^p$, and (one uses the triangle inequality for $\|\cdot\|_p$ in the above computation)

Convergence in norm:

• $f \in L^\infty$. Hence $f \in L^1$ too, and $\|f - f_n\|_p \rightarrow 0$ a.e. for $p < \infty$.

Also: $|\bar{f}(x)| \leq \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} |f(T^k x)| \leq \|f\|_\infty$ a.e.

$$\Rightarrow \|f_n(x) - \bar{f}(x)\|_p \leq 2^p \|f\|_\infty^p \xrightarrow[\text{thm.}]{\text{dom. conv.}} f_n \xrightarrow{L^p} \bar{f}$$

• Fix $f \in L^p$, $\varepsilon > 0$. Let $f_0 \in L^\infty$ s.t. $\|f - f_0\|_p \leq \varepsilon/3$, and $N > 0$ s.t.

$$\left\| f_0 - \frac{1}{n} \sum_{k=0}^{n-1} f_0 \circ T^k \right\|_p \leq \varepsilon/3 \quad \forall n \geq N$$

Then

$$\| \bar{f} - f_n \|_p \leq \underbrace{\| \bar{f} - f_0 \|_p}_I + \underbrace{\left\| f_0 - \frac{1}{n} \sum_{k=0}^{n-1} f_0 \circ T^k \right\|_p}_{II} + \underbrace{\left\| \frac{1}{n} \sum_{k=0}^{n-1} (f - f_0) \circ T^k \right\|_p}_{III}$$

To I: $\bar{f} - \bar{f}_0 = \overline{(f - f_0)}$, hence by \ast 24

$$\|\bar{f} - \bar{f}_0\|_p \leq \| \overline{(f - f_0)} \|_p \leq \|f - f_0\|_p \leq \varepsilon/3$$

To II: For $m \geq N$, $\| \bar{f} - \bar{f}_m \|_p \leq \varepsilon/3$ by the assumption above

$$\text{To III: } \left\| \frac{1}{m} \sum_{k=0}^{m-1} (f - f_0) \circ T^k \right\|_p \stackrel{\mu \text{ inv.}}{=} \left\| \frac{1}{m} \sum_{k=0}^{m-1} f - f_0 \right\|_p \leq \|f - f_0\|_p \leq \varepsilon/3$$

Adding up proves $f_m \xrightarrow{L^p} \bar{f}$.

(c): Simple consequence of

$$\| \bar{f} - \bar{f}_m \circ T \|_p \leq \| \bar{f} - f_m \|_p + \| f_m - f_m \circ T \|_p + \| f_m \circ T - \bar{f}_m \circ T \|_p$$

for every $m \in \mathbb{N}$ and $f_m \rightarrow \bar{f}$ in L^p .

(d): Follows from $f_m \rightarrow \bar{f}$ in L^1 . ▀

Remark: The BET is also called *individual ergodic theorem*, because it shows the a.e. convergence of the means $f_m(x)$ for individual paths of the system (property (a) in the thm), while von Neumann's MET only shows L^2 convergence.

2.5 CONSEQUENCES OF THE BET

Corollary: If (X, \mathcal{B}, μ, T) is an ergodic ppt $f \in L^1(X, \mu)$, then

$$\lim_{m \rightarrow \infty} \frac{1}{m} \sum_{k=0}^{m-1} f(T^k x) = \int f d\mu \quad \text{for a.e. } x \in X$$

Examples:

1) If $f = \chi_A$, then $\mu(A)$ is the relative amount of time the trajectory x, Tx, T^2x, \dots spends in A (holds for a.e. x)

2) *Borel's normal number theorem*: $x \in [0, 1)$ is called normal, if

$x = 0.d_1d_2d_3\dots$, and

$$\lim_{m \rightarrow \infty} \frac{1}{m} \# \{ 0 \leq i \leq m-1 \mid d_i = k \} = \frac{1}{10} \quad \text{for every } k = 0, 1, \dots, 9$$

Almost every (w.r.t. Lebesgue) $x \in [0,1)$ is normal.

Proof: $T(x) = 10x \bmod 1$, preserves Lebesgue meas. μ .

Let $f = \chi_{\left[\frac{k}{10}, \frac{k+1}{10}\right)}$, then

$$f(T^m x) = 1 \iff d_m = k$$

Use the BET. ■

3) $T: [0,1] \rightarrow [0,1], x \mapsto \sqrt{x}$

• $\mu = \delta_1$ invariant. f continuous, then $\int f d\mu$
 $\frac{1}{m} \sum_{k=0}^{m-1} f(T^k x) = \frac{1}{m} \sum_{k=0}^{m-1} f(x^{2^{-k}}) \xrightarrow{m \rightarrow \infty} f(1)$ for $x \neq 0$,
since $x^{2^{-m}} \rightarrow 1$ for $x \neq 0$.

• $\mu = \delta_0$ also invariant! $\int f d\delta_0 = f(0) \neq f(1)$ in general!

Skill: BET holds for δ_0 -a.e. x , i.e. only for $x=0$

↪ restricted classes of invariant measures are interesting!
↳ (more informative)

2.6 APPLICATION TO MARKOV CHAINS (cf. [Wa])

A stochastic matrix is **irreducible**, if for $\forall i, j \in S$ exists $m \in \mathbb{N}$ s.t.

$$0 < p_{ij}^{(m)}, \quad \text{where } P^m = (p_{ij}^{(m)})_{i, j \in S}$$

Lemma: P stoch. matrix, $p > 0$ prob. vector with $p^T P = p^T$. Then

$$Q = \lim_{m \rightarrow \infty} \frac{1}{m} \sum_{k=0}^{m-1} P^k \text{ exists. Also}$$

(i) Q is stochastic

(ii) $QP = PQ = Q$

(iii) $v^T P = v^T \implies v^T Q = v^T$

(iv) $Q^2 = Q$

Proof: Let μ be the (p, P) -Markov measure, T the one-sided shift. 26

Let $\chi_i := \chi_{[i]}$, $[i] = \{x \in \Sigma^{\mathbb{N}} \mid x_0 = i\}$ cylinder

$$\text{BET} \Rightarrow \frac{1}{m} \sum_{k=0}^{m-1} \chi_j(T^k x) \xrightarrow{m \rightarrow \infty} \overline{\chi_j}(x) \quad \mu\text{-a.e.}$$

Multiply by χ_i , and integrate ($|\chi_{i,j}| < 1$, use mon. conv. thm)

$$\int \frac{1}{m} \sum_{k=0}^{m-1} \chi_i(x) \chi_j(T^k x) d\mu \rightarrow \int \chi_i \overline{\chi_j} d\mu$$

$$\frac{1}{m} \sum_{k=0}^{m-1} \mu([i] \cap T^{-k}[j]) = \frac{1}{m} \sum_{k=0}^{m-1} P_i P_{ij}^{(k)}$$

$$\Rightarrow \left(\frac{1}{m} \sum_{k=0}^{m-1} P_{ij}^{(k)} \right)_{ij} = \frac{1}{m} \sum_{k=0}^{m-1} P_{ij}^{(k)} \rightarrow \frac{1}{P_i} \int \chi_i \overline{\chi_j} d\mu =: q_{ij}$$

(i) - (iv) follows by direct computation.

Thm: Let T denote the (p, P) -Markov shift, $p > 0$. Let Q be as in the Lemma above. Then the following are equivalent:

- (i) T is ergodic
- (ii) All rows of Q are identical
- (iii) Every entry in Q is strictly positive
- (iv) P is irreducible
- (v) 1 is a simple eigenvalue of P

Proof: (i) \Rightarrow (ii) As in Lemma: $\frac{1}{m} \sum_{k=0}^{m-1} \mu([i] \cap T^{-k}[j]) = P_i q_{ij}$

ergodicity \parallel Cor. 2.2

$$\mu[i] \mu[j] = P_i P_j$$

$$\Rightarrow q_{ij} = P_j$$

$$(ii) \Rightarrow (iii) \quad P^T P = P^T \xrightarrow{\text{Lemma}} P^T Q = P^T \xrightarrow[\text{ident}]{\text{rows}} q_{ij} = P_j > 0 \quad \forall i, j$$

$$(iii) \Rightarrow (iv) \quad \frac{1}{m} \sum_{k=0}^{m-1} P_{ij}^{(k)} \rightarrow q_{ij} > 0 \Rightarrow \exists k \in \mathbb{N} : P_{ij}^{(k)} > 0$$

(iv) \Rightarrow (iii) Lemma: $QP = Q$. Then: $QP^m = Q \forall m$. |27

Let $S_i = \{j \in S \mid q_{ij} > 0\}$. Then $QP^m = Q \Leftrightarrow q_{ij} = \sum_{\ell} q_{i\ell} p_{\ell j}^{(m)}$

$\Rightarrow q_{i\ell} p_{\ell j}^{(m)} \leq q_{ij} \forall \ell$, hence if $\ell \in S_i \wedge p_{\ell j}^{(m)} > 0 \Rightarrow q_{ij} > 0$

Now for fixed i, j , pick any $\ell \in S_i$ (S_i not empty, since Q stoch.) and choose m s.t. $p_{\ell j}^{(m)} > 0$. Thus $q_{ij} > 0$.

(iii) \Rightarrow (ii) Fix j and let $q_j := \max_i q_{ij} > 0$. If $q_{ij} < q_j$ for some i , then by $Q^2 = Q$:

$$q_{ij} = \sum_m q_{im} q_{mj} < \sum_m q_{im} q_j = q_j \quad \forall i, \text{ a contradiction}$$

(ii) \Rightarrow (i) Since cylinders generate the whole Borel- σ -algebra, by Cor. 2.2 it suffices to show that

$$\frac{1}{n} \sum_{\mathbf{z}} \mu(A \cap T^{-k} B) = \mu(A) \mu(B), \quad \begin{array}{l} A = [a_0, \dots, a_i] \\ B = [b_0, \dots, b_j] \end{array}$$

For k suff. large (i.e. $k > i$)

$$\mu(A \cap T^{-k} B) = p_{a_0} p_{a_0 a_1} \dots p_{a_{i-1} a_i} p_{a_i b_0} p_{b_0 b_1} \dots p_{b_{j-1} b_j} \quad (k-i)$$

\Rightarrow

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{\mathbf{z}} \mu(A \cap T^{-k} B) &= \underbrace{p_{a_0} p_{a_0 a_1} \dots p_{a_{i-1} a_i}}_{=\mu(A)} \cdot \underbrace{p_{a_i b_0} p_{b_0 b_1} \dots p_{b_{j-1} b_j}}_{=\mu(B)/p_{b_0}} \cdot \underbrace{\frac{1}{n} \sum_{\mathbf{z}} p_{a_i b_0}^{(k-i)}}_{\rightarrow q_{a_i b_0} \stackrel{(ii)}{=} p_{b_0}} \\ &= \mu(A) \mu(B) \end{aligned}$$

(ii) \Rightarrow (v) Since (ii) $\Rightarrow q_{ij} = p_j \Rightarrow v^T Q = v^T$ iff $v = \lambda p$, $\lambda \in \mathbb{C}$.

Lemma \Rightarrow every left eigenvector of P at eigenvalue 1 is colinear with p .

(v) \Rightarrow (i) $QP = Q$, hence each row of Q is left eigenvector of P at 1. Q stochastic and P has a simple eigenvalue 1 \Rightarrow all rows of Q are identical ■

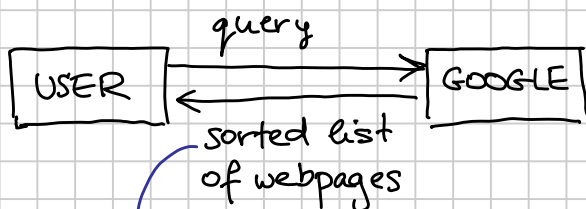
Remark: If P is aperiodic, the (p_i, P) -Markov shift is mixing

$$\hookrightarrow : \Leftrightarrow \forall i: \gcd \{ n \in \mathbb{N} \mid p_{ii}^{(n)} > 0 \} = 1$$

\hookrightarrow greatest common divisor

2.7 INTERNET SEARCH

Search engine: Google



\hookrightarrow sorted by:

- relevance (frequency and position of keywords, ...)

- PageRank

PageRank:

Webpages are nodes, enumerated $1, 2, \dots, N$

Node 0: "Google node"

Edge from i to j ($i \rightarrow j$), iff webpage i links to webpage j . Add $i \rightarrow 0$ and $0 \rightarrow i \forall i = 1, \dots, N$.

Let d_i denote the number of outgoing edges from node i ; $0 < \beta < 1$ param.

Define $P \in \mathbb{R}^{(N+1) \times (N+1)}$ by

$$P_{0i} = \frac{1}{N} \quad P_{i0} = \begin{cases} 1 & , d_i = 1 \\ 1-\beta & , d_i > 1 \end{cases} \quad P_{ij} = \begin{cases} \frac{\beta}{d_i} & , i \rightarrow j \quad (i \neq j) \\ 0 & , \text{otherwise} \end{cases}$$

The PageRank of webpage i is π_i , where $\rho^T P = \rho^T$ (*)

BET: The rank of a page is the relative time a random walker on the world wide web spends on that page. $1-\beta$ is the probability to restart the walk in node 0.

Direct computation of (*) unthinkable: $\sim 50 \cdot 10^9$ indexed webpages

<http://www.worldwidewebsite.com/>

By dynamical systems (BET) approach \rightarrow vector iteration

(cf. numerical eigenvalue computation)

II.3 THE ERGODIC DECOMPOSITION

2.8 CONDITIONAL EXPECTATION

(X, \mathcal{B}, μ) probability space, $\mathcal{C} \subset \mathcal{B}$ sub- σ -algebra

Def: Let $f \in L^1(X, \mathcal{B}, \mu)$, then the **conditional expectation** of f w.r.t. \mathcal{C} is the unique function $\mathbb{E}(f|\mathcal{C}) \in L^1(X, \mathcal{C}, \mu)$ s.t.

$$\int \varphi f d\mu = \int \varphi \mathbb{E}(f|\mathcal{C}) d\mu \quad \forall \varphi \in L^\infty(X, \mathcal{C}, \mu),$$

or equivalently

$$\int_A f d\mu = \int_A \mathbb{E}(f|\mathcal{C}) d\mu \quad \forall A \in \mathcal{C}.$$

Proof of existence and uniqueness:

For $A \in \mathcal{C}$, $\nu_f(A) := \int_A f d\mu$ is a finite measure $\nu_f: \mathcal{C} \rightarrow [0, \infty)$, with $\nu_f \ll \mu$

\Rightarrow by Radon-Nikodym: $\exists! g \in L^1(X, \mathcal{C}, \mu)$ s.t. $g = \frac{d\nu_f}{d\mu}$ ▣

Example: Fix $A \in \mathcal{B}$, and let $\mathcal{C} = \{\emptyset, A, A^c, X\}$, μ prob. measure

Then $g \in L^\infty(X, \mathcal{C}, \mu) \iff g = c_1 \chi_A + c_2 \chi_{A^c}$ for $c_1, c_2 \in \mathbb{R}$

$$\Rightarrow \mathbb{E}(f|\mathcal{C}) = \left[\int_A f d\mu \right] \chi_A + \left[\int_{A^c} f d\mu \right] \chi_{A^c}$$

2.9 ERGODIC LIMIT

• MET: $f \in L^2 \Rightarrow \bar{f}$ is L^2 -projection of f on the space of inv. functions

• BET: $\frac{1}{n} \sum_{k=0}^{n-1} f \circ T^k \rightarrow \bar{f}$ in L^1

Q: How does $f \in L^1$ relate to the invariant function \bar{f} ?

Observe:

μ ergodic $\Rightarrow \bar{f} = \text{const}$, and the const. fcn's are the only inv. fcn's

$f \equiv c$ m'ble wrt $\{\emptyset, X\}$, the only "inv. sets" if μ ergodic

$\mathcal{I}_T := \{ E \in \mathcal{B} \mid T^{-1}E = E \}$ the σ -alg of inv. sets

Note: $\varphi \in L^1(X, \mathcal{I}_T, \mu)$:

$$\int_E \varphi \circ T \, d\mu = \int_{T^{-1}E} \varphi \circ T \, d\mu = \int_E \varphi \, d\mu \iff \varphi = \varphi \circ T$$

\uparrow E inv. \uparrow $T^{-1}E$ \uparrow μ inv.

Step 1: \bar{f} from BET: $\bar{f} = \bar{f} \circ T \implies \bar{f}$ is \mathcal{I}_T -measurable and clearly integrable

Step 2: $\varphi \in L^\infty(X, \mathcal{I}_T, \mu)$:

$$\int \varphi \underbrace{\frac{1}{n} \sum_{k=0}^{n-1} f \circ T^k}_{\xrightarrow[n \rightarrow \infty]{L^1} \bar{f}} \, d\mu = \frac{1}{n} \sum_{k=0}^{n-1} \int (\varphi f) \circ T^k \, d\mu \xrightarrow[T \text{ mpt}]{} \int \varphi f \, d\mu = \underline{\underline{\int \varphi f \, d\mu}}$$

\uparrow φ inv. \uparrow T mpt

$$\int \varphi \bar{f} \, d\mu \xrightarrow[n \rightarrow \infty]{} \underline{\underline{\int \varphi f \, d\mu}} \implies \boxed{\bar{f} = \mathbb{E}(f | \mathcal{I}_T)} \text{ "projection"}$$

$\bar{f} \in L^1(X, \mathcal{I}_T, \mu)$

Note:

If there is $A \in \mathcal{I}_T$ s.t. $\nexists E \in \mathcal{I}_T$ with $E \subset A$ and $\mu(A|E) > 0$, then $\bar{f}|_A$ constant, and A invariant $\implies T|_A$ ergodic

Q: Can we decompose μ into ergodic parts?

2.10 **CONDITIONAL PROBABILITIES**

Standard prob. space: (X, \mathcal{B}, μ)
 Complete metric separable \nearrow
 Borel \uparrow

Thm: (X, \mathcal{B}, μ) : std. prob. space, and $\mathcal{C} \subset \mathcal{B}$ a σ -algebra.

Then there exists a family of prob-measures $\{\mu_x\}_{x \in X}$ s.t.:

(a) $x \mapsto \mu_x(A)$ is \mathcal{C} -measurable for every $A \in \mathcal{B}$;

(b) if f is μ -integrable, then $x \mapsto \int f d\mu_x$ is μ -integrable, and

$$\int \int f d\mu_x d\mu = \int f d\mu$$

(c) if f is μ -integrable then $\int f d\mu_x = \mathbb{E}(f|\mathcal{C})(x)$ for μ -a.e. x

If $\mathcal{C} = \mathcal{I}_T$ for a mpt T , then

(d) μ_x is invariant and ergodic for μ -a.e. x

Proof: See [Sa, Section 2.3]

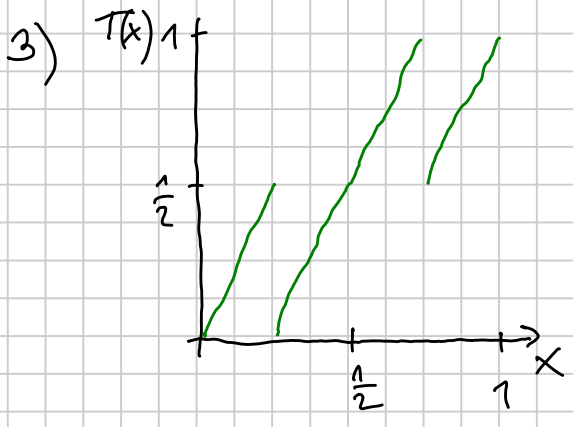
The idea: $f \in C^0(X) \rightarrow$ define $L_x(f) = \bar{f}(x)$

$\xrightarrow{\text{Riesz representation}} \exists! \mu_x$ prob. measure with $L_x(f) = \int f d\mu_x$

Examples:

1) (X, \mathcal{B}, μ, T) ergodic $\Rightarrow L_x(f) = \bar{f}(x) \stackrel{\text{a.e.}}{=} \int f d\mu =: \int f d\mu_x$
 $\Rightarrow \mu_x = \mu$ a.e.

2) $T = \text{id} \Rightarrow L_x(f) = f(x) = \int f d\delta_x \Rightarrow \mu_x = \delta_x$
 \uparrow Dirac measure $\rightarrow \delta_x(A) = \begin{cases} 1, & x \in A \\ 0, & x \notin A \end{cases}$



two indep. angle doubling maps

$$L_x(f) = \begin{cases} \int_0^{1/2} f(x) dx & , x \in [0, 1/2) \\ \int_{1/2}^1 f(x) dx & , x \in [1/2, 1) \end{cases}$$

$$\Rightarrow \mu_x = \begin{cases} 2m_{[0, 1/2)} & , x \in [0, 1/2) \\ 2m_{[1/2, 1)} & , x \in [1/2, 1) \end{cases}$$

m : Lebesgue measure

1) When does a map have a unique ergodic measure?

→ "unique ergodicity", cf. [Ma], [BS]

2) Given a measure ν with $\nu \neq \nu \circ T^{-1}$, will $\nu \circ T^{-m}$ converge in some weak or strong sense? (Note: with the right notion of convergence, the limit — if it exists — will be an inv. measure)

Ex: ergodicity is not enough! → Markov chains: irreducibility $\Rightarrow \exists_1$ stationary distribution p , but $q^T P^m \rightarrow p^T$ does not have to hold prob. vectors q . Take $P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

Aperiodicity of P (\Rightarrow mixingness of the corr. Markov shift) guarantees $q^T P^m \rightarrow p^T$ ($m \rightarrow \infty$).