Topics in Ergodic Theory

==

(4th October 2018, Thursday)

Measure preserving system : $(X, \mathcal{B}, \mu, T)$, where X is a set, $\mathcal{B}$ is a σ-algebra, μ is a probability measure with $\mu(A) \geq 0\ \forall A \in \mathcal{B}$, $\mu(X) = 1$, and T is a measure-preserving transformation. That is, $T : X \to X$ is measurable s.t. $\mu(T^{-1}(A)) = \mu(A)\ \forall A \in \mathcal{B}$.

If Y is a random element of X with distribution μ, then $T(Y)$ also has distribution μ.

Example)

- (Circle rotation) Let $X = \mathbb{R}/\mathbb{Z}$, $\mathcal{B}$ be the Borel sets, μ be the Lebesgue measure and $T = R_\alpha$ where $R_\alpha(x) = x + \alpha$, and $\alpha \in \mathbb{R}/\mathbb{Z}$ is parameter.
- (Times 2 map) $X = \mathbb{R}/\mathbb{Z}$, $\mathcal{B}$ be the Borel sets, μ is a Lebesgue measure, $T = T_2$ where $T_2(x) = 2x$.

  (proof that T_2 is measure preserving) First prove for intervals : let $I = (a, b)$. Then $\mu(I) = b - a$ and $\mu(T_2^{-1}I) = \mu((\frac{a}{2}, \frac{b}{2}) \cup (\frac{a+1}{2}, \frac{b+1}{2})) = b/2 - a/2 + b/2 - a/2 = b - a$.(Just use Dynkin's lemma to conclude... Or,)
  Now let $U \subset \mathbb{R}/\mathbb{Z}$ be open. Then $U = \sqcup_i I_i$ is a disjoint union of intervals, so

  $$\mu(T^{-1}U) = \mu(\sqcup_j T^{-1}I_j) = \sum_j \mu(T^{-1}I_j) = \sum_j \mu(I_j) = \mu(U)$$

  Let $K \subset \mathbb{R}/\mathbb{Z}$ be a compact set. Then

  $$\mu(T^{-1}K) = 1 - \mu((T^{-1}K)^c) = 1 - \mu(T^{-1}(K^c)) = 1 - \mu(K^c) = \mu(K)$$

  Let A be an arbitrary Borel set and let $\epsilon > 0$. Then $\exists U$ open and $\exists K$ compact such that $K \subset A \subset A$ and $\mu(U \backslash K) < \epsilon$, so

  $$\mu(K) = \mu(T^{-1}K) \leq \mu(T^{-1}A) \leq \mu(T^{-1}U) = \mu(U)$$

  We also have $\mu(K) \leq \mu(A)\mu(U)$. Since $\mu(U) - \mu(K) < \epsilon$, $|\mu(A) - \mu(T^{-1}A)| < \epsilon$. Since ϵ was arbitrary, so $\mu(A) = \mu(T^{-1}A)$.

  (End of proof) $\square$

The **orbit** $x \in X$ is the sequence $x, Tx, T^2x, \cdots$.

Some Questions:

- Let $A \in \mathcal{B}$ and $x \in A$. Doest the orbit of x visit A infinitely often?
- What is the proportion of the times n such that $T^n x$ is in A?
- (Mixing property) What is $\mu(\{x \in A : T^n x \in A\})$ if n is large

Example) Let $A = [0, \frac{1}{4}) \subset \mathbb{R}/\mathbb{Z}$ and $T = T_2$. Then $T^n x \in A \Leftrightarrow (n+1)^{\text{st}}$ and $(n+2)^{\text{nd}}$ binary digits of x are 0.
For example, $x = 1/6 = 0.00101010\ldots_{(2)}$ never comes back to A.
Another interesting fact : $\mu(\{x : x \in A, T_2^n x \in A\}) = 1/16$ if $n \geq 2$.(Circle rotation has very different property.)

Markov Shift

- Let $(p_1, p_2, \cdots, p_n)^T$ be a probability vector. Let $A \in \mathbb{R}_{\geq 0}^{n\times n}$ be the **matrix of transition probabilities**.
  Assumptions : (1) $A(1,\cdots,1)^T = (1,\cdots,1)^T$; (2) $(p_1,\cdots,p_n)A = (p_1,\cdots,p_n)$
- Let $X = \{1,\cdots,n\}^{\mathbb{Z}}$, $\mathcal{B}$ be the Borel σ-algebra generated by the product topology of the discrete topology on $\{1,\cdots,n\}$, and $T = \sigma$ is the shift map $(\sigma X)_m = X_{m+1}$.
- Let $\mu(\{x \in X : x_m = i_0, \cdots, x_{m+n} = i_n\}) = p_{i_0} a_{i_0 i_1} \cdots a_{i_{n-1} i_n}$.

Theorem) (Szemerédi) Let $S \subset \mathbb{Z}$ of positive upper Banach density. That is:

$$\bar{d}(S) = \limsup_{N,M:M-N\to\infty} \frac{1}{M-N} |S \cap [N, M-1]| > 0.$$

Then S contains arbitrary long arithmetic progressions. That is, $\forall l$, $\exists a \in \mathbb{Z}$, $d \in \mathbb{Z}_{>0}$ such that $a, a+d, \cdots, a+(l-1)d \in S$.

Theorem) (Furstenberg) (Multiple recurrence) Let $(X, \mathcal{B}, \mu, T)$ be a MPS(Measure preserving system). Let $A \in \mathcal{B}$ b s.t. $\mu(A) > 0$. Let $l \in \mathbb{Z}_{>0}$. Then

$$\liminf_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \mu(A \cap T^{-n}(A) \cap \cdots \cap T^{-(l-1)n}(A)) > 0$$

==
(6th October 2018, Saturday)

Theorem) (Szemerédi) Let $S \subset \mathbb{Z}$ of positive upper Banach density. Then S contains arbitrary long arithmetic progressions.

Theorem) (Furstenberg) Let $(X, \mathcal{B}, \mu, T)$ be a MPS. Let $A \in \mathcal{B}$ be s.t. $\mu(A) > 0$. Then for $\forall l \in \mathbb{Z}_{>0}$

$$\liminf_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \mu(A \cap T^{-n}(A) \cap \cdots \cap T^{-(l-1)n}(A)) > 0$$

Let $X = \{0,1\}^{\mathbb{Z}}$, $\mathcal{B}$ be the Borel σ-algebra, $T = \sigma$ be the shift map.
For a set $S \subset \mathbb{Z}_{\geq 0}$, Let $x^S \in X$ be defined by

$$x_n^S = \begin{cases} 1 & \text{if} \quad n \in S \\ 0 & \text{if} \quad \text{o/w} \end{cases}$$

Let $A \in \mathcal{B}$ and $A = \{x \in X : x_0 = 1\}$
Observation : $n \in S \Leftrightarrow \sigma^n x^S \in A \Leftrightarrow (\sigma^n x^S)_0 = 1 \Leftrightarrow x_n^S = 1$.
Let $\{M_m\}$ and $\{N_m\}$ be sequences s.t.

$$\bar{d}(S) = \lim_{m\to\infty} \frac{1}{M_m - N_m} \big| S \cap [N_m, M_m - 1] \big|$$

Let $\mu_m = \frac{1}{M_m - N_m} \sum_{n=N_m}^{M_m-1} \delta_{\sigma^n x^S}$, where δ_x is a measure on X defined as

$$\delta_x(B) = \begin{cases} 1 & x \in B \\ 0 & \text{o/w} \end{cases}$$

Let μ be the weak limit of a subsequence of μ_m.

(Reminder)

- **Weak Limits) :** (In fact, weak-* limits) Let X be a compact metric space, Let μ_m be a sequence of Borel measures on X, and let μ be another Borel measure. Then μ_m weakly converges to μ. In notation,

$$\lim -w_{m\to\infty}\mu_m = \mu$$

  if $\int f d\mu_m \to \int f d\mu \ \forall f \in C(X)$

- **Theorem)** (Banach-Alaoglu/Helly) Let X be a compact metric space. Then $\mathcal{M}(X)$, the set of Borel probability measures endowed with the topology of weak convergence, is compact and metrizable.

  In particular, there is a weakly convergent subsequence in any sequence of Borel probability measures.

Lemma) Let $(X, \mathcal{B}, \mu, \sigma)$ be as defined above is a measure preserving system.

proof sketch) Let $B \in \mathcal{B}$ Then:

$$\begin{aligned}
\mu_m(B) &= \frac{1}{M_m - N_m}|\{n \in [N_m, M_m - 1] : \sigma^n x^S \in B\}| \\
\mu_m(\sigma^{-1}B) &= \frac{1}{M_m - N_m}|\{n \in [N_m, M_m - 1] : \sigma^n x^S \in \sigma^{-1}B\}| \\
&= \frac{1}{M_m - N_m}|\{n \in [N_m + 1, M_m] : \sigma^n x^S \in B\}| \\
|\mu_m(B) - \mu_m(\sigma^{-1}B)| &\le \frac{1}{M_m - N_m} \to 0 \quad \text{as } m \to \infty
\end{aligned}$$

It can be shown that we can pass to the limit and conclude $\mu(B) = \mu(\sigma^{-1}B)$.

•**Remark :** If B is a cylinder set, i.e. $\exists L \in \mathbb{Z}_{>0}$ and $\tilde{B} \subset \{0,1\}^{2L+1}$ s.t.

$$B = \{x \in X : (x_{-L}, \cdots, x_L) \in \tilde{B}\}$$

then B is both closed and open. Therefore χ_B, the characteristic function of B, is continuous. Hence, the limit

$$\lim_{m\to\infty} \mu_m(B) = \mu(B)$$

Proposition) Let $S \subset \mathbb{Z}$, let x^S, A, $(X, \mathcal{B}, \mu, \sigma)$ be as defined above. Let $l \in \mathbb{Z}_{>0}$. Suppose that $\exists n \in \mathbb{Z}_{>0}$ s.t.

$$\mu(A \cap \sigma^{-n}(A) \cap \cdots \cap \sigma^{-n(l-1)}(A)) > 0$$

Then S contains an arithmetic progression of length l.

proof) Without loss of generality, we may assume that $\mu = \lim_{m\to\infty} \mu_m$(if this is not the case, we just replace μ_m with its converging subsequence). Let $B = A \cap \sigma^{-n}(A) \cap \cdots \cap \sigma^{-n(l-1)}(A)$ and observe that B is a cylinder set. Then $\mu(B) = \lim \mu_m(B)$ hence $\exists m$ s.t. $\mu_m(B) > 0$.

By definition of μ_m, $\exists k \in [N_m, M_m - 1]$ such that $\sigma^k x^S \in B$. Hence

$$\begin{aligned}
&\sigma^k x^S \in A,\ \sigma^k x^S \in \sigma^{-n}(A), \cdots, \sigma^k x^S \in \sigma^{-n(l-1)}(A) \\
\Rightarrow &\sigma^k x^S \in A,\ \sigma^{k+n} x^S \in A, \cdots, \sigma^{k+n(l-1)} x^S \in A
\end{aligned}$$

and so $k, k+n, \cdots, k+n(l-1) \in S$ by earlier observation.

(End of proof) □

Note A is also a cylinder set. Then $\mu(A) = \lim_m \mu_m(A)$ and

$$\mu(A) = \lim_m \mu_m(A) = \lim_m \frac{1}{M_m - N_m}\big|\{n \in [N_m, M_m - 1] : n \in S\}\big| = \bar{d}(S) > 0$$

by assumption that S is of positive upper Banach density, and therefore we can prove Szemerédi when assuming Furstenberg.

==

(9th October, Tuesday)

Lemma) Let $(X, \mathcal{B}, \mu, T)$ be MPS. Let $A \in \mathcal{B}$ with $\mu(A) > 0$. Then $\exists n \in \mathbb{Z}_{>0}$ s.t. $\mu(A \cap T^{-n}A) > 0$.

proof) Suppose $\mu(A \cap T^{-n}A) = 0$ for all $n > 0$. Then

$$\mu(T^{-k}A \cap T^{-n}A) = \mu(A \cap T^{-(n-l)}A) = 0$$

for all $n > k \geq 0$. Hence the sets $A, T^{-1}A, \cdots$ are "almost pairwise disjoint". Then

$$\begin{aligned}\mu(A \cup T^{-1}A \cup \cdots \cup T^{-n}A) &= \mu(A) + \big(\mu(T^{-1}A) - \mu(T^{-1}A \cap A)\big) \\ &+ \big(\mu(T^{-2}A) - \mu(T^{-2}A \cap (A \cup T^{-1}A)) + \cdots \\ &+ \big(\mu(T^{-n}A) - \mu(T^{-n}A \cap (A \cup T^{-1}A \cup \cdots \cup T^{-(n-1)}A))\big) \\ &= (n+1)\mu(A),\end{aligned}$$

a contradiction if $n + 1 > \mu(A)^{-1}$.

(End of proof) $\square$

Theorem) (Poincaré recurrence) Let $(X, \mathcal{B}, \mu, T)$ be MPS. Let $A \in \mathcal{B}$ with $\mu(A) > 0$. Then a.e. $x \in A$ returns to A infinitely often. That is,

$$\mu(A \backslash \bigcap_{N=1}^{\infty} \bigcup_{n=N}^{\infty} T^{-n}A) = 0$$

Remark : $x \in T^{-n}A \Leftrightarrow T^n x \in A$. So $\cup_{n=N}^{\infty} T^{-n}A$ are the points that visit A at least once after time N.

proof) Let A_0 be the set of point in A that never returns to A. We first show $\mu(A_0) = 0$. Note that $\mu(A_0 \cap T^{-n}A_0) \leq \mu(A_0 \cap T^{-n}A) = \mu(\phi) = 0$ for all $n > 0$. By the previous lemma, we have $\mu(A_0) = 0$. Note that if $x \in A \backslash \big(\bigcap_{N=1}^{\infty} \bigcup_{n=N}^{\infty} T^{-n}A\big)$, then there is a maximal $m \in \mathbb{Z}_{\geq 0}$ such that $T^m x \in A_0$. This means that

$$A \backslash \cap \cup T^{-n}A \subset \cup_{m=0}^{\infty} T^{-m}A_0$$

and since $T^{-m}A_0$ has measure 0 for each $m \geq 0$, $A \backslash \cap \cup T^{-n}A$ also has measure 0.

(End of proof) $\square$

However, if we are aim to show that any point of X(or almost every) visits a set A with $\mu(A)$ infinitely often, we should prevent elements of X being partitioned by orbits of T^{-1}. Assumption of ergodicity turns out to be enough for this.(In fact, we can make 'ergodic decomposition' for T to to satisfy ergodicity on each partition - but not lecturing on this; bit tricky)

Definition) A MPS $(X, \mathcal{B}, \mu, T)$ is called **ergodic** if $A = T^{-1}A$ implies $\mu(A) = 0$ or 1 for all $A \in \mathcal{B}$.

If the MPS is not ergodic, and $A \in \mathcal{B}$ with $\mu(A) \in (0, 1)$ s.t. $T^{-1}A = A$, then we can restrict the MPS to A. That is, we consider the MPS:
$(A, \mathcal{B}_A, \mu_A, T\big|_A)$ where $\mathcal{B}_A = \{B \in \mathcal{B} : B \subset A\}$, $\mu_A(B) = \mu(B)/\mu(A)$ for all $B \in \mathcal{B}_A$.

Theorem) The following are equivalent for a MPS $(X, \mathcal{B}, \mu, T)$:

(1) $(X, \mathcal{B}, \mu, T)$ is ergodic.

(2) $\mu(\bigcap_{N=1}^{\infty} \bigcup_{n=N}^{\infty} T^{-n}A) = 1$ for all $A \in \mathcal{B}$ with $\mu(A) > 0$.

(3) $\mu(A \triangle T^{-1}A) = 0$ implies $\mu(A) = 0$ or 1 for all $A \in \mathcal{B}$.

(4) For all bounded measurable functions $f : X \to \mathbb{R}$, $f = f \circ T$ a.e. implies f is constant a.e.

(5) For all measurable functions $f : X \to \mathbb{C}$, $f = f \circ T$ a.e. implies f is constant a.e.

Each condition show different perspective to view ergodicity. The second item shows that for ergodic systems Poincaré recurrence holds in a stronger form: not only almost every point in A but also almost every point in X visits A infinitely often. The last three conditions are often used in practice.

proof)

(1)⇒(2) Let $A \in \mathcal{B}$ with $\mu(A) > 0$ Let $B = \bigcap \bigcup T^{-n}A$, the set of points that visit A infinitely often. By Poicaré recurrence(or P-recurrence), $\mu(B) \geq \mu(A) > 0$. So if we show that $B = T^{-1}B$, then $\mu(B) = 1$ follows by ergodicity.

While, $x \in B \Leftrightarrow x$ visits A i.o. $\Leftrightarrow Tx$ visits A i.o. $\Leftrightarrow Tx \in B$. So we proved $B = T^{-1}B$.

(2)⇒(3) Let $A \in \mathcal{B}$ s.t. $\mu(A \triangle T^{-1}A) = 0$. If $\mu(A) = 0$, there is nothing to prove, so assume $\mu(A) > 0$. Let $B = \cap \cup T^{-n}A$. By (2), we know that $\mu(B) = 1$. We show $\mu(B \backslash A) = 0$, which completes the proof.

Let $x \in B \backslash A$, then there is a first time $m > 0$ s.t. $T^m x \in A$, hence $x \in T^{-m}A \backslash T^{-(m-1)}A$. This shows $B \backslash A \subset \bigcup T^{-m}A \backslash T^{-(m-1)A}$. But $T^{-m}A \backslash T^{-(m-1)A}$ has measure 0 because $\mu(T^{-m}A \backslash T^{-(m-1)}A) = \mu(T^{-1}A \backslash A) = 0$.

So we conclude $\mu(B \backslash A) = 0$.

(3)⇒(4) Let $f : X \to \mathbb{R}$ be a bounded measurable function s.t. $f = f \circ T$ almost everywhere. For any $t \in \mathbb{R}$, define $A_t = \{x \in A : f(x) \leq t\}$. Then

$$\mu(A_t \triangle T^{-1}A_t) = \mu(\{x \in A : f(x) \leq t\} \triangle \{x \in A : f \circ T(x) \leq t\}) = 0$$

By (3), we have $\mu(A_t) \in \{0, 1\}$ for all t. Since f was bounded, if t is very small, then $\mu(A_t) = 0$ and if t is very large $\mu(A_t) = 1$. But $t \mapsto \mu(A_t)$ is a monotone function, we have $\exists c \in \mathbb{R}$ s.t. $\mu(A_t) = 0$ for all $t < c$ and $\mu(A_t) = 1$ for all $t > c$. Therefore we have $f(x) = c$.

(4)⇒(1) Let $A \in \mathcal{B}$ with $A = T^{-1}A$. Then $\chi_A = \chi_A \circ T$ everywhere, so χ_A is constant a.e.

Example : The circle rotation $(\mathbb{R}/\mathbb{Z}, \mathcal{B}, \mu, R_\alpha)$ is ergodic if and only if α is irrational.

proof) Let $f : X \to \mathbb{R}$ be measurable, and let $f(x) = \sum_{n \in \mathbb{Z}} a_n \exp(2\pi i n x)$. Then

$$\begin{aligned} f \circ R_\alpha(x) = f(x + \alpha) &= \sum_{n \in \mathbb{Z}} a_n \exp(2\pi i n(x + \alpha)) \\ &= \sum_{n \in \mathbb{Z}} a_n \exp(2\pi i n \alpha) \exp(2\pi i n x) \end{aligned}$$

so $f = f \circ R_\alpha$ is equivalent to having $a_n = a_n \exp(2\pi i n \alpha)$ for all n. If α is irrational, then $\exp(2\pi i n \alpha) \neq 1$ for all $n \neq 0$ so $a_n = 0$ for all $n \neq 0$.

(End of proof) $\square$

==

(11th October, Thursday)

Theorem) (Mean ergodic theorem, von Neumann) Let $(x, \mathcal{B}, \mu, T)$ be a MPS. Write

$$I = \{f \in L^2(X) : f \circ T = f \text{ a.e.}\} \subset L^2(X)$$

for the closed subspace of T-invariant functions. Write $P_T : L^2(X) \to I$ for the orthogonal projection. Then for every $f \in L^2(X)$, we have

$$\frac{1}{N} \sum_{n=0}^{N-1} f \circ T^n \to P_T f \quad \text{in } L^2(X)$$

Here, $\frac{1}{N} \sum_{n=0}^{N-1} f \circ T^n$ called the **ergodic average**.

There are two proofs for this theorem : one uses spectral theory and the the other does not. We would prove using the second approach, and sketch the first proof in the example sheet.

Theorem) (Pointwise ergodic theorem, Birkhoff) Let $(X, \mathcal{B}, \mu, T)$ be a MPS. Then for all $f \in L^1(X)$, $\exists f^* \in L^1(X)$ s.t. $f^* = f^* \circ T$ a.e. and

$$\frac{1}{N} \sum_{n=0}^{N-1} f \circ T^n(x) \to f^*(x) \quad \text{a.e. in } X$$

Comments

(1) If $f \in L^2 \cap L^1$, then $f^* = P_T f$.

(2) There is an L^p version of convergence in norm. That is, if $f \in L^p$, then

$$\frac{1}{N}\sum_{n=0}^{N-1} f \circ T^n \to f^* \quad \text{in } L^p \text{ norm}$$

This will be proved in the example sheet.

(3) If $(X, \mathcal{B}, \mu, T)$ is ergodic, then f^*(or $P_T f$) is constant a.e., because it is T-invariant.

Note : $f^*(x) = \int f^* d\mu$ a.e. By L^1 norm convergence, we also have

$$\lim_{N\to\infty} \int \frac{1}{N}\sum_{n=0}^{N-1} f \circ T^n d\mu \to \int f^* d\mu$$

By a lemma that would follow,

$$\int f \circ T^n d\mu = \int f d\mu \quad \forall n, \text{ hence } \quad \int f d\mu = \int f^* d\mu$$

Then

$$\frac{1}{N}\sum_{n=0}^{N-1} f(T^n x) \to \int f d\mu$$

Can be interpreted as ”time average(LHS) converges to spatial average(RHS)”.

Lemma) Let $T : X \to X$ be a measurable transformation and let μ be a probability measure. Then μ is T-invariant *if and only if*

$$\int f \circ T d\mu = \int f d\mu \quad \forall f \in L^1(X, \mu) \tag{1}$$

proof) (*(1) ⇒ measure preserving property*) : Let $A \in \mathcal{B}$. Then

$$\mu(T^{-1}A) = \int \chi_{T^{-1}A} d\mu = \int \chi_A \circ T d\mu = \int \chi_A d\mu = \mu(A)$$

(*MPP ⇒ (1)*) : Let $f \in L^1(X)$. If $f = \chi_A$ for some $A \in \mathcal{B}$, then

$$\int f \circ T d\mu = \mu(T^{-1}A) = \mu(A) = \int f d\mu$$

(1) hold for such f. Then (1) also holds for simple functions by linearity of integration. In the case where f is non-negative, let f_n be a monotone increasing sequence of simple functions such that $\lim_n f_n = f$(e.g. $f_n = f \wedge n$),

$$\int f \circ T d\mu = \lim_n \int f_n \circ T d\mu = \lim_n \int f_n d\mu = \int f d\mu$$

In the general case, separate f into positive and negative parts and conclude the proof.

(End of proof) □

Definition) Let $(X, \mathcal{B}, \mu, T)$ be a MPS. Then the **Koopman operator** is defined as : $U_T f = f \circ T$ acting on functions on X.

Lemma) The Koopman operator is an isometry on $L^2(X)$. That is,

$$\langle f, g\rangle = \langle U_T f, U_T g\rangle$$

proof) Apply the previous lemma for the function $f \circ \bar{g}$.

$$\mu(U_T f \cdot U_T \bar{g}) = \mu(U_T(f\bar{g})) = \mu(f\bar{g})$$

(End of proof) □

Definition) A MPS $(X, \mathcal{B}, \mu, T)$ is called invertible if $\exists S : X \to X$, measure preserving, s.t.

$$S \circ T = T \circ S = id_X \quad \text{a.e.}$$

If such a map exists, we denote it by $T^{-1} = S$. (such operator is unique up to a.s. equality)

Lemma) If $(X, \mathcal{B}, \mu, T)$ is invertible, then U_T is unitary, and $U_T^* = U_{T^{-1}}$.

proof) Note : $U_{T^{-1}} \circ U_T = U_T \circ U_{T^{-1}} = id_{L^2(X)}$, so it is enough to show that $U_T^* = U_{T^{-1}}$. To do this, we need to show :

$$\langle U_{T^{-1}} f, g \rangle = \langle f, U_T g \rangle \quad \forall f, g, \in L^2$$

and

$$\langle U_{T^{-1}} f, g \rangle = \int f \circ T^{-1} \cdot \bar{g} d\mu = \int \left(f \circ T^{-1} \cdot \bar{g} \right) \circ T d\mu = \int f \cdot \bar{g} \circ T d\mu = \langle f, U_T g \rangle$$

(End of proof) □

Both von Neumann's and Birkhoff's theorems are easy for certain special kinds of functions. For instance, if $f \in I$:

$$\frac{1}{N} \sum_{n=0}^{N-1} f \circ T^n = f$$

Also, if $f = g \circ T - g$ for some g, then

$$\frac{1}{N} \sum_{n=0}^{N-1} f \circ T^n = \frac{1}{N} \left(g \circ T^N - g \right)$$

It turns out that these two are the only functions that we have to worry about (in the case of von Neumann's theorem) - as presented in the following lemma.

Lemma) Write $B = \{ g \circ T - g : g \in L^2(x) \}$. Then $B^\perp = I$.

Caution ! : B is not close in L^2. So we get $L^2 = I \oplus \bar{B}$, but not $L^2 = I \oplus B$.

proof) Let $f \in L^2(X)$. Then

$$\begin{aligned} & f \in B^\perp \quad \Leftrightarrow \quad \langle f, g \circ T - g \rangle = 0 \quad \forall g \in L^2 \\ \Leftrightarrow \quad & \langle f, g \circ T \rangle = \langle f, g \rangle \quad \forall g \in L^2 \\ \Leftrightarrow \quad & \langle U_T^* f, g \rangle = \langle f, g \rangle \quad \forall g \in L^2 \\ \Leftrightarrow \quad & U_T^* f = f \end{aligned}$$

Now we only need to see that $U_T^* f = f \Leftrightarrow U_T f = f$:

$$\begin{aligned} & U_T f = f \\ \Leftrightarrow \quad & \| f - U_T f \|^2 = 0 \\ \Leftrightarrow \quad & \| f \|^2 + \| U_T f \|^2 - \langle f, U_T f \rangle - \langle U_T f, f \rangle = 0 \\ \Leftrightarrow \quad & \| f \|^2 + \| U_T^* f \|^2 - \langle f, U_T^* f \rangle - \langle U_T^* f, f \rangle + \Big(\| U_T f \|_- \| U_T^* f \|^2 \Big) = 0 \\ \Leftrightarrow \quad & \| f - U_T^* f \|^2 + \Big(\| U_T f \|^2 - \| U_T^* f \|^2 \Big) = 0 \end{aligned}$$

Since $\| f - U_T^* f \|^2 \geq 0$, $\| U_T f \|^2 - \| U_T^* f \|^2 \geq 0$(note that we do not know that U_T^* is unitary, since we do not know if T is invertible, but we know that $\| U_T^* \|_{op} \leq 1$), this statement is equivalent to having $f = U_T^* f$.

Now we are ready to prove the mean ergodic theorem.

proof of MET) Fix $\epsilon > 0$. Let $f \in L^2$. By the lemma, $\exists g, e \in L^2$ s.t.

$$f = P_T f + (g \circ T - g) + e$$

with $\| e \| < \epsilon$ and

$$\frac{1}{N}\sum_{n=0}^{N-1} f \circ T^n = P_T f + \frac{1}{N}(g \circ T^N - g) + \frac{1}{N}\sum_{n=0}^{N-1} e \circ T^n$$

This gives bound

$$\| \frac{1}{N}\sum_{n=0}^{N-1} f \circ T^n - P_T f \| \leq \frac{2 \| g \|}{N} + \epsilon$$

Taking $N \to \infty$ gives

$$\limsup_{N\to\infty} \| \frac{1}{N}\sum_{n=0}^{N-1} f \circ T^n - P_T f \|_{\leq} \epsilon$$

(End of proof) $\square$

==
(13th October, Saturday)

Now we start proving mean ergodic theorem, starting with the following theorem.

Theorem) *(Maximal Ergodic Theorem, Wiener)* Let $(X, \mathcal{B}, \mu, T)$ be a MPS. Let $f \in L^1$, $\alpha \in \mathbb{R}_{>0}$. Let

$$E_\alpha = \{x \in X : \sup_N \frac{1}{N}\sum_{n=0}^{N-1} f(T^n x) > \alpha\}$$

Then $\mu(E_\alpha) \leq \frac{1}{\alpha} \| f \|_1$.

-the theorem is useful, because we can bound some set of particular irregularity depending on the parameter α.
-usually, these kinds of maximal inequalities are prove using covering lemmas, e.g. using Vitalli covering lemma. This is also possible in this case, but the proof gets too long.

The proof of the theorem depends on the following proposition

Proposition) Let $(X, \mathcal{B}, \mu, T)$ be a MPS. Let $f \in L^1$. Let

$$\begin{aligned} &f_0 = 0, \quad f_1 = f, \quad f_2 = f \circ T + f, \quad \cdots \\ &f_n = f \circ T^{n-1} + \cdots + f \circ T + f = f_{n-1} \circ T + f \\ &F_N = \max_{n=0,\cdots,N} f_n \end{aligned}$$

Then $\int_{x:F_N(x)>0} f(x) d\mu(x) \geq 0$ for all N.

proof) Suppose that $F_N(x) > 0$. Then $F_N(x) = f_n(x)$ for some $n \in \{1, \cdots N\}$. Then $F_N(x) = f_{n-1}(Tx) + f(x) \leq F_N(Tx) + f(x)$, hence $f(x) \leq F_N(x) - F_N(Tx)$.

$$\int_{\{x:F_N(x)>0\}} f(x) d\mu \geq \int_{\{x:F_N(x)>0\}} (F_N(x) - F_N \circ T(x)) d\mu(x)$$

Note, if $F_N(x) \leq 0$, then we have $F_N(x) = 0$ and $F_N(x) - F_N(Tx) \leq 0$, so we have $F_N(x) - F_N \circ T(x) \geq 0$ on the domain $\{x : F_N(x) > 0\}$.

(End of proof) $\square$

We now prove maximal ergodic theorem.

proof of Maximal E.T.) Define

$$E_{\alpha,N} = \{x \in X : \max_{m=0\cdots,N} \frac{1}{m}\sum_{n=0}^{m-1} f(T^n x) > \alpha\}$$
$$= \{x \in X : \max_{m=0,\cdots,N} \sum_{n=0}^{m-1} \big(f(T^n x) - \alpha\big) > 0\}$$

(with convention that the sum is just 0 in the case $m = 0$) We apply the proposition for the function $f - \alpha$. Then

$$\int_{E_{\alpha,N}} \big(f(x) - \alpha\big) d\mu \geq 0$$

Then

$$\parallel f \parallel_1 \geq \int_{E_{\alpha,N}} f(x) \geq \alpha\mu(E_{\alpha,N})$$

Note that $E_\alpha = \bigcup_M E_{\alpha,M}$ is an increasing union and the inequality holds for any N, so $\parallel f \parallel_1 \geq \mu(E_\alpha)$.

(End of proof) □

Theorem) (Pointwise ergodic theorem) Let $(X, \mathcal{B}, \mu, T)$ be a MPS. Let $f \in L^1$. Then $\exists f^* \in L^1$, T-invariant s.t.

$$\frac{1}{N}\sum_{n=0}^{N-1} f(T^n x) \to f^*(x) \quad \text{pointwise a.e.}$$

proof) Fix $\epsilon > 0$. Then $\exists f_\epsilon \in L^2$, $e_{\epsilon,1} \in L^1$ s.t.

$$f = f_\epsilon + e_{\epsilon,1} \quad \text{and} \quad \parallel e_{\epsilon,1} \parallel_1 < \epsilon.$$

Also $\exists g_\epsilon \in L^2$, $e_{\epsilon,2} \in L^2$ s.t.

$$f_\epsilon = P_T f_\epsilon + g_\epsilon \circ T - g_\epsilon + e_{\epsilon,2} \quad \text{and} \quad \parallel e_{\epsilon,2} \parallel_1 < \epsilon$$

and $\exists h_\epsilon \in L^\infty$, $e_{\epsilon,3} \in L^1$ s.t.

$$g_\epsilon = h_\epsilon + e_{\epsilon,3} \quad \text{and} \quad \parallel e_{\epsilon,3} \parallel_1 < \epsilon$$

So $f = P_T f_\epsilon + h_\epsilon \circ T - h_\epsilon + e_\epsilon$, where $e_\epsilon \in L^1$ with$\parallel e_\epsilon \parallel_1 < \epsilon$.

$$\frac{1}{N}\sum_{n=0}^{N-1} f(T^n x) = P_T f_\epsilon(x) + \frac{1}{N}\big(h_\epsilon(T^n x) - h_\epsilon(x)\big) + \frac{1}{N}\sum_{n=0}^{N_1} e_\epsilon(T^n x)$$

Let

$$E_{\epsilon,\alpha} = \{x \in: \limsup_{N\to\infty} \Big|\frac{1}{N}\sum_{n=0}^{N-1} f(T^n x) - P_T f_\epsilon(x) > \alpha\Big|\}$$

(Not same as $E_{\alpha,N}$ defined earlier) Applying the Maximal ergodic theorem for the f_n gives

$$\mu(E_{\epsilon,\alpha}) \leq \frac{1}{\alpha} \parallel e_\epsilon \parallel_1 \leq \frac{h\epsilon}{\alpha}$$

Let F be the set of points x s.t. $\frac{1}{N}\sum_{n=0}^{N_1} f(T^n x)$ does not converge at x. Then $F \subset \bigcup_\alpha F_\alpha$, where

$$F_\alpha = \{x \in X : \limsup_{N_1,N_2\to\infty} \Big|\frac{1}{N}\sum_{n=0}^{N_1-1} f(T^n x) - \frac{1}{N_2}\sum_{n=0}^{N_2-1} f(T^n x)\Big| > 2\alpha\}$$

Notice, $F_\alpha \subset E_{\epsilon,\alpha}$ for all $\epsilon > 0$(?????), so $\mu(F_\alpha) \leq \mu(E_{\epsilon,\alpha}) \leq \frac{h\epsilon}{\alpha}$. Therefore $\mu(F_\alpha) = 0$. We can take a countable sequence of α's(e.g. $(1/k)_{k\in\mathbb{N}}$) and conclude $\mu(F) = 0$.

We proved that $\frac{1}{N}\sum_{n=0} f(T^n x) \to f^*(x)$ for some function f^*. By Fatou's lemma, we have $f^* \in L^1$, and it remains to prove $f^*(x) = f^*(Tx)$ a.e.

For almost every x,

$$\begin{aligned}
f^*(x) &= \lim_{N\to\infty} \frac{1}{N}\sum_{n=0}^{N-1} f(T^n x) \\
f^*(Tx) &= \lim_{N\to\infty} \frac{1}{N}\sum_{n=0}^{N-1} f(T^{n-1} x) \\
&= \lim_{N\to\infty} \frac{1}{N}\sum_{n=0}^{N-1} f(T^n x) \\
&= \lim_{N\to\infty} \frac{1}{N-1}\sum_{n=1}^{N-1} f(T^n x) \\
&= \lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N-1} f(T^n x)
\end{aligned}$$

Therefore $f^*(x) - f^*(Tx) = \lim_{N\to\infty} \frac{1}{N} f(x) = 0$

(End of proof) $\square$

==

(this proof has faults. will repair, or delete,,,,,,)

way more elegant proof) For simplicity, let $S_N(x)\frac{1}{N}\sum_{n=0}^{N-1} f\circ T^n(x)$.

(1) First, let f be a positive L^1 function. Then we may find a positive measurable function $\bar{f}(x)$ such that

$$\bar{f}(x) = \liminf_{N\to\infty} S_N(x) \quad \text{a.e.}$$

Note that $\bar{f}$ is T-invariant, since

$$S_N \circ T = \frac{1}{N}(f\circ T + \cdots + f\circ T^N) = \frac{N+1}{N} S_{N+1} - \frac{1}{N} f$$

Also, by Fatou's lemma,

$$\mu(|\bar{f}|) \le \liminf_N \mu(|\frac{1}{N}\sum_{n=0}^{N-1} f\circ T^n|) = \| f \|_1 < \infty$$

and therefore $\bar{f} \in L^1$ with $\int \bar{f} d\mu = \int f d\mu$. Now let $g(x) = f(x) - \bar{f}(x)$, then again $g \in L^1$, with $\int g d\mu = 0$ and

$$\liminf_{N\to\infty} \frac{1}{N}\sum_{n=0}^{N-1} g\circ T^n(x) = 0 \quad \text{a.e.}$$

Now consider the set F_q, defined for $q \in \mathbb{Q}_{>0}$.

$$F_q = \{x : \limsup_N \frac{1}{N}\sum_{n=0}^{N-1} g\circ T^n(x) > q\}$$

Observe that F_q is a T-invariant set, so $\mu(g1_{F_q}) = \mu(f1_{F_q}) - \mu(\bar{f}1_{F_q}) = 0$. But by maximal ergodic theorem, we have

$$q\mu(F_q) \le \int_{F_q} g d\mu = 0$$

hence $\mu(F_q) = 0$, and

$$\mu(\{x : |\limsup_N \frac{1}{N}\sum_{n=0}^{N-1} g\circ T^n(x)| > 0\}) = \mu(\cap_q F_q) = 0$$

We may conclude that

$$\limsup_N \frac{1}{N}\sum_{n=0}^{N-1} g\circ T^n(x) = \liminf_N \frac{1}{N}\sum_{n=0}^{N-1} g\circ T^n(x) = 0 \quad \text{a.e.}$$

and we see that

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=0}^{N-1} f\circ T^n(x) \to \bar{f} \quad \text{a.e.}$$

with $\| \bar{f} \|_1 \leq \| f \|_1$.

(2) For the general case, just divide f into a non-negative part and a negative part, e.g. $f = f^+ - f^-$ and apply part (1) to show that

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=0}^{N-1} f^+\circ T^n(x) \to \bar{f}^+ \quad \text{a.e.}$$
$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=0}^{N-1} f^-\circ T^n(x) \to \bar{f}^- \quad \text{a.e.}$$

and put $f^* = \bar{f}^+ - \bar{f}^-$, then we have the desired result.

(End of proof) □

==
(not done in the lecture. a question in example sheet.)
Theorem) (Pointwise ergodic theorem, L^p-version) Let $(X, \mathcal{B}, \mu, T)$ be a MPS, that is σ-finite. Let $f \in L^p$. Then $\exists f^* \in L^p$, T-invariant s.t.

$$\frac{1}{N}\sum_{n=0}^{N-1} f(T^n x) \to f^*(x) \quad \text{pointwise a.e.}$$

proof) First, assume that f is a positive function. Let $(f_n)_n$ be a increasing sequence of L^1 functions s.t. $f_n \to f$ in L^p and almost everywhere. Then by pointwise ergodic theorem for L^1 functions, we may find $(f_n^*)_n \subset L^1$ s.t.

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=0}^{N-1} f_n\circ T^n(x) \to f_n^* \quad \text{a.e.}$$

Then $(f_n^*)_n$ also forms an increasing sequence, and therefore converges almost everywhere, say $f_n^* \to f^*$ a.e. Now, by Fatou's lemma, we have, for $n \geq m$,

$$\mu((f_n^* - f_m^*)^p)^{1/p} \leq \liminf_n \left(\mu(\frac{1}{N}\sum_{k=0}^{N-1}(f_n - f_m)\circ T^k)\right)^{1/p} \leq \liminf_n \mu((f_n - f_m)^p)^{1/p}$$

where the last inequality follows from Minkowski's inequality. Therefore, $(f_n^*)_n$ forms a Cauchy sequence in L^p, and in fact converges to f^* in L^p. Also, again by Minkowski's inequality, we have $\| f_n^* \|_p \leq \| f_n \|_p$ and

$$\| \frac{1}{N}\sum_{m=0}^{N-1} f_n\circ T^m \|_p \leq \| f_n \|_p$$

so by dominated convergence theorem, we realize that $\frac{1}{N}\sum_{n=0}^{N-1} f_n \to f^*$ is in fact in L^p. Putting these results together, we conclude that

$$\frac{1}{N}\sum_{n=0}^{N-1} f\circ T^n \to f^* \quad \text{in } L^p \text{ and a.e.}$$

$$\begin{array}{ccc} f_n & \xrightarrow{L^p} & f \\ \Big\downarrow{\scriptstyle L^p} & & \Big\downarrow \\ f_n^* & \xrightarrow{L^p} & f^* \end{array}$$

For general functions(not necessarily positive), divide it into a non-negative part and a negative part, and find a.e. and L^p converging functions separately and add them.

==

(16th October, Tuesday)

Definition) A number $x \in [0, 1]$ is called **normal** in base K, if for every $b_1, b_2, \cdots, b_M \in \{0, \cdots, K\}$, we have :

$$\frac{1}{N}\big|\{n \in \{0, \cdots, N-1\} : x_{n+1} = b+1, \cdots, x_{n+M} = b_M\}\big| \to \frac{1}{K^M}$$

where $x = 0.x_1x_2\cdots_{(K)}$ is a base K expansion.

Theorem) Almost every number (w.r.t. Lebesgue measure) is normal in any base $K \geq 2$.

proof) Consider the MPS $(\mathbb{R}/\mathbb{Z}, \mathcal{B}, m, T_K)$($\mathcal{B}$ the Borel σ-algebra, m the Lebesgue measure) where $T_K(x) = K \cdot x$. From the example sheet, this is an ergodic MPS. Now fix M and $b_1, \cdots, b_M$ as in the definition and consider the set

$$A = [(0.b_1b_2\cdots b_M)_{(K)}, (0.b_1\cdots b_M)_{(K)} + \frac{1}{K^M})$$

•Note : $T^n x \in A \quad \Leftrightarrow \quad x_{n+1} = b_1, c \ldots, x_{n+M} = b_M$

To see that x is normal, we need

$$\frac{1}{N}\sum_{n=0}^{N-1} \chi_A(T^n x) \to \frac{1}{K^M}$$

This holds by the pointwise ergodic theorem for almost every x. Since there are countably many choices for K, M, and $b, \cdots, b_M$, the theorem follows.

(End of proof) □

The problem with these ergodic theorems is that the converging point might differ depending on the selection of measure. To study in which cases this can be prevented, we study the uniqueness of measures that is preserved under a fixed map T.

Definition) A **topological dynamical system** is a tuple (X, T), where X is a compact metric space and $T : X \to X$ is a continuous map. We say that (X, T) is **uniquely ergodic**, if there is only one T-invariant Borel probability measure on X.

Theorem) Let (X, T) be a topological dynamical system. The followings are equivalent :

(1) (X, T) is uniquely ergodic.

(2) For every $f \in C(X)$, there is $c_f \in \mathbb{C}$ s.t.

$$\frac{1}{N}\sum_{n=0}^{N-1} f(T^n x) \to c_f \quad \text{uniformly on } X$$

(3) There is a dense $A \subset C(X)$ and for each $f \in A$ there is $c_f \in \mathbb{C}$ s.t.

$$\frac{1}{N}\sum_{n=0}^{N-1} f(T^n x) \to c_f \quad \text{not necessarily uniformly } \forall x \in X$$

Theorem) *(Riesz representation theorem)* Let X be a compact metric space. Then to each finite Borel measure on X, we associate bounded linear functional on $C(X)$ as follows :

$$L_\mu f = \int f d\mu$$

Then $\mu \mapsto L_\mu$ is a bijection from the space of finite Borel measures on X, $\mathcal{M}(X)$, to bounded linear functional on $C(X)$.

Corollary) Let $\mu_1 \neq \mu_2$ be two Borel measures on a compact metric space. Then $\mu_1 = \mu_2$ if and only if

$$\int f d\mu_1 = \int f d\mu_2 \quad \forall f \in C(X)$$

Definition) Let X, T be as above, let μ be a Borel measure. The push-forward of μ via T is the measure

$$T_*\mu(A) = \mu(T^{-1}(A)) \quad \forall A \in \mathcal{B}$$

-This indeed defines a measure.

Lemma) Let X, T, μ be as above. Then

$$\int f dT_*\mu = \int f \circ T d\mu$$

for every bounded measurable function f.

> **proof)** First prove this for characteristic functions of sets. Let $A \in \mathcal{B}$.
>
> $$\int \chi_A dT_*\mu = T_*\mu(A) = \mu(T^{-1}A) = \int \chi_{T^{-1}A} d\mu = \int \chi_A \circ T d\mu$$
>
> Now use uniform class theorem to complete the proof.
>
> *(End of proof)* $\square$

•**Remark :** μ is T-invariant *iff* $\mu = T_*\mu$.

Lemma) Let X, T, μ be as above. Then μ is T-invariant *iff*

$$\int f d\mu = \int f \circ T d\mu \quad \forall f \in C(X) \quad \cdots\cdots (\star)$$

(we are talking about continuous functions in place of measurable functions - so is in fact enough to work with only continuous functions.)

> **proof)** We have already seen that μ being T-invariant implies $(\star)$.
>
> For the other direction, note the following : suppose that $(\star)$ holds. Then $\int f dT_* d\mu = \int f \circ T d\mu = \int f d\mu$ for all $f \in C(X)$. Now by the corollary before, we have $\mu = T_*\mu$.
>
> *(End of proof)* $\square$

Theorem) Let (X, T) be a topological dynamical system. Let $(\nu_j)_j$ be a sequence of Borel probability measures on X. Let $(N_j) \subset \mathbb{Z}_{>0}$ be sequence s.t. $N_j \to \infty$ as $j \to \infty$. Let μ be the weak limit of a subsequence of

$$\frac{1}{N_j} \sum_{n=0}^{N_j - 1} T_*^n \nu_j$$

Then μ is T-invariant.

proof) Fix $f \in C(X)$. Wlog, assume $w - \lim \frac{1}{N_j} \sum_{n=0}^{N_j-1} T_*^n \nu_j = \mu$.

$$\begin{aligned}
\int f \circ T d\mu &= \lim_{j\to\infty} \int f \circ T d\Big(\frac{1}{N_j} \sum_{n=0}^{N_j 01} T^n) * \nu_j\Big) \\
&= \lim_{j\to\infty} \frac{1}{N_j} \sum_{n=0}^{N_j-1} \int f \circ T dT_*^n \nu_j \\
&= \lim_{j\to\infty} \frac{1}{N_j} \sum_{n=0}^{N_j-1} \int f \circ T^{n+1} dT \nu_j \\
&= \lim_{j\to\infty} \frac{1}{N_j} \sum_{n=1}^{N_j} \int f \circ T^n dT \nu_j
\end{aligned}$$

Now we can expand $\int f d\mu$ similarly

$$\int f d\mu = \lim_{j\to\infty} \frac{1}{N_j} \sum_{n=0}^{N_j-1} \int f \circ T^n d\nu_j$$

then

$$\Big| \int f d\mu - \int f \circ T d\mu \Big| \leq \limsup_{j\to\infty} \frac{\| f \|_\infty + \| f \|_\infty}{N_j} = 0$$