Part III – Topics in Ergodic Theory (Ongoing course, rough)

Based on lectures by Dr. P. Varjú
Notes taken by Bhavik Mehta

Michaelmas 2018

Contents

Ergodic theory is all about measure preserving systems.

Definition (Measure preserving system)**.** A **measure preserving system** $(X, \mathcal{B}, \mu, T)$ with X a set, $\mathcal{B}$ a σ-algebra, μ a probability measure ($\mu(A) \geq 0\ \forall A \in \mathcal{B}$ and $\mu(X) = 1$) and T is a measure preserving transformation. Recall a measure preserving transformation $T : X \to X$ is a measurable function such that $\mu(T^{-1}(A)) = \mu(A)\ \forall A \in \mathcal{B}$.

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

Example. For example, consider a circle rotation. We have $X = \mathbb{R}/\mathbb{Z}$, $\mathcal{B}$ is the Borel sets, μ the Lebesgue measure, and $T = R_\alpha$, with $x \mapsto x + \alpha$ and $\alpha \in \mathbb{R}/\mathbb{Z}$ is a parameter.

We also have the 'times 2 map', with the same $X, \mathcal{B}, \mu$ and $T = T_2$, $x \mapsto 2 \cdot x$.

Proof that T_2 is measure preserving. First check for intervals: Let $I = (a, b)$, then $\mu(I) = b - a$. Also, $\mu(T_2^{-1} I) = \mu\left((\frac{a}{2}, \frac{b}{2}) \cup (\frac{a}{2} + \frac{1}{2}, \frac{b}{2} + \frac{1}{2})\right) = \frac{b}{2} - \frac{a}{2} + \frac{b}{2} - \frac{a}{2} = b - a$, as required.

Now, let $U \subset \mathbb{R}/\mathbb{Z}$ be open. Then $U = I_1 \sqcup I_2 \sqcup \cdots$ is a disjoint union of intervals:

$$\begin{aligned}
\mu(T^{-1}U) &= \mu\left(\bigcup T^{-1} I_j\right) \\
&= \sum \mu(T^{-1} I_j) \\
&= \sum \mu(I_j) \\
&= \mu(U).
\end{aligned}$$

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

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

Now let $A \in \mathcal{B}$ be arbitrary. Let $\epsilon > 0$. $\exists U$ open and $\exists K$ compact such that $K \subset A \subset U$ and $\mu(U \setminus K) < \epsilon$.

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

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

The two examples generalise to the Haar measure on a topological group and to endomorphisms respectively.

In ergodic theory, we study the long term behaviour of orbits.

Definition (Orbit)**.** The orbit of $x \in X$ is the sequence

$$x, Tx, T^2x, \ldots$$

Some questions we might ask are:

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

Example. Let $A = [0, \frac{1}{4}) \subset \mathbb{R}/\mathbb{Z}$. Then ${T_2}^n x \in A \iff$ the $n+1$th and $n+2$th 'binary digits' of x are 0.

For some $x = 0.x_1 x_2 x_3 \ldots_2$, $x \in A$ corresponds to x_1, x_2 both being 0 and the doubling map sends x to $T_2 x = x_2 x_3 \ldots_2$, giving the required property above.

For example, $x = \frac{1}{6} = 0.00101010\ldots_2$ starts in A but never comes back to A. Also, we have $\mu(\{\, x \in A \mid T_2^n x \,\}) = \frac{1}{16}$ if $n \geq 2$.

Example (Markov shift)**.** Let $P_1, P_2, \ldots, P_n$ be a probability vector. Let $A \in \mathbb{R}_{\geq 0}^{n \times n}$ be the 'matrix of transition probabilities'. Assume

$$A \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}, \begin{pmatrix} P_1 & P_2 & \ldots & P_n \end{pmatrix} A = \begin{pmatrix} P_1 & P_2 & \ldots & P_n \end{pmatrix}$$

Take $X = \{1, \ldots, n\}^{\mathbb{Z}}$, $\mathcal{B}$ the Borel σ-algebra generated by the product topology of the discrete topology on $\{1, \ldots, n\}$, $T = \sigma$ the shift map: $(\sigma x)_m = x_{m+1}$. Finally, set the measure

$$\mu(\{\, x \in X \mid x_m = i_0, x_{m+1} = i_1, \ldots, x_{m+n} = i_n \,\}) = P_{i_0} a_{i_0 i_1} \cdots a_{i_{n-1} i_n}.$$

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

$$\bar{d}(S) \coloneqq \limsup_{N,M : M-N \to \infty} \frac{1}{M-N} \big| S \cap [N, M-1] \big|$$

and $\bar{d}(S) > 0$. Then S contains arbitrarily long arithmetic progressions. That is, $\forall l, \exists a \in \mathbb{Z}, d \in \mathbb{Z}_{>0}$,

$$a, a+d, \ldots, a+(l-1)d \in S.$$

Theorem (Furstenberg, multiple recurrence)**.** Let $(X, \mathcal{B}, \mu, T)$ be a measure preserving system. Let $A \in \mathcal{B}$ such that $\mu(A) > 0$. Let $l \in \mathbb{Z}_{>0}$. Then

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

Let

- $X = \{0, 1\}^{\mathbb{Z}}$
- $\mathcal{B} =$ Borel σ-algebra
- $\sigma =$ the shift map $\mathbf{x} \mapsto (x_{n+1})_n$

Let $\mathbf{x}^S \in X$ be defined by

$$\mathbf{x}_n^S = \begin{cases} 1 & n \in S \\ 0 & n \notin S. \end{cases}$$

Also let $A \in \beta$ be given by $A = \{\, x \in X \mid x_0 = 1 \,\}$. Observe then that

$$\mathbf{x}_n^S = 1 \iff n \in S \iff \sigma^n \mathbf{x}^S \in A \iff (\sigma^n \mathbf{x}^S)_0 = 1.$$

Let $\{M_m\}$ and $\{N_m\}$ be sequences s.t. $M_n - N_m \to \infty$ and

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

Let

$$\mu_m = \frac{1}{M_m - N_m} \sum_{n=N_m}^{M_m-1} \delta_{\sigma^n \mathbf{x}^S}$$

where δ_x is a measure on X defined as

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

Let μ be the weak limit of *a* subsequence of μ_m. Note how the μ could be different dependent on subsequence choice.

Definition (Weak limit)**.** 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 converges weakly to μ if for any $f \in C(X)$, we have

$$\lim_{n\to\infty} \int_X f \, d\mu_n = \int_X f \, d\mu.$$

Theorem. (Banach-Alaoglu, or Helly) Let X be a compact metric space. Then $\mathcal{M}(X)$, the set of Borel probability measures on X, endowed with the topology of weak convergence, is compact and metrizable. That is, there is a weakly convergent subsequence in any sequence of Borel probability measures.

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

Proof sketch. Let $B \in \mathcal{B}$. Then

$$\begin{aligned} \mu_m(B) &= \frac{1}{M_m - N_m} \left|\{\, n \in [N_m, M_m - 1] \mid \sigma^n \mathbf{x}^S \in \mathcal{B} \,\}\right| \\ \mu_m(\sigma^{-1}B) &= \frac{1}{M_m - N_m} \left|\{\, n \in [N_m, M_m - 1] \mid \sigma^n \mathbf{x}^S \in \sigma^{-1}\mathcal{B} \,\}\right| \\ &= \frac{1}{M_m - N_m} \left|\{\, n \in [N_m + 1, M_m] \mid \sigma^n \mathbf{x}^S \in \mathcal{B} \,\}\right| \end{aligned}$$

So the difference is such that

$$\left|\mu_m(B) - \mu_m(\sigma^{-1}B)\right| \le \frac{1}{M_m - N_m} \to 0$$

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

Remark. If B is a cylinder set, i.e. $\exists L \in \mathbb{Z}_{>0}$ and $\tilde{B} \subseteq \{0,1\}^{2L+1}$ such that

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

then B is both closed and open. Therefore χ_B, the characteristic function of B is continuous. Hence $\lim_{n\to\infty} \mu_m(B) = \mu(B)$, since $\mu_m(B) = \int \chi_B \, d\mu_m$ and $\mu(B) = \int \chi_B \, d\mu$.

Approximating any Borel set by such cylinder sets would help complete the proof, but we in fact can get this result on spaces where χ is not continuous on nice set of sets. So we leave full proof till a more general theorem.

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

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

Then S contains an arithmetic progression of length l.

Proof. Without loss of generality, we can assume $\mu = \lim \mu_m$ - if not, pass to a subsequence. Let $B = A \cap \sigma^{-n} A \cap \cdots \cap \sigma^{-n(l-1)}(A)$. Observe that B is a cylinder set. Then by the earlier remark, $\mu(B) = \lim \mu_m(B)$, hence $\exists m$ such that $\mu_m(B) > 0$.

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

$$\sigma^k \mathbf{x}^S \in A, \sigma^k \mathbf{x}^S \in \sigma^{-n}(A), \ldots, \sigma^k \mathbf{x}^S \in \sigma^{-n(l-1)}(A).$$

Thus, $k, k+n, \ldots, k+n(l-1) \in S$. □

Returning to the overall proof, we note A is a cylinder set. Then $\mu_m(A) \to \mu(A)$, i.e.

$$\mu(A) = \underbrace{\lim_{m \to \infty} \frac{1}{M_m - N_m} \left|\{\, n \in [N_m, M_m - 1] : n \in S \,\}\right|}_{\bar{d}(S)} > 0$$

where the inequality comes from satisfying the conditions of Furstenberg.

Index