Lecture 23

Basic ergodic theory
Furstenberg correspondence principle and Sárközy theorem

Graduate Real Analysis

Fall Semester

Measure-preserving systems

Definition

Suppose that $(X_1, \mathcal{M}_1, \mu_1)$ and $(X_2, \mathcal{M}_2, \mu_2)$ are two measure spaces.

1. A transformation $T : X_1 \to X_2$ is **measurable** if

$$T^{-1}[E] \in \mathcal{M}_1 \quad \text{for every} \quad E \in \mathcal{M}_2.$$

2. A transformation $T : X_1 \to X_2$ is **measure-preserving** if T is measurable and

$$\mu_1(T^{-1}[E]) = \mu_2(E) \quad \text{for every} \quad E \in \mathcal{M}_2.$$

3. We say that $T : X_1 \to X_2$ is an **invertible measure-preserving transformation** if T is measure-preserving, bijective, and T^{-1} is also measure-preserving.

Examples of measure-preserving systems

1. **The integer shift system:** $(\mathbb{Z}, \mathcal{P}(\mathbb{Z}), |\cdot|, S)$ with $S : \mathbb{Z} \to \mathbb{Z}$ given by
$$S(x) = x + 1.$$
2. **The circle rotation system:** $(\mathbb{T}, \mathcal{L}(\mathbb{T}), \mathrm{d}x, R_\alpha)$ with the rotation map $R_\alpha : \mathbb{T} \to \mathbb{T}$ by $R_\alpha(x) = x + \alpha \pmod 1$ for $\alpha \in \mathbb{R} \setminus \mathbb{Q}$.
3. **The circle-doubling system:** $(\mathbb{T}, \mathcal{L}(\mathbb{T}), \mathrm{d}x, D_2)$ with the doubling map $D_2 : \mathbb{T} \to \mathbb{T}$ given by $D_2(x) = 2x \pmod 1$.
4. **The continued fraction system:** $([0,1), \mathcal{L}([0,1)), \mu, T)$ with the Gauss measure
$$\mu(A) = \frac{1}{\log 2} \int_A \frac{\mathrm{d}x}{1+x},$$
and continued fraction map $T : [0,1) \to [0,1)$ given by $T(0) = 0$ and
$$T(x) = \frac{1}{x} \pmod 1, \qquad \text{when } x \neq 0.$$

Bernoulli shift

- Define the measure ν on $\{0,1\}$ by setting $\nu(\{0\}) = \nu(\{1\}) = 1/2$.
- Let $X = \{0,1\}^{\mathbb{N}}$ with $\mu = \bigotimes_{n\in\mathbb{N}} \nu$. The space $(X, \mathrm{Bor}(X), \mu)$ is a natural model for the set of possible outcomes of an infinitely repeated toss of a fair coin. The left shift map $\sigma : X \to X$ defined by $\sigma(x_0, x_1, \ldots) = (x_1, x_2, \ldots)$ preserves μ.
- The map $\phi : X \to \mathbb{T}$ defined by

$$\phi(x_0, x_1, \ldots) = \sum_{n=0}^{\infty} \frac{x_n}{2^{n+1}}$$

  is measure-preserving from $(X, \mathrm{Bor}(X), \mu)$ to $(\mathbb{T}, \mathcal{L}(\mathbb{T}), dx)$ and

$$\phi(\sigma(x)) = D_2(\phi(x)) = \{2\phi(x)\}.$$

- The map ϕ has a measurable inverse defined on all but the countable set of dyadic rationals, so this shows that $(X, \mathrm{Bor}(X), \mu, \sigma)$ and $(\mathbb{T}, \mathcal{L}(\mathbb{T}), dx, D_2)$ are measurably isomorphic.

Unitary operators

- Let $(X, \mathcal{M}, \mu, T)$ be a measure-preserving system.
- Define $U_T : L^2(X, \mu) \to L^2(X, \mu)$ by $U_T f = f \circ T$. Then

$$\langle U_T f_1, U_T f_2 \rangle = \int_X f_1 \circ T \cdot \overline{f_2 \circ T} d\mu = \int_X f_1 \overline{f_2} d\mu = \langle f_1, f_2 \rangle$$

  since μ is T-invariant.
- Thus, U_T is an isometry from $L^2(X, \mu)$ into $L^2(X, \mu)$ whenever $(X, \mathcal{M}, \mu, T)$ is a measure preserving system.
- If $U : H_1 \to H_2$ is a continuous linear operator from one Hilbert space to another, then the relation $\langle Uf, g \rangle = \langle f, U^* g \rangle$ defines the adjoint operator $U^* : H_2 \to H_1$.
- The operator U is an isometry, i.e., $\|Uh\|_{H_2} = \|h\|_{H_1}$ for all $h \in H_1$ if and only if $U^* U = I_{H_1}$ and $UU^* = P_{R(U)}$ (the projection onto $R(U)$).
- In conclusion, for any measure-preserving transformation T, the associated operator U_T is an isometry, and, if T is invertible, then U_T is a unitary operator, sometimes called the **Koopman operator** of T.

von Neumann ergodic theorem, revised

Theorem

Let $(X, \mathcal{M}, \mu, T)$ be a measure-preserving system and let P_T be the orthogonal projection onto the closed subspace

$$I = \{g \in L^2(X, \mu) : U_T g = g\}.$$

Then, for any $f \in L^2(X, \mu)$, one has

$$\frac{1}{N} \sum_{n=0}^{N-1} U_T^n f = \frac{1}{N} \sum_{n=0}^{N-1} f \circ T^n \xrightarrow[N \to \infty]{} P_T f \quad \text{in} \quad L^2(X, \mu).$$

Proof.

We apply the von Neumann ergodic theorem for the Koopman operator, which is the unitary operator $Uf = U_T f = f \circ T$. $\square$

Poincaré recurrence theorem

Theorem (Poincaré recurrence theorem)

Let $(X, \mathcal{M}, \mu, T)$ *be a measure-preserving system with* $\mu(X) < \infty$. *Let*

$$E \in \mathcal{M} \quad \textit{with} \quad \mu(E) > 0.$$

Then almost all points of E *return to* E *infinitely often under positive iteration by* T*: there is* $F \in \mathcal{M}$ *with* $F \subseteq E$ *and* $\mu(E) = \mu(F)$ *such that, for each* $x \in F$*, there is a sequence* $n_1 < n_2 < \dots$ *with* $T^{n_i}x \in F$ *for all* $i \in \mathbb{N}$. *In particular,* $\mu(E \cap T^{-n}[E]) > 0$ *for infinitely many* $n \in \mathbb{N}$.

Proof. For $N \geq 0$, let $E_N = \bigcup_{n=N}^{\infty} T^{-n}[E]$. Then $\bigcap_{N\geq 0} E_N$ is the set of all points of X that enter E infinitely often under positive iteration by T. Define

$$F = E \cap \bigcap_{N \geq 0} E_N.$$

We will show that $\mu(E) = \mu(F)$, and the other properties are clear.

Proof

- Since $E_N = \bigcup_{n=N}^{\infty} T^{-n}[E]$, note that

$$T^{-1}[E_N] = \bigcup_{n=N}^{\infty} T^{-n-1}[E] = E_{N+1} \subseteq E_N.$$

- Moreover, $\mu(E_{N+1}) = \mu(T^{-1}[E_N]) = \mu(E_N)$, since T preserves μ.
- Hence, $\mu(E_N \setminus E_{N+1}) = \mu(E_N) - \mu(E_{N+1}) = 0$.
- By the continuity and finiteness of μ we have

$$\mu(F) = \mu\Big(\bigcap_{N=0}^{\infty} (E \cap E_N) \Big) = \lim_{N\to\infty} \mu(E \cap E_N).$$

- But $\mu(E \cap E_N) = \mu(E \cap E_{N+1})$ since $\mu(E_N \setminus E_{N+1}) = 0$ and $E_{N+1} \subseteq E_N$. Thus, $\mu(E \cap E_N) = \mu(E \cap E_0)$ and

$$\mu(F) = \lim_{N\to\infty} \mu(E \cap E_N) = \mu(E \cap E_0) = \mu(E),$$

since $E \subseteq E_0 = \bigcup_{n=0}^{\infty} T^{-n}[E]$. $\square$

Ergodicity

Definition

A measure-preserving transformation $T : X \to X$ of a measure space $(X, \mathcal{M}, \mu)$ is called **ergodic** if, for any $B \in \mathcal{M}$, the following holds

$$\mu(T^{-1}[B] \triangle B) = 0 \quad \Longrightarrow \quad \mu(B) = 0 \quad \text{or} \quad \mu(B^c) = 0.$$

Remark

Ergodicity says that it is impossible to split X into two subsets of positive measure, each of which is invariant under T, if T is ergodic.

Theorem

Let $(X, \mathcal{M}, \mu, T)$ be a measure-preserving system. Then the following statements are equivalent:

(i) *T is ergodic.*

(ii) *For $f : X \to \mathbb{C}$ measurable, $f \circ T = f$ almost everywhere implies that $f \equiv \mathrm{const}$ almost everywhere.*

Proof

Proof (ii) $\implies$ (i). Suppose that $A \in \mathcal{M}$ and $\mu(T^{-1}[A] \triangle A) = 0$. Then

$$\mu(T^{-1}[A] \triangle A) = 0 \quad \iff \quad \mathbb{1}_A \circ T = \mathbb{1}_A \quad \text{almost everywhere.}$$

By **(ii)**, $\mathbb{1}_A$ is T-invariant almost everywhere, so $\mathbb{1}_A \equiv \mathrm{const} \in \{0, 1\}$ almost everywhere, so $\mu(A) = 0$ or $\mu(A^c) = 0$. □

Proof (i) $\implies$ (ii). Let $f : X \to \mathbb{C}$ be measurable and $f \circ T = f$ almost everywhere. We may assume without loss of generality that $f : X \to \mathbb{R}$.

- Suppose for a contradiction that f is not constant, which means that there is $r \in \mathbb{R}$ such that

$$\mu(E_r)\mu(E_r^c) > 0, \quad \text{where} \quad E_r = \{x \in X : f(x) > r\} \in \mathcal{M}.$$

- Since $f \circ T = f$ almost everywhere, then $\mu(T^{-1}[E_r] \triangle E_r) = 0$ for each $r \in \mathbb{R}$. Thus $\mu(E_r) = 0$ or $\mu(E_r^c) = 0$ for each $r \in \mathbb{R}$, which is impossible! Thus f must be constant almost everywhere. □

Equivalent forms of ergodicity

Theorem

The following are equivalent properties for a measure-preserving transformation T of a probability measure space $(X, \mathcal{M}, \mu)$.

1. *For $A \in \mathcal{M}$, $T^{-1}[A] = A$ implies $\mu(A) \in \{0, 1\}$.*
2. *For $B \in \mathcal{M}$, $\mu(T^{-1}[B] \triangle B) = 0$ implies $\mu(B) \in \{0, 1\}$. (Ergodicity).*
3. *For $A \in \mathcal{M}$, $\mu(A) > 0$ implies that*
$$\mu\left(\bigcup_{n=1}^{\infty} T^{-n}[A]\right) = 1.$$
4. *For $A, B \in \mathcal{M}$, $\mu(A)\mu(B) > 0$ implies that there exists $n \in \mathbb{N}$ with*
$$\mu(T^{-n}[A] \cap B) > 0.$$

Proof

Proof (1) $\Longrightarrow$ (2). Assume that T is ergodic and let $B \in \mathcal{M}$ with

$$\mu(T^{-1}[B] \triangle B) = 0.$$

Although B may not be T-invariant, the idea is to construct a T-invariant set C with the same measure as B.

- Let $C_N = \bigcup_{n=N}^{\infty} T^{-n}[B]$ and $C = \bigcap_{N=0}^{\infty} C_N$. For any $N \geq 0$, we have

$$B \triangle C_N \subseteq \bigcup_{n=N}^{\infty} B \triangle T^{-n}[B].$$

- Using that $U \triangle V \subseteq U \triangle W \cup W \triangle V$, we write

$$B \triangle T^{-n}[B] \subseteq \bigcup_{i=0}^{n-1} T^{-i}[B] \triangle T^{-(i+1)}[B],$$

  and this has measure zero, since

$$\mu(T^{-i}[B] \triangle T^{-(i+1)}[B]) = \mu(B \triangle T^{-1}[B]) = 0.$$

Proof

- Therefore, $\mu(B \triangle C_N) = 0$ for each $N \geq 0$. Since $C_0 \supseteq C_1 \supseteq C_2 \supseteq \ldots$, we conclude $\mu(B \triangle C) = 0$, and consequently $\mu(B) = \mu(C)$. Then

$$T^{-1}[C] = \bigcap_{N=0}^{\infty} \bigcup_{n=N}^{\infty} T^{-(n+1)}[B] = \bigcap_{N=0}^{\infty} \bigcup_{n=N+1}^{\infty} T^{-n}[B] = C.$$

- Hence, $T^{-1}[C] = C$, so $\mu(C) \in \{0, 1\}$ by ergodicity of T. Therefore, $\mu(B) \in \{0, 1\}$ as desired. □

Proof (2) $\Longrightarrow$ (3). Let A be a set with $\mu(A) > 0$ and let

$$B = \bigcup_{n=1}^{\infty} T^{-n}[A].$$

Then $T^{-1}[B] \subseteq B$ and $\mu(T^{-1}[B]) = \mu(B)$, so $\mu(T^{-1}[B] \triangle B) = 0$, implying $\mu(B) \in \{0, 1\}$. Since $T^{-1}[A] \subseteq B$, we deduce that $\mu(B) = 1$. □

Proof

Proof (3) $\Longrightarrow$ (4). Let $A, B \in \mathcal{M}$ be sets of positive measure. By **(3)**, $\mu(\bigcup_{n=1}^{\infty} T^{-n}[A]) = 1$, so

$$0 < \mu(B) = \mu(\bigcup_{n=1}^{\infty} B \cap T^{-n}[A]) \leq \sum_{n=1}^{\infty} \mu(B \cap T^{-n}[A]).$$

Thus, there must be some $n \in \mathbb{N}$ such that

$$\mu(B \cap T^{-n}[A]) > 0. \qquad \square$$

Proof (4) $\Longrightarrow$ (1). Let $A \in \mathcal{M}$ and $T^{-1}[A] = A$. Then

$$0 = \mu(A \cap A^c) = \mu(T^{-n}[A] \cap A^c) \quad \text{for all} \quad n \geq 1.$$

Thus, $\mu(A) = 0$ or $\mu(A^c) = 0$. Otherwise, this contradicts **(4)**. $\square$

Examples of ergodic maps

Theorem

The integer shift, circle-rotation, circle-doubling, Bernoulli shift and continued fraction systems, are all ergodic.

Proof. Here we will use $e(z) = e^{2\pi i z}$ for $z \in \mathbb{R}$.

- **The integer shift:** If $f : \mathbb{Z} \to \mathbb{C}$ is such that $f(x+1) = f(x)$ for all $x \in \mathbb{Z}$, then is f is constant, since

$$f(x) = f(x+y) = f(y) \quad \text{for all} \quad x, y \in \mathbb{Z}.$$

- **The circle-rotation system:** Take $f \in L^2(\mathbb{T})$ with $f(x+\alpha) = f(x)$ for $x \in \mathbb{T}$ and $\alpha \in \mathbb{R} \setminus \mathbb{Q}$. Then expanding f as a Fourier series

$$f(x) = \sum_{n \in \mathbb{Z}} c_n e(nx) = \sum_{n \in \mathbb{Z}} c_n e(\alpha n) e(nx) = f(x+\alpha).$$

  Thus, $c_n = e(\alpha n) c_n$ for all $n \in \mathbb{Z}$, but $\alpha \in \mathbb{R} \setminus \mathbb{Q}$, so $e(\alpha n) \neq 1$ for all $n \in \mathbb{Z} \setminus \{0\}$, so $c_n = 0$ for all $n \in \mathbb{Z} \setminus \{0\}$ and $f(x) = c_0$ as desired.

Proof

- **The circle-doubling system:** Similarly, take $f \in L^2(\mathbb{T})$ with $f(2x) = f(x)$ for $x \in \mathbb{T}$. Then expanding f as a Fourier series

$$f(x) = \sum_{n \in \mathbb{Z}} c_n e(nx) = \sum_{n \in \mathbb{Z}} c_n e(2nx) = f(2x).$$

Thus, $c_{2n} = c_n$ for all $n \in \mathbb{Z}$. If there is $c_n \neq 0$ for some $n \neq 0$, then

$$\sum_{n \in \mathbb{Z}} |c_n|^2 = \infty,$$

but on the other hand, by Parseval's identity we know that

$$\sum_{n \in \mathbb{Z}} |c_n|^2 = \|f\|_{L^2}^2 < \infty.$$

Thus, $f(x) = c_0$ as desired.

- **The Bernoulli shift system:** This system is ergodic since it is measurably isomorphic to the circle-doubling system.
- **The continued fraction system:** Exercise! $\square$

Total ergodicity

Definition

A measure-preserving system $(X, \mathcal{M}, \mu, T)$ is **totally ergodic** if, for every $n \in \mathbb{N}$, the measure-preserving system $(X, \mathcal{M}, \mu, T^n)$ is ergodic.

1. The circle rotation system $(\mathbb{T}, \mathcal{L}(\mathbb{T}), dx, R_\alpha)$, $R_\alpha(x) = (x + \alpha) \bmod 1$ is totally ergodic if and only if $\alpha \in \mathbb{R} \setminus \mathbb{Q}$.
2. Let $X = \{0, 1\}$ be a two point space, $\mathcal{M} = \{\emptyset, \{0\}, \{1\}, \{0, 1\}\}$, $\mu(x) = \frac{1}{2}(\delta_0(x) + \delta_1(x))$, and $T(x) = x + 1 \bmod 2$. This system is ergodic since the only sets with measure in $(0, 1)$ are $\{0\}$ and $\{1\}$, and neither is T-invariant. However, this system is not totally ergodic since T^2 is the identity map, so $\{0\}$ and $\{1\}$ are T^2-invariant sets.

Remark

While this counterexample seems trivial, it turns out that finite systems are, in some sense, the only obstruction to total ergodicity.

Total ergodicity

Theorem

A measure-preserving system $(X, \mathcal{M}, \mu, T)$ is totally ergodic if and only if the Koopman operator $U_T f = f \circ T$ admits no rational eigenvalues of the form $e(p/q)$ other than 1.

Proof of ($\Longrightarrow$). Assume that $(X, \mathcal{M}, \mu, T)$ is totally ergodic and suppose that there are $p \in \mathbb{Z}$ and $q \in \mathbb{Z}$ such that $e(p/q)$ is an eigenvalue for the function f. Then

$$U_T f = f \circ T = e(p/q) f.$$

Thus, $U_T^q f = f \circ T^q = f$, so $f \equiv \mathrm{const}$, so $e(p/q) = 1$. □

Proof of ($\Longleftarrow$). Conversely, suppose that $(X, \mathcal{M}, \mu, T)$ is not totally ergodic. Then we let f to be a non-constant function such that $U_T^q f = f \circ T^q = f$ for some $q \in \mathbb{N}$. Consider

$$F_j = \sum_{k=1}^{q} e(-jk/q) f \circ T^k.$$

von Neumann's ergodic theorem for arithmetic progressions

Then

$$\sum_{j=1}^{q} F_j = \sum_{k=1}^{q}\sum_{j=1}^{q} e(-jk/q) f \circ T^k = qf \circ T^q = qf.$$

Thus, there is $1 \le j \le q$ such that F_j is non-constant. Then

$$F_j \circ T = \sum_{k=1}^{q} e(-jk/q) f \circ T^{k+1} = e(j/q) \sum_{k=2}^{q+1} e(-jk/q) f \circ T^k = e(j/q) F_j.$$

Thus, F_j is an eigenfunction with eigenvalue $e(j/q)$. □

Theorem

Let $(X, \mathcal{M}, \mu, T)$ be a measure-preserving system with $\mu(X) = 1$. Then it is totally ergodic if and only if, for every $f \in L^2(X)$ and $q, r \in \mathbb{N}$, we have

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} f(T^{qn+r}x) = \int_X f d\mu \quad \text{in} \quad L^2(X). \qquad (*)$$

Proof

Proof of ($\Longleftarrow$). If the system is not totally ergodic, then there exists $q \in \mathbb{N}$ and a non-constant $f \in L^2(X)$ such that $f(T^q x) = f(x)$. Then (*) implies that

$$f(x) = \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} f(T^{qn}x) = \int_X f d\mu,$$

giving a contradiction!

Proof of ($\Longrightarrow$). Conversely, suppose that $(X, \mathcal{M}, \mu, T)$ is totally ergodic and let $f \in L^2(X)$ and $q, r \in \mathbb{N}$. Applying the mean ergodic theorem, we see that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} f(T^{qn+r}x) = \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} f(T^{qn}(T^r x))$$
$$= \int_X f(T^r x) d\mu(x) = \int_X f d\mu.$$

This completes the proof. $\square$

van der Corput's inequality in Hilbert spaces

Theorem (van der Corput's inequality in Hilbert spaces)

Let $(u_n)_{n\in\mathbb{N}}$ be a bounded sequence in a Hilbert space $(H, \|\cdot\|)$. Define a sequence $(s_n)_{n\in\mathbb{N}}$ of real numbers by

$$s_m = \limsup_{N\to\infty} \left| \frac{1}{N} \sum_{n=1}^{N} \langle u_{n+m}, u_n \rangle \right|.$$

If

$$\lim_{H\to\infty} \frac{1}{H} \sum_{h=0}^{H-1} s_h = 0,$$

then

$$\lim_{N\to\infty} \left\| \frac{1}{N} \sum_{n=1}^{N} u_n \right\| = 0.$$

Polynomial ergodic theorem

Theorem

Let $(X, \mathcal{M}, \mu, T)$ be a totally ergodic measure-preserving system with $\mu(X) = 1$. Let $P \in \mathbb{Z}[n]$, and assume that $T : X \to X$ is invertible if P takes negative values. Then, for any $f \in L^2(X)$, we have

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} f(T^{P(n)}x) = \int_X f d\mu \quad \text{in} \quad L^2(X). \qquad (*)$$

Proof. We proceed by induction on the degree of P. If P is linear, then the result follows from the previous theorem. Assume that $\deg(\mathrm{P}) \geq 2$.

- Note that (*) holds for f if and only if it holds for $f - c$ for any $c \in \mathbb{C}$.
- In particular, taking $c = \int_X f d\mu$, we can assume that $\int_X f d\mu = 0$.
- Let $u_n = f(T^{P(n)}x)$. We will use the van der Corput lemma to prove

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} u_n = 0 \quad \text{in} \quad L^2(X).$$

Proof

- Fix $h \in \mathbb{N}$ and note that
$$\langle u_{n+h}, u_n \rangle = \int_X f(T^{P(n+h)}x)\overline{f(T^{P(n)}x)}d\mu = \langle f \circ T^{P(n+h)-P(n)}, f \rangle$$
- Since $Q(n) = P(n+h) - P(n)$ is of lower degree than P, by the induction hypothesis, we have for every $h \in \mathbb{N}$ that
$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} f \circ T^{P(n+h)-P(n)} = 0 \quad \text{in} \quad L^2(X).$$
- By the Cauchy–Schwarz inequality
$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \langle u_{n+h}, u_n \rangle = \lim_{N\to\infty} \left\langle \frac{1}{N} \sum_{n=1}^{N} f \circ T^{P(n+h)-P(n)}, f \right\rangle = 0.$$
- This establishes the hypothesis of the van der Corput lemma, so
$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} u_n = 0 \quad \text{in} \quad L^2(X). \qquad \square$$

Polynomial recurrence theorem for totally ergodic systems

Corollary (Polynomial recurrence theorem for totally ergodic systems)

For any totally ergodic system $(X, \mathcal{M}, \mu, T)$ *with* $\mu(X) = 1$ *and* T *invertible, any* $A \in \mathcal{M}$, $P \in \mathbb{Z}[n]$, *and* $\varepsilon > 0$, *there exists* $n \in \mathbb{N}$ *such that*

$$\mu(A \cap T^{-P(n)}[A]) > \mu(A)^2 - \varepsilon.$$

Proof. From the previous theorem, we have

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \mathbb{1}_A(T^{P(n)}x) = \mu(A) = \int_X \mathbb{1}_A d\mu$$

Thus, by the **(DCT)** we conclude

$$\lim_{N\to\infty} \langle \frac{1}{N} \sum_{n=1}^{N} \mathbb{1}_A(T^{P(n)}x), \mathbb{1}_A(x) \rangle = \int_X \mu(A)\mathbb{1}_A(x) d\mu = \mu(A)^2.$$

Polynomial Poincaré recurrence theorem

But
$$\mu(T^{-P(n)}[A]\cap A)=\langle \mathbb{1}_A\circ T^{P(n)},\mathbb{1}_A\rangle.$$
Hence,
$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\mu(T^{-P(n)}[A]\cap A)=\mu(A)^2.$$
Consequently, for any $\varepsilon>0$, there is $n\in\mathbb{N}$ such that
$$\mu(T^{-P(n)}[A]\cap A)>\mu(A)^2-\varepsilon. \qquad \square$$

Theorem (Polynomial Poincaré recurrence theorem)

For any measure-preserving system $(X,\mathcal{M},\mu,T)$ with $\mu(X)=1$, any $A\in\mathcal{M}$, $\varepsilon>0$, and $P\in\mathbb{Z}[n]$ with $P(0)$, there exists $n\in\mathbb{N}$ such that
$$\mu(A\cap T^{-P(n)}[A])>\mu(A)^2-\varepsilon,$$
and we assume that $T:X\to X$ is invertible if P takes negative values.

Decomposing $L^2(X)$

Before we give a proof, consider the following subspaces of $L^2(X)$:

1. The rational spectrum component

$$H_{rat} = \overline{\{f \in L^2(X) : f \circ T^k = f \text{ for some } k \in \mathbb{N}\}}.$$

2. The totally ergodic component

$$H_{te} = \left\{ f \in L^2(X) : \lim_{N \to \infty} \left\| \frac{1}{N} \sum_{n=1}^{N} f \circ T^{nk} \right\|_{L^2(X)} = 0 \text{ for all } k \in \mathbb{N} \right\}.$$

Observe that, for a totally ergodic system, H_{rat} consists only of constant functions and H_{te} contains every function with zero integral.

Proposition

For any measure-preserving system $(X, \mathcal{M}, \mu, T)$, the spaces H_{rat} and H_{te} are orthogonal and

$$L^2(X) = H_{rat} \oplus H_{te}.$$

Proof

Proof. Let $f \in L^2(X)$ with $f \circ T^k = f$ for some $k \in \mathbb{N}$ and let $g \in H_{te}$. Then for all $n \in \mathbb{N}$, we have

$$\langle f, g \rangle = \langle f \circ T^k, g \circ T^k \rangle = \langle f, g \circ T^k \rangle = \ldots = \langle f, g \circ T^{kn} \rangle.$$

Averaging over n, we have

$$\langle f, g \rangle = \lim_{N \to \infty} \left\langle f, \frac{1}{N} \sum_{n=1}^{N} g \circ T^{kn} \right\rangle = 0,$$

by the Cauchy–Schwarz inequality, which implies that $H_{te} \subseteq H_{rat}^{\perp}$.

Conversely, let $f \in H_{rat}^{\perp}$. For every $k \in \mathbb{N}$, H_{rat} contains the invariant subspace I_k of functions invariant under T^k. Then f is orthogonal to I_k for every $k \in \mathbb{N}$, in particular $\langle f, 1 \rangle = 0$, so, by the mean ergodic theorem,

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} f \circ T^{kn} = \int_X f d\mu = \langle f, 1 \rangle = 0,$$

so $f \in H_{te}$, so $H_{rat}^{\perp} \subseteq H_{te}$. Hence, $H_{rat}^{\perp} = H_{te}$ as claimed. $\square$

Proof of the polynomial Poincaré recurrence theorem

Theorem (Polynomial Poincaré recurrence theorem)

For any measure-preserving system $(X, \mathcal{M}, \mu, T)$ *with* $\mu(X) = 1$*, any* $A \in \mathcal{M}$*,* $\varepsilon > 0$*, and* $P \in \mathbb{Z}[n]$ *with* $P(0)$*, there exists* $n \in \mathbb{N}$ *such that*

$$\mu(A \cap T^{-P(n)}[A]) > \mu(A)^2 - \varepsilon,$$

and we assume that $T : X \to X$ *is invertible if* P *takes negative values.*

Proof. Let $A \in \mathcal{M}$. Decompose $\mathbb{1}_A = f + g$ for some $f \in H_{rat},\ g \in H_{te}$ with $f \perp g$. Since H_{rat} contains the constants we have $1 \perp g$. Thus by the Cauchy–Schwarz inequality we obtain

$$\langle \mathbb{1}_A, f \rangle = \langle f + g, f \rangle = \|f\|^2 \geq \langle f, 1 \rangle^2 = \langle \mathbb{1}_A - g, 1 \rangle^2 = \mu(A)^2.$$

Find $h \in H_{rat}$ such that $h \circ T^k = h$ for some $k \in \mathbb{N}$ and $\|f - h\| < \varepsilon/2$ (since H_{rat} is defined as a closure). It follows that

$$\langle \mathbb{1}_A, h \rangle > \mu(A)^2 - \varepsilon/2.$$

Proof

Since $P(0) = 0$, then $Q(n) = P(kn) \equiv 0 \mod k$, and $h \circ T^{Q(n)} = h$ for all $k \in \mathbb{N}$. By the van der Corput lemma, we have

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} g \circ T^{Q(n)} = 0,$$

since $\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \langle f \circ T^{Q(n+m)}, f \circ T^{Q(n)} \rangle = 0$ for any $m \in \mathbb{N}$. Finally,

$$\begin{aligned}
\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \mu(A \cap T^{-Q(n)}[A]) &= \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \langle \mathbb{1}_A, \mathbb{1}_A \circ T^{Q(n)} \rangle \\
&= \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \langle \mathbb{1}_A, h + (f-h) \circ T^{Q(n)} + g \circ T^{Q(n)} \rangle \\
&= \langle \mathbb{1}_A, h \rangle - \lim_{N\to\infty} \left\langle \mathbb{1}_A, \frac{1}{N} \sum_{n=1}^{N} (f-h) \circ T^{Q(n)} + g \circ T^{Q(n)} \right\rangle \\
&\geq \langle \mathbb{1}_A, h \rangle - \varepsilon/2 \geq \mu(A)^2 - \varepsilon. \qquad \square
\end{aligned}$$

Furstenberg correspondence principle

Definition

Given a set $A \subseteq \mathbb{N}$, its **upper density** is

$$\overline{d}(A) = \limsup_{N\to\infty} N^{-1}|A \cap [1, N]|.$$

1. $\overline{d}(a\mathbb{N} + b) = \frac{1}{a}$ for any $a, b \in \mathbb{N}$.
2. $\overline{d}(\mathbb{P}) = 0$ where $\mathbb{P}$ is the set of prime numbers.
3. $\overline{d}(S) = \frac{6}{\pi^2}$ where S is the set of square-free numbers.

Theorem (Furstenberg correspondence principle)

Given $E \subseteq \mathbb{N}$, there exists a measure-preserving system $(X, \mathcal{M}, \mu, T)$ with $\mu(X) = 1$ and a set $A \in \mathcal{M}$ with $\mu(A) = \overline{d}(E)$ such that, for every finite $F \subseteq \mathbb{N}$, we have

$$\mu\Big(\bigcap_{n\in F} T^{-n}[A]\Big) \leq \overline{d}\Big(\bigcap_{n\in F} S^{-n}[E]\Big),$$

where $S : \mathbb{Z} \to \mathbb{Z}$ is the integer shift defined by $S(x) = x + 1$ for all $x \in \mathbb{Z}$.

Proof

Proof. Let $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$, and define $X = \{0,1\}^{\mathbb{N}_0}$ with the Borel σ-algebra $\mathrm{Bor}(X)$. Let $T : X \to X$ be the shift map

$$T(x_n)_{n\in\mathbb{N}_0} = (x_{n+1})_{n\in\mathbb{N}_0}.$$

Let $A = \{(x_n)_{n\in\mathbb{N}_0} \in X : x_0 = 1\}$. All that remains is to construct the measure. This will be a consequence of the Banach–Alaoglu theorem.

- Let $(N_j)_{j\in\mathbb{N}}$ be a sequence such that $\overline{d}(E) = \lim_{j\to\infty} \frac{|E\cap[1,N_j]|}{N_j}$. Let $\omega = (\omega_n)_{n\in\mathbb{N}_0} \in X$ be the indicator function of E, so

$$\omega_n = \begin{cases} 1 & \text{if } n \in E, \\ 0 & \text{if } n \notin E. \end{cases}$$

- For each $j \in \mathbb{N}$, let μ_j be the probability measure on $(X, \mathrm{Bor}(X))$ defined by

$$\mu_j = \frac{1}{N_j} \sum_{n=1}^{N_j} \delta_{T^n\omega},$$

where δ_y is the Diract point mass at the point $y \in X$.

Proof

- Observe that

$$\overline{d}(E) = \lim_{j\to\infty} \frac{|E \cap [1, N_j]|}{N_j} = \lim_{j\to\infty} \frac{1}{N_j} \sum_{n=1}^{N_j} \omega_n = \lim_{j\to\infty} \mu_j(A)$$

- The sequence of probability measures $(\mu_j)_{j\in\mathbb{N}}$ defines bounded linear functionals on the space $C(X)$ of continuous functions on X by

$$\Lambda_j(f) = \int_X f(x) d\mu_j(x) \quad \text{for any} \quad f \in C(X), \text{ and } j \in \mathbb{N}.$$

- It is easy to see that $|\Lambda_j(f)| \le |\mu_j| \|f\|_\infty$, where $|\mu_j|$ denotes the variation of the measure μ_j and $\|f\|_\infty = \sup_{x\in X} |f(x)|$.
- Then we obtain that $\sup_{j\in\mathbb{N}} \|\Lambda_j\|_{(C(X))^*} \le 1$, since $\sup_{j\in\mathbb{N}} |\mu_j| = 1$.
- Finally, by the Banach–Alaoglu theorem there is a subsequence $(j_m)_{m\in\mathbb{N}} \subseteq \mathbb{N}$ and a linear functional $\Lambda \in (C(X))^*$ such that

$$\Lambda(f) = \lim_{m\to\infty} \Lambda_{j_m}(f) \quad \text{for any} \quad f \in C(X).$$

Proof

- By the Riesz representation theorem, (which will be proved in 502), we may deduce
$$\Lambda(f) = \int_X f(x) d\mu(x) \quad \text{for any} \quad f \in C(X),$$
for some probability measure μ on X.
- In particular,
$$\mu(A) = \overline{d}(E).$$
- To show that T preserves the measure μ, recall that a cylinder set in X is a set of the form $\{(x_n)_{n\in\mathbb{N}_0} \in X : x_0 = a_0, \ldots, x_n = a_n\}$ for some $n \in \mathbb{N}$ and $a_0, \ldots, a_n \in \{0, 1\}$.
- By the Stone–Weierstrass theorem, (which will be proved in 502), finite linear combinations of indicator functions of cylinder sets form a dense subset of $C(X)$, so it suffices to show that
$$\mu(T^{-1}[B]) = \mu(B)$$
for every cylinder set $B \subseteq X$.

Proof

- One can directly compute that $|\mu_j(B) - \mu_j(T^{-1}[B])| \leq \frac{2}{N_j}$ when B is a cylinder set, and this proves that $(X, \mathrm{Bor}(X), \mu, T)$ is a probability measure-preserving system. We now prove that
$$\mu\Big(\bigcap_{n\in F} T^{-n}[A]\Big) \leq \overline{d}\Big(\bigcap_{n\in F} S^{-n}[E]\Big).$$
- For $n \in \mathbb{N}$, we have $T^{-n}[A] = \{(x_n)_{n\in\mathbb{N}_0} : x_n = 1\}$. Thus, for a finite $F \subseteq \mathbb{N}$ and $j \in \mathbb{N}$, we have
$$\mu_j\Big(\bigcap_{n\in F} T^{-n}[A]\Big) = \frac{1}{N_j}|\{m \in [1, N_j] : (T^m\omega)_n = 1 \text{ for all } n \in F\}|$$
$$= \frac{1}{N_j}|\{m \in [1, N_j] : m + n \in E \text{ for all } n \in F\}|$$
$$= \frac{1}{N_j}|\{m \in [1, N_j] : m \in \bigcap_{n\in F}(E - n)\}| = \frac{1}{N_j}\Big|\bigcap_{n\in F} S^{-n}[E] \cap [1, N_j]\Big|.$$

Sárköky–Furstenberg theorem

This completes the Furstenberg correspondence principle, since

$$\overline{d}\Big(\bigcap_{n\in F} S^{-n}[E]\Big) \geq \lim_{j\to\infty} \frac{1}{N_j}\Big|\bigcap_{n\in F} S^{-n}[E]\cap[1,N_j]\Big|$$
$$= \lim_{j\to\infty} \mu_j\Big(\bigcap_{n\in F} T^{-n}[A]\Big) = \mu\Big(\bigcap_{n\in F} T^{-n}[A]\Big). \qquad \square$$

Theorem (Sárköky-Furstenberg theorem)

Let $E\subseteq\mathbb{N}$ be a set with positive upper Banach density and let $P\in\mathbb{Z}[n]$ with $P(0)=0$. Then there exist $x,y\in E$ and $n\in\mathbb{N}$ with $x-y=P(n)$.

Proof. Fix $P\in\mathbb{Z}[n]$ with $P(0)=0$. By the Furstenberg correspondence principle, there exists a measure-preserving system $(X,\mathcal{M},\mu,T)$ with $\mu(X)=1$, and an invertible $T:X\to X$ and $A\in\mathcal{M}$ such that

$$\mu(A)=\overline{d}(E)>0.$$

Proof

Moreover, we have

$$\overline{d}(S^{-P(n)}[E] \cap E) \geq \mu(A \cap T^{-P(n)}[A]).$$

Now using the polynomial recurrence theorem with $\varepsilon = \frac{1}{2}\mu(A)^2 > 0$, i.e.

Theorem (Polynomial Poincaré recurrence theorem)

For any measure-preserving system $(X, \mathcal{M}, \mu, T)$ with $\mu(X) = 1$, any $A \in \mathcal{M}$, $\varepsilon > 0$, and $P \in \mathbb{Z}[n]$ with $P(0)$, there exists $n \in \mathbb{N}$ such that

$$\mu(A \cap T^{-P(n)}[A]) > \mu(A)^2 - \varepsilon,$$

and we assume that $T : X \to X$ is invertible if P takes negative values.

we obtain that

$$\overline{d}(S^{-P(n)}[E] \cap E) \geq \mu(A \cap T^{-P(n)}[A]) > \frac{1}{2}\mu(A)^2 > 0$$

for some $n \in \mathbb{N}$, which in turn proves that there are $x, y \in E$ such that

$$x - y = P(n). \qquad \square$$