MAGIC: Ergodic Theory Lecture 5 - Recurrence & Ergodicity

Charles Walkden

February 20th 2013

In this lecture we discuss recurrence properties of measure preserving transformations.
We will state Birkhoff's Ergodic Theorem and study some simple applications.

Recurrence

Recurrence

Birkhoff's Ergodic Theorem implies that given an ergodic measure-preserving transformation T of a probability space $(X, \mathcal{B}, \mu)$ and a set $A \in \mathcal{B}$ with $\mu(A) > 0$, the frequency with which the orbit of almost every point lands in A is equal to $\mu(A)$. We begin with a simpler theorem:

Poincaré's Recurrence Theorem

Recurrence

Birkhoff's Ergodic Theorem implies that given an ergodic measure-preserving transformation T of a probability space $(X, \mathcal{B}, \mu)$ and a set $A \in \mathcal{B}$ with $\mu(A) > 0$, the frequency with which the orbit of almost every point lands in A is equal to $\mu(A)$. We begin with a simpler theorem:

Poincaré's Recurrence Theorem

Let $T : X \to X$ be a m.p.t. of a probability space $(X, \mathcal{B}, \mu)$. Let $A \in \mathcal{B}$, $\mu(A) > 0$. Then the orbit of μ-a.e. point of A returns to A infinitely often.

Recurrence

Birkhoff's Ergodic Theorem implies that given an ergodic measure-preserving transformation T of a probability space $(X, \mathcal{B}, \mu)$ and a set $A \in \mathcal{B}$ with $\mu(A) > 0$, the frequency with which the orbit of almost every point lands in A is equal to $\mu(A)$. We begin with a simpler theorem:

Poincaré's Recurrence Theorem

Let $T : X \to X$ be a m.p.t. of a probability space $(X, \mathcal{B}, \mu)$. Let $A \in \mathcal{B}$, $\mu(A) > 0$. Then the orbit of μ-a.e. point of A returns to A infinitely often.

Proof (sketch):

Let

$$A_n = \bigcup_{k \geq n} T^{-k} A = \{x \in X \mid \text{the orbit of } x \text{ lies in } A \text{ at some time} \geq n\} .$$

Then

$$\{x \in A \mid \text{orbit of } x \text{ returns to } A \text{ i.o.}\} = \bigcap_{n=0}^{\infty} A_n \cap A.$$

Then

$$\{x \in A \mid \text{orbit of } x \text{ returns to } A \text{ i.o.}\} = \bigcap_{n=0}^{\infty} A_n \cap A.$$

Now

$$A \subset A_0$$

Then

$$\{x \in A \mid \text{orbit of } x \text{ returns to } A \text{ i.o.}\} = \bigcap_{n=0}^{\infty} A_n \cap A.$$

Now

$$A \subset A_0 \supset A_1 \supset \cdots \supset A_n \supset \cdots . \tag{1}$$

Then

$$\{x \in A \mid \text{orbit of } x \text{ returns to } A \text{ i.o.}\} = \bigcap_{n=0}^{\infty} A_n \cap A.$$

Now

$$A \subset A_0 \supset A_1 \supset \cdots \supset A_n \supset \cdots . \tag{1}$$

Moreover, $T^{-n}A_n = A_0$. As T is a m.p.t. we have $\mu(A_n) = \mu(A_0)$.

Then

$$\{x \in A \mid \text{orbit of } x \text{ returns to } A \text{ i.o.}\} = \bigcap_{n=0}^{\infty} A_n \cap A.$$

Now

$$A \subset A_0 \supset A_1 \supset \cdots \supset A_n \supset \cdots . \tag{1}$$

Moreover, $T^{-n}A_n = A_0$. As T is a m.p.t. we have $\mu(A_n) = \mu(A_0)$. Hence (1) implies:

$$A_0 = A_1 = \cdots = A_n = \cdots \text{ a.e.}$$

Then

$$\{x \in A \mid \text{orbit of } x \text{ returns to } A \text{ i.o.}\} = \bigcap_{n=0}^{\infty} A_n \cap A.$$

Now

$$A \subset A_0 \supset A_1 \supset \cdots \supset A_n \supset \cdots . \tag{1}$$

Moreover, $T^{-n}A_n = A_0$. As T is a m.p.t. we have $\mu(A_n) = \mu(A_0)$. Hence (1) implies:

$$A_0 = A_1 = \cdots = A_n = \cdots \text{ a.e.}$$

Hence

$$\bigcap_{n=0}^{\infty} A_n = A_0 \text{ a.e.}$$

Then

$$\{x \in A \mid \text{orbit of } x \text{ returns to } A \text{ i.o.}\} = \bigcap_{n=0}^{\infty} A_n \cap A.$$

Now

$$A \subset A_0 \supset A_1 \supset \cdots \supset A_n \supset \cdots . \tag{1}$$

Moreover, $T^{-n}A_n = A_0$. As T is a m.p.t. we have $\mu(A_n) = \mu(A_0)$. Hence (1) implies:

$$A_0 = A_1 = \cdots = A_n = \cdots \text{ a.e.}$$

Hence

$$\bigcap_{n=0}^{\infty} A_n = A_0 \text{ a.e.}$$

Hence

$$\bigcap_{n=0}^{\infty} A_n \cap A = A_0 \cap A = A \text{ a.e.}$$

Remarks

Remarks

1. Note that we hardly used anything in this proof!
   'Measure-preserving' was used just once.

Remarks

1. Note that we hardly used anything in this proof! 'Measure-preserving' was used just once.
2. The assumption that X is a probability space is essential (why?).

Remarks

1. Note that we hardly used anything in this proof! 'Measure-preserving' was used just once.
2. The assumption that X is a probability space is essential (why?).
3. It is *not* true that the orbit of a.e. point in B lands in A i.o. if $B \neq A$.

Remarks

1. Note that we hardly used anything in this proof! 'Measure-preserving' was used just once.
2. The assumption that X is a probability space is essential (why?).
3. It is *not* true that the orbit of a.e. point in B lands in A i.o. if $B \neq A$.

Before we discuss ergodic theorems we need the notion of conditional expectation.

Conditional Expectation

Conditional Expectation

Let $(X, \mathcal{B}, \mu)$ be a probability space. Let $\mathcal{A} \subset \mathcal{B}$ be a sub-σ-algebra.

Conditional Expectation

Let $(X, \mathcal{B}, \mu)$ be a probability space. Let $\mathcal{A} \subset \mathcal{B}$ be a sub-σ-algebra.
One can show (see notes) that there is a unique map

$$\mathbb{E}(\cdot \mid \mathcal{A}) : L^1(X, \mathcal{B}, \mu) \longrightarrow L^1(X, \mathcal{A}, \mu)$$

such that

$$\int_A \mathbb{E}(f|\mathcal{A})\, d\mu = \int_A f\, d\mu \;\; \forall f \in L^1(X, \mathcal{B}, \mu),\ \forall A \in \mathcal{A}.$$

Conditional Expectation

Let $(X, \mathcal{B}, \mu)$ be a probability space. Let $\mathcal{A} \subset \mathcal{B}$ be a sub-σ-algebra.
One can show (see notes) that there is a unique map

$$\mathbb{E}(\cdot \mid \mathcal{A}) : L^1(X, \mathcal{B}, \mu) \longrightarrow L^1(X, \mathcal{A}, \mu)$$

such that

$$\int_A \mathbb{E}(f|\mathcal{A})\, d\mu = \int_A f\, d\mu \ \ \forall f \in L^1(X, \mathcal{B}, \mu),\ \forall A \in \mathcal{A}.$$

Idea: $\mathbb{E}(f|\mathcal{A})$ - the conditional expectation of f given $\mathcal{A}$ - is the unique $\mathcal{A}$-measurable function that best approximates f.

Example

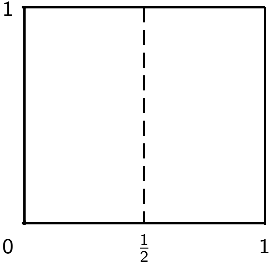
Let $X = [0, 1]$, $\mathcal{B} =$ Borel σ-algebra.

Example

Let $X = [0,1]$, $\mathcal{B} =$ Borel σ-algebra.
Let $\mathcal{A} = \left\{\emptyset, X, \left[0, \frac{1}{2}\right], \left[\frac{1}{2}, 1\right]\right\}$.

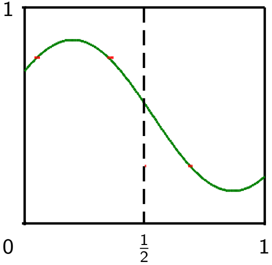
Example

Let $X = [0,1]$, $\mathcal{B} =$ Borel σ-algebra.
Let $\mathcal{A} = \left\{\emptyset, X, \left[0, \frac{1}{2}\right], \left[\frac{1}{2}, 1\right]\right\}$.



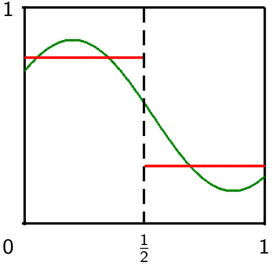
Example

Let $X = [0,1]$, $\mathcal{B} =$ Borel σ-algebra.
Let $\mathcal{A} = \left\{\emptyset, X, \left[0, \frac{1}{2}\right], \left[\frac{1}{2}, 1\right]\right\}$.



Example

Let $X = [0, 1]$, $\mathcal{B} =$ Borel σ-algebra.
Let $\mathcal{A} = \left\{\emptyset, X, \left[0, \frac{1}{2}\right], \left[\frac{1}{2}, 1\right]\right\}$.



Ergodic Theorems

Ergodic Theorems

An ergodic theorem concerns the limiting behaviour of sequences of the form

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

Ergodic Theorems

An ergodic theorem concerns the limiting behaviour of sequences of the form

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

We get different ergodic theorems depending on

1. the class of function f (L^1, L^2, continuous,...)
2. the notion of convergence (a.e., L^2-convergence, uniform, ...)

Ergodic Theorems

An ergodic theorem concerns the limiting behaviour of sequences of the form

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

We get different ergodic theorems depending on

1. the class of function f (L^1, L^2, continuous,...)
2. the notion of convergence (a.e., L^2-convergence, uniform, ...)

Here we state von Neumann's (mean) Ergodic Theorem and Birkhoff's (pointwise) Ergodic Theorem.

Ergodic Theorems

An ergodic theorem concerns the limiting behaviour of sequences of the form

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

We get different ergodic theorems depending on

1. the class of function f (L^1, L^2, continuous,...)
2. the notion of convergence (a.e., L^2-convergence, uniform, ...)

Here we state von Neumann's (mean) Ergodic Theorem and Birkhoff's (pointwise) Ergodic Theorem.

Let $T : X \to X$ be a mpt of a probability space $(X, \mathcal{B}, \mu)$.

Ergodic Theorems

An ergodic theorem concerns the limiting behaviour of sequences of the form

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

We get different ergodic theorems depending on

1. the class of function f (L^1, L^2, continuous,...)
2. the notion of convergence (a.e., L^2-convergence, uniform, ...)

Here we state von Neumann's (mean) Ergodic Theorem and Birkhoff's (pointwise) Ergodic Theorem.
Let $T : X \to X$ be a mpt of a probability space $(X, \mathcal{B}, \mu)$.
Let $\mathcal{I} = \left\{ B \in \mathcal{B} \mid T^{-1}B = B \text{ a.e. } \right\}$ be the sub σ-algebra of invariant sets.

Ergodic Theorems

An ergodic theorem concerns the limiting behaviour of sequences of the form

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

We get different ergodic theorems depending on

1. the class of function f (L^1, L^2, continuous,...)
2. the notion of convergence (a.e., L^2-convergence, uniform, ...)

Here we state von Neumann's (mean) Ergodic Theorem and Birkhoff's (pointwise) Ergodic Theorem.
Let $T : X \to X$ be a mpt of a probability space $(X, \mathcal{B}, \mu)$.
Let $\mathcal{I} = \left\{ B \in \mathcal{B} \mid T^{-1}B = B \text{ a.e.} \right\}$ be the sub σ-algebra of invariant sets.
Note: if T is ergodic then $\mathcal{I} = \{\emptyset, X\}$ a.e. and $E(f \mid \mathcal{I}) = \int f \, d\mu$.

Von Neumann's Ergodic Theorem

Let $(X, \mathcal{B}, \mu)$ be a probability space and let $T : X \to X$ be an mpt. Let $\mathcal{I}$ denote the σ-algebra of invariant sets. Let $f \in L^2(X, \mathcal{B}, \mu)$.

Von Neumann's Ergodic Theorem

Let $(X, \mathcal{B}, \mu)$ be a probability space and let $T : X \to X$ be an mpt. Let $\mathcal{I}$ denote the σ-algebra of invariant sets. Let $f \in L^2(X, \mathcal{B}, \mu)$. Then

$$\frac{1}{n}\sum_{j=0}^{n-1} f \circ T^j \longrightarrow \mathbb{E}(f \mid \mathcal{I}) \text{ in } L^2.$$

Von Neumann's Ergodic Theorem

Let $(X, \mathcal{B}, \mu)$ be a probability space and let $T : X \to X$ be an mpt. Let $\mathcal{I}$ denote the σ-algebra of invariant sets. Let $f \in L^2(X, \mathcal{B}, \mu)$. Then

$$\frac{1}{n}\sum_{j=0}^{n-1} f \circ T^j \longrightarrow \mathbb{E}(f \mid \mathcal{I}) \text{ in } L^2.$$

Corollary

If, in addition, T is ergodic, then

$$\frac{1}{n}\sum_{j=0}^{n-1} f \circ T^j \longrightarrow \int f \, d\mu \textit{ in } L^2.$$

Sketch of a proof for von Neumann's Ergodic Theorem

The mpt T of the probability space $(X, \mathcal{B}, \mu)$ gives an operator on L^2:

$$U : L^2 \to L^2 : f \mapsto f \circ T.$$

Sketch of a proof for von Neumann's Ergodic Theorem

The mpt T of the probability space $(X, \mathcal{B}, \mu)$ gives an operator on L^2:

$$U : L^2 \to L^2 : f \mapsto f \circ T.$$

This is an isometry as

$$\langle Uf, Ug \rangle = \int f \circ T \overline{g \circ T} \, d\mu = \int (f\overline{g}) \circ T \, d\mu = \int f\overline{g} \, d\mu = \langle f, g \rangle.$$

Sketch of a proof for von Neumann's Ergodic Theorem

The mpt T of the probability space $(X, \mathcal{B}, \mu)$ gives an operator on L^2:

$$U : L^2 \to L^2 : f \mapsto f \circ T.$$

This is an isometry as

$$\langle Uf, Ug\rangle = \int f \circ T \overline{g \circ T}\, d\mu = \int (f\overline{g}) \circ T\, d\mu = \int f\overline{g}\, d\mu = \langle f, g\rangle.$$

Von Neumann's Ergodic Theorem can be phrased in terms of a linear operator U on a Hilbert space H.

Sketch of a proof for von Neumann's Ergodic Theorem

The mpt T of the probability space $(X, \mathcal{B}, \mu)$ gives an operator on L^2:

$$U : L^2 \to L^2 : f \mapsto f \circ T.$$

This is an isometry as

$$\langle Uf, Ug \rangle = \int f \circ T \overline{g \circ T}\, d\mu = \int (f\overline{g}) \circ T\, d\mu = \int f\overline{g}\, d\mu = \langle f, g \rangle.$$

Von Neumann's Ergodic Theorem can be phrased in terms of a linear operator U on a Hilbert space H.
Let $I = \{v \in H \mid Uv = v\}$ and let P_I be orthogonal projection onto I.

Sketch of a proof for von Neumann's Ergodic Theorem

The mpt T of the probability space $(X, \mathcal{B}, \mu)$ gives an operator on L^2:

$$U : L^2 \to L^2 : f \mapsto f \circ T.$$

This is an isometry as

$$\langle Uf, Ug \rangle = \int f \circ T \overline{g \circ T} \, d\mu = \int (f \overline{g}) \circ T \, d\mu = \int f \overline{g} \, d\mu = \langle f, g \rangle.$$

Von Neumann's Ergodic Theorem can be phrased in terms of a linear operator U on a Hilbert space H.
Let $I = \{v \in H \mid Uv = v\}$ and let P_I be orthogonal projection onto I. We claim

$$\frac{1}{n} \sum_{j=0}^{n-1} U^j v \to P_I v. \tag{2}$$

Sketch of a proof for von Neumann's Ergodic Theorem

The mpt T of the probability space $(X, \mathcal{B}, \mu)$ gives an operator on L^2:

$$U : L^2 \to L^2 : f \mapsto f \circ T.$$

This is an isometry as

$$\langle Uf, Ug \rangle = \int f \circ T \overline{g \circ T} \, d\mu = \int (f\overline{g}) \circ T \, d\mu = \int f\overline{g} \, d\mu = \langle f, g \rangle.$$

Von Neumann's Ergodic Theorem can be phrased in terms of a linear operator U on a Hilbert space H.
Let $I = \{v \in H \mid Uv = v\}$ and let P_I be orthogonal projection onto I. We claim

$$\frac{1}{n} \sum_{j=0}^{n-1} U^j v \to P_I v. \tag{2}$$

Write $H = I \oplus \overline{\{Uw - w \mid w \in H\}}$.

Sketch of a proof for von Neumann's Ergodic Theorem

The mpt T of the probability space $(X, \mathcal{B}, \mu)$ gives an operator on L^2:

$$U : L^2 \to L^2 : f \mapsto f \circ T.$$

This is an isometry as

$$\langle Uf, Ug \rangle = \int f \circ T \overline{g \circ T} \, d\mu = \int (f\overline{g}) \circ T \, d\mu = \int f \overline{g} \, d\mu = \langle f, g \rangle.$$

Von Neumann's Ergodic Theorem can be phrased in terms of a linear operator U on a Hilbert space H.

Let $I = \{v \in H \mid Uv = v\}$ and let P_I be orthogonal projection onto I. We claim

$$\frac{1}{n} \sum_{j=0}^{n-1} U^j v \to P_I v. \tag{2}$$

Write $H = I \oplus \overline{\{Uw - w \mid w \in H\}}$. Use this decomposition to prove (2).

Birkhoff's Ergodic Theorem

Birkhoff's Ergodic Theorem

Let $(X, \mathcal{B}, \mu)$ be a probability space and let $T : X \to X$ be an mpt. Let $\mathcal{I}$ denote the σ-algebra of invariant sets. Let $f \in L^1(X, \mathcal{B}, \mu)$.

Birkhoff's Ergodic Theorem

Let $(X, \mathcal{B}, \mu)$ be a probability space and let $T : X \to X$ be an mpt. Let $\mathcal{I}$ denote the σ-algebra of invariant sets. Let $f \in L^1(X, \mathcal{B}, \mu)$. Then

$$\frac{1}{n}\sum_{j=0}^{n-1} f(T^j x) \longrightarrow \mathbb{E}(f \mid \mathcal{I})(x) \ \mu\text{-a.e.}$$

Birkhoff's Ergodic Theorem

Let $(X, \mathcal{B}, \mu)$ be a probability space and let $T : X \to X$ be an mpt. Let $\mathcal{I}$ denote the σ-algebra of invariant sets. Let $f \in L^1(X, \mathcal{B}, \mu)$. Then

$$\frac{1}{n}\sum_{j=0}^{n-1} f(T^j x) \longrightarrow \mathbb{E}(f \mid \mathcal{I})(x) \ \mu\text{-a.e.}$$

Corollary

If, in addition, T is ergodic, then

$$\frac{1}{n}\sum_{j=0}^{n-1} f(T^j x) \longrightarrow \int f \, d\mu \ \mu\text{-a.e.}$$

Birkhoff's Ergodic Theorem

Let $(X, \mathcal{B}, \mu)$ be a probability space and let $T : X \to X$ be an mpt. Let $\mathcal{I}$ denote the σ-algebra of invariant sets. Let $f \in L^1(X, \mathcal{B}, \mu)$. Then

$$\frac{1}{n}\sum_{j=0}^{n-1} f(T^j x) \longrightarrow \mathbb{E}(f \mid \mathcal{I})(x) \ \mu\text{-a.e.}$$

Corollary

If, in addition, T is ergodic, then

$$\frac{1}{n}\sum_{j=0}^{n-1} f(T^j x) \longrightarrow \int f \, d\mu \ \mu\text{-a.e.}$$

The proof is hard analysis!

Kac's Lemma

Kac's Lemma

Recall: Poincaré's Recurrence Theorem: given an mpt of a probability space $(X, \mathcal{B}, \mu)$ and a set A of positive measure, a.e. point of A returns to A infinitely often.

Kac's Lemma

Recall: Poincaré's Recurrence Theorem: given an mpt of a probability space $(X, \mathcal{B}, \mu)$ and a set A of positive measure, a.e. point of A returns to A infinitely often.

What is the expected first return time?

Let $T : X \to X$ be an mpt of a probability space $(X, \mathcal{B}, \mu)$. Let $A \in \mathcal{B}$, $\mu(A) > 0$.

Let $T : X \to X$ be an mpt of a probability space $(X, \mathcal{B}, \mu)$. Let $A \in \mathcal{B}$, $\mu(A) > 0$. Define

$$n_A(x) = \text{first return time to } A$$

Let $T : X \to X$ be an mpt of a probability space $(X, \mathcal{B}, \mu)$. Let $A \in \mathcal{B}$, $\mu(A) > 0$. Define

$$n_A(x) = \text{first return time to } A = \inf\{n \geq 1 \mid T^n x \in A\}.$$

Let $T : X \to X$ be an mpt of a probability space $(X, \mathcal{B}, \mu)$. Let $A \in \mathcal{B}$, $\mu(A) > 0$. Define

$$n_A(x) = \text{first return time to } A = \inf\{n \geq 1 \mid T^n x \in A\}.$$

Then Poincaré's Recurrence Theorem says $n_A(x) < \infty$, μ-a.e. $x \in A$.

Let $T : X \to X$ be an mpt of a probability space $(X, \mathcal{B}, \mu)$. Let $A \in \mathcal{B}$, $\mu(A) > 0$. Define

$$n_A(x) = \text{first return time to } A = \inf\{n \geq 1 \mid T^n x \in A\}.$$

Then Poincaré's Recurrence Theorem says $n_A(x) < \infty$, μ-a.e. $x \in A$.

Kac's Lemma

Suppose T is an ergodic mpt of a probability space $(X, \mathcal{B}, \mu)$, $A \in \mathcal{B}$, $\mu(A) > 0$.

Let $T : X \to X$ be an mpt of a probability space $(X, \mathcal{B}, \mu)$. Let $A \in \mathcal{B}$, $\mu(A) > 0$. Define

$$n_A(x) = \text{first return time to } A = \inf\{n \geq 1 \mid T^n x \in A\}.$$

Then Poincaré's Recurrence Theorem says $n_A(x) < \infty$, μ-a.e. $x \in A$.

Kac's Lemma

Suppose T is an ergodic mpt of a probability space $(X, \mathcal{B}, \mu)$, $A \in \mathcal{B}$, $\mu(A) > 0$. Then

$$\int_A n_A \, d\mu = 1.$$

Let $T : X \to X$ be an mpt of a probability space $(X, \mathcal{B}, \mu)$. Let $A \in \mathcal{B}$, $\mu(A) > 0$. Define

$$n_A(x) = \text{first return time to } A = \inf\{n \geq 1 \mid T^n x \in A\}.$$

Then Poincaré's Recurrence Theorem says $n_A(x) < \infty$, μ-a.e. $x \in A$.

Kac's Lemma

Suppose T is an ergodic mpt of a probability space $(X, \mathcal{B}, \mu)$, $A \in \mathcal{B}$, $\mu(A) > 0$. Then

$$\int_A n_A \, d\mu = 1.$$

Remark

The expected return time of a point $x \in A$ to A is

$$\frac{1}{\mu(A)} \int_A n_A \, d\mu = \frac{1}{\mu(A)}.$$

Proof (sketch):

Proof (sketch):

By Birkhoff's Ergodic Theorem, almost every point of X lands in A.

Proof (sketch):

By Birkhoff's Ergodic Theorem, almost every point of X lands in A. Let

$$A_n = \{x \in A \mid \text{the first return time of } x \text{ to } A \text{ is } n\}$$

Proof (sketch):

By Birkhoff's Ergodic Theorem, almost every point of X lands in A. Let

$$\begin{aligned} A_n &= \{x \in A \mid \text{the first return time of } x \text{ to } A \text{ is } n\} \\ &= A \cap T^{-1}A^c \cap \cdots \cap T^{-(n-1)}A^c \cap T^{-n}A. \end{aligned}$$

Proof (sketch):

By Birkhoff's Ergodic Theorem, almost every point of X lands in A. Let

$$\begin{aligned} A_n &= \{x \in A \mid \text{the first return time of } x \text{ to } A \text{ is } n\} \\ &= A \cap T^{-1}A^c \cap \cdots \cap T^{-(n-1)}A^c \cap T^{-n}A. \end{aligned}$$

Then $A = \bigcup_n A_n$ a.e.

Proof (sketch):

By Birkhoff's Ergodic Theorem, almost every point of X lands in A. Let

$$\begin{aligned} A_n &= \{x \in A \mid \text{the first return time of } x \text{ to } A \text{ is } n\} \\ &= A \cap T^{-1}A^c \cap \cdots \cap T^{-(n-1)}A^c \cap T^{-n}A. \end{aligned}$$

Then $A = \bigcup_n A_n$ a.e. Consider the following diagram

Proof (sketch):

By Birkhoff's Ergodic Theorem, almost every point of X lands in A. Let

$$\begin{aligned} A_n &= \{x \in A \mid \text{the first return time of } x \text{ to } A \text{ is } n\} \\ &= A \cap T^{-1}A^c \cap \cdots \cap T^{-(n-1)}A^c \cap T^{-n}A. \end{aligned}$$

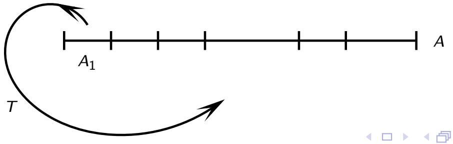
Then $A = \bigcup_n A_n$ a.e. Consider the following diagram

Proof (sketch):

By Birkhoff's Ergodic Theorem, almost every point of X lands in A. Let

$$\begin{aligned} A_n &= \{x \in A \mid \text{the first return time of } x \text{ to } A \text{ is } n\} \\ &= A \cap T^{-1}A^c \cap \cdots \cap T^{-(n-1)}A^c \cap T^{-n}A. \end{aligned}$$

Then $A = \bigcup_n A_n$ a.e. Consider the following diagram

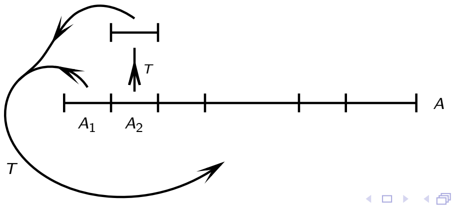


Proof (sketch):

By Birkhoff's Ergodic Theorem, almost every point of X lands in A. Let

$$\begin{aligned} A_n &= \{x \in A \mid \text{the first return time of } x \text{ to } A \text{ is } n\} \\ &= A \cap T^{-1}A^c \cap \cdots \cap T^{-(n-1)}A^c \cap T^{-n}A. \end{aligned}$$

Then $A = \bigcup_n A_n$ a.e. Consider the following diagram

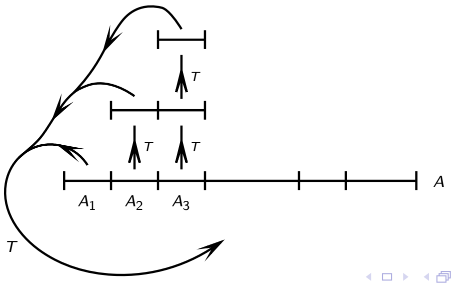


Proof (sketch):

By Birkhoff's Ergodic Theorem, almost every point of X lands in A. Let

$$\begin{aligned} A_n &= \{x \in A \mid \text{the first return time of } x \text{ to } A \text{ is } n\} \\ &= A \cap T^{-1}A^c \cap \cdots \cap T^{-(n-1)}A^c \cap T^{-n}A. \end{aligned}$$

Then $A = \bigcup_n A_n$ a.e. Consider the following diagram

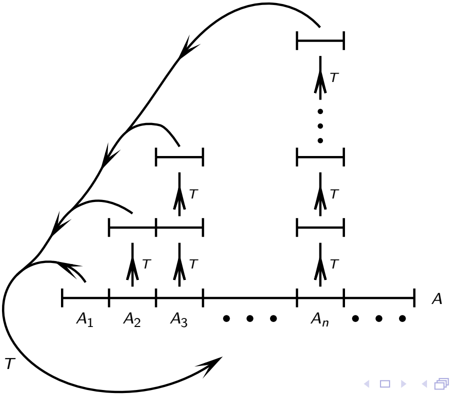


Proof (sketch):

By Birkhoff's Ergodic Theorem, almost every point of X lands in A. Let

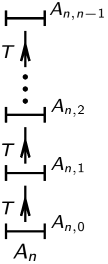
$$\begin{aligned} A_n &= \{x \in A \mid \text{the first return time of } x \text{ to } A \text{ is } n\} \\ &= A \cap T^{-1}A^c \cap \cdots \cap T^{-(n-1)}A^c \cap T^{-n}A. \end{aligned}$$

Then $A = \bigcup_n A_n$ a.e. Consider the following diagram

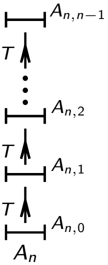


As a.e. point of X eventually enters A, this diagram is a partition of almost all of X.

As a.e. point of X eventually enters A, this diagram is a partition of almost all of X. Consider the n^{th} tower:



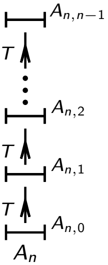
As a.e. point of X eventually enters A, this diagram is a partition of almost all of X. Consider the n^{th} tower:



Then $A_{n,0} = A_n$

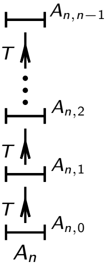
$$T^{-k}A_{n,k} = A_{n,0}$$

As a.e. point of X eventually enters A, this diagram is a partition of almost all of X. Consider the n^{th} tower:



$$\begin{aligned} \text{Then } A_{n,0} &= A_n \\ T^{-k}A_{n,k} &= A_{n,0} \\ \text{Hence } \mu(A_{n,k}) &= \mu(A_{n,0}) \\ &= \mu(A_n) \end{aligned}$$

As a.e. point of X eventually enters A, this diagram is a partition of almost all of X. Consider the n^{th} tower:



$$\begin{aligned} \text{Then } A_{n,0} &= A_n \\ T^{-k}A_{n,k} &= A_{n,0} \\ \text{Hence } \mu(A_{n,k}) &= \mu(A_{n,0}) \\ &= \mu(A_n) \end{aligned}$$

Note that $n_A(x) = n$ iff $x \in A_n$. (So the preceding diagram is essentially the graph of n_A.)

As a.e. point of X eventually enters A, this diagram is a partition of almost all of X. Consider the n^{th} tower:

$$\begin{tikzcd}
A_{n,n-1} \\
\vdots \arrow[u, "T"] \\
A_{n,2} \\
A_{n,1} \arrow[u, "T"] \\
A_{n,0} \arrow[u, "T"]
\end{tikzcd}$$

A_n

Then $A_{n,0} = A_n$

$$T^{-k}A_{n,k} = A_{n,0}$$

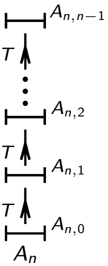
Hence $\mu(A_{n,k}) = \mu(A_{n,0})$

$$= \mu(A_n)$$

Note that $n_A(x) = n$ iff $x \in A_n$. (So the preceding diagram is essentially the graph of n_A.) Hence

$$1 = \mu(X) \;=\; \sum_{n=1}^{\infty}\sum_{k=0}^{n-1}\mu(A_{n,k})$$

As a.e. point of X eventually enters A, this diagram is a partition of almost all of X. Consider the n^{th} tower:



Then $A_{n,0} = A_n$

$T^{-k}A_{n,k} = A_{n,0}$

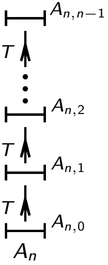
Hence $\mu(A_{n,k}) = \mu(A_{n,0})$

$= \mu(A_n)$

Note that $n_A(x) = n$ iff $x \in A_n$. (So the preceding diagram is essentially the graph of n_A.) Hence

$$1 = \mu(X) \;=\; \sum_{n=1}^{\infty}\sum_{k=0}^{n-1}\mu(A_{n,k}) = \sum_{n=1}^{\infty}\sum_{k=0}^{n-1}\mu(A_n)$$

As a.e. point of X eventually enters A, this diagram is a partition of almost all of X. Consider the n^{th} tower:

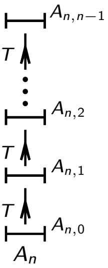


$$\begin{aligned} \text{Then } A_{n,0} &= A_n \\ T^{-k}A_{n,k} &= A_{n,0} \\ \text{Hence } \mu(A_{n,k}) &= \mu(A_{n,0}) \\ &= \mu(A_n) \end{aligned}$$

Note that $n_A(x) = n$ iff $x \in A_n$. (So the preceding diagram is essentially the graph of n_A.) Hence

$$\begin{aligned} 1 = \mu(X) &= \sum_{n=1}^{\infty}\sum_{k=0}^{n-1}\mu(A_{n,k}) = \sum_{n=1}^{\infty}\sum_{k=0}^{n-1}\mu(A_n) \\ &= \sum_{n=1}^{\infty} n\mu(A_n) \end{aligned}$$

As a.e. point of X eventually enters A, this diagram is a partition of almost all of X. Consider the n^{th} tower:



$$
\begin{aligned}
\text{Then } A_{n,0} &= A_n \\
T^{-k}A_{n,k} &= A_{n,0} \\
\text{Hence } \mu(A_{n,k}) &= \mu(A_{n,0}) \\
&= \mu(A_n)
\end{aligned}
$$

Note that $n_A(x) = n$ iff $x \in A_n$. (So the preceding diagram is essentially the graph of n_A.) Hence

$$
\begin{aligned}
1 = \mu(X) &= \sum_{n=1}^{\infty} \sum_{k=0}^{n-1} \mu(A_{n,k}) = \sum_{n=1}^{\infty} \sum_{k=0}^{n-1} \mu(A_n) \\
&= \sum_{n=1}^{\infty} n\mu(A_n) = \int_A n_A(x)\, d\mu. \qquad \square
\end{aligned}
$$

Ehrenfest's Example

We give an example to show that return times can be very large.

Ehrenfest's Example

We give an example to show that return times can be very large. Suppose we have two urns:

Ehrenfest's Example

We give an example to show that return times can be very large. Suppose we have two urns:

- one urn contains 100 balls, numbered 1 to 100

Ehrenfest's Example

We give an example to show that return times can be very large. Suppose we have two urns:

- one urn contains 100 balls, numbered 1 to 100
- the other urn is empty.

Ehrenfest's Example

We give an example to show that return times can be very large. Suppose we have two urns:

- one urn contains 100 balls, numbered 1 to 100
- the other urn is empty.

Suppose we have a random number generator, picking each number between 1 and 100 with probability $\frac{1}{100}$.

Ehrenfest's Example

We give an example to show that return times can be very large. Suppose we have two urns:

- one urn contains 100 balls, numbered 1 to 100
- the other urn is empty.

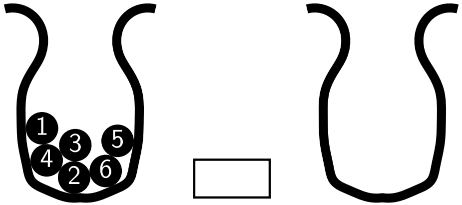
Suppose we have a random number generator, picking each number between 1 and 100 with probability $\frac{1}{100}$.
Each second a random number is chosen. The ball corresponding to this number is then moved from one urn to the other.

Ehrenfest's Example

We give an example to show that return times can be very large. Suppose we have two urns:

- one urn contains 100 balls, numbered 1 to 100
- the other urn is empty.

Suppose we have a random number generator, picking each number between 1 and 100 with probability $\frac{1}{100}$.
Each second a random number is chosen. The ball corresponding to this number is then moved from one urn to the other.

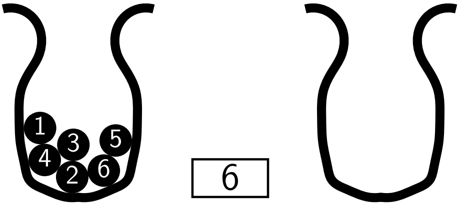


Ehrenfest's Example

We give an example to show that return times can be very large. Suppose we have two urns:

- one urn contains 100 balls, numbered 1 to 100
- the other urn is empty.

Suppose we have a random number generator, picking each number between 1 and 100 with probability $\frac{1}{100}$.
Each second a random number is chosen. The ball corresponding to this number is then moved from one urn to the other.

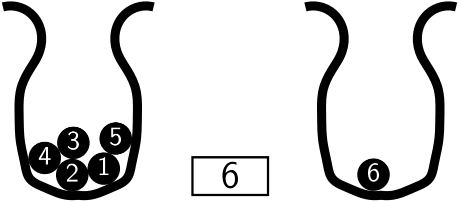


Ehrenfest's Example

We give an example to show that return times can be very large.
Suppose we have two urns:

- one urn contains 100 balls, numbered 1 to 100
- the other urn is empty.

Suppose we have a random number generator, picking each number between 1 and 100 with probability $\frac{1}{100}$.
Each second a random number is chosen. The ball corresponding to this number is then moved from one urn to the other.

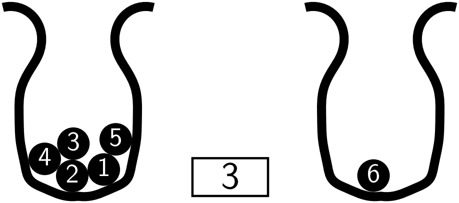


Ehrenfest's Example

We give an example to show that return times can be very large.
Suppose we have two urns:

- one urn contains 100 balls, numbered 1 to 100
- the other urn is empty.

Suppose we have a random number generator, picking each number between 1 and 100 with probability $\frac{1}{100}$.
Each second a random number is chosen. The ball corresponding to this number is then moved from one urn to the other.

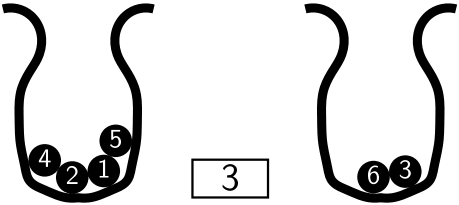


Ehrenfest's Example

We give an example to show that return times can be very large.
Suppose we have two urns:

- one urn contains 100 balls, numbered 1 to 100
- the other urn is empty.

Suppose we have a random number generator, picking each number between 1 and 100 with probability $\frac{1}{100}$.
Each second a random number is chosen. The ball corresponding to this number is then moved from one urn to the other.

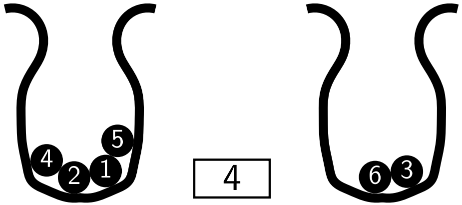


Ehrenfest's Example

We give an example to show that return times can be very large.
Suppose we have two urns:

- one urn contains 100 balls, numbered 1 to 100
- the other urn is empty.

Suppose we have a random number generator, picking each number between 1 and 100 with probability $\frac{1}{100}$.
Each second a random number is chosen. The ball corresponding to this number is then moved from one urn to the other.

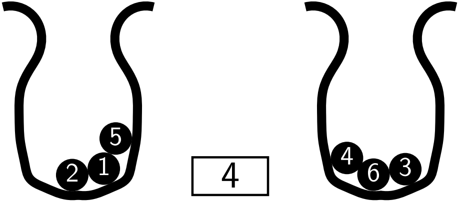


Ehrenfest's Example

We give an example to show that return times can be very large. Suppose we have two urns:

- one urn contains 100 balls, numbered 1 to 100
- the other urn is empty.

Suppose we have a random number generator, picking each number between 1 and 100 with probability $\frac{1}{100}$.
Each second a random number is chosen. The ball corresponding to this number is then moved from one urn to the other.

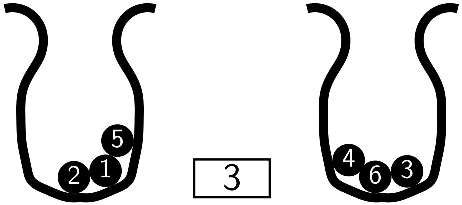


Ehrenfest's Example

We give an example to show that return times can be very large.
Suppose we have two urns:

- one urn contains 100 balls, numbered 1 to 100
- the other urn is empty.

Suppose we have a random number generator, picking each number between 1 and 100 with probability $\frac{1}{100}$.
Each second a random number is chosen. The ball corresponding to this number is then moved from one urn to the other.

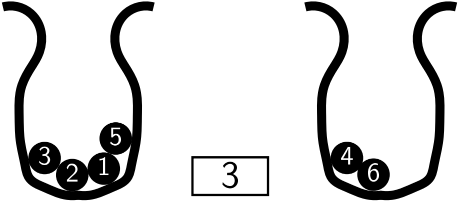


Ehrenfest's Example

We give an example to show that return times can be very large.
Suppose we have two urns:

- one urn contains 100 balls, numbered 1 to 100
- the other urn is empty.

Suppose we have a random number generator, picking each number between 1 and 100 with probability $\frac{1}{100}$.
Each second a random number is chosen. The ball corresponding to this number is then moved from one urn to the other.

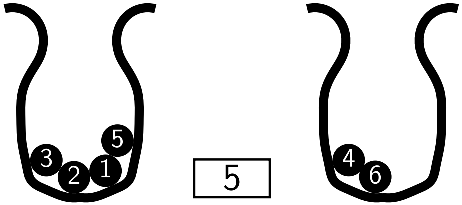


Ehrenfest's Example

We give an example to show that return times can be very large.
Suppose we have two urns:

- one urn contains 100 balls, numbered 1 to 100
- the other urn is empty.

Suppose we have a random number generator, picking each number between 1 and 100 with probability $\frac{1}{100}$.
Each second a random number is chosen. The ball corresponding to this number is then moved from one urn to the other.

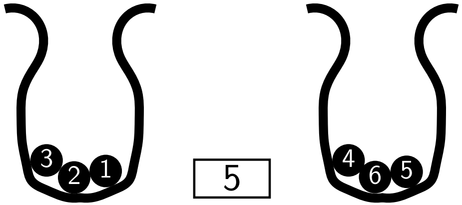


Ehrenfest's Example

We give an example to show that return times can be very large.
Suppose we have two urns:

- one urn contains 100 balls, numbered 1 to 100
- the other urn is empty.

Suppose we have a random number generator, picking each number between 1 and 100 with probability $\frac{1}{100}$.
Each second a random number is chosen. The ball corresponding to this number is then moved from one urn to the other.



We would expect the system to settle down to about 50 balls in each urn, subject to some fluctuations.

We would expect the system to settle down to about 50 balls in each urn, subject to some fluctuations.
It would appear unlikely that all 100 balls should ever return to the first urn.

We would expect the system to settle down to about 50 balls in each urn, subject to some fluctuations.
It would appear unlikely that all 100 balls should ever return to the first urn.
However, Poincaré's Recurrence Theorem tells us that this will, almost surely, happen and Kac's Lemma tells us how long we should expect to wait.

We would expect the system to settle down to about 50 balls in each urn, subject to some fluctuations.
It would appear unlikely that all 100 balls should ever return to the first urn.
However, Poincaré's Recurrence Theorem tells us that this will, almost surely, happen and Kac's Lemma tells us how long we should expect to wait.

We model the system as a shift of finite type on 101 symbols with an appropriate Markov measure.

We would expect the system to settle down to about 50 balls in each urn, subject to some fluctuations.
It would appear unlikely that all 100 balls should ever return to the first urn.
However, Poincaré's Recurrence Theorem tells us that this will, almost surely, happen and Kac's Lemma tells us how long we should expect to wait.

We model the system as a shift of finite type on 101 symbols with an appropriate Markov measure.
Regard $x_j \in \{0, 1, \ldots, 100\}$ as the number of balls in the first urn after j seconds.

As the number of balls in the first urn either increases or decreases by 1, $(x_j)_{j=0}^{\infty} \in \Sigma_A^+$ where

$$A = \begin{pmatrix} 0 & 1 & 0 & \multicolumn{2}{c}{\dots\dots\dots} & 0 \\ 1 & 0 & 1 & \ddots & & \vdots \\ 0 & 1 & 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & 1 & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 0 & 1 \\ 0 & \multicolumn{2}{c}{\dots\dots\dots} & 0 & 1 & 0 \end{pmatrix}.$$

As the number of balls in the first urn either increases or decreases by 1, $(x_j)_{j=0}^{\infty} \in \Sigma_A^+$ where

$$A = \begin{pmatrix} 0 & 1 & 0 & \multicolumn{2}{c}{\dots\dots\dots} & 0 \\ 1 & 0 & 1 & \ddots & & \vdots \\ 0 & 1 & 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & 1 & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 0 & 1 \\ 0 & \multicolumn{2}{c}{\dots\dots\dots} & 0 & 1 & 0 \end{pmatrix}.$$

The probability of there being i balls in the first urn is

$$p_i = \frac{1}{2^{100}} \binom{100}{i}$$

As the number of balls in the first urn either increases or decreases by 1, $(x_j)_{j=0}^{\infty} \in \Sigma_A^+$ where

$$A = \begin{pmatrix} 0 & 1 & 0 & \cdots\cdots & 0 \\ 1 & 0 & 1 & \ddots & & \vdots \\ 0 & 1 & 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & 1 & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 0 & 1 \\ 0 & \cdots\cdots & & 0 & 1 & 0 \end{pmatrix}.$$

The probability of there being i balls in the first urn is

$$p_i = \frac{1}{2^{100}}\binom{100}{i} = \frac{\#\text{ of configurations of } i \text{ balls}}{\text{total }\#\text{ of configurations of balls}}.$$

From state i we can

From state i we can

- move to state $i - 1$ - this has probability $P_{i,i-1} = \frac{i}{100}$

From state i we can

- move to state $i-1$ - this has probability $P_{i,i-1} = \frac{i}{100}$
- move to state $i+1$ - this has probability $P_{i,i+1} = \frac{100-i}{100}$.

From state i we can

- move to state $i-1$ - this has probability $P_{i,i-1} = \frac{i}{100}$
- move to state $i+1$ - this has probability $P_{i,i+1} = \frac{100-i}{100}$.

Let $P = (P_{ij})$ a stochastic matrix.

From state i we can

- move to state $i-1$ - this has probability $P_{i,i-1} = \frac{i}{100}$
- move to state $i+1$ - this has probability $P_{i,i+1} = \frac{100-i}{100}$.

Let $P = (P_{ij})$ a stochastic matrix.

One can easily check $pP = p$.

From state i we can

- move to state $i-1$ - this has probability $P_{i,i-1} = \frac{i}{100}$
- move to state $i+1$ - this has probability $P_{i,i+1} = \frac{100-i}{100}$.

Let $P = (P_{ij})$ a stochastic matrix.

One can easily check $pP = p$.

Hence we get a Markov measure μ_P.

From state i we can

- move to state $i-1$ - this has probability $P_{i,i-1} = \frac{i}{100}$
- move to state $i+1$ - this has probability $P_{i,i+1} = \frac{100-i}{100}$.

Let $P = (P_{ij})$ a stochastic matrix.

One can easily check $pP = p$.

Hence we get a Markov measure μ_P.

Let $A = [100]$ - a cylinder of length 1. This represents there being 100 balls in the first urn.

From state i we can

- move to state $i-1$ - this has probability $P_{i,i-1} = \frac{i}{100}$
- move to state $i+1$ - this has probability $P_{i,i+1} = \frac{100-i}{100}$.

Let $P = (P_{ij})$ a stochastic matrix.
One can easily check $pP = p$.
Hence we get a Markov measure μ_P.
Let $A = [100]$ - a cylinder of length 1. This represents there being 100 balls in the first urn. By Kac's Lemma, the expected first return time to A is

$$\frac{1}{\mu_P(A)} \quad = \quad 2^{100} \text{ seconds}$$

From state i we can

- move to state $i-1$ - this has probability $P_{i,i-1} = \frac{i}{100}$
- move to state $i+1$ - this has probability $P_{i,i+1} = \frac{100-i}{100}$.

Let $P = (P_{ij})$ a stochastic matrix.
One can easily check $pP = p$.
Hence we get a Markov measure μ_P.
Let $A = [100]$ - a cylinder of length 1. This represents there being 100 balls in the first urn. By Kac's Lemma, the expected first return time to A is

$$\begin{aligned} \frac{1}{\mu_P(A)} &= 2^{100} \text{ seconds} \\ &= 4 \times 10^{22} \text{ years} \end{aligned}$$

From state i we can

- move to state $i-1$ - this has probability $P_{i,i-1} = \frac{i}{100}$
- move to state $i+1$ - this has probability $P_{i,i+1} = \frac{100-i}{100}$.

Let $P = (P_{ij})$ a stochastic matrix.
One can easily check $pP = p$.
Hence we get a Markov measure μ_P.
Let $A = [100]$ - a cylinder of length 1. This represents there being 100 balls in the first urn. By Kac's Lemma, the expected first return time to A is

$$\begin{aligned}
\frac{1}{\mu_P(A)} &= 2^{100} \text{ seconds} \\
&= 4 \times 10^{22} \text{ years} \\
&= \text{about } 3 \times 10^{12} \times \text{ age of the Universe!}
\end{aligned}$$

From state i we can

- move to state $i-1$ - this has probability $P_{i,i-1} = \frac{i}{100}$
- move to state $i+1$ - this has probability $P_{i,i+1} = \frac{100-i}{100}$.

Let $P = (P_{ij})$ a stochastic matrix.
One can easily check $pP = p$.
Hence we get a Markov measure μ_P.
Let $A = [100]$ - a cylinder of length 1. This represents there being 100 balls in the first urn. By Kac's Lemma, the expected first return time to A is

$$\begin{aligned} \frac{1}{\mu_P(A)} &= 2^{100} \text{ seconds} \\ &= 4 \times 10^{22} \text{ years} \\ &= \text{about } 3 \times 10^{12} \times \text{ age of the Universe!} \end{aligned}$$

Open problem: how are return times distributed?

From state i we can

- move to state $i-1$ - this has probability $P_{i,i-1} = \frac{i}{100}$
- move to state $i+1$ - this has probability $P_{i,i+1} = \frac{100-i}{100}$.

Let $P = (P_{ij})$ a stochastic matrix.
One can easily check $pP = p$.
Hence we get a Markov measure μ_P.
Let $A = [100]$ - a cylinder of length 1. This represents there being 100 balls in the first urn. By Kac's Lemma, the expected first return time to A is

$$\begin{aligned}\frac{1}{\mu_P(A)} &= 2^{100} \text{ seconds} \\ &= 4 \times 10^{22} \text{ years} \\ &= \text{about } 3 \times 10^{12} \times \text{ age of the Universe!}\end{aligned}$$

Open problem: how are return times distributed? We would expect a Poisson distribution.

From state i we can

- move to state $i-1$ - this has probability $P_{i,i-1} = \frac{i}{100}$
- move to state $i+1$ - this has probability $P_{i,i+1} = \frac{100-i}{100}$.

Let $P = (P_{ij})$ a stochastic matrix.
One can easily check $pP = p$.
Hence we get a Markov measure μ_P.
Let $A = [100]$ - a cylinder of length 1. This represents there being 100 balls in the first urn. By Kac's Lemma, the expected first return time to A is

$$\begin{aligned} \frac{1}{\mu_P(A)} &= 2^{100} \text{ seconds} \\ &= 4 \times 10^{22} \text{ years} \\ &= \text{about } 3 \times 10^{12} \times \text{ age of the Universe!} \end{aligned}$$

Open problem: how are return times distributed? We would expect a Poisson distribution. This is known for a number of examples (hyperbolic dynamics,...)

Normality of numbers

Normality of numbers

Let $r \geq 2$ be a fixed integer base.

Normality of numbers

Let $r \geq 2$ be a fixed integer base. Recall that every $x \in [0,1]$ can be expanded as a base r "decimal":

$$x = \sum_{j=1}^{\infty} \frac{x_j}{r^j}; \qquad x_j \in \{0, 1, \ldots, r-1\}.$$

Normality of numbers

Let $r \geq 2$ be a fixed integer base. Recall that every $x \in [0,1]$ can be expanded as a base r "decimal":

$$x = \sum_{j=1}^{\infty} \frac{x_j}{r^j}; \qquad x_j \in \{0, 1, \ldots, r-1\}.$$

This expansion is unique unless the sequence $(x_j)_{j=1}^{\infty}$ ends in infinitely repeated 0s or $(r-1)s$.

Normality of numbers

Let $r \geq 2$ be a fixed integer base. Recall that every $x \in [0, 1]$ can be expanded as a base r "decimal":

$$x = \sum_{j=1}^{\infty} \frac{x_j}{r^j}; \qquad x_j \in \{0, 1, \ldots, r-1\}.$$

This expansion is unique unless the sequence $(x_j)_{j=1}^{\infty}$ ends in infinitely repeated $0s$ or $(r-1)s$.

Definition

$x \in [0, 1]$ is *(simply) normal in base r* if it has a unique base r expansion and for each digit $k \in \{0, 1, \ldots, r-1\}$, the frequency with which the digit k occurs in the base r expansion of x is $\frac{1}{r}$.

Proposition

Lebesgue-a.e. $x \in [0, 1]$ is simultaneously normal in every base $r \geq 2$.

Proposition

Lebesgue-a.e. $x \in [0, 1]$ is simultaneously normal in every base $r \geq 2$.

Proof:

Let $\mu =$ Lebesgue measure.

Proposition

Lebesgue-a.e. $x \in [0, 1]$ is simultaneously normal in every base $r \geq 2$.

Proof:

Let μ = Lebesgue measure.
Fix $r \geq 2$ and define $T_r : [0, 1] \to [0, 1]$ by $T_r x = rx \bmod 1$.

Proposition

Lebesgue-a.e. $x \in [0,1]$ is simultaneously normal in every base $r \geq 2$.

Proof:

Let $\mu =$ Lebesgue measure.
Fix $r \geq 2$ and define $T_r : [0,1] \to [0,1]$ by $T_r x = rx \bmod 1$.
Then μ is an ergodic invariant probability measure for T_r.

Proposition

Lebesgue-a.e. $x \in [0,1]$ is simultaneously normal in every base $r \geq 2$.

Proof:

Let $\mu =$ Lebesgue measure.
Fix $r \geq 2$ and define $T_r : [0,1] \to [0,1]$ by $T_r x = rx \bmod 1$.
Then μ is an ergodic invariant probability measure for T_r.
Note: If

$$x = \sum_{j=1}^{\infty} \frac{x_j}{r^j}$$

then

$$x_j = k \Leftrightarrow T^{j-1} x \in \left[\frac{k}{r}, \frac{k+1}{r}\right).$$

Hence

$$\text{freq. with which } k \text{ occurs} \quad = \quad \lim_{n\to\infty} \frac{1}{n}\mathrm{card}\{1 \le j \le n \mid x_j = k\}$$

Hence

$$\begin{aligned} \text{freq. with which } k \text{ occurs} &= \lim_{n\to\infty} \frac{1}{n} \mathrm{card}\{1 \le j \le n \mid x_j = k\} \\ &= \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \chi_{[\frac{k}{r}, \frac{k+1}{r})}(T^j x) \end{aligned}$$

Hence

$$\begin{aligned}
\text{freq. with which } k \text{ occurs} &= \lim_{n\to\infty} \frac{1}{n}\mathrm{card}\{1 \le j \le n \mid x_j = k\} \\
&= \lim_{n\to\infty} \frac{1}{n}\sum_{j=0}^{n-1} \chi_{\left[\frac{k}{r}, \frac{k+1}{r}\right)}(T^j x) \\
&= \mu\left[\frac{k}{r}, \frac{k+1}{r}\right) \ \mu\text{-a.e.}
\end{aligned}$$

Hence

$$\begin{aligned}
\text{freq. with which } k \text{ occurs} &= \lim_{n\to\infty} \frac{1}{n}\mathrm{card}\{1 \le j \le n \mid x_j = k\} \\
&= \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \chi_{\left[\frac{k}{r}, \frac{k+1}{r}\right)}(T^j x) \\
&= \mu\left[\frac{k}{r}, \frac{k+1}{r}\right) \ \mu\text{-a.e.} \\
&= \frac{1}{r} \ \mu\text{-a.e.}
\end{aligned}$$

Hence

$$\begin{aligned}
\text{freq. with which } k \text{ occurs} &= \lim_{n\to\infty} \frac{1}{n}\mathrm{card}\{1 \le j \le n \mid x_j = k\} \\
&= \lim_{n\to\infty} \frac{1}{n}\sum_{j=0}^{n-1} \chi_{[\frac{k}{r},\frac{k+1}{r})}(T^j x) \\
&= \mu\left[\frac{k}{r},\frac{k+1}{r}\right) \ \mu\text{-a.e.} \\
&= \frac{1}{r} \ \mu\text{-a.e.}
\end{aligned}$$

Hence μ-a.e. $x \in [0,1]$ is normal in base r.

Hence

$$\begin{aligned}
\text{freq. with which } k \text{ occurs} &= \lim_{n\to\infty} \frac{1}{n} \mathrm{card}\{1 \le j \le n \mid x_j = k\} \\
&= \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \chi_{[\frac{k}{r},\frac{k+1}{r})}(T^j x) \\
&= \mu\left[\frac{k}{r}, \frac{k+1}{r}\right) \ \mu\text{-a.e.} \\
&= \frac{1}{r} \ \mu\text{-a.e.}
\end{aligned}$$

Hence μ-a.e. $x \in [0,1]$ is normal in base r.
Taking the union of the exceptional sets of measure zero gives μ-a.e. $x \in [0,1]$ is normal in every base $r \geq 2$. $\square$

Hence

$$\begin{aligned}
\text{freq. with which } k \text{ occurs} &= \lim_{n\to\infty} \frac{1}{n}\mathrm{card}\{1 \le j \le n \mid x_j = k\} \\
&= \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \chi_{[\frac{k}{r},\frac{k+1}{r})}(T^j x) \\
&= \mu\left[\frac{k}{r},\frac{k+1}{r}\right) \ \mu\text{-a.e.} \\
&= \frac{1}{r} \ \mu\text{-a.e.}
\end{aligned}$$

Hence μ-a.e. $x \in [0,1]$ is normal in base r.
Taking the union of the exceptional sets of measure zero gives μ-a.e. $x \in [0,1]$ is normal in every base $r \ge 2$. $\square$

Remark

It is easy to give an example of an $x \in [0,1]$ that is (simply) normal in a given base r.

Hence

$$\begin{aligned}
\text{freq. with which } k \text{ occurs} &= \lim_{n\to\infty} \frac{1}{n} \mathrm{card}\{1 \le j \le n \mid x_j = k\} \\
&= \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \chi_{[\frac{k}{r}, \frac{k+1}{r})}(T^j x) \\
&= \mu \left[\frac{k}{r}, \frac{k+1}{r}\right) \ \mu\text{-a.e.} \\
&= \frac{1}{r} \ \mu\text{-a.e.}
\end{aligned}$$

Hence μ-a.e. $x \in [0,1]$ is normal in base r.
Taking the union of the exceptional sets of measure zero gives μ-a.e. $x \in [0,1]$ is normal in every base $r \ge 2$. $\square$

Remark

It is easy to give an example of an $x \in [0,1]$ that is (simply) normal in a given base r. However, there is no known example of an x that is simultaneously normal in every base.

Continued fractions

Continued fractions

Let $Tx = \frac{1}{x} \bmod 1$ be the continued fraction map.

Continued fractions

Let $Tx = \frac{1}{x} \bmod 1$ be the continued fraction map. Recall:

1. If $x = \cfrac{1}{x_0 + \cfrac{1}{x_1 + \cfrac{1}{x_2 + \dots}}}$ then $Tx = \cfrac{1}{x_1 + \cfrac{1}{x_2 + \cfrac{1}{x_3 + \dots}}}$.

Continued fractions

Let $Tx = \frac{1}{x} \bmod 1$ be the continued fraction map. Recall:

1. If $x = \cfrac{1}{x_0 + \cfrac{1}{x_1 + \cfrac{1}{x_2 + \dots}}}$ then $Tx = \cfrac{1}{x_1 + \cfrac{1}{x_2 + \cfrac{1}{x_3 + \dots}}}$.
2. T is ergodic w.r.t. Gauss' measure μ:

$$\mu(B) = \frac{1}{\log 2} \int_B \frac{dx}{1+x}.$$

Continued fractions

Let $Tx = \frac{1}{x} \bmod 1$ be the continued fraction map. Recall:

1. If $x = \cfrac{1}{x_0 + \cfrac{1}{x_1 + \cfrac{1}{x_2 + \dots}}}$ then $Tx = \cfrac{1}{x_1 + \cfrac{1}{x_2 + \cfrac{1}{x_3 + \dots}}}$.
2. T is ergodic w.r.t. Gauss' measure μ:

$$\mu(B) = \frac{1}{\log 2} \int_B \frac{dx}{1+x}.$$

3. Gauss' measure μ and Lebesgue measure λ are equivalent: they have the same sets of measure zero.

Proposition

For Lebesgue a.e. $x \in [0, 1]$, the frequency with which digit $k \in \mathbb{N}$ occurs in the continued fraction expansion of x is

$$\frac{1}{\log 2} \log\left(\frac{(k+1)^2}{k(k+2)}\right).$$

Proposition

For Lebesgue a.e. $x \in [0,1]$, the frequency with which digit $k \in \mathbb{N}$ occurs in the continued fraction expansion of x is

$$\frac{1}{\log 2} \log\left(\frac{(k+1)^2}{k(k+2)}\right).$$

Proof:

Recall: if $x = \cfrac{1}{x_0 + \cfrac{1}{x_1 + \dots}}$ then $x_n = \left[\frac{1}{T^n x}\right]$.

Proposition

For Lebesgue a.e. $x \in [0,1]$, the frequency with which digit $k \in \mathbb{N}$ occurs in the continued fraction expansion of x is

$$\frac{1}{\log 2} \log\left(\frac{(k+1)^2}{k(k+2)}\right).$$

Proof:

Recall: if $x = \cfrac{1}{x_0 + \cfrac{1}{x_1 + \dots}}$ then $x_n = \left[\frac{1}{T^n x}\right]$.

Hence:

$$x_n = k \quad \Longleftrightarrow \quad \left[\frac{1}{T^n x}\right] = k$$

Proposition

For Lebesgue a.e. $x \in [0, 1]$, the frequency with which digit $k \in \mathbb{N}$ occurs in the continued fraction expansion of x is

$$\frac{1}{\log 2} \log\left(\frac{(k+1)^2}{k(k+2)}\right).$$

Proof:

Recall: if $x = \cfrac{1}{x_0 + \cfrac{1}{x_1 + \ldots}}$ then $x_n = \left[\frac{1}{T^n x}\right]$.

Hence:

$$x_n = k \quad \Longleftrightarrow \quad \left[\frac{1}{T^n x}\right] = k \Longleftrightarrow k \leq \frac{1}{T^n x} < k+1$$

Proposition

For Lebesgue a.e. $x \in [0,1]$, the frequency with which digit $k \in \mathbb{N}$ occurs in the continued fraction expansion of x is

$$\frac{1}{\log 2} \log\left(\frac{(k+1)^2}{k(k+2)}\right).$$

Proof:

Recall: if $x = \cfrac{1}{x_0 + \cfrac{1}{x_1 + \dots}}$ then $x_n = \left[\frac{1}{T^n x}\right]$.

Hence:

$$\begin{aligned} x_n = k \quad &\iff \quad \left[\frac{1}{T^n x}\right] = k \iff k \leq \frac{1}{T^n x} < k+1 \\ &\iff \quad T^n x \in \left(\frac{1}{k+1}, \frac{1}{k}\right]. \end{aligned}$$

Hence:

$$\text{freq. with which } k \text{ occurs} \quad = \quad \lim_{n\to\infty} \frac{1}{n} \mathrm{card}\{0 \le j \le n-1 \mid x_j = k\}$$

Hence:

$$\begin{aligned}
\text{freq. with which } k \text{ occurs} &= \lim_{n\to\infty} \frac{1}{n}\mathrm{card}\{0 \le j \le n-1 \mid x_j = k\} \\
&= \lim_{n\to\infty} \frac{1}{n}\sum_{j=0}^{n-1} \chi_{(\frac{1}{k+1},\frac{1}{k}]}(T^j x)
\end{aligned}$$

Hence:

$$\begin{aligned} \text{freq. with which } k \text{ occurs} &= \lim_{n\to\infty} \frac{1}{n} \mathrm{card}\{0 \le j \le n-1 \mid x_j = k\} \\ &= \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \chi_{\left(\frac{1}{k+1}, \frac{1}{k}\right]}(T^j x) \\ &= \int \chi_{\left(\frac{1}{k+1}, \frac{1}{k}\right]}\, d\mu \quad \mu\text{-a.e.} \end{aligned}$$

Hence:

$$\begin{aligned}
\text{freq. with which } k \text{ occurs} &= \lim_{n\to\infty} \frac{1}{n}\mathrm{card}\{0 \le j \le n-1 \mid x_j = k\} \\
&= \lim_{n\to\infty} \frac{1}{n}\sum_{j=0}^{n-1} \chi_{\left(\frac{1}{k+1},\frac{1}{k}\right]}(T^j x) \\
&= \int \chi_{\left(\frac{1}{k+1},\frac{1}{k}\right]}\, d\mu \quad \mu\text{-a.e.} \\
&= \frac{1}{\log 2}\int_{\frac{1}{k+1}}^{\frac{1}{k}} \frac{dx}{1+x} \quad \mu\text{-a.e}
\end{aligned}$$

Hence:

$$\begin{aligned}
\text{freq. with which } k \text{ occurs} &= \lim_{n\to\infty} \frac{1}{n}\mathrm{card}\{0 \le j \le n-1 \mid x_j = k\} \\
&= \lim_{n\to\infty} \frac{1}{n}\sum_{j=0}^{n-1} \chi_{\left(\frac{1}{k+1},\frac{1}{k}\right]}(T^j x) \\
&= \int \chi_{\left(\frac{1}{k+1},\frac{1}{k}\right]}\, d\mu \ \ \mu\text{-a.e.} \\
&= \frac{1}{\log 2}\int_{\frac{1}{k+1}}^{\frac{1}{k}} \frac{dx}{1+x} \ \ \mu\text{-a.e} \\
&= \frac{1}{\log 2}\log\left[\frac{(k+1)^2}{k(k+2)}\right] \ \ \mu\text{-a.e.}
\end{aligned}$$

Hence:

$$
\begin{aligned}
\text{freq. with which } k \text{ occurs} &= \lim_{n\to\infty} \frac{1}{n} \mathrm{card}\{0 \le j \le n-1 \mid x_j = k\} \\
&= \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \chi_{(\frac{1}{k+1},\frac{1}{k}]}(T^j x) \\
&= \int \chi_{(\frac{1}{k+1},\frac{1}{k}]}\, d\mu \quad \mu\text{-a.e.} \\
&= \frac{1}{\log 2} \int_{\frac{1}{k+1}}^{\frac{1}{k}} \frac{dx}{1+x} \quad \mu\text{-a.e} \\
&= \frac{1}{\log 2} \log \left[\frac{(k+1)^2}{k(k+2)} \right] \quad \mu\text{-a.e.}
\end{aligned}
$$

As Gauss' measure and Lebesgue measure are equivalent, this holds Lebesgue a.e.

$\square$

Proposition

For Lebesgue a.e. $x \in [0, 1]$

Proposition

For Lebesgue a.e. $x \in [0, 1]$

1. the arithmetic mean of the continued fraction digits of x is infinite

Proposition

For Lebesgue a.e. $x \in [0, 1]$

1. the arithmetic mean of the continued fraction digits of x is infinite:

$$\frac{1}{n}(x_0 + x_1 + \cdots + x_{n-1}) \longrightarrow \infty \; a.e.$$

Proposition

For Lebesgue a.e. $x \in [0,1]$

1. the arithmetic mean of the continued fraction digits of x is infinite:
$$\frac{1}{n}(x_0 + x_1 + \cdots + x_{n-1}) \longrightarrow \infty \; a.e.$$

2. the geometric mean of the continued fraction digits of x is:
$$\sqrt[n]{x_0 \ldots x_{n-1}} \longrightarrow \prod_{k=1}^{\infty} \left(1 + \frac{1}{k(k+2)}\right)^{\frac{\log k}{\log 2}} < \infty \text{ a.e.}$$

Proof.

Proof.

1. is an exercise

Proof.

1. is an exercise
2. Define $f(x) = \log k$ for $x \in \left(\frac{1}{k+1}, \frac{1}{k}\right]$.

Proof.

1. is an exercise
2. Define $f(x) = \log k$ for $x \in \left(\frac{1}{k+1}, \frac{1}{k}\right]$. Then

$$\frac{1}{n}(\log x_0 + \cdots + \log x_{n-1}) = \frac{1}{n}\sum_{j=0}^{n-1} f(T^j x)$$

Proof.

1. is an exercise
2. Define $f(x) = \log k$ for $x \in \left(\frac{1}{k+1}, \frac{1}{k}\right]$. Then

$$\frac{1}{n}(\log x_0 + \cdots + \log x_{n-1}) = \frac{1}{n}\sum_{j=0}^{n-1} f(T^j x)$$

$$\longrightarrow \int f \, d\mu \quad \mu\text{-a.e.}$$

Proof.

1. is an exercise
2. Define $f(x) = \log k$ for $x \in \left(\frac{1}{k+1}, \frac{1}{k}\right]$. Then

$$\begin{aligned}\frac{1}{n}(\log x_0 + \quad \cdots \quad + \log x_{n-1}) = \frac{1}{n}\sum_{j=0}^{n-1} f(T^j x) \\ \longrightarrow \quad \int f \, d\mu \ \mu\text{-a.e.} \\ = \quad \frac{1}{\log 2}\int \frac{f(x)}{1+x}\, d\mu \ \mu\text{-a.e.}\end{aligned}$$

Proof.

1. is an exercise
2. Define $f(x) = \log k$ for $x \in \left(\frac{1}{k+1}, \frac{1}{k}\right]$. Then

$$\begin{aligned}
\frac{1}{n}(\log x_0 + \quad \cdots \quad + \log x_{n-1}) &= \frac{1}{n}\sum_{j=0}^{n-1} f(T^j x) \\
&\longrightarrow \quad \int f \, d\mu \ \ \mu\text{-a.e.} \\
&= \quad \frac{1}{\log 2}\int \frac{f(x)}{1+x} \, d\mu \ \ \mu\text{-a.e.} \\
&= \quad \frac{1}{\log 2}\sum_{k=1}^{\infty}\int_{\frac{1}{k+1}}^{\frac{1}{k}} \frac{\log k}{1+x} \, dx \ \ \mu\text{-a.e.}
\end{aligned}$$

Proof.

1. is an exercise
2. Define $f(x) = \log k$ for $x \in \left(\frac{1}{k+1}, \frac{1}{k}\right]$. Then

$$
\begin{aligned}
\frac{1}{n}(\log x_0 + \cdots + \log x_{n-1}) &= \frac{1}{n}\sum_{j=0}^{n-1} f(T^j x) \\
&\longrightarrow \int f \, d\mu \quad \mu\text{-a.e.} \\
&= \frac{1}{\log 2}\int \frac{f(x)}{1+x}\, d\mu \quad \mu\text{-a.e.} \\
&= \frac{1}{\log 2}\sum_{k=1}^{\infty}\int_{\frac{1}{k+1}}^{\frac{1}{k}} \frac{\log k}{1+x}\, dx \quad \mu\text{-a.e.} \\
&= \sum_{k=1}^{\infty}\frac{\log k}{\log 2}\log\left(1+\frac{1}{k(k+2)}\right) \quad \mu\text{-a.e.}
\end{aligned}
$$

Next lecture

Given a measure-preserving transformation T of a probability space $(X, \mathcal{B}, \mu)$, we have been asking whether the given measure μ is invariant, and if so, if it is ergodic.

Next lecture

Given a measure-preserving transformation T of a probability space $(X, \mathcal{B}, \mu)$, we have been asking whether the given measure μ is invariant, and if so, if it is ergodic.

Is it the case that a given transformation of a measure space $(X, \mathcal{B})$ possesses an ergodic measure?

Next lecture

Given a measure-preserving transformation T of a probability space $(X, \mathcal{B}, \mu)$, we have been asking whether the given measure μ is invariant, and if so, if it is ergodic.

Is it the case that a given transformation of a measure space $(X, \mathcal{B})$ possesses an ergodic measure?

In the next lecture, we study this question in the context of continuous transformations of compact metric spaces.