Introduction to Dynamical Systems

George Voutsadakis[1]

[1]Mathematics and Computer Science
Lake Superior State University

LSSU Math 500

1 Ergodic Theory
- Notions from Measure Theory
- Invariant Measures
- Nontrivial Recurrence
- The Ergodic Theorem
- Metric Entropy
- Proof of the Ergodic Theorem

Subsection 1

Notions from Measure Theory

σ-Algebras

- Let X be a set.
- Let $\mathcal{A}$ be a family of subsets of X.

Definition

$\mathcal{A}$ is said to be a σ-**algebra** in X if:

1. $\emptyset, X \in \mathcal{A}$;
2. $X \backslash B \in \mathcal{A}$ when $B \in \mathcal{A}$;
3. $\bigcup_{n=1}^{\infty} B_n \in \mathcal{A}$ when $B_n \in \mathcal{A}$, for all $n \in \mathbb{N}$.

The σ-**algebra generated by a family** $\mathcal{A}$ of subsets of X is the smallest σ-algebra in X containing all elements of $\mathcal{A}$.

Measure Spaces

Definition

Let $\mathcal{A}$ be a σ-algebra in X. A function $\mu : \mathcal{A} \to [0, +\infty]$ is called a **measure on** X (**with respect to** $\mathcal{A}$) if:

1. $\mu(\emptyset) = 0$;
2. Given pairwise disjoint sets $B_n \in \mathcal{A}$, for $n \in \mathbb{N}$, we have

$$\mu\left(\bigcup_{n=1}^{\infty} B_n\right) = \sum_{n=1}^{\infty} \mu(B_n).$$

We then say that $(X, \mathcal{A}, \mu)$ is a **measure space**.

- When the σ-algebra is understood from the context, we still refer to the pair (X, μ) as a "measure space".

Examples

- Let $\mathcal{A}$ be the σ-algebra in X containing all subsets of X.

  We define a measure $\mu : \mathcal{A} \to \mathbb{N}_0 \cup \{\infty\}$ on X by

  $$\mu(B) = \text{card} B.$$

  We call μ the **counting measure** on X.
- Let $\mathcal{B}$ be the **Borel σ-algebra** in $\mathbb{R}$.

  This is the σ-algebra generated by the open intervals.

  Then there exists a unique measure $\lambda : \mathcal{B} \to [0, +\infty]$ on $\mathbb{R}$, such that

  $$\lambda((a, b)) = b - a, \quad \text{for } a < b.$$

  We call λ the **Lebesgue measure** on $\mathbb{R}$.

Example

- Let $\mathcal{B}$ be the **Borel σ-algebra** in $\mathbb{R}^n$.

  This is the σ-algebra generated by the open rectangles

  $$\prod_{i=1}^{n}(a_i, b_i), \quad \text{with } a_i < b_i, \text{ for } i = 1, \ldots, n.$$

  Then there exists a unique measure $\lambda : \mathcal{B} \to [0, +\infty]$ on $\mathbb{R}^n$, such that

  $$\lambda\left(\prod_{i=1}^{n}(a_i, b_i)\right) = \prod_{i=1}^{n}(b_i - a_i),$$

  for any $a_i < b_i$ and $i = 1, \ldots, n$.

  We call λ the **Lebesgue measure** on $\mathbb{R}^n$.

Measurable Functions

- Let X be a set.
- Let $\mathcal{A}$ be a σ-algebra in the set X.

Definition

A function $\varphi : X \to \mathbb{R}$ is said to be $\mathcal{A}$-**measurable** or simply **measurable** if

$$\varphi^{-1}B \in \mathcal{A}, \quad \text{for all } B \in \mathcal{B},$$

where $\mathcal{B}$ is the Borel σ-algebra in $\mathbb{R}$.

Characteristic and Simple Functions

- The **characteristic function** of a set $B \subseteq X$, $\chi_B : X \to \{0,1\}$ is defined by

$$\chi_B(x) = \begin{cases} 1, & \text{if } x \in B, \\ 0, & \text{if } x \notin B. \end{cases}$$

Definition

Given sets $B_1, \ldots, B_n \in \mathcal{A}$ and numbers $a_1, \ldots, a_n \in \mathbb{R}$, the function

$$s = \sum_{k=1}^{n} a_k \chi_{B_k}$$

is called a **simple function**.

- Clearly, all simple functions are measurable.

The Lebesgue Integral of a Positive Measurable Function

Definition

Let (X, μ) be a measure space. The (**Lebesgue**) **integral** of a measurable function $\varphi : X \to \mathbb{R}_0^+$ is defined by

$$\int_X \varphi d\mu = \sup \left\{ \sum_{k=1}^{n} a_k \mu(B_k) : \sum_{k=1}^{n} a_k \chi_{B_k} \leq \varphi \right\}.$$

The Lebesgue Integral of an Integrable Function

Definition

Let (X,μ) be a measure space. Let $\varphi : X \to \mathbb{R}$ be a measurable function. φ is μ-**integrable** if

$$\int_X \varphi^+ d\mu < \infty \quad \text{and} \quad \int_X \varphi^- d\mu < \infty,$$

where $\varphi^+ = \max\{\varphi, 0\}$ and $\varphi^- = \max\{-\varphi, 0\}$.

Definition

The (**Lebesgue**) **integral** of a μ-integrable function $\varphi : X \to \mathbb{R}$ is defined by

$$\int_X \varphi d\mu = \int_X \varphi^+ d\mu - \int_X \varphi^- d\mu.$$

Subsection 2

Invariant Measures

Measurable Functions and Invariant Measures

- Let $(X, \mathcal{A}, \mu)$ be a measure space.

Definition

A map $f : X \to X$ is said to be $\mathcal{A}$-**measurable** or simply **measurable** if

$$f^{-1}B \in \mathcal{A}, \quad \text{for every } B \in \mathcal{A},$$

where $f^{-1}B = \{x \in X : f(x) \in B\}$.

Definition

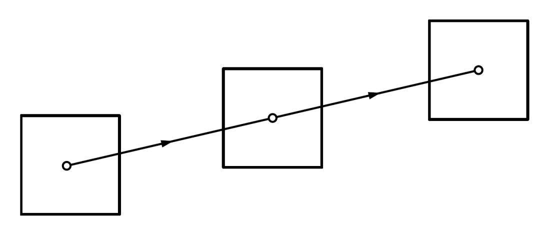
Given a measurable map $f : X \to X$, we say that μ is f-**invariant** and that f **preserves** μ if

$$\mu(f^{-1}B) = \mu(B), \quad \text{for } B \in \mathcal{A}.$$

- We note that when f is an invertible map with measurable inverse, f-invariance is equivalent to $\mu(f(B)) = \mu(B)$, for $B \in \mathcal{A}$.

Example: Translations and Lebesgue Measure

- Given $v \in \mathbb{R}^n$, let $f : \mathbb{R}^n \to \mathbb{R}^n$ be the translation
$$f(x) = x + v.$$



Clearly, f is invertible.

We consider the Lebesgue measure λ on $\mathbb{R}^n$.

Example (Cont'd)

- For all $B \in \mathcal{B}$, we have

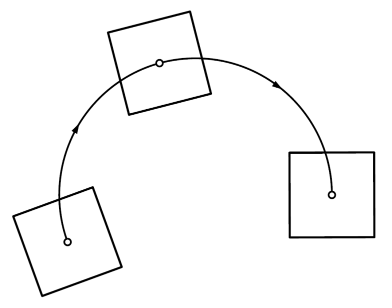
$$\begin{aligned} \lambda(f(B)) &= \int_{f(B)} 1 d\lambda \\ &= \int_B |\det d_x f| d\lambda(x) \\ &= \int_B 1 d\lambda \\ &= \lambda(B). \end{aligned}$$

  So the measure λ is f-invariant.

  So the translations of $\mathbb{R}^n$ preserve Lebesgue measure.

Example: Rotations and Lebesgue Measure

- Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be a rotation.



Then there exists an $n \times n$ orthogonal matrix A (i.e., such that $A^\top A = \text{Id}$, where $A^\top$ is the transpose of A), such that

$$f(x) = Ax.$$

Example (Cont'd)

- Orthogonal matrices have determinant ± 1.

  Let λ be the Lebesgue measure on $\mathbb{R}^n$.

  Then, for each $B \in \mathcal{B}$, we have

$$\begin{array}{rcl} \lambda(f(B)) & = & \int_{f(B)} 1 d\lambda \\ & = & \int_B |\det d_x f| d\lambda(x) \\ & = & \int_B |\det A| d\lambda \\ & = & \int_B 1 d\lambda \\ & = & \lambda(B). \end{array}$$

  Rotations are invertible maps.

  This shows that λ is f-invariant.

  Thus, the rotations of $\mathbb{R}^n$ preserve Lebesgue measure.

Example: Rotations of the Circle

- Consider a rotation of the circle $R_\alpha : S^1 \to S^1$.

  Without loss of generality, we assume that $\alpha \in [0,1]$.

  We first introduce a measure μ on S^1.

  For each set $B \subseteq [0,1]$ in the Borel σ-algebra in $\mathbb{R}$, we define

  $$\mu(B) = \lambda(B).$$

  Then μ is a measure on S^1 with $\mu(S^1) = 1$.

  We also have $R_\alpha^{-1} B = B - \alpha$, where

  $$B - \alpha = \{x - \alpha : x \in B\} \subseteq \mathbb{R}.$$

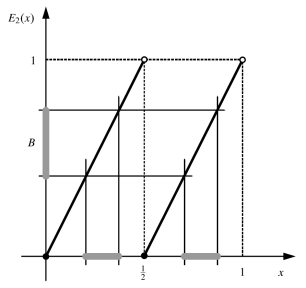
  By the invariance of Lebesgue measure under translations,

  $$\mu(R_\alpha^{-1} B) = \lambda(B - \alpha) = \lambda(B) = \mu(B).$$

  This shows that the rotations of the circle preserve the measure μ.

Example: Expanding Maps

- Consider the expanding map $E_m : S^1 \to S^1$.
  We show that the measure μ of the previous slide is E_m-invariant.
  Let $B \subseteq [0,1]$ be a set in the Borel σ-algebra in $\mathbb{R}$.



Then $E_m^{-1}B = \bigcup_{i=1}^m B_i$, where

$$B_i = \left\{ \frac{x+i}{m} : x \in B \right\} \quad \text{mod } 1.$$

Example (Cont'd)

- Note that the sets B_i are pairwise disjoint.

  So we have

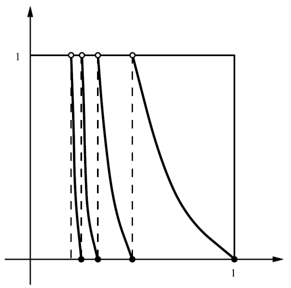
$$\begin{array}{rcl}\mu(E_m^{-1}B) & = & \sum_{i=1}^m \lambda(B_i) \\ & = & \sum_{i=1}^m \frac{\lambda(B+i)}{m} \\ & = & \sum_{i=1}^m \frac{\lambda(B)}{m} \\ & = & \lambda(B) \\ & = & \mu(B).\end{array}$$

  Thus, the μ measure is E_m-invariant.

Example: The Gauss Map

- The **Gauss map** $f : [0,1] \to [0,1]$ is defined by

$$f(x) = \begin{cases} \frac{1}{x} \mod 1, & \text{if } x \neq 0, \\ 0, & \text{if } x = 0. \end{cases}$$



The map f is related to the theory of continued fractions.

Let $x \in (0,1)$ be an irrational number with continued fraction

$$x = \frac{1}{n_1 + \frac{1}{n_2 + \cdots}}.$$

Then $n_j = \left\lfloor \frac{1}{f^{j-1}(x)} \right\rfloor$, for $j \in \mathbb{N}$.

Example (Cont'd)

- We show the Gauss map preserves the measure μ in $[0,1]$ given by
$$\mu(A) = \int_A \frac{1}{1+x} dx.$$
We note that it is sufficient to consider the intervals of the form $(0,b)$, with $b \in (0,1)$ (they generate the Borel σ-algebra). Since $f^{-1}(0,b) = \bigcup_{n=1}^{\infty}(\frac{1}{n+b}, \frac{1}{n})$ is a disjoint union, we obtain
$$\begin{array}{rcl} \mu(f^{-1}(0,b)) & = & \sum_{n=1}^{\infty} \mu((\frac{1}{n+b}, \frac{1}{n})) \\ & = & \sum_{n=1}^{\infty} \int_{1/(n+b)}^{1/n} \frac{1}{1+x} dx \\ & = & \sum_{n=1}^{\infty} \log \frac{1+\frac{1}{n}}{1+\frac{1}{n+b}} \\ & = & \sum_{n=1}^{\infty} (\log \frac{n+1}{n+1+b} - \log \frac{n}{n+b}) \\ & = & -\log \frac{1}{1+b}. \end{array}$$
Thus, the measure μ is f-invariant.

Subsection 3

Nontrivial Recurrence

Poincaré's Recurrence Theorem

- We show that for a finite invariant measure, almost every point of a given set returns infinitely often to this set.

Theorem (Poincaré's Recurrence Theorem)

Let $f : X \to X$ be a measurable map. Let μ be a finite f-invariant measure on X. For each set $A \in \mathcal{A}$, we have

$$\mu(\{x \in A : f^n(x) \in A \text{ for infinitely many values of } n\}) = \mu(A).$$

Set

$$A_n = \bigcup_{k=n}^{\infty} f^{-k}A.$$

Define

$$B = \{x \in A : f^n(x) \in A \text{ for infinitely many values of } n\}.$$

Poincaré's Recurrence Theorem (Cont'd)

- We have
$$B = A \cap \bigcap_{n=1}^{\infty} A_n = A \setminus \bigcup_{n=1}^{\infty} (A \setminus A_n).$$
We note that
$$A \setminus A_n \subseteq A_0 \setminus A_n = A_0 \setminus f^{-n} A_0.$$
Now $A_0 \supseteq A_n = f^{-n} A_0$ and the measure μ is finite.
It follows, using that μ is f-invariant, that
$$0 \leq \mu(A \setminus A_n) \leq \mu(A_0 \setminus f^{-n} A_0) = \mu(A_0) - \mu(f^{-n} A_0) = 0.$$
Thus, $\mu(B) = \mu(A)$.

Example: Rotations of the Circle

- Consider a rotation of the circle
$$R_\alpha : S^1 \to S^1, \quad \alpha \in [0,1].$$
Consider the Borel σ-algebra $\mathcal{B}$ in $\mathbb{R}$.
Let μ be the measure S^1 defined by
$$\mu(B) = \lambda(B), \quad \text{for every } B \subseteq [0,1] \text{ in } \mathcal{B}.$$
By a previous example, μ is R_α-invariant.
By the theorem, given $c \in [0,1]$, the set
$$\{x \in [-c,c] : |R_\alpha^n(x)| \leq c \text{ for infinitely many values of } n\}$$
has measure $\mu([-c,c]) = 2c$.
So almost all points in $[-c,c]$ return infinitely often to $[-c,c]$.
- The property established in this example is trivial for $\alpha \in \mathbb{Q}$.
- When $\alpha \in \mathbb{R} \backslash \mathbb{Q}$, it also follows from the density of the orbits of R_α.

Example: Expanding Maps

- Consider the expanding map

$$E_m : S^1 \to S^1.$$

By a previous example, the measure μ on S^1 defined as in the preceding example is E_m-invariant.

By the theorem, for each interval $[a, b] \subseteq [0, 1]$, the set

$$\{x \in [a, b] : E_m^n(x) \in [a, b] \text{ for infinitely many values of } n\}$$

has measure $\mu([a, b]) = b - a$.

Subsection 4

The Ergodic Theorem

Birkhoff's Ergodic Theorem

- Poincaré's Recurrence Theorem says that for a finite invariant measure almost all points of a given set return infinitely often to this set.
- Birkhoff's Ergodic Theorem establishes a frequency of return.

Theorem (Birkhoff's Ergodic Theorem)

Let $f : X \to X$ be a measurable map. Let μ be a finite f-invariant measure on X. Given a μ-integrable function $\varphi : X \to \mathbb{R}$, the limit

$$\varphi_f(x) = \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \varphi(f^k(x))$$

exists for almost every point $x \in X$, the function φ_f is μ-integrable, and

$$\int_X \varphi_f d\mu = \int_X \varphi d\mu.$$

- The proof is given in the last section of this set.

Example

- Let $f : X \to X$ be a measurable map.
  Let μ be a finite f-invariant measure on X.
  Given a set $B \in \mathcal{A}$, consider the μ-integrable function $\varphi = \chi_B$.
  We have

$$\begin{array}{rcl} \int_X \varphi d\mu & = & \int_X \chi_B d\mu = \mu(B); \\ \varphi_f(x) & = & \lim\limits_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \chi_B(f^k(x)) \\ & = & \lim\limits_{n\to\infty} \frac{1}{n} \text{card}\{k \in \{0, \ldots, n-1\} : f^k(x) \in B\}. \end{array}$$

  By Birkhoff's Theorem,

$$\int_X \lim_{n\to\infty} \frac{1}{n} \text{card}\{k \in \{0, \ldots, n-1\} : f^k(x) \in B\} d\mu(x) = \mu(B).$$

  The number $\varphi_f(x)$ can be described as the frequency with which the orbit of x visits the set B. So, Birkhoff's Ergodic Theorem describes in quantitative terms how often each orbit returns to the set B.

Lyapunov Exponent

- Let $f : M \to M$ be a differentiable map.

Definition

Given $x \in M$ and $v \in T_x M$, the **Lyapunov exponent** of the pair (x, v) is defined by

$$\lambda(x, v) = \limsup_{n \to \infty} \frac{1}{n} \log \|d_x f^n v\|.$$

Lyapunov Exponent and Birkhoff's Theorem

- Now we consider the particular case of the maps of the circle.

Theorem

Let $f : S^1 \to S^1$ be a C^1 map. Let μ be a finite f-invariant measure on S^1. Then $\lambda(x, v)$ is a limit for almost every x, that is,

$$\lambda(x, v) = \lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} \varphi(f^k(x)),$$

for almost every $x \in S^1$ and any $v \neq 0$, where $\varphi(x) = \log \|d_x f\|$.

- Since the circle S^1 has dimension 1, we have

$$\|d_x f^n v\| = \|d_x f^n\| \cdot \|v\|.$$

Prof of the Theorem

- Thus, for each $v \neq 0$,

$$\lambda(x, v) = \limsup_{n\to\infty} \frac{1}{n} \log \|d_x f^n\|.$$

Recall the identity

$$d_x f^n = d_{f^{n-1}(x)} f \circ \cdots d_{f(x)} f \circ d_x f.$$

It follows that

$$\|d_x f^n\| = \prod_{k=0}^{n-1} \|d_{f^k(x)} f\|.$$

Thus, if $\varphi(x) = \log \|d_x f\|$,

$$\frac{1}{n} \log \|d_x f^n\| = \frac{1}{n} \sum_{k=0}^{n-1} \log \|d_{f^k(x)} f\| = \frac{1}{n} \sum_{k=0}^{n-1} \varphi(f^k(x)).$$

Prof of the Theorem (Cont'd)

- We saw that
$$\frac{1}{n}\log\|d_x f^n\| = \frac{1}{n}\sum_{k=0}^{n-1}\varphi(f^k(x)).$$
Thus, we get, for each $v \neq 0$,
$$\lambda(x,v) = \limsup_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1}\varphi(f^k(x)).$$
Now the function φ is continuous.

  So the result follows from Birkhoff's Ergodic Theorem.

Subsection 5

Metric Entropy

Partitions of a Measure Space

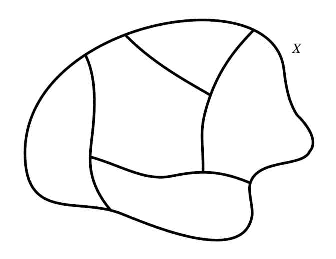
- Let $(X, \mathcal{A}, \mu)$ be a measure space, with $\mu(X) = 1$.

Definition

A finite set $\xi \subseteq \mathcal{A}$ is called a **partition** of X (with respect to μ) if:

1. $\mu(\bigcup_{C \in \xi} C) = 1$;
2. For all $C, D \in \xi$, with $C \neq D$,

$$\mu(C \cap D) = 0.$$



Derived Partitions

- Let $f : X \to X$ be a measurable map preserving the measure μ.
- Let ξ be a partition of X and $n \in \mathbb{N}$.
- We construct a new partition ξ_n formed by the sets

$$C_1 \cap f^{-1} C_2 \cap \cdots \cap f^{-(n-1)} C_n,$$

with $C_1, \ldots, C_n \in \xi$.

Definition

We define

$$h_\mu(f, \xi) = \inf_{n \in \mathbb{N}} -\frac{1}{n} \sum_{C \in \xi_n} \mu(C) \log \mu(C),$$

with the convention that $0 \log 0 = 0$.

Example

- Let $f = \mathrm{Id}$.

  Let ξ be a partition of X and $n \in \mathbb{N}$.

  Then we have

  $$\xi_n = \xi.$$

  Thus,

  $$h_\mu(f, \xi) = \inf_{n \in \mathbb{N}} -\frac{1}{n} \sum_{C \in \xi} \mu(C) \log \mu(C) = 0.$$

Example

- Consider the expanding map $E_2 : S^1 \to S^1$.

  Let μ be the E_2-invariant measure defined previously.

  Consider the partition $\xi = \{[0, \frac{1}{2}], [\frac{1}{2}, 1]\}$ and let $n \in \mathbb{N}$.

  Then we have

$$\xi_n = \left\{ \left[\frac{i}{2^n}, \frac{i+1}{2^n} \right] : i = 0, \ldots, 2^n - 1 \right\}.$$

  Thus,

$$\begin{array}{rcl} h_\mu(E_2, \xi) & = & \inf_{n \in \mathbb{N}} -\frac{1}{n} \sum_{C \in \xi_n} \mu(C) \log \mu(C) \\ & = & \inf_{n \in \mathbb{N}} -\frac{1}{n} \cdot 2^n \cdot \frac{1}{2^n} \log \frac{1}{2^n} \\ & = & \log 2. \end{array}$$

Metric Entropy

Definition

The **metric entropy** of f with respect to μ is defined by

$$h_\mu(f) = \sup_{n \in \mathbb{N}} h_\mu(f, \xi^{(n)}),$$

where $\xi^{(n)}$ is any sequence of partitions, such that:

1. Given $n \in \mathbb{N}$ and $C \in \xi^{(n)}$, there exist $C_1, \ldots, C_m \in \xi^{(n+1)}$, such that
$$\mu\left(C \setminus \bigcup_{i=1}^m C_i\right) = \mu\left(\bigcup_{i=1}^m C_i \setminus C\right) = 0;$$
2. The union of all partitions $\xi^{(n)}$ generates the σ-algebra $\mathcal{A}$.

- One can show that the definition of $h_\mu(f)$ indeed does not depend on the particular sequence $\xi^{(n)}$.

Example: Rotations of the Circle

- Let $R_\alpha : S^1 \to S^1$ be a rotation of the circle.

  Let μ be the R_α-invariant measure defined previously.

  Let ξ be a partition of X by intervals and $n \in \mathbb{N}$.

  The endpoints of the intervals in the preimages $f^{-i}\xi$, for $i = 0, \ldots, n-1$, determine at most $n\text{card}\xi$ points in S^1.

  It follows that

  $$\text{card}\xi_n \leq n\text{card}\xi.$$

  Note that

  $$-\sum_{C \in \xi_n} \mu(C) \log \mu(C) = \sum_{C \in \xi_n} \varphi(\mu(C)).$$

  where

  $$\varphi(x) = \begin{cases} -x \log x, & \text{if } x \in (0, 1], \\ 0, & \text{if } x = 0. \end{cases}$$

Example: Rotations of the Circle (Cont'd)

- We have $\varphi''(x) = -\frac{1}{x} < 0$, for $x \in (0,1)$.
  So the function φ is strictly concave.
  Thus, we get

$$\begin{array}{rcl} -\sum_{C \in \xi_n} \mu(C) \log \mu(C) & = & \sum_{C \in \xi_n} \frac{1}{\text{card}\xi_n} \varphi(\mu(C)) \text{card}\xi_n \\ & \leq & \varphi\left(\sum_{C \in \xi_n} \frac{\mu(C)}{\text{card}\xi_n}\right) \text{card}\xi_n \\ & = & \varphi\left(\frac{1}{\text{card}\xi_n}\right) \text{card}\xi_n \\ & = & -\log \frac{1}{\text{card}\xi_n} \\ & = & \log \text{card}\xi_n. \end{array}$$

  Hence, it follows that

$$h_\mu(f,\xi) \leq \inf_{n \in \mathbb{N}} \frac{1}{n} \log \text{card}\xi_n \leq \inf_{n \in \mathbb{N}} \frac{1}{n} \log(n\text{card}\xi) = 0.$$

  Thus, $h_\mu(f) = 0$.

Example

- Consider the expanding map $E_2 : S^1 \to S^1$.

  Let μ be the E_2-invariant measure defined previously.

  Consider the partition

  $$\xi^{(m)} = \left\{ \left[\frac{i}{2^m}, \frac{i+1}{2^m} \right] : i = 0, \ldots, 2^m - 1 \right\}.$$

  For each $m, n \in \mathbb{N}$, we obtain

  $$\begin{aligned} \xi_n^{(m)} &= \left\{ \left[\tfrac{i}{2^{m+n-1}}, \tfrac{i+1}{2^{m+n-1}} \right] : i = 0, \ldots, 2^{m+n-1} - 1 \right\} \\ &= \xi^{(m+n-1)}. \end{aligned}$$

Example (Cont'd)

- Now we get

$$\begin{aligned} h_\mu(E_2,\xi^{(m)}) &= \inf_{n\in\mathbb{N}} -\frac{1}{n}\sum_{C\in\xi_n^{(m)}} \mu(C)\log\mu(C) \\ &= \inf_{n\in\mathbb{N}} -\frac{1}{n}\sum_{C\in\xi^{(m+n-1)}} \mu(C)\log\mu(C) \\ &= \inf_{n\in\mathbb{N}} -\frac{1}{n}\cdot 2^{m+n-1}\cdot\frac{1}{2^{m+n-1}}\log\frac{1}{2^{m+n-1}} \\ &= \inf_{n\in\mathbb{N}} \frac{m+n-1}{n}\log 2 \\ &= \log 2. \end{aligned}$$

The partitions $\xi^{(m)}$ satisfy the hypotheses of the definition.

We conclude that

$$h_\mu(E_2) = \sup_{m\in\mathbb{N}} h_\mu(E_2,\xi^{(m)}) = \log 2.$$

Subsection 6

Proof of the Ergodic Theorem

Lemma 1

Lemma

Given a μ-integrable function $\psi : X \to \mathbb{R}$, the function $\psi \circ f$ is also μ-integrable and

$$\int_X (\psi \circ f) d\mu = \int_X \psi d\mu.$$

- Given a set $B \in \mathcal{A}$, f-invariance means

$$\int_X \chi_{f^{-1}B} d\mu = \int_X \chi_B d\mu.$$

Equivalently, since $\chi_B \circ f = \chi_{f^{-1}B}$,

$$\int_X (\chi_B \circ f) d\mu = \int_X \chi_B d\mu.$$

Lemma 1 (Cont'd)

- We obtained
$$\int_X (\chi_B \circ f) d\mu = \int_X \chi_B d\mu.$$
For a simple function
$$s = \sum_{k=1}^{n} a_k \chi_{B_k},$$
it follows that
$$\int_X (s \circ f) d\mu = \int_X s d\mu.$$
In general, we have
$$\psi = \psi^+ - \psi^- \quad \text{and} \quad \psi^+, \psi^- \geq 0.$$
By the definition of Lebesgue integrals, it is sufficient to establish the result for nonnegative functions.

Lemma 1 (Cont'd)

- Let $\psi : X \to \mathbb{R}_0^+$ be a μ-integrable function.

  By the definition of the integral, there exists a sequence of simple functions $(s_n)_{n\in\mathbb{N}}$, such that:

  1. $0 \leq s_n \leq s_{n+1} \leq \psi$ for $n \in \mathbb{N}$, with $\lim_{n\to\infty} s_n(x) = \psi(x)$, for $x \in X$;
  2. $\lim_{n\to\infty} \int_X s_n d\mu = \int_X \psi d\mu$.

  By Fatou's Lemma,

$$\begin{aligned}\int_X \lim_{n\to\infty}(s_n \circ f)d\mu &\leq \liminf_{n\to\infty} \int_X (s_n \circ f)d\mu \\ &= \liminf_{n\to\infty} \int_X s_n d\mu \\ &= \lim_{n\to\infty} \int_X s_n d\mu \\ &= \int_X \psi d\mu \\ &< \infty.\end{aligned}$$

  Hence, the function $\lim_{n\to\infty}(s_n \circ f) = \psi \circ f$ is μ-integrable.

Lemma 1 (Conclusion)

- By construction, $s_n \circ f \nearrow \psi \circ f$ when $n \to \infty$.

  By the Monotone Convergence Theorem, the limit exists,

$$\lim_{n\to\infty} \int_X (s_n \circ f) d\mu = \int_X \lim_{n\to\infty} (s_n \circ f) d\mu = \int_X (\psi \circ f) d\mu.$$

  Finally, we obtain

$$\begin{aligned} \int_X (\psi \circ f) d\mu &= \lim_{n\to\infty} \int_X (s_n \circ f) d\mu \\ &= \lim_{n\to\infty} \int_X s_n d\mu \\ &= \int_X \psi d\mu. \end{aligned}$$

  This completes the proof of the lemma.

Lemma 2

- Now we consider the set

$$A = \left\{ x \in X : \sup_{n \in \mathbb{N}} \sum_{k=0}^{n-1} \varphi(f^k(x)) > 0 \right\}.$$

Lemma

We have

$$\int_A \varphi d\mu \geq 0.$$

- The functions $s_0(x) = 0$ and $s_n(x) = \sum_{k=0}^{n-1} \varphi(f^k(x))$, for $n \in \mathbb{N}$, satisfy the identity

$$s_n(f(x)) = \sum_{k=0}^{n-1} \varphi(f^k(f(x))) = \sum_{k=0}^{n-1} \varphi(f^{k+1}(x)) = s_{n+1}(x) - \varphi(x).$$

Lemma 2 (Cont'd)

- Write $t_n(x) = \max\{s_1(x), \ldots, s_n(x)\}$ and $r_n(x) = \max\{0, t_n(x)\}$. Then we obtain
$$r_n(f(x)) = t_{n+1}(x) - \varphi(x).$$
On the set $A_n = \{x \in X : t_n(x) > 0\}$, we have $t_n(x) = r_n(x)$. Thus,
$$\int_{A_n} t_{n+1} d\mu \geq \int_{A_n} t_n d\mu = \int_{A_n} r_n d\mu = \int_X r_n d\mu.$$
We now have, taking into account Lemma 1,
$$\begin{array}{rcl}\int_{A_n} \varphi d\mu & = & \int_{A_n} t_{n+1} d\mu - \int_{A_n} (r_n \circ f) d\mu \\ & \geq & \int_X r_n d\mu - \int_X (r_n \circ f) d\mu = 0.\end{array}$$
Now we note that $A_n \subseteq A_{n+1}$, for all $n \in \mathbb{N}$, and $\bigcup_{n=1}^{\infty} A_n = A$. Hence, letting $n \to \infty$, we obtain $\int_A \varphi d\mu \geq 0$.

Proof of the Ergodic Theorem

- Given $a, b \in \mathbb{Q}$ with $a < b$, consider the set

$$B = B_{a,b} = \left\{ x \in X : \liminf_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} \varphi(f^k(x)) < a < b < \limsup_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} \varphi(f^k(x)) \right\}.$$

Consider also the function

$$\psi(x) = \begin{cases} \varphi(x) - b, & \text{if } x \in B, \\ 0, & \text{if } x \notin B. \end{cases}$$

Define

$$\begin{aligned} A_\psi &= \{x \in X : \sup_{n \in \mathbb{N}} \tfrac{1}{n} \textstyle\sum_{k=0}^{n-1} \psi(f^k(x)) > 0\} \\ &= \{x \in X : \sup_{n \in \mathbb{N}} \tfrac{1}{n} \textstyle\sum_{k=0}^{n-1} \varphi(f^k(x)) > b\}. \end{aligned}$$

By Lemma 2, $\int_{A_\psi} \psi d\mu \geq 0$.

Proof of the Ergodic Theorem (Cont'd)

- We note that $A_\psi \supseteq B$.
  Since $f^{-1}B = B$, we also have, for $x \notin B$,
  $$\sum_{k=0}^{n-1} \psi(f^k(x)) = 0.$$
  That is
  $$X \backslash B \subseteq X \backslash A_\psi.$$
  This shows that $A_\psi = B$.
  So the inequality $\int_{A_\psi} \psi d\mu \geq 0$ is equivalent to
  $$\int_B \varphi d\mu \geq b\mu(B).$$

Proof of the Ergodic Theorem (Cont'd)

- Similarly, we may consider

$$\overline{\psi}(x) = \begin{cases} a - \varphi(x), & \text{if } x \in B, \\ 0, & \text{if } x \notin B. \end{cases}$$

By similar reasoning, we can show that

$$\int_B \varphi d\mu \leq a\mu(B).$$

Since $a < b$, it follows that $\mu(B_{a,b}) = \mu(B) = 0$.

But the union of the sets $B_{a,b}$, for $a, b \in \mathbb{Q}$, with $a < b$, coincides with the set of points $x \in X$, such that

$$\liminf_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \varphi(f^k(x)) < \limsup_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \varphi(f^k(x)).$$

So $\varphi_f(x) = \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \varphi(f^k(x))$ exists for almost every $x \in X$.

Proof of the Ergodic Theorem (Cont'd)

- It remains to establish the integrability of the function φ_f and the identity
$$\int_X \varphi_f d\mu = \int_X \varphi d\mu.$$
Write $\varphi = \varphi^+ - \varphi^-$.
The functions φ^+ and φ^- are μ-integrable.
By the previous argument, the limits
$$\begin{aligned} \varphi_f^+(x) &= \lim_{n\to\infty} \tfrac{1}{n} \textstyle\sum_{k=0}^{n-1} \varphi^+(f^k(x)), \\ \varphi_f^-(x) &= \lim_{n\to\infty} \tfrac{1}{n} \textstyle\sum_{k=0}^{n-1} \varphi^-(f^k(x)) \end{aligned}$$
exist for almost every $x \in X$.

Proof of the Ergodic Theorem (Cont'd)

- By Fatou's Lemma, together with Lemma 1,

$$\begin{array}{rcl} \int_X \varphi_f^+ d\mu & \leq & \liminf_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} (\varphi^+ \circ f^k) d\mu \\ & = & \liminf_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \varphi^+ d\mu \\ & = & \int_X \varphi^+ d\mu \\ & < & \infty. \end{array}$$

Analogously,

$$\int_X \varphi_f^- d\mu \leq \int_X \varphi^- d\mu < \infty.$$

Thus, the functions φ_f^+ and φ_f^- are μ-integrable.

It follows that φ_f is also μ-integrable.

Proof of the Ergodic Theorem (Cont'd)

- Consider, for each $a, b \in \mathbb{Q}$, with $a < b$, the set

$$D_{a,b} = \{x \in X : a \leq \varphi_f(x) \leq b\}.$$

One can repeat the former argument to show that

$$a\mu(D_{a,b}) \leq \int_{D_{a,b}} \varphi d\mu \leq b\mu(D_{a,b}).$$

We also have

$$a\mu(D_{a,b}) \leq \int_{D_{a,b}} \varphi_f d\mu \leq b\mu(D_{a,b}).$$

Thus,

$$\left| \int_{D_{a,b}} \varphi_f d\mu - \int_{D_{a,b}} \varphi d\mu \right| \leq (b-a)\mu(D_{a,b}).$$

Proof of the Ergodic Theorem (Conclusion)

- We obtained

$$\left|\int_{D_{a,b}} \varphi_f d\mu - \int_{D_{a,b}} \varphi d\mu\right| \leq (b-a)\mu(D_{a,b}).$$

Hence, given $r > 0$, and setting $E_n = D_{nr,(n+1)r}$, we obtain

$$\begin{array}{rcl} \left|\int_X \varphi_f d\mu - \int_X \varphi d\mu\right| & \leq & \sum_{n\in\mathbb{Z}} \left|\int_{E_n} \varphi_f d\mu - \int_{E_n} \varphi d\mu\right| \\ & \leq & \sum_{n\in\mathbb{Z}} r\mu(E_n) \\ & = & r. \end{array}$$

Letting $r \to 0$, we conclude that

$$\int_X \varphi_f d\mu = \int_X \varphi d\mu.$$