Entropy and Dynamics
Introduction to Fractal Geometry and Chaos

Matilde Marcolli

MAT1845HS Winter 2020, University of Toronto
M 5-6 and T 10-12 BA6180

References:

- Yakov Pesin and Vaughn Climenhaga, *Lectures on fractal geometry and dynamical systems*, American Mathematical Society, 2009
- Yitzhak Katznelson and Benjamin Weiss, *A simple proof of some ergodic theorems*, Israel Journal of Math. 42 (1982) N.4, 291–296.
- Yakov Pesin and Howard Weiss, *The multifractal analysis of Birkhoff averages and large deviations*, Global analysis of dynamical systems, 419–431, Inst. Phys., Bristol, 2001.

Equidimensional measures (recall)

- finite measure μ in $\mathbb{R}^N$ support on some $E \subset \mathbb{R}^N$
- pointwise dimension

$$d_\mu(x) = \lim_{r \to 0} \frac{\log \mu(B(x,r))}{\log r}$$

balls of radius r centered at $x \in E$

- the measure μ is exact-dimensional if $\exists \alpha$ such that

$$d_\mu(x) = \alpha, \quad \mu\text{-almost all } x \in E$$

- Hausdorff dimension of the measure $\dim_H(\mu) := \alpha$ in this case
- for μ exact-dimensional and $Z \subset \mathbb{R}^N$ with $\mu(Z) = 1$ have $\dim_H(Z) \geq \dim_H(\mu)$
- example of Cantor set with Bernoulli measures μ_P

$$\dim_H(C) \geq \dim_H(\mu_P) = -\frac{p \log p + (1-p)\log(1-p)}{\log 3}$$

right-hand-side has max at uniform distribution where

$$\dim_H(C) = \dim_H(\mu_{P_{unif}}) = \frac{\log 2}{\log 3}$$

Bowen balls

- dynamical system $f : X \to X$ continuous on metric space (X, d)
- for $x \in X$, $n \in \mathbb{N}$, $\delta > 0$ Bowen ball

$$B_f(x, n, \delta) := \{ y \in X \mid d(f^{\circ j}(x), f^{\circ j}(y)) < \delta \; \forall j = 0, \ldots, n \}$$

- length of (discrete) time during which orbits under iterations of f remain close
- size of Bowen balls $B_f(x, n, \delta)$ shrinks for larger n
- analog of local dimension computation replacing balls by Bowen balls

$$\mu(B(x,r)) \sim r^{d_\mu(x)} \quad \text{and} \quad \mu(B_f(x,n,\delta)) \sim e^{-n\alpha}$$

Note: asymptotic behavior in the length of the orbit (long time) rather than in the size of the ball (small spatial scale)

Local Entropy

- local entropy (when limits exist)

$$h_{\mu,f}(x) := \lim_{\delta\to 0}\lim_{n\to\infty} -\frac{1}{n}\log \mu(B_f(x,n,\delta))$$

- Case of Cantor sets with Bernoulli measures and shift map
  metric $d_a(x,y) = \sum_\ell a^{-\ell}|x_\ell - y_\ell|$

$$B(w,r) = \mathcal{C}(w_1,\ldots,w_n) = B_\sigma(x,n,\delta)$$

  with $\delta = 1/a$, cylinder sets; uniform Bernoulli measure $p = 1/2 = 1-p$ has
  $\mu(B_\sigma(x,n,\delta)) = \mu(\mathcal{C}(w_1,\ldots,w_n)) = 2^{-n} = e^{-n\log 2}$

Bowen balls and Entropy on shift spaces (Σ_k^+, σ)

- shift space Σ_k^+ with Bernoulli measure $P = (p_1, \ldots, p_k)$
- choose $\delta > 0$ such that $a^{-N} \leq \delta < a^{-(N-1)}$ (because of metric)
- then have $B_\sigma(w, n, \delta) = \mathcal{C}(w_1, \ldots, w_{n+N})$ cylinder set

$$B_\sigma(w, n, \delta) = \{x \in \Sigma_k^+ \,|, d_a(\sigma^{\circ j}(x), \sigma^{\circ j}(w)) < \delta, \, \forall j \leq n\}$$

- measure of Bowen balls

$$\mu(B_\sigma(w, n, \delta)) = \mu(\mathcal{C}(w_1, \ldots, w_{n+N})) = p_1^{a^1_{n+N}(w)} \cdots p_k^{a^k_{n+N}(w)}$$

with $a^i_m(w) =$ number of occurrences of digit i among the first m letters of the word w

- then for entropy

$$-\frac{1}{n} \log \mu(B_\sigma(w, n, \delta)) = -\sum_{\ell=1}^{k} \frac{a^\ell_{n+N}(w)}{n} \log p_i$$

- as in the case of binary shift, almost everywhere limit for these

$$\lim_{n\to\infty} \frac{a_n^\ell(w)}{n} =^{\mu-a.e.} p_\ell$$

- then local entropy is almost everywhere

$$h_{\mu,\sigma}(w) = \lim_{\delta\to 0} \lim_{n\to\infty} -\sum_{\ell=1}^{k} \frac{a_{n+N}^\ell(w)}{n+N} \frac{N+n}{n} \log p_i = -\sum_\ell p_\ell \log p_\ell$$

  but it does not have this value everywhere: exceptional sets
- coding map from shift space to Cantor set: can transfer this computation of entropy? invariance under topological conjugacy

Invariance of local entropy under topological conjugacy

- topological conjugacy
  - compact metric spaces (X, d_X) and (Y, d_Y)
  - continuous maps $f : X \to X$ and $g : Y \to Y$
  - homeomorphism $\phi : X \to Y$ that conjugates them $g \circ \phi = \phi \circ f$
- invariant measures and local entropy
  - μ a g-invariant measure on Y: pullback $\mu^*(E) := \mu(\phi(E))$ to an f-invariant measure μ^* on X
  - local entropies satisfy

$$h_{\mu^*,f}(x) = h_{\mu,g}(\phi(x))$$

- invariance under topological conjugacy
    - compact spaces $\Rightarrow$ continuity is uniform continuity

$$\forall \epsilon > 0 \, \exists \delta = \delta(\epsilon) > 0 \quad d(x, x') < \delta \Rightarrow d(\phi(x), \phi(x')) < \epsilon$$

$$\phi(B(x, \delta)) \subset B(\phi(x), \epsilon)$$

    - this holds uniformly for all iterates $f^n(x), g^n(\phi(x))$ as well

$$\phi(B_f(x, n\delta)) \subset B_g(\phi(x), n, \epsilon)$$

    - thus measures

$$\mu^*(B_f(x, n\delta)) \leq \mu(B_g(\phi(x), n, \epsilon))$$

    - limit then gives

$$\lim_{n\to\infty} -\frac{1}{n} \log \mu^*(B_f(x, n\delta)) \geq \lim_{n\to\infty} -\frac{1}{n} \log \mu(B_g(\phi(x), n, \epsilon))$$

    - taking limit for $\epsilon \to 0$ implies also $\delta \to 0$ then

$$h_{\mu^*, f}(x) \geq h_{\mu, g}(\phi(x))$$

    - for opposite inequality $\leq$ use ϕ^{-1} and same argument

- Ergodic measures
    - dynamical system $f : X \to X$ and an f-invariant probability measure μ on X: the measure μ is ergodic for f if

$$\forall E \subset X : \quad f^{-1}(E) = E \quad \Rightarrow \quad \mu(E) = 0 \quad \text{or} \quad \mu(E) = 1$$

    - equivalently measurable functions h with $h \circ f = h$ are μ-almost everywhere constant

- Kolmogorov-Sinai Entropy
    - local entropy depends on the point x, but if the measure μ is ergodic then almost everywhere constant $h(\mu, f)$: this constant is called Kolmogorov-Sinai Entropy

- General idea: replacing the "static" metric balls $B(x, r)$ with the "dynamical" Bowen balls $B_f(x, n, \delta)$; the first give at the local level pointwise dimension, the second give pointwise entropy; globally the first give $\dim_H(\mu)$ almost everywhere, the second the Kolmogorov-Sinai Entropy $h(\mu, f)$

- equivalent property for ergodic measures:
  - $A \subset X$ measurable set not necessarily f-invariant

$$N(x, A, n) := \#\{k \in \{0, \ldots, n\} \, : \, f^k(x) \in A\}$$

  - ergodicity of μ equivalent to

$$\lim_{N \to \infty} \frac{N(x, A, n)}{n} = \mu(A)$$

  - this measures how often the orbit of x passes through the set A, for ergodic measure this is proportional to the size of A: the dynamics visits all subsets with frequency proportional to their size
  - ergodic measures detect how "mixing" a dynamical system $f : X \to X$ is

Topological Entropy: in this analogy between dimension (static metric balls) and entropy (dynamical Bowen balls) what is the analog of the relation between Hausdorff dimension and topological dimension

- Topological Entropy
  - (X, d) metric space $f : X \to X$ dynamical system, $Z \subset X$, for all $N \in \mathbb{N}$ and $\delta > 0$ set $\mathcal{P}(Z, N, \delta)$ of all (countable) coverings by Bowen balls $B_f(x, n, \delta)$ with $x \in Z$ and $n \geq N$

$$m_h(Z, \delta, s) = \lim_{N\to\infty} \inf_{\mathcal{U}\in\mathcal{P}(Z,N,\delta)} \sum_{U_i\in\mathcal{U}} e^{-n_i \alpha s}$$

  - similar argument shows it behaves like the Hausdorff measure: jump from ∞ to 0
  - $h_{top}(Z, f) := \lim_{\delta\to 0} h_{top}(Z, f, \delta)$

$$h_{top}(Z, f, \delta) := \sup\{s \in \mathbb{R}_+ \,|\, m_h(Z, \delta, s) = \infty\}$$

$$= \inf\{s \in \mathbb{R}_+ \,|\, m_h(Z, \delta, s) = 0\}$$

    monotonicity shows limit exists
  - $h_{top}(Z, f)$ also invariant under topological conjugacy

- Topological and Kolmogorov–Sinai entropy
  - X compact metric space $f : X \to X$ dynamical system

$$h_{top}(Z, f) = \sup\{h(\mu, f) \,|\, \mu \text{ ergodic invariant measure } \}$$

- Example: Shift spaces Σ_k^+ with shift map

$$h_{top}(\Sigma_k^+, \sigma) = \log k$$

  - maximum of the Shannon entropy over probabilities P, hence of the entropies of μ_P Bernoulli measures
  - for Cantor set obtain

$$\dim_H(C) = \frac{\log 2}{\log 3} = \frac{h_{top}(C, f)}{\log 3}$$

for f topologically conjugate to the shift map $\sigma : \Sigma_2^+ \to \Sigma_2^+$

Entropy for Markov Measures

- $\Sigma_A^+ \subset \Sigma_k^+$ subshift of finite type with admissibility matrix A
- Markov measure $\mu_{P,\pi}$ with support on Σ_A^+ with data $\pi = (\pi_i)_{i=1}^k$ and $P = (p_{ij})_{i,j=1}^k$ probability and stochastic matrix with $\pi P = \pi$
- Bowen balls for the shift map are cylinder sets

$$B_\sigma(w, n, \delta) = \mathcal{C}(w_1, \ldots, w_{n+N}), \qquad \text{for } a^{-N} \leq \delta < a^{-(N-1)}$$

- Markov measure of the Bowen balls

$$\mu_{P,\pi}(B_\sigma(w, n, \delta)) = \pi_{w_1} p_{w_1 w_2} \cdots p_{w_{n+N-1} w_{n+N}}$$

$$= \pi_{w_1} \prod_{i=1}^k \prod_{j=1}^k p_{ij}^{a_{n+N}^{ij}(w)}$$

with $a_m^{ij}(w)$ = number of indices $m' < m$ with $w_{m'} = i$ and $w_{m'+1} = j$

- under the assumption that A is primitive (A^m is positive for some $m \in \mathbb{N}$)

$$\lim_{m\to\infty} \frac{a_m^{ij}(w)}{m} \overset{\mu_{P,\pi}-a.e.}{=} \pi_i p_{ij}$$

- then entropy

$$h_{\mu_{P,\pi},\sigma}(w) \overset{\mu_{P,\pi}-a.e.}{=} -\sum_{i=1}^{k} \pi_i \sum_{j=1}^{k} p_{ij} \log p_{ij}$$

equal to Kolmogorov-Sinai Entropy $h(\mu_{P,\pi}, \sigma)$

Measures of maximal entropy for subshifts

- Parry measure
  - $\Sigma_A^+ \subset \Sigma_k^+$ subshift of finite type with admissibility matrix $A = (a_{ij})$ with $a_{ij} \in \{0, 1\}$
  - χ largest positive eigenvalue of A (Perron–Frobenius)
  - $u = (u_1, \dots, u_k)$ left eigenvector and $v = (v_1, \dots, v_k)^t$ right eigenvector of A with χ eigenvalue
  - both have positive entries and normalization $\sum_i u_i v_i = 1$
  - take $\mu_{P,\pi}$ Markov measure with $\pi_i = u_i v_i$ and $p_{ij} = \chi^{-1} a_{ij} \frac{v_i}{v_j}$
  - measure of cylinder sets: if $(w_1, \dots, w_n)$ admissible for A

$$\mu_{P,\pi}(\mathcal{C}(w_1, \dots, w_n)) = \pi_{w_1} p_{w_1 w_2} \cdots p_{w_{n-1} w_n} = \chi^{-n} u_{w_1} v_{w_n}$$

- Entropy of the Parry measure (shift map)
    - cylinder sets = Bowen balls
    - local entropy (constant for all $w \in \Sigma_A^+$):

$$h_{\mu_{P,\pi},\sigma}(w) = \lim_{n\to\infty} -\frac{1}{n} \log \mu_{P,\pi}(\mathcal{C}(w_1, \ldots, w_n)) =$$

$$\lim_{n\to\infty} -\frac{1}{n}(-n \log \chi + \log u_{w_1} + \log v_{w_n}) = \log \chi$$

    - topological entropy $h_{top}(\Sigma_A^+) = \log \chi$ Perron-Frobenius eigenvalue
    - Parry measure has max entropy

- piecewise linear Markov maps
  - uniform contraction ratio λ for construction of Cantor set C_A in $[0,1]$ associated to Σ_A^+
  - f piecewise linear Markov map $|f'(x)| = \lambda^{-1}$
  - $A =$ transition matrix for f

$$a_{ij} = \begin{cases} 1 & f(\mathcal{I}_i) \cap \mathcal{I}_j \neq \emptyset \\ 0 & \text{otherwise} \end{cases}$$

  - C_A domain of all iterates of f
  - χ Perron-Frobenius eigenvalue of A (largest positive eigenvalue)
  - then Hausdorff dimension

$$\dim_H(C_A) = -\frac{\log \chi}{\log \lambda}$$

  - because $|f'(x)| = \lambda^{-1}$ constant, Bowen balls are ordinary metric balls (hence Hausdorff dim as entropy up to a constant factor $\log \lambda$)

$$B_f(x, n, \delta) = B(x, \delta\lambda^{-n})$$

What when $|f'(x)|$ not constant?

- entropy functions $h_{\nu,f}(x)$, $h(\mu, f)$, $h_{top}(Z, f)$ invariant under topological conjugacy but dimensions $d_\mu(x)$, $\dim_H(\mu)$, $\dim_H(Z)$ depend on the metric structure
- what happens when the factor $\log|f'(x)|$ is not constant?
- compare Bowen balls $B_f(x, n, \delta)$ with metric balls $B(x, r)$
- how fast orbits of nearby points diverge under iterates of f

$$f(y) = f(x) + f'(x)(y - x) + o(y - x)$$

$$d(f(y), f(x)) = |f(x) - f(y)| \sim |f'(x)|\, d(x, y)$$

up to higher order terms in $d(x, y)$

$$d(f^2(x), f^2(y)) \sim |f'(f(x))|\,|f'(x)|\, d(x, y)$$

$$d(f^{\circ n}(x), f^{\circ n}(y)) \sim \prod_{i=0}^{n-1} |f'(f^i(x))|\, d(x, y)$$

Lyapunov exponent

- asymptotic behavior:

$$\lambda_f(x) := \lim_{n\to\infty} \frac{1}{n} \log \prod_{i=0}^{n-1} |f'(f^i(x))|$$

  when the limit exists

- asymptotic rate of expansion: how quickly distance between points grows under iterates

Example

- $f : \mathcal{I}_1 \cup \mathcal{I}_2 \to [0,1]$ piecewise linear with $\ell(\mathcal{I}_1) = \lambda_1$ and $\ell(\mathcal{I}_2) = \lambda_2$
- Bernoulli measure μ_P with $P = (p, q = 1 - p)$ on Σ_2^+
- Lyapunov exponents exists μ_P-almost everywhere

$$\lambda_f(x) \stackrel{\mu_P - a.e.}{=} -(p \log \lambda_1 + (1-p) \log \lambda_2)$$

- check by considering intervals

$$\mathcal{I}_{w_1,\dots,w_n} = \{x \,|\, f^{j-1}(x) \in \mathcal{I}_{w_j}, \, j = 1, \dots, n\}$$

$$x \in \mathcal{I}_{w_1,\dots,w_n} \;\Rightarrow\; d_n(x) := \prod_{i=0}^{n-1} |f'(f^i(x))| =$$

$$\prod_{j=1}^{n} \lambda_{w_j}^{-1} = \lambda_1^{-a_n(w)} \lambda_2^{-(n-a_n(w))}$$

- know that μ_P-almost everywhere

$$\lim_{n\to\infty} \frac{a_n(w)}{n} \stackrel{\mu_P - a.e.}{=} p$$

- then Lyapunov exponent

$$\lambda_f(x) = \lim_{n\to\infty} \frac{1}{n} \log d_n(x)$$

$$= \lim_{n\to\infty} \frac{1}{n} \log(\lambda_1^{-a_n(w)} \lambda_2^{-(n-a_n(w))})$$

$$= -\lim_{n\to\infty} \frac{a_n(w)}{n} \log \lambda_1 + \frac{n - a_n(w)}{n} \log \lambda_2$$

$$\stackrel{\mu_P - a.e.}{=} -(p \log \lambda_1 + (1-p) \log \lambda_2)$$

- Ergodicity
  - μ ergodic measure for f then $\lambda_f(x)$ exists and is μ-almost everywhere constant
  - value μ-almost everywhere: Lyapunov exponent of μ $\lambda(\mu, f)$
- Pointwise dimension of Bernoulli measures
  - as before

$$\log \ell(\mathcal{I}_{w_1,\dots,w_n}) = \log(\lambda_1^{-a_n(w)} \lambda_2^{-(n-a_n(w))})$$
$$= a_n(w) \log \lambda_1 + (n - a_n(w)) \log \lambda_2$$

  - then pointwise dimension almost everywhere

$$d_\mu(x) = \frac{p \log p + (1-p) \log(1-p)}{p \log \lambda_1 + (1-p) \log \lambda_2}$$

  - Hausdorff dimension of the measure

$$\dim_H(\mu) = \frac{h(\mu, f)}{\lambda(\mu, f)}$$

entropy divided by Lyapunov exponent

- relation of metric and Bowen balls

$$B_f(x, n, \delta) \sim B(x, \delta e^{-n\lambda_f(x)})$$

- pointwise dimension, local entropy, Laypunov exponent

$$d_\mu(x) = \frac{h_{\mu,f}(x)}{\lambda_f(x)}$$

by the relation of metric and Bowen balls

- Hausdorff dimension estimate

$$\dim_H(\mu) = \frac{h(\mu, f)}{\lambda(\mu, f)} \Rightarrow \dim_H(C) \geq \frac{h(\mu, f)}{\lambda(\mu, f)}$$

can optimize by searching for measures of maximal entropy

Multifractal decomposition

- Cantor set C

$$\dim_H(C) = s = \sup_{p\in[0,1]} \phi(p)$$

$$\phi(p) = \frac{h(\mu_P, f)}{\lambda(\mu_P, f)} = \frac{p\log p + (1-p)\log(1-p)}{p\log\lambda_1 + (1-p)\log\lambda_2}$$

- for each $p \in [0,1]$ subset $C_p \subset C$ of full measure for μ_P
- decomposition

$$C = \cup_{p\in[0,1]} C_p \cup C'$$

with C' = exceptional set where limit may not exist

- level sets of local dimension: everywhere on C_p

$$d_{\mu_P}(x) = \frac{h(\mu_P, f)}{\lambda(\mu_P, f)}$$

- $\phi'(p) = 0$ max at $p = \lambda_1^s$ with s = self-similairty $\lambda_1^s + \lambda_2^s = 1$

Birkhoff Ergodic Theorem

- (X, Σ, μ) probability measure space, $T : X \to X$ dynamical system, measure preserving
- for any integrable function $f \in L^1(X, \mu)$ limit μ-almost everywhere exists

$$f^*(x) := \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} f(T^j(x))$$

and f^* is a T-invariant measurable function with $\int f^*(x) d\mu(x) = \int f(x) d\mu(x)$

- in particular, if μ is ergodic for T then measurable T-invariant functions are constant μ-almost everywhere so

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} f(T^j(x)) \stackrel{\mu-a.e.}{=} \int_X f(x) d\mu(x)$$

temporal average equals spatial average

- can show for non-negative functions, more general by linearity
- take liminf and limsup

$$\underline{f}(x) := \liminf_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} f(T^j(x)), \quad \overline{f}(x) := \limsup_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} f(T^j(x))$$

  both are T-invariant by construction; show that

$$\int_X \overline{f}(x) d\mu(x) \leq \int_X f(x) d\mu(x) \leq \int_X \underline{f}(x) d\mu(x)$$

- this implies equality $\overline{f}(x) = \underline{f}(x)$ holds μ-almost everywhere (if it fails on a positive measure set cannot have first $\leq$ last above) and it also gives $\int f^*(x) d\mu(x) = \int f(x) d\mu(x)$

- fix some $M > 0$ and some $\epsilon > 0$ and take

$$\overline{f}_M(x) := \min\{\overline{f}(x), M\}$$

- let $n(x)$ be least integer in $\mathbb{N}$ such that

$$\overline{f}_M(x) \leq \frac{1}{n} \sum_{j=0}^{n-1} f(T^j(x)) + \epsilon$$

- $\overline{f}_M$ is also T-invariant so average

$$\sum_{j=0}^{n(x)-1} \overline{f}_M(T^j(x)) \leq \sum_{j=0}^{n(x)-1} f(T^j(x)) + n(x)\epsilon$$

- $n(x)$ is everywhere finite, so for some $N > 0$ the set $A = \{x \mid n(x) > N\}$ has μ-measure less than ϵ/M

- define the functions

$$\tilde{f}(x) := \begin{cases} f(x) & x \notin A \\ \max\{f(x), M\} & x \in A \end{cases} \qquad \tilde{n}(x) := \begin{cases} n(x) & x \notin A \\ 1 & x \in A \end{cases}$$

- now the advantage is that $\tilde{n}(x)$ is everywhere bounded by same N
- still have inequality

$$\sum_{j=0}^{n(x)-1} \overline{f}_M(T^j(x)) \leq \sum_{j=0}^{n(x)-1} \tilde{f}(T^j(x)) + \tilde{n}(x)\epsilon$$

- also have inequality (by $\mu(A) \leq \epsilon/M$)

$$\int \tilde{f}(x) d\mu(x) \leq \int f(x) d\mu(x) + \int_A M \, d\mu(x) \leq \int f(x) d\mu(x) + \epsilon$$

- choose L such that $NM/L < \epsilon$ and define inductively:

$$n_0(x) = 0, \qquad n_k(x) = n_{k-1}(x) + \tilde{n}(T^{n_{k-1}(x)}(x))$$

- for $k(x)$ maximal k for which $n_k(x) \leq L - 1$

$$\sum_{j=0}^{L-1} \overline{f}_M(T^j(x)) = \sum_{k=1}^{k(x)} \sum_{j=n_{k-1}(x)}^{n_k(x)-1} \overline{f}_M(T^j(x)) + \sum_{j=n_{k(x)}(x)}^{L-1} \overline{f}_M(T^j(x))$$

- apply to each of the $k(x)$ terms the estimate

$$\sum_{j=0}^{n(x)-1} \overline{f}_M(T^j(x)) \leq \sum_{j=0}^{n(x)-1} \tilde{f}(T^j(x)) + \tilde{n}(x)\epsilon$$

and estimate by M the last $L - n_{k(x)}(x) \leq N - 1$ terms: obtain (using non-negative function to sum to $L - 1$)

$$\sum_{j=0}^{L-1} \overline{f}_M(T^j(x)) \leq \sum_{j=0}^{L-1} \tilde{f}(T^j(x)) + L\epsilon + (N-1)M$$

- integrate and divide by L to get

$$\int_X \overline{f}_N(x) d\mu(x) \leq \int_X \tilde{f}(x) d\mu(x) + \epsilon + \frac{(N-1)M}{L}$$

$$\leq \int_X f(x) d\mu(x) + 3\epsilon$$

- so for $M \to \infty$ and $\epsilon \to 0$ get one side of inequality

$$\int \overline{f}(x) d\mu(x) \leq \int_X f(x) d\mu(x)$$

- to get other side fix $\epsilon > 0$ and take $n(x)$ now least integer with

$$\frac{1}{n} \sum_{j=0}^{n-1} f(T^j(x)) \leq \underline{f}(x) + \epsilon$$

- as before take set $A = \{x \mid n(x) > N\}$ where N is such that $\int_A f(x) d\mu(x) < \epsilon$
- define functions

$$\tilde{n}(x) := \begin{cases} n(x) & x \notin A \\ 1 & x \in A \end{cases} \qquad \tilde{f}(x) := \begin{cases} f(x) & x \notin A \\ 0 & x \in A \end{cases}$$

- then same kind of proof as before works for the inequality

$$\int_X f(x) d\mu(x) \leq \int_X \underline{f}(x) d\mu(x)$$

Ergodicity of Bernoulli measures for the shift map $\sigma : \Sigma_k^+ \to \Sigma_k^+$

- any μ_P-measurable set A approximated by a finite union of cylinders $C_1, \ldots, C_N$

$$\mu(A \triangle E) < \epsilon, \quad E = \cup_{i=1}^N C_i$$

- for sufficiently large n the rectangles $F := \sigma^{-n}(E)$ depend on different coordinates

$$\mu_P(E \cap F) = \mu_P(E)\mu_P(F) = \mu_P(E)\mu_P(\sigma^{-n}(E)) = \mu_P(E)^2$$

- suppose $A = \sigma^{-1}(A)$ want to check $\mu_P(A) = 0$ or $\mu_P(A) = 1$: have

$$\mu(A \triangle F) = \mu(\sigma^{-n}(A) \triangle \sigma^{-n}(F)) = \mu(\sigma^{-n}(A \triangle F)) = \mu(A \triangle E) < \epsilon$$

- estimate then using $\mu(A \triangle F) < \epsilon$

$$|\mu(A) - \mu(A)^2| \leq |\mu(A) - \mu(E \cap F)| + |\mu(E \cap F) - \mu(A)^2|$$

$$\leq \mu(A \triangle (E \cap F)) + |\mu(E)^2 - \mu(A)^2|$$

$$\leq \mu(A \triangle E) + \mu(A \triangle F) + |\mu(E) - \mu(A)|\,|\mu(E) + \mu(A)| < 4\epsilon$$

- arbitrary $\epsilon > 0$ so $\mu(A) = \mu(A)^2$ either 0 or 1

• Markov measures $\mu_{P,\pi}$ also ergodic for $\sigma : \Sigma_A^+ \to \Sigma_A^+$ subshift of finite type with $A = (a_{ij})$ admissibility matrix of stochastic matrix $P = (p_{ij})$

Consequences of the ergodic theorem

- the stated characterization of ergodicity: proportion of time spent by orbits in a set A is equal to the mass of A

$$N(x, A, n) := \#\{0 \leq k \leq n \,|\, T^k(x) \in A\}$$

$$\lim_{n\to\infty} \frac{N(x, A, n)}{n} \stackrel{\mu - ae}{=} \mu(A)$$

  ergodic theorem applied to the characteristic function $f = \chi_A$

- law of large numbers and Bernoulli measure (more complicated proof earlier)

$$\lim_{n\to\infty} \frac{a_n(w)}{n} \stackrel{\mu_P - ae}{=} p = \mu(\mathcal{C}_1)$$

  from previous using $a_n(w) = N(w, \mathcal{C}_1, n)$

- Markov measures: $a_n^{i,j}(w)$ = number of indices $\ell < n$ such that $w_\ell = i$ and $w_{\ell+1} = j$; shift map, function $f = \chi_{\mathcal{C}_{ij}}$ characteristic function of cylinder set

$$\lim_{n\to\infty} \frac{a_n^{i,j}(w)}{n} = \lim_{n\to\infty} \frac{1}{n} \sum_{\ell=0}^{n-1} \chi_{\mathcal{C}_{ij}}(\sigma^\ell(w)) = \mu(\mathcal{C}_{ij}) = \pi_i \pi_{ij}$$

- behaviour outside of the full measure set where limit of the Birkhoff average is equal to $\alpha_\mu(f) = \int_X f(x) d\mu(x)$:

$$X = \{x \,|\, f^*(x) = \alpha_\mu(f)\} \cup \bigcup_{\alpha \neq \alpha_\mu(f)} \{x \,|\, f^*(x) = \alpha\} \cup \{x \,|\, \text{ no limit } \}$$

$B_\alpha = \{x \,|\, f^*(x) = \alpha\}$ level sets, last term exceptional set

Birkhoff spectrum: multifractal decomposition (for $\sigma : \Sigma_A^+ \to \Sigma_A^+$)

- Birkhoff spectrum: $b_f(\alpha) := \dim_H(B_\alpha)$
- μ_{max} measure of maximal entropy
- μ_f equilibrium measure: shift invariant probability measure that achieves maximum for pressure functional

$$P(f) := \sup_{\mu}\{h_{\sigma,\mu} + \int_X f(x) d\mu(x)\}$$

Kolmogorov–Sinai entropy $h_{\sigma,\mu}$

- $\exists$ interval $[a, b]$ such that:
  1. if $\mu_f \neq \mu_{max}$ then $b_f(\alpha)$ real analytic and strictly convex on (a, b)
  2. for all $\alpha \in [a, b]$ level set B_α uncountable dense subset of Σ_A^+
  3. interval $[a, b]$ maximal: no values of f^* outside
  4. if $\mu_f \neq \mu_{max}$ exceptional set has maximal Hausdorff dimension $= \dim_H \Sigma_A^+$
- level sets B_α complicated but Birkhoff spectrum $b_f(\alpha)$ smooth and convex
- B_α negligible in measure but large topologically
- exceptional set large in dimension but negligible in measure
- proof for case of subshifts of finite type in
  - Yakov Pesin and Howard Weiss, *The multifractal analysis of Birkhoff averages and large deviations*, Global analysis of dynamical systems, 419–431, Inst. Phys., Bristol, 2001.