Entropy and Expansive Map

Elmar Lasa(elasa@clemson.edu)

Our references include Michael Brin's Introduction to Dynamical System ([2]), James Brown Ergodic Theory and Topological dynamics [3] and Goodman's article [1].

1 Positioning the problem

Let (X, d) be a compact metric space. A homeomorphism $f : X \to X$ is *expansive* if there exist $\delta_0 > 0$ such that for any two distinct points $x \neq y$ and for some $n \in \mathbb{Z}$, $d\left(f^n\left(x\right), f^n\left(y\right)\right) \geqslant \delta_0$. Given a dynamical system, there are two notions of entropy: topological entropy and measure-theoretic entropy also called (we shall call) *metric entropy*. It is well known (cf.[1]) that for Hausdorff topological space, therefore for all metric spaces, the metric entropy is less than the topological entropy. Using variational principle, it is also shown that the topological entropy of a homeomorphism on a compact metric space is the suppremum of its metric entropies over all Borel probability measures. We shall prove in this work that this suppremum is actually attained for expansive maps. In this case, f is said to have a *measure of maximal entropy*.

Metric Entropy

$(X, \mathcal{B}, \mu)$ is Borel space such that $\mu\left[X\right] = 1$.

1.1 Entropy for a partition

Let $\zeta = \{C_i : 1 \leqslant i \leqslant m\}$ for $m \in \mathbb{N}$ be a partition for X of essentially disjoint subsets of X, that is for all $i \neq j$, we have $\mu\left[C_i \cap C_j\right] = 0$. We have to set up some vocabularies[2].

- *Refinement*. The partition ζ' is said to be a refinement of ζ, denoted $\zeta' \geqslant \zeta$ if $\forall C \in \zeta\, \exists C' \in \zeta' \left(C' \subseteq C \left(\mathrm{mod} 0\right)\right)$, that is every element in ζ is essentially contained in an element in ζ';
- *Equivalence*. Two partitions are said to be equivalent if they are refinement of each other;
- *Independence*. Two partitions ζ and ζ' are independent if $\mu\left[C \cap C'\right] = \mu\left[C\right] \cdot \mu\left[C'\right]$ whenever $C \in \zeta$ and $C' \in \zeta'$;
- If $T : X \to X$ is a transformation, we shall write $T^{-1}\left(X\right) = \{T^{-1}\left(C_i\right) : 1 \leqslant i \leqslant m\}$.

We can define a metric on the space of partitions we shall denote $\wp(X)$. Define $d : \wp(X) \to \mathbb{R}^+$ by $d(\zeta,\zeta') = \min_{\sigma\in S_m}\sum_{i=1}^m \mu\left[C_i\Delta C'_{\sigma(i)}\right]$ where Δ denotes symmteric difference, and $\zeta = \{C_i : 1 \leqslant i \leqslant m\}$, $\eta = \{D_i : 1 \leqslant i \leqslant m\}$, and finally S_m is the group of permution of $\{1,\dots,m\}$. Indeed, d is a distance. (i) Given $\zeta = \{C_i : 1 \leqslant i \leqslant m\}$ and $\zeta' = \{C'_i : 1 \leqslant i \leqslant m\}$, if $d(\zeta,\zeta') = 0$ then $\sum_{i=1}^m \mu\left[C_i\Delta C'_{\sigma_0(i)}\right] = 0$ for some $\sigma_0 \in S_m$. By positivity, this must imply that $\mu\left[C_i\Delta D_{\sigma_0(i)}\right] = 0$ for all $1 \leqslant i \leqslant m$. This means that for each i, there exists $j = \sigma_0(i)$ such that C_i is essentially subset an element in ζ', namely C'_j. Similarly, given i, there exists $j = \sigma_0^{-1}(i)$ such that C'_i is essentially subset of C_j. Hence ζ and ζ' are equivalent. Now, let us prove the triangle inequality. Given three partitions ζ, ζ' and ζ'' whose elements are denoted C_i, C'_i, C''_i respectively we have for some $\sigma_0 \in S_m$

$$d(\zeta,\zeta') = \min_{\sigma\in S_m}\sum_{i=1}^m \mu\left[C_i\Delta C'_{\sigma(i)}\right] = \sum_{i=1}^m \mu\left[C_i\Delta C'_{\sigma_0(i)}\right]$$

Given $\sigma,\sigma' \in S_m$, we have $\sum_{i=1}^m \mu\left[C_i\Delta C'_{\sigma(i)}\right] = \sum_{i=1}^m\left(\mu[C_i] + \mu\left[C'_{\sigma(i)}\right] - 2\mu\left[C_i\cap C'_{\sigma(i)}\right]\right) = 2\left(1-\sum_{i=1}^m \mu\left[C_i\cap C'_{\sigma(i)}\right]\right)$ and $\sum_{i=1}^m \mu\left[C'_i\Delta C''_{\sigma'(i)}\right] = 2\left(1-\sum_{i=1}^m \mu\left[C'_i\cap C''_{\sigma'(i)}\right]\right)$. Let i_0 and $\ell_0 = \sigma(i_0)$ be two distinct indices such that $\mu\left[C_i\cap C'_{\sigma(i)}\right] \leqslant \mu\left[P_i\cap C'_{\ell_0}\right]$ and $\mu\left[C'_i\cap C''_{\sigma'(i)}\right] \leqslant \mu\left[C'_{i_0}\cap Q_i\right]$ for all i and for some partitions $\{P_i : 1 \leqslant i \leqslant m\}$ and $\{Q_i : 1 \leqslant i \leqslant m\}$. Then, with $\sum_j \mu\left[C_j\Delta C''_{\tau(j)}\right] = 2\left(1-\sum_j \mu\left[C_j\cap C''_{\tau(j)}\right]\right)$, this gives

$$\begin{aligned}\sum_i \mu\left[C_i\Delta C'_{\sigma(i)}\right]+\sum_i \mu\left[C'_i\Delta C''_{\sigma'(i)}\right]-\sum_i \mu\left[C_i\Delta C''_{\tau(i)}\right] &\geqslant 2\Bigg(1-\sum_i \mu\left[P_i\cap C'_{\ell_0}\right]-\sum_i \mu\left[C'_{i_0}\cap Q_i\right]\\ &\qquad+\sum_j \mu\left[C_j\cap C''_{\tau(j)}\right]\Bigg)\\ &= 2\left(1-\mu\left[C'_{\ell_0}\right]-\mu\left[C'_{i_0}\right]+\sum_j \mu\left[C_j\cap C''_{\tau(j)}\right]\right) \geqslant 0\end{aligned}$$

The last inequality is from the fact that $\zeta' = \{C'_i\}_{1\leqslant i\leqslant m}$ is a partition of X, implying that $\mu\left[C'_{i_0}\right] + \mu\left[C'_{\ell_0}\right] \leqslant \sum_i \mu\left[C'_i\right] = 1$.

Since this is true for all σ, σ' and τ, we can take the infimums which are attained at some σ_0, σ'_0 and τ_0 therefore establishing the triangle inequality.

What motivates the definition of metric entropy follows from the Bernouilli automorphism of the shift space Σ_m. The metric entropy for a partition $\zeta = \{C_1,\dots,C_m\}$ is defined by

$$H(\zeta) = -\sum_{i=1}^m \mu[C_i]\log\mu[C_i]$$

If for $x \in X$ we consider $m(x,\zeta)$ the element C_x of ζ containing x, then $H(\zeta) = -\sum_{i=1}^{m} \int_{x \in C_i} \log \mu[C_i]\, d\mu = -\sum_{i=1}^{m} \int_{x \in C_i} \log m(x,\zeta)\, d\mu = -\int_X \log m(x,\zeta)\, d\mu.$

Proposition. *[2] Let ξ and η be finite partitions. Then*

(i) $H(\xi) \geqslant 0$, $H(\xi)$ if and only if ξ is the trivial partition;

(ii) if $\xi \leqslant \eta$ then $H(\xi) \leqslant H(\eta)$ and equality holds if and only if ξ and η are equivalent;

(iii) If ξ has n elements, then $H(\xi) \leqslant \log n$, and equality holds if and only if each element has measure $\frac{1}{n}$;

(iv) $H(\xi \vee \eta) \leqslant H(\xi) + H(\eta)$ with equality if and only if $\xi \perp \eta$.

Proof. We only prove the first statement, the rest is found in [2]. Let $\xi = \{C_i : 1 \leqslant i \leqslant m\}$ be a partition for X. It is obvious that $H(\xi) \geqslant 0$. If $H(\xi) = 0$ then $\int_X -\log m(x,\xi)\, d\mu = 0$ with $-\log m(x,\xi) \geqslant 0$. Since $\mu[X] = 1$, it must be true that $-\log m(x,\xi) = 0$ for all $x \in X$, then $m(x,\xi) = 1$ that is, any element of ξ contains essentially the whole set X. In other words, $\nu := \{X\}$ is a refinement of ξ. Therefore, $\xi \sim \nu$. $\square$

1.2 Conditional Entropy

Condiditional entropy for a finite partition ξ, $|\xi| = I$, with respect to finite partition η, $|\eta| = J$ is defined as

$$H(\xi|\eta) = -\sum_{j \in J} \mu[D_j] \sum_{i \in I} \mu[C_i|D_j] \log \mu[C_i|D_j] = -\int_X \log \mu[C(x)\,|D(x)]\, d\mu[x]\,.$$

Remark: $\rho(\xi,\eta) = H(\xi|\eta) + H(\eta|\xi)$, is called Rokhlin metric.

1.3 Entropy of a measure-preserving transformation

Let $T : X \to X$ be a measure-preserving transformation and $\zeta = \{C_\alpha : \alpha \in I\}$ be a partition of X with finite entropy. Consider $T^{-k}(\zeta) = \left\{T^{-k}(C_\alpha) : \alpha \in I\right\}$ and

$$\zeta^n = \zeta \vee T^{-1}(\zeta) \vee \cdots \vee T^{-n+1}(\zeta)$$

Since $\mu\left[T^{-k}(C_i)\right] = \mu[C_i]$ (measure preserving), also since $H\left(T^{-k}(\zeta)\right) = H(\zeta)$, and $H(\xi \vee \eta) \leqslant H(\xi) + H(\eta)$, it follows that

$$
\begin{aligned}
H\left(\zeta^{m+n}\right) &\leqslant H\left(\zeta \vee T^{-1} \vee \cdots \vee T^{-m+1}(\zeta) \vee T^{-m}(\zeta) \vee \cdots \vee T^{-m-n+1}(\zeta)\right) \\
&\leqslant H\left(\zeta^{m}\right)+H\left(\bigvee_{i=m}^{m+n-1} T^{-i}(\zeta)\right) \\
&= H\left(\zeta^{m}\right)+H\left(\bigvee_{i=0}^{n-1} T^{-i-m}(\zeta)\right) \\
&\leqslant H\left(\zeta^{m}\right)+\bigvee_{i=0}^{n-1} H\left(T^{-i-m}(\zeta)\right) \\
&= H\left(\zeta^{m}\right)+\bigvee_{i=0}^{n-1} H\left(T^{-i}(\zeta)\right) \\
&= H\left(\zeta^{m}\right)+H\left(\zeta^{n}\right)
\end{aligned}
$$

By Fekete's subadditivity lemma, we know

$$
h(T, \zeta)=\lim_{n \rightarrow \infty} \frac{1}{n} H\left(\zeta^{n}\right)
$$

exists, and is called *metric entropy*(or *measure-theoretic entropy*) of T with respect to the partition ζ. Note by subadditivity that for all $n \in \mathbb{N}$, $H\left(\zeta^{n}\right) \leqslant n H(\zeta)$, therefore $\frac{1}{n} H\left(\zeta^{n}\right) \leqslant H(\zeta)$ so that $H(T, \zeta) \leqslant H(\zeta)$.

Proposition. $h(T, \zeta)=\lim_{n \rightarrow \infty} H\left(\zeta \mid T^{-1}\left(\zeta^{n}\right)\right)$.

The metric entropy of T is defined as the suppremum of $h(T, \zeta)$ over all finite measurable partitions ζ of X:

$$
h(T):=\sup \{h(T, \zeta): \zeta \text{ is measurable finite partition of } X\}.
$$

Entropy is an invariant for isomorphic dynamical systems[3]: it was first introduced by Kolmogorov in 1958 before Sinaï brought slight modifications. In 1970, D.S. Ornstein showed that entropy is a complete invariant for invertible Bernoulli shifts, that is, two Bernoulli shifts are measure theretic isomoprhic if and only if they have the same entropy.

1.4 Kolmogorov-Sinaï Theorem

Definition. *[Refining and generating partition]*

- $\{\zeta_n\}$ *is refining if* $\zeta_{n+1} \geqslant \zeta_n$;
- $\{\zeta_n\}$ *is generating if* $\forall \xi \forall \delta>0 \exists\, n_0$ *such that for all* $n \geqslant n_0$, *there exists* $\xi_n \leqslant \bigvee_{k=-n}^{n} \zeta_n$ *with* $d\left(\xi_n, \xi\right)<\delta$;
- *A generator is a finite partition* ξ *such that the sequence* $\bigvee_{k=0}^{n} T^{k}(\xi)$ *is generating.*

Lebesgue space has a generating sequence of finite partitions.

Theorem (Kolmogorov-Sinaï, [2])**.** *Let* ξ *be a generator for* T*. Then* $h(T)=h(T, \xi)$.

2 Elements of Solution

2.1 Variational Principle

in this section $f : X \to X$ is a homeomorphism, X is compact and $\mathcal{M}$ is the space of Borel probability measure on X, that is the space of all finite Borel measures μ with $\mu[X] = 1$.

Lemma. *Let $\mu, \nu \in \mathcal{M}$. For any measurable partition ξ of X*

$$tH_\mu(\xi) + (1-t) H_\nu(\xi) \leqslant H_{t\mu+(1-t)\nu}(\xi)$$

Proof. For each $n \in \mathbb{N}$, $tH_\mu(\xi^n)+(1-t) H_\nu(\xi^n) = -t\sum_i \mu[C_{i,n}]\log\mu[C_{i,n}]-(1-t)\sum_i \nu[C_{i,n}]\log\nu[C_{i,n}] = -\sum_i (t\mu[C_{i,n}]\log\mu[C_{i,n}] + (1-t)\nu[C_{i,n}]\log\nu[C_{i,n}])$. Since $x \mapsto x\log x$ is convex, it follows that $t\mu[C_{i,n}]\log\mu[C_{i,n}]+(1-t)\nu[C_{i,n}]\log\nu[C_{i,n}] \geqslant (t\mu[C_{i,n}] + (1-t)\nu[C_{i,n}])\log(t\mu[C_{i,n}] + (1-t)\nu[C_{i,n}])$. Then $tH_\mu(\xi^n) + (1-t) H_\nu(\xi^n) \leqslant -\sum_i (t\mu[C_{i,n}] + (1-t)\nu[C_{i,n}])\log(t\mu[C_{i,n}] + (1-t)\nu[C_{i,n}]) = H_{t\mu+(1-t)\nu}(\xi^n)$. Therefore $tH_\mu(\xi) + (1-t) H_\nu(\xi) \leqslant H_{t\mu+(1-t)\nu}(\xi)$. $\square$

Given a partition $\xi = \{A_1, \ldots, A_k\}$, define its boundary as $\partial\xi = \bigcup_{1\leqslant i\leqslant k} \partial A_i$ where $\partial A = \overline{A} \setminus \operatorname{int}(A) = \overline{A} \cap \overline{X \setminus A}$ is the topological boundary in the ordinary sense.

Lemma. *Let $\mu \in \mathcal{M}$.*

1. *For any $x \in X$ and any $\delta > 0$, there exists $\delta' \in (0, \delta)$ such that $\mu[\partial B(0, \delta')] = 0$;*
2. *For any $\delta > 0$, there is a finite measurable partition $\xi = \{C_1, \ldots, C_k\}$ with $\operatorname{diam}(C_i) < \delta$ for all i and $\mu[\partial\xi] = 0$;*
3. *If $\{\mu_n\} \subseteq \mathcal{M}$ is a sequence of Borel probability measure that converges to μ in the weak* topology*[1]*, and A is a measurable set with $\mu[\partial A] = 0$, then $\mu[A] = \lim_{n\to\infty} \mu_n[A]$.*

Let $\mathcal{M}_f$ represents the set of all $f-$invariant Borel probability measures. Then:

Theorem (Variational principle, [2])**.** *Let f be a homeomorphism of a compact metric space X. Then $h_{\text{top}}(f) = \sup\{h_\mu(f) : \mu \in \mathcal{M}_f\}$.*

2.2 Entropy for expansive map

From variational principle, $h_\lambda(f) \leqslant h_{\text{top}}(f)$ for all $\lambda \in \mathcal{M}_f$. Let E_n be an $(n, \epsilon)-$ separated set with $\epsilon \leqslant \delta_0$, where δ_0 is the expansiveness constant of f. By expansiveness, for distinct x, y in X there exists $i \in \mathbb{Z}$ such that $d(f^i(x), f^i(y)) \geqslant \delta_0$. Define $\nu_n = \frac{1}{|E_n|}\sum_{x\in E_n}\delta_x$ and $\mu_n = \frac{1}{n}\sum_{i=0}^{n-1} f_*^i\nu_n$. By compactness, some $\{\mu_{n_k}\}_{k\in\mathbb{N}}$ converges. Let $\mu := \lim_k \mu_{n_k}$, which an accumulation point for $\{\mu_n\}_{n\in\mathbb{N}}$, and which is obviously $f-$invariant. Let ξ be a measurable partition such that $\operatorname{diam}(X) < \epsilon$ for all $C \in \xi$ and $\partial\xi = 0$.

Consider the sequence $\left\{\bigvee_{k=-n}^{n} f^k(\xi)\right\}_{n\in\mathbb{N}}$. For each $x, y \in \bigvee_{k=-\infty}^{\infty} f^k(\xi)$, we have $d\left(f^k(x), f^k(y)\right) < \epsilon \leqslant \delta_0$ for all $k \in \mathbb{Z}$. By expansiveness, it must be true that $x = y$. Then the maximal diameter of

[1]meaning $\displaystyle\int h(x)\,\mu_n(x) \to \int h(x)\,d\mu(x)$ for all h measurable.

element of $\left\{\bigvee_{-n\leqslant i\leqslant n} T^i\left(\xi\right)\right\}$ goes to 0. Hence, $\left\{\bigvee_{k=-n}^{n} f^k\left(\xi\right)\right\}$ is generating. Then ξ is generator and by Kolmogorov-Sinaï, we know $h_\mu\left(f\right) = h_\mu\left(f,\xi\right)$.

Now, for each $C \in \xi^n$, $\nu_n\left[C\right] = 0$ or $\frac{1}{|E_n|}$. Then $H_{\nu_n}\left(\xi^n\right) = \log|E_n|$. Then, as in a lemma in [2], we can prove that $\overline{\lim}_{n\to\infty}\log|E_n| \leqslant \lim_{q\to\infty}\lim_{n\to\infty}\frac{1}{q}H_{\mu_n}\left(\xi^q\right) = h_\mu\left(f,\xi\right)$. Therefore $h_{\text{top}}\left(f\right) \leqslant h_\mu\left(f\right)$.

References

[1] T.N.T. Goodman. Relating topological entropy and measure entropy. *Bulletin London Mathematical Society*, (3):176–180, 1971.

[2] Garrett Stuck Michael Brin. *Introduction to Dynamical System*. Cambrigde University Press, 2002.

[3] Ergodic Theory and Topological Dyanamics. *James R. Brown*. Academic Press, Columbia University, New York, 1976.