Lecture 10

Let us consider some examples of calculation of topological entropy.

Example 1. (Full shift over finite states). Let (Σ,σ) be the one-sided full shift over $\{1,2,\cdots,k\}$. Consider a partition $\mathscr{P}=\{[i]:1\le i\le k\}$, note that $\mathscr{P}$ is also an open cover of Σ and

$$\bigvee_{i=0}^{n-1}\sigma^{-i}\mathscr{P}=\{[x_1x_2\cdots x_n]:x_1,\cdots,x_n\in\{1,\cdots,k\}\}.$$

Since open sets in $\bigvee_{i=0}^{n-1}\sigma^{-i}\mathscr{P}$ are disjoint, $N\big(\bigvee_{i=0}^{n-1}\sigma^{-i}\mathscr{P}\big)=k^n$. Since $\mathrm{diam}\big(\bigvee_{i=0}^{n-1}\sigma^{-i}\mathscr{P}\big)\to 0$ as $n\to\infty$, we have

$$h(\sigma)=h(\sigma,\mathscr{P})=\lim_{n\to\infty}\frac{1}{n}\log N\big(\bigvee_{i=0}^{n-1}\sigma^{-i}\mathscr{P}\big)=\lim_{n\to\infty}\frac{1}{n}\log k^n=\log k.$$

Example 2. (One-sided subshift over finite states). Let (Σ,σ) be the one-sided full shift over $\{1,2,\cdots,k\}$. Let $X\subseteq\Sigma$ be compact and σ-invariant, that is $\sigma X\subseteq X$, then we call (X,σ_X) a subshift. Let $\mathscr{P}$ be defined as in the above example, define

$$\mathscr{P}_X=\{[i]\cap X:i=1,\cdots,k\},$$

then

$$\bigvee_{i=0}^{n-1}\sigma^{-i}\mathscr{P}_X=\{[x_1x_2\cdots x_n]\cap X\neq\emptyset:x_1,\cdots,x_n\in\{1,\cdots,k\}\}.$$

Also we have $\mathrm{diam}\big(\bigvee_{i=0}^{n-1}\sigma^{-i}\mathscr{P}_X\big)\to 0$ as $n\to\infty$, hence

$$h(\sigma_X)=h(\sigma_X,\mathscr{P}_X)=\lim_{n\to\infty}\frac{\log l_n}{n},$$

where $l_n:=\sharp\{[x_1\cdots x_n]\cap X\neq\emptyset:x_1,\cdots,x_n\in\{1,\cdots,k\}\}$. Those $[x_1\cdots x_n]\cap X\neq\emptyset$ are called admissible words.

Remark: For any $\alpha\in[0,\log k]$, there exists a subshift $X\subseteq\Sigma$ such that $h(\sigma_X)=\alpha$.

Example 3. (One-sided subshift of finite type over $\{1,\cdots,k\}$). Let $A=(a_{ij})_{k\times k}$ with $a_{ij}=0$ or 1. Define

$$\Sigma_A=\{(x_i)_{i=1}^{\infty}\in\Sigma:a_{x_ix_{i+1}}=1\forall i\geq 1\}.$$

Then

$$\begin{aligned} l_n &= \sharp\{x_1 \cdots x_n : a_{x_i x_{i+1}} = 1 \text{ for } i = 1, \cdots, n-1\} \\ &= \sum_{x_1 \cdots x_n} a_{x_1 x_2} a_{x_2 x_3} \cdots a_{x_{n-1} x_n} \\ &= \sum_{x_1, x_n} \sum_{x_2 \cdots x_{n-1}} a_{x_1 x_2} a_{x_2 x_3} \cdots a_{x_{n-1} x_n} \\ &= \sum_{x_1, x_n} (A^{n-1})_{x_1, x_n} \approx ||A^{n-1}||. \end{aligned}$$

Hence

$$h(\sigma_A) = \lim_{n \to \infty} \frac{\log l_n}{n} = \lim_{n \to \infty} \frac{\log ||A^{n-1}||}{n} = \log \rho(A),$$

where $\rho(A)$ is the spectral radius of A defined by $\rho(A) := \max_i |\lambda_i|$, where λ_i are eigenvalues of A.

Proposition 5.2. *Let (X, T) be a TDS. Assume that $d(Tx, Ty) \leq d(x, y)$ for $x, y \in X$, then $h(T) = 0$.*

Proof. Recall $d_n(x, y) = \max_{1 \leq i \leq n-1} d(T^i x, T^i y)$, $B_n(x, \epsilon) = \{y \in X : d_n(x, y) < \epsilon\}$ and $N_n(\epsilon) = \inf\{k : \exists x_1, \cdots, x_k, s.t. \cup_{i=1}^k B_n(x_i, \epsilon) \supseteq X\}$. Then

$$h(T) = \lim_{\epsilon \to 0} \overline{\lim_{n \to \infty}} \frac{\log N_n(\epsilon)}{n}.$$

Since T is contractive, $d_n(x, y) = d(x, y)$ for all n, hence $N_n(\epsilon)$ is independent of n, by above formula $h(T) = 0$. □

As an application, the rotation on the circle defined by $T : \mathbb{T} \to \mathbb{T}, x \mapsto x + \alpha(\text{mod}1)$ has entropy zero. Also the rotation on $\mathbb{T}^d$ defined by $(x_1, \cdots, x_d) \mapsto (x_1 + \alpha, \cdots, x_d + \alpha_d)(\text{mod}1)$ has entropy zero.

Example 4. (Affine map on $\mathbb{T}^d$). Let A be a $d \times d$ integral matrix. Consider $f_A : \mathbb{T}^d \to \mathbb{T}^d, \vec{x} \mapsto A\vec{x}(\text{mod}1)$, then

$$h(f_A) = \sum_{|\lambda_i| > 1} \log |\lambda_i|,$$

where $\lambda_1, \cdots, \lambda_d$ are eigenvalues of A.

6 Relations between topological and measure-theoretical entropies

6.1 Entropy map

Let (X, T) be a TDS. Let $\mathscr{M}(X)$ be the collection of all Borel probability measures on X, recall it is compact in the weak-$*$ topology (that is the topology defined by $\mu_n \to \mu$ if $\int_X f d\mu_n \to \int_X f d\mu$ for all $f \in C(X)$).

Let (M, T) be the collection of T-invariant Borel probability measures on X, then $M(X, T)$ is a convex compact subspace of $\mathscr{M}(X)$. Notice that $M(X, T) \neq \emptyset$ by Krylov–Bogolyubov theorem. We define the entropy map by

$$\mu \mapsto h_\mu(T), \ \mu \in M(X, T).$$

Proposition 6.1. *The entropy map is affine, that is for any μ and $m \in M(X, T)$ and $p \in (0, 1)$,*

$$h_{p\mu+(1-p)m}(T) = ph_\mu(T) + (1 - p)h_m(T).$$

Lemma 6.1. *Let $\phi(x) = -x \log x$ for $x \in [0, 1]$. Then for any $x_1, x_2, \cdots, x_k \in [0, 1]$ and any probability vector $\vec{p} = (p_1, p_2, \cdots, p_k)$,*

$$\sum_{i=1}^{k} p_i\phi(x_i) \leq \phi(\sum_{i=1}^{k} p_i x_i) \leq \sum_{i=1}^{k} [p_i\phi(x_i) + x_i\phi(p_i)].$$

Proof. The first "≤" follows from the concavity of ϕ. For the second,

$$\begin{aligned}\phi(\sum_{i=1}^{k} p_i x_i) &= -(\sum_{i=1}^{k} p_i x_i) \log (\sum_{i=1}^{k} p_i x_i) \leq -\sum_{i=1}^{k} p_i x_i \log p_i x_i \\ &= -\sum_{i=1}^{k}(p_i x_i \log x_i + x_i p_i \log p_i) = \sum_{i=1}^{k}[p_i\phi(x_i) + x_i\phi(p_i)].\end{aligned}$$

□

Corollary 6.1.1. *Let $\mu_1, \cdots, \mu_k \in M(X, T)$ and $\vec{p} = (p_1, \cdots, p_k)$ be a probability vector. Let $\xi = \{A_1, \cdots, A_n\}$ be a finite Borel partition of X. Then*

$$\sum_{i=1}^{k} p_i H_{\mu_i}(\xi) \leq H_{\sum_i p_i\mu_i}(\xi) \leq \sum_{i=1}^{k} p_i H_{\mu_i}(\xi) - \sum_{i=1}^{k} p_i \log p_i.$$

Proof. By the above lemma,

$$\begin{aligned}H_{\sum_i p_i\mu_i}(\xi) &= \sum_{A\in\xi} \phi(\sum_{i} p_i\mu_i(A)) \leq \sum_{A\in\xi} \sum_{i=1}^{k}[p_i\phi(\mu_i(A)) + \mu_i(A)\phi(p_i)] \\ &= \sum_{i=1}^{k} p_i H_{\mu_i}(\xi) + \sum_{i=1}^{k} \phi(p_i).\end{aligned}$$

□

Now we can prove that the entropy map is affine.

Proof of Proposition 6.1. Let ξ be a finite partition of X. Let $\mu, m \in M(X,T)$ and $p \in (0,1)$. By the above corollary,

$$\begin{aligned} pH_\mu(\bigvee_{i=0}^{n-1} T^{-i}\xi) + (1-p)H_m(\bigvee_{i=0}^{n-1} T^{-i}\xi) &\leq H_{p\mu+(1-p)m}(\bigvee_{i=0}^{n-1} T^{-i}\xi) \\ &\leq pH_\mu(\bigvee_{i=0}^{n-1} T^{-i}\xi) + (1-p)H_m(\bigvee_{i=0}^{n-1} T^{-i}\xi) \\ &- p\log p - (1-p)\log(1-p). \end{aligned}$$

Dividing by n, letting $n \to \infty$, we have

$$ph_\mu(T,\xi) + (1-p)h_m(T,\xi) = h_{p\mu+(1-p)m}(T,\xi).$$

Hence $\sup_\xi h_{p\mu+(1-p)m}(T,\xi) \leq ph_\mu(T) + (1-p)h_m(T)$, that is $h_{p\mu+(1-p)m}(T) \leq ph_\mu(T) + (1-p)h_m(T)$. For the converse inequality, if one of $h_\mu(T)$, $h_m(T)$ is ∞, then clearly $h_{p\mu+(1-p)m}(T) = \infty$. In the other case, for any $\epsilon > 0$, we can find ξ_1, ξ_2 such that

$$h_\mu(T,\xi_1) > h_\mu(T) - \epsilon,\ h_m(T,\xi_2) > h_m(T) - \epsilon.$$

Take $\eta = \xi_1 \vee \xi_2$, then

$$\begin{aligned} h_{p\mu+(1-p)m}(T,\eta) &= ph_\mu(T,\eta) + (1-p)h_m(T,\eta) \\ &\geq ph_\mu(T,\xi_1) + (1-p)h_m(T,\xi_2) \\ &\geq ph_\mu(T) + (1-p)h_m(T) - 2\epsilon. \end{aligned}$$

This completes the proof. □

Our target is the following famous variational principle.

Theorem 6.2.

$$h_{top}(T) = \sup\{h_\mu(T) : \mu \in M(X,T)\}.$$

We first establish some lemmas.

Lemma 6.3. *Let X be a compact metric space, $\mu \in \mathscr{M}(X)$.*

(i) For $x \in X$ and $\delta > 0$, $\exists\ 0 < \epsilon < \delta$, s.t. $\mu(\partial B(x,\epsilon)) = 0$.

(ii) $\forall \epsilon > 0$, $\exists$ a Borel partition $\xi = \{A_1, \cdots, A_k\}$ of X, s.t. diam$(\xi) < \epsilon$ and $\mu(\partial(A_i)) = 0, i = 1, \cdots, k$.

Proof. (i) Fix $x \in X$, consider the map $r \mapsto \mu(\partial B(x,r))$. Since μ is a probability measure, $\mu(\partial B(x,r)) > 0$ for at most countable many r, which proves (i). (ii) $\forall \epsilon > 0$, by (i) we can find finite open balls $\{B_1, \cdots, B_k\}$ covering X with $\text{diam}(B_i) < \epsilon$ and $\mu(\partial B_i) = 0$. Take $A_1 = B_1,\ A_2 = B_2 \setminus B_1, \cdots, A_k = B_k \setminus (B_1 \cup B_2 \cup \cdots B_k)$, then $\xi = \{A_1, \cdots, A_k\}$ is a partition with $\text{diam}(\xi) < \epsilon$, moreover, $\partial A_i \subseteq \cup_{j=1}^k \partial B_j$, hence $\mu(\partial A_i) = 0$ for $i = 1, \cdots, k$. □