Lecture 9

We continue to consider some more examples.

Example 1. (Bernoulli shifts). Consider $(p_1, p_2, \cdots, p_k)$-shift on $\{1, 2, \cdots, k\}$. Recall $\Sigma = \{1, 2, \cdots, k\}^{\mathbb{N}}$, σ is the left shift on Σ, $\mu([i_1 i_2 \cdots i_n]) = p_{i_1} p_{i_2} \cdots p_{i_n}$. Consider partition $\mathscr{P} = \{[i] : i = 1, 2, \cdots, k\}$, then

$$\bigvee_{i=0}^{n-1} \sigma^{-i} \mathscr{P} = \{[i_1 i_2 \cdots i_n] : i_1, \cdots, i_n \in \{1, 2, \cdots, k\}\}.$$

Since $\text{diam}(\bigvee_{i=0}^{n-1} \sigma^{-i} \mathscr{P}) = 2^{-n} \to 0$ as $n \to \infty$, we have

$$\begin{aligned}
h(\sigma) = h(\sigma, \mathscr{P}) &= \lim_{n \to \infty} \frac{1}{n} H(\bigvee_{i=0}^{n-1} \sigma^{-i} \mathscr{P}) \\
&= \lim_{n \to \infty} \frac{1}{n} \sum_{i_1 i_2 \cdots i_n} -\mu([i_1 i_2 \cdots i_n]) \log \mu([i_1 i_2 \cdots i_n]) \\
&= \lim_{n \to \infty} \frac{1}{n} \sum_{i_1 i_2 \cdots i_n} -p_{i_1} p_{i_2} \cdots p_{i_n} (\log p_{i_1} + \log p_{i_2} + \cdots + \log_{p_{i_n}}) \\
&= \lim_{n \to \infty} \sum_{i_1 i_2 \cdots i_n} -p_{i_1} p_{i_2} \cdots p_{i_n} \log p_{i_1} = \sum_{i_1} -p_{i_1} \log p_{i_1} \sum_{i_2 i_3 \cdots i_n} p_{i_2} \cdots p_{i_n} \\
&= \sum_{i=1}^{k} -p_i \log p_i.
\end{aligned}$$

Example 2. (Markov shifts). Consider $(\vec{p}, P)$-shift on $\{1, 2, \cdots, k\}$. Recall that P is a stochastic matrix (p_{ij}), $\vec{p}$ is a probability vector with all entries positive such that $\vec{p}P = \vec{p}$, σ is the left shift, and $\mu([p_{i_1} p_{i_2} \cdots p_{i_n}]) = p_{i_1} p_{i_1 i_2} \cdots p_{i_{n-1} i_n}$. By the same argument in the above example, $\mathscr{P} = \{[i] : 1 \leq i \leq k\}$ is a partition and we have $h(\sigma) = h(\sigma, \mathscr{P})$. Hence

$$\begin{aligned}
h(\sigma) &= \lim_{n \to \infty} \frac{1}{n} \sum_{i_1 i_2 \cdots i_n} -\mu([i_1 i_2 \cdots i_n]) \log \mu([i_1 i_2 \cdots i_n]) \\
&= \lim_{n \to \infty} \frac{1}{n} \sum_{i_1 i_2 \cdots i_n} -p_{i_1} p_{i_1 i_2} \cdots p_{i_{n-1} i_n} \log p_{i_1} p_{i_1 i_2} \cdots p_{i_{n-1} i_n} \\
&= \lim_{n \to \infty} \frac{1}{n} \sum_{i_1 i_2 \cdots i_n} -p_{i_1} p_{i_1 i_2} \cdots p_{i_{n-1} i_n} (\log p_{i_1} + \log p_{i_1 i_2} + \cdots + \log p_{i_{n-1} i_n}).
\end{aligned}$$

Notice that

$$\sum_{i_1} \sum_{i_2 \cdots i_n} -p_{i_1} p_{i_1 i_2} \cdots p_{i_{n-1} i_n} \log p_{i_1} = \sum_{i=1}^{k} -p_i \log p_i,$$

and

$$\sum_{i_1 i_2\cdots i_n} -p_{i_1}p_{i_1 i_2}\cdots p_{i_{n-1}i_n}\log p_{i_1 i_2} = \sum_{i_1 i_2} -p_{i_1}p_{i_1 i_2}\log p_{i_1 i_2}\sum_{i_3\cdots i_n} p_{i_2 i_3}\cdots p_{i_{n-1}i_n} = \sum_i\sum_j -p_i p_{ij}\log p_{ij},$$

similarly for $l = 1, 2, \cdots, n-1$, we have

$$\sum_{i_1 i_2\cdots i_n} -p_{i_1}p_{i_1 i_2}\cdots p_{i_{n-1}i_n}\log p_{i_l i_{l+1}} = \sum_i\sum_j -p_i p_{ij}\log p_{ij},$$

hence

$$h(T) = \lim_{n\to\infty}\frac{1}{n}[-\sum_i p_i\log p_i - (n-1)\sum_i\sum_j p_i p_{ij}\log p_{ij}] = -\sum_i\sum_j p_i p_{ij}\log p_{ij}.$$

The motivation of introducing the notion of entropy is to clarify MPSs. There was an open problem before 1958:

Open problem (before 1958) Is $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$-shift isomorphic to $(\frac{1}{2}, \frac{1}{2})$-shift?

Kolmogorov showed that the answer is NO, by showing that entropy is an isomorphism invariant and the two systems have different entropy.

4.5 Entropy as an isomorphism invariant

Definition 4.8. *Let $(X_i, \mathscr{B}_i, \mu_i, T_i)$ $(i = 1, 2)$ be two MPSs. Say they are isomorphic if there exists a map $\varphi : X_1 \to X_2$ satisfies the following properties:*

(i) φ is bijective (after removing some sets of measure zero).

(ii) φ is measurable, i.e. $\varphi^{-1}\mathscr{B}_2 \subseteq \mathscr{B}_1$ and $\varphi\mathscr{B}_1 \subseteq \mathscr{B}_2$.

(iii) $\mu_2 = \mu_1 \circ \varphi^{-1}$, $\mu_1 = \mu_2 \circ \varphi$, that is φ preserves measures.

(iv) $\varphi \circ T_1 = T_2 \circ \varphi$, that is the following diagram commutes,

$$\begin{array}{ccc} X_1 & \xrightarrow{T_1} & X_2 \\ \big\downarrow{\scriptstyle\varphi} & & \big\downarrow{\scriptstyle\varphi} \\ X_2 & \xrightarrow{T_2} & X_2. \end{array}$$

Theorem 4.14. *If $(X_1, \mathscr{B}_1, \mu_1, T_1)$, $(X_2, \mathscr{B}_2, \mu_2, T_2)$ are isomorphic, then $h_{\mu_1}(T_1) = h_{\mu_2}(T_2)$.*

Proof. We prove $h_{\mu_2}(T_2) \le h_{\mu_1}(T_1)$, the reverse inequality will hold symmetrically. Let $\alpha = \{A_1, \cdots, A_k\}$ be a partition of X_2, then $\varphi^{-1}\alpha = \{\varphi^{-1}A_1, \cdots, \varphi^{-1}A_k\}$ is a partition of X_1. Then

$$H_{\mu_2}\big(\bigvee_{i=0}^{n-1} T_2^{-i}\alpha\big) = H_{\mu_1\circ\varphi^{-1}}\big(\bigvee_{i=0}^{n-1} T_2^{-1}\alpha\big) = H_{\mu_1}\big(\varphi^{-1}\bigvee_{i=0}^{n-1} T_2^{-i}\alpha\big) = H_{\mu_1}\big(\bigvee_{i=0}^{n-1} T_1^{-i}(\varphi^{-1}\alpha)\big),$$

hence

$$h_{\mu_2}(T_2, \alpha) = \lim_{n\to\infty}\frac{1}{n}H_{\mu_2}\big(\bigvee_{i=0}^{n-1} T_2^{-i}\alpha\big) = \lim_{n\to\infty}\frac{1}{n}H_{\mu_1}\big(\bigvee_{i=0}^{n-1} T_1^{-i}(\varphi^{-1}\alpha)\big) = h_{\mu_1}(T_1, \varphi^{-1}\alpha) \le h_{\mu_1}(T_1),$$

taking supremum over all finite partitions, we complete the proof. □

Now we see that the two Bernoulli shifts $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$-shift and $(\frac{1}{2}, \frac{1}{2})$-shift are not isomorphic since they have entropy $\log 3$ and $\log 2$ respectively. In 1969, Ornstein proved the following deep theorem.

Theorem 4.15 (Ornstein)**.** *For any two Bernoulli shifts both on finite state spaces, they are isomorphic iff they have the same entropy.*

4.6 Ergodic theory of information

The following theorem is called Shannon-McMillan-Breiman theorem, for a proof see William Parry's book.

Theorem 4.16. *Let $(X, \mathscr{B}, \mu, T)$ be a MPS. Let $\xi = \{A_1, \cdots, A_k\}$ be a finite partition of X. For $n \in \mathbb{N}$ and $x \in X$, let $\xi_n(x)$ be the member of $\bigvee_{i=0}^{n-1} T^{-i}\xi$ that contains x. If T is ergodic, then*

$$\lim_{n\to\infty} -\frac{1}{n} \log \mu(\xi_n(x)) = h(T, \xi), \text{ for } \mu\text{-a.e. } x \in X.$$

That is $\mu(\xi_n(x)) \sim e^{-nh(T,\xi)}$ for μ-a.e. $x \in X$.

5 Topological entropy

5.1 Conjugacy problem in TDS

Recall (X, T) is a TDS if X is a compact metric space and $T : X \to X$ is continuous.

Definition 5.1. *Two TDSs (X, T) and (Y, S) are said to be topological conjugate if there is a homeomorphism $\phi : X \to Y$ such that $\phi \circ T = S \circ \phi$, that is the following diagram commutes,*

$$\begin{array}{ccc} X & \xrightarrow{T} & X \\ \downarrow{\scriptstyle\phi} & & \downarrow{\scriptstyle\phi} \\ Y & \xrightarrow{S} & Y. \end{array}$$

Question: How can we determine whether two TDSs are topological conjugate?

Just as in the situation of MPS, we expect to find some conjugacy invariant.

5.2 Definition of topological entropy

The notion of topological entropy was first introduced by Adler, Konheim and McAndrew in 1965.

Let (X, T) be a TDS. Say $\alpha = \{A_i : i \in \mathcal{I}\}$ is an open cover of X if $\bigcup_{i\in\mathcal{I}} A_i = X$ and A_i are open.

Definition 5.2. *Let α be an open cover of X. Define*

$$N(\alpha) = \inf\{k : \exists A_1, \cdots, A_k \in \alpha, s.t. X \subseteq \bigcup_{i=1}^{k} A_i\},$$

and define $H(\alpha) := \log N(\alpha)$.

Let α, β be two open covers of X. Say β is a refinement of α if every member of β is a subset of some member of α, we write $\alpha < \beta$. For instance, let $X = \mathbb{T}$, set $\beta = \{(0, \frac{1}{2}), (\frac{1}{3}, \frac{3}{4}), (\frac{2}{3}, \frac{5}{4})\}$ and $\alpha = \{(0, \frac{1}{2}), (\frac{1}{4}, \frac{3}{4}), (0, \frac{2}{3}), (\frac{2}{3}, \frac{5}{4})\}$, then $\alpha < \beta$.

Remark.

(i) $N(\alpha) \geq 1$.

(ii) If $\alpha < \beta$, then $N(\alpha) \leq N(\beta)$.

(iii) $N(T^{-1}\alpha) \leq N(\alpha)$, where $T^{-1}\alpha := \{T^{-1}A : A \in \alpha\}$.

Proof. (i) is clear. For (ii), let $t = N(\beta)$, then $\exists A_1, \cdots, A_k \in \beta$ s.t. $X = \bigcup_{i=1}^{k} A_i$. Since $\alpha < \beta$, $\exists B_i \in \alpha$ s.t. $A_i \subseteq B_i$, hence $X \subseteq \bigcup_{i=1}^{k} B_i$, then $N(\alpha) \leq N(\beta)$. (iii) be can seen in the same way. □

Definition 5.3. *Let α, β be two open covers of X. Define*

$$\alpha \vee \beta = \{A \cap B : A \in \alpha, B \in \beta\}.$$

It's clear that $\alpha \vee \beta$ is an open cover of X and $\alpha < \alpha \vee \beta$, $\beta < \alpha \vee \beta$.

Lemma 5.1. *$N(\alpha \vee \beta) \leq N(\alpha)N(\beta)$ and $H(\alpha \vee \beta) \leq H(\alpha) + H(\beta)$.*

Proof. Suppose $X = \bigcup_{i=1}^{N(\alpha)} A_i = \bigcup_{j=1}^{N(\beta)} B_j$, with $A_i \in \alpha$, $B_j \in \beta$, then $X = \bigcup_{i,j} A_i \cap B_j$, hence $N(\alpha \vee \beta) \leq N(\alpha)N(\beta)$, the second inequality follows after taking logarithm. □

Definition 5.4 (Entropy of an open cover)**.** *Let α be an open cover of X. Define*

$$h(T, \alpha) := \lim_{n\to\infty} \frac{1}{n} H\big(\bigvee_{i=0}^{n-1} T^{-i}\alpha\big),$$

we call $h(T, \alpha)$ the topological entropy of T w.r.t α.

The existence of the above limit is guarantee by the following lemma.

Lemma 5.2. *Set $a_n = H\big(\bigvee_{i=0}^{n-1} T^{-i}\alpha\big)$, then $a_{n+m} \leq a_n + a_m$ and hence*

$$\lim_{n\to\infty} \frac{a_n}{n} = \inf_n \frac{a_n}{n}.$$

Proof.

$$\begin{aligned} a_{n+m} = H\big(\bigvee_{i=0}^{n+m-1} T^{-i}\alpha\big) &= H\Big(\big(\bigvee_{i=0}^{n-1} T^{-i}\alpha\big) \bigvee T^{-n}\big(\bigvee_{j=0}^{m-1} T^{-j}\alpha\big)\Big) \\ &\leq a_n + H\Big(T^{-n}\big(\bigvee_{j=0}^{m-1} T^{-j}\alpha\big)\Big) \leq a_n + a_m. \end{aligned}$$

□

Definition 5.5 (Topological entropy of T)**.**

$$h(T) := \sup_{\alpha} h(T, \alpha),$$

where the supremum is taking over all open covers of X.

The definition of topological entropy is quite similar to that of measure-theoretical entropy, it turns out to be an invariant of topological conjugacy.

Theorem 5.3. *Suppose (X, T) and (Y, S) are topological conjugate, then $h(T) = h(S)$.*

Proof. We show that $h(S) \leq h(T)$. It suffices to show $h(S, \alpha) \leq h(T)$ for any open cover α of Y. Let $\phi : X \to Y$ be the conjugacy map, we have

$$\begin{aligned} h(S, \alpha) &= \lim_{n \to \infty} \frac{1}{n} H\big(\bigvee_{i=0}^{n-1} S^{-i} \alpha \big) = \lim_{n \to \infty} \frac{1}{n} H\big(\phi^{-1} \bigvee_{i=0}^{n-1} S^{-i} \alpha\big) \\ &= \lim_{n \to \infty} \frac{1}{n} H\big(\bigvee_{i=0}^{n-1} T^{-i} (\phi^{-1} \alpha)\big) = h(T, \phi^{-1} \alpha) \leq h(T), \end{aligned}$$

notice that in the second equality we have used the fact that $H(\beta) = H(\phi^{-1}\beta)$ for any open cover β of Y. □

5.3 Calculation of topological entropy

For any open cover α of X, define

$$\operatorname{diam}(\alpha) := \sup_{A \in \alpha} \operatorname{diam}(A).$$

A Lebesgue number of α is a value $\delta > 0$ such that for any $x \in X$, the open ball $B(x, \delta)$ is a subset of some member of α. Lebesuge number of an open cover always exists due to the compactness of X.

Claim. Any open cover has a Lebesgue number.

Proof. Suppose α is an open cover of X which does not have a Lebesgue number, then for any n, $\exists x_n \in X$, s.t. $B(x_n, \frac{1}{n})$ is not contained in any member of α. By compactness, $\exists$ subsequence (n_k) and some $x \in X$, s.t. $x_{n_k} \to x$ as $k \to \infty$. But since $x \in A$ for some open set $A \in \alpha$, $\exists r > 0$, s.t. $B(x, r) \subseteq A$, however $B(x_{n_k}, \frac{1}{n_k}) \subseteq B(x, r)$ when k is large, which contradicts with our assumption. □

Lemma 5.4. *Let α, β be two open covers of X. If diam(β) is a Lebesgue number of α, then $\alpha < \beta$ and $h(T, \alpha) \leq h(T, \beta)$.*

Proof. Let $B \in \beta$, pick $x \in B$, then $B \subseteq B(x, \operatorname{diam}(\beta)) \subseteq A$ for some $A \in \alpha$, hence $\alpha < \beta$. The second inequality follows from the definition of entropy. □

Lemma 5.5. *Let (α_n) be a sequence of open covers of X with diam$(\alpha_n) \to 0$ as $n \to \infty$, then*

$$h(T) = \lim_{n\to\infty} h(T, \alpha_n).$$

Proof. It suffices to prove for any open cover α,

$$h(T, \alpha) \leq \varliminf_{n\to\infty} h(T, \alpha_n).$$

Since $\mathrm{diam}(\alpha_n) \to 0$ as $n \to \infty$, when n is large, $\mathrm{diam}(\alpha_n)$ is a Lebesgue number of α, hence $h(T, \alpha) \leq h(T, \alpha_n)$ for n large, this completes the proof. □

Lemma 5.6. *If diam$\big(\bigvee_{i=0}^{n-1} T^{-i}\alpha\big) \to 0$ as $n \to \infty$, then $h(T) = H(T, \alpha)$.*

Proof. We first check an identity

$$h(T, \alpha) = h\big(T, \bigvee_{i=0}^{n-1} T^{-i}\alpha\big) \text{ for all } n \in \mathbb{N}.$$

By definition

$$\begin{aligned} h\big(T, \bigvee_{i=0}^{n-1} T^{-i}\alpha\big) &= \lim_{m\to\infty} \frac{1}{m} H\big(\bigvee_{j=0}^{m-1} T^{-j}\big(\bigvee_{i=0}^{n-1} T^{-i}\alpha\big)\big) \\ &= \lim_{m\to\infty} \frac{1}{m} H\big(\bigvee_{i=0}^{m+n-2} T^{-i}\alpha\big) \\ &= \lim_{m\to\infty} \frac{1}{m+n-1} H\big(\bigvee_{i=0}^{m+n-2} T^{-i}\alpha\big) = h(T, \alpha). \end{aligned}$$

Applying the above lemma, we complete the proof. □

Notice that the above definition of topological entropy is completely topological, it is Rufus Bowen who found an equivalent definition which may have more apparent dynamical interpretation.

Let d be the metric on X. For $n \in \mathbb{N}$, define

$$d_n(x, y) := \max_{0\leq i\leq n-1} d(T^i x, T^i y) \text{ for } x, y \in X,$$

then d_n is again a metric. For $x \in X$ and $\epsilon > 0$, define

$$B_n(x, \epsilon) := \{y \in X : d_n(x, y) < \epsilon\},$$

and call it a Bowen ball. Define

$$N_n(\epsilon) := \inf\{k : \exists x_1, x_2, \cdots, x_k \text{ s.t. } \bigcup_{i=1}^{k} B_n(x_i, \epsilon) \supseteq X\}.$$

Proposition 5.1. $h(T) = \lim_{\epsilon\to 0} \overline{\lim}_{n\to\infty} \frac{1}{n}\log N_n(\epsilon)$.

For convenience, let us denote the right hand side by $h_B(T)$, to mean the definition of Bowen. Before we prove this proposition, we mention another dual but also equivalent definition as follows.

Define

$$S_n(\epsilon) = \sup\{k : \exists x_1, x_2, \cdots, x_k \in X, \text{s.t. } B_n(x_i,\epsilon) \text{ are pairwisely disjoint}\}.$$

Remark: $N_n(2\epsilon) \le S_n(\epsilon) \le N_n(\epsilon)$.

Proof. Assume $x_1, \cdots, x_{S_n(\epsilon)} \in X$ such that $B_n(x_i,\epsilon)$ are pairwisely disjoint. We claim that $\{B_n(x_1, 2\epsilon), \cdots, B_n(x_{S_n(\epsilon)}, 2\epsilon)\}$ is an open cover of X. Otherwise if $\tilde{x} \in X\backslash\bigcup_{i=1}^{S_n(\epsilon)} B_n(x_i, 2\epsilon)$, then $B_n(\tilde{x},\epsilon)$ is disjoint with $B_n(x_1,\epsilon), \cdots, B_n(x_{S_n(\epsilon)},\epsilon)$, contradicting with the definition of $S_n(\epsilon)$, this proves the first inequality. On the other hand, assume $B_n(y_1,\epsilon), \cdots, B_n(y_k,\epsilon)$ are Bowen balls such that $X \subseteq \bigcup_{i=1}^{k} B_n(y_i,\epsilon)$, then each $B_n(y_i,\epsilon)$ can contain at most one x_j since $d_n(x_j, x_{j'}) \ge 2\epsilon$ if $j \ne j'$, hence $S_n(\epsilon) \le k$, taking infimum over all such k, we have $S_n(\epsilon) \le N_n(\epsilon)$. □

Write $S(\epsilon) = \overline{\lim}_{n\to\infty} \frac{1}{n}\log S_n(\epsilon)$, it is clear if $\epsilon_1 < \epsilon_2$, then $S(\epsilon_2) \le S(\epsilon_1)$. Combining this fact with the above remark, we immediately have

$$\lim_{\epsilon\to 0} \overline{\lim}_{n\to\infty} \frac{1}{n}\log N_n(\epsilon) = \lim_{\epsilon\to 0} \overline{\lim}_{n\to\infty} \frac{1}{n}\log S_n(\epsilon).$$

We will use the following lemma to relate our previous definition of the entropy of an open cover and Bowen's notation.

Lemma 5.7. *Let (X,T) be a TDS, then*

(i) Let α be an open cover of X. Let δ be a Lebesgue number of α, then

$$N\big(\bigvee_{i=0}^{n-1} T^{-i}\alpha\big) \le N_n(\delta).$$

(ii) Let β be an open cover of X with diam$(\beta) < \epsilon$, then

$$N_n(\epsilon) \le N\big(\bigvee_{i=0}^{n-1} T^{-i}\beta\big).$$

Proof. (i) Assume that $X \subseteq \bigcup_{i=1}^{N_n(\delta)} B_n(x_i,\delta)$ for some $x_1, \cdots, x_{N_n(\delta)} \in X$. Notice that for $x \in X$, $B_n(x,\delta) = \bigcap_{i=0}^{n-1} T^{-i}B(T^i x,\delta)$. Since δ is a Lebesgue number of α, we have $B(T^i x,\delta)$ is a subset of some element of α, hence $B_n(x,\delta)$ is a subset of some element of $\bigvee_{i=0}^{n-1} T^{-i}\alpha$. In particular, $B_n(x_i,\delta) \subseteq A_i \in \bigvee_{j=0}^{n-1} T^{-j}\alpha$ for $i = 1, \cdots, N_n(\delta)$, hence $X \subseteq \bigcup_{i=1}^{N_n(\delta)} A_i$ with $A_i \in \bigvee_{j=0}^{n-1} T^{-j}\alpha$, therefore by definition

$$N\big(\bigvee_{i=0}^{n-1} T^{-i}\alpha\big) \le N_n(\delta).$$

(ii) Write $l = N\big(\bigvee_{i=0}^{n-1} T^{-i}\beta\big)$ and assume $A_1, \cdots, A_l \in \bigvee_{i=0}^{n-1} T^{-i}\beta$ is an open cover of X. For $i = 1, \cdots, l$, pick $x_i \in A_i$, then it's easy to see $A_i \subseteq B_n(x_i, \epsilon)$, hence $X \subseteq \bigcup_{i=1}^{l} B_n(x_i, \epsilon)$, which implies $N_n(\epsilon) \leq l$. □

Corollary 5.7.1. *Let (X, T) be a TDS. For $\epsilon > 0$, let $\alpha_\epsilon = \{$all open balls of radius $\epsilon\}$, then*

$$N\big(\bigvee_{i=0}^{n-1} T^{-i}\alpha_\epsilon\big) \leq N_n(\epsilon) \leq N\big(\bigvee_{i=0}^{n-1} T^{-i}\alpha_{\frac{\epsilon}{3}}\big).$$

Proof. Notice that α_ϵ and $\alpha_{\frac{\epsilon}{3}}$ both are open covers of X and ϵ itself is a Lebesgue number of α_ϵ, then the corollary follows by applying the above lemma. □

Now we can prove that the two definitions of topological entropy coincide, that is $h(T) = h_B(T)$.

Proof of Proposition 5.1. By the above corollary, we have

$$\frac{1}{n}\log N\big(\bigvee_{i=0}^{n-1} T^{-i}\alpha_\epsilon\big) \leq \frac{1}{n}\log N_n(\epsilon) \leq \frac{1}{n}\log N\big(\bigvee_{i=0}^{n-1} T^{-i}\alpha_{\frac{\epsilon}{3}}\big),$$

letting $n \to \infty$, we have

$$h(T, \alpha_\epsilon) \leq \overline{\lim_{n\to\infty}} \frac{1}{n}\log N_n(\epsilon) \leq h(T, \alpha_{\frac{\epsilon}{3}}),$$

taking $\epsilon = \frac{1}{n}$ and letting $n \to \infty$, by Lemma 5.5, we complete the proof. □