Far East J. Dynamical Systems **8**(2) (2006), 193-208
This paper is available online at http://www.pphmj.com

CONDITIONAL ENTROPY WITH A COSET PARTITION

WEN-CHIAO CHENG

Department of Mathematics
National Chung Cheng University
Min-Hsiung, Chia-Yi, 621, Taiwan
e-mail: wccheng@math.ccu.edu.tw

Abstract

In this paper we consider the relationship of conditional entropy and topological entropy on a compact metric topological group. According to recent studies, the variance of conditional entropy with right coset partition on a measurable closed T-invariant set is restricted by the topological entropy of the original set and its complement set. Here we prove more advanced results of such a relationship using Misiurewicz's method. The variational reformation of topological entropy is necessary when we show the variational principle. Finally we give a similar version of the Shannon-McMillan-Breiman theorem in the case of a right coset T-invariant action.

1. Introduction

In 1964, Adler, Konheim and McAndrew [1] introduced the topological entropy of a continuous self-mapping T on a compact space as analogous to the measure-theoretic entropy. Later, Bowen and Dinaburg presented an equivalent approach to the notion of entropy when the domain of considered transformation is a metrizable space. The problem of its relations with measure-theoretic entropy, defined in 1958 by

2000 Mathematics Subject Classification: 37A35, 37B40.

Keywords and phrases: conditional entropy, variational principle, Shannon-McMillan-Breiman theorem.

Received January 25, 2006

Kolmogorov, was progressively solved by various authors during the first half of the seventies: the variational principle states that the topological entropy of (X, T) is equal to the supremum of the entropies of invariant measures.

When a considered mapping T is invertible we can show that $h(T^{-1})$ equals $h(T)$ according to Bowen's result. However, if the mapping T is non-invertible, then there are several other possibilities which lead to several entropy-like invariants for non-invertible maps. For example, Hurley [4], Langevin and Przytycki [5], Langevin and Walczak [6] and Nitecki and Przytycki [7] studied different entropy-like invariants. In an earlier paper with Cheng and Newhouse [3], we defined two other such invariants, one "topological" and the other measure-theoretic in nature, and proved a number of interesting results, in particular a "variational principle" relating these two invariants.

With this motivation, we were interested in the situation when a measurable closed T-invariant subgroup B exists. We concentrate on conditional entropy under right coset partition with kernel B, $h_\mu(T \mid [B])$ when T is a continuous and endomorphism self-map of compact metric topological group X. These definitions enable us to obtain a local version of the relations. We called it *variational inequality* and denoted

$$h_{top}(T \mid B) \le \sup_{\mu \in M(X, T)} h_\mu(T \mid [B]) \le h_{top}(T \mid B) + h_{top}(T \mid cl(X \setminus B)),$$

where $cl(X \setminus B)$ denotes the closure of $X \setminus B$.

The aim of this paper is a generalization of the above mentioned entropy-like invariant defined by author [2] to the same space. A more accurate definition is introduced which leads to the variational principle. The same technique of topological dynamics is sufficiently applied to incorporate variational inequality into the variational principle. Since the Shannon-McMillan-Breiman theorem is used in many problems related to the metric entropy map of an ergodic measure, we show the analogue of this well-known result to the case of right coset T-invariant action. Our proof follows the method of Petersen's book [8].

2. Conditional Entropy

Assume that $(X, d, \cdot)$ is a compact metric topological group with metric d and group action $\cdot$ and $T : X \to X$ is continuous and an endomorphism. Assume B is a measurable closed T-invariant subgroup of X, and T-invariant means $T^{-1}B = TB = B$, then we can define the topological entropy of T for $B \cdot x$ to be

$$h_{top}(T \mid B \cdot x) = \lim_{\varepsilon \to 0} \limsup_{n \to \infty} \frac{1}{n} \log r(n, \varepsilon, B \cdot x)$$

$$= \sup_{\text{open cover } \beta} \limsup_{n \to \infty} \frac{1}{n} \log \aleph\left(\bigvee_{i=0}^{n-1} T^{-i}\beta \mid_{B \cdot x}\right),$$

where $r(n, \varepsilon, B \cdot x)$ is the minimal cardinality of (n, ε) spanning set for $B \cdot x$ or the maximal cardinality of (n, ε) separating set for $B \cdot x$ and $\aleph(\beta \mid_{B \cdot x})$ is the minimal cardinality of subcover of β which can cover $B \cdot x$. For those concepts and notations of (n, ε) spanning set or separating set and the refinement of open cover, we recommend [10] and [8].

Lemma 1. $h_{top}(T^m \mid B \cdot x) = m \cdot h_{top}(T \mid B \cdot x)$ *for all positive integers* m.

Assume that $(X_i, d_i, \cdot)$ is a compact metric topological group and $T_i : X_i \to X_i$ is continuous and an endomorphism. Let B_i be a measurable closed T_i-invariant subgroup of X_i, for $i = 1, 2$. Similar result of product rule of topological entropy, we have

Lemma 2. *Under the same assumption as above, we get*

$$h_{top}(T_1 \times T_2 \mid B_1 \cdot x_1 \times B_2 \cdot x_2) = h_{top}(T_1 \mid B_1 \cdot x_1) + h_{top}(T_2 \mid B_2 \cdot x_2).$$

Next, let $T : (X, \mathcal{B}, \mu) \to (X, \mathcal{B}, \mu)$ be a measure preserving transformation (m.p.t.) of probability space $(X, \mathcal{B}, \mu)$ (i.e., if $A \in \mathcal{B}$, then $T^{-1}A \in \mathcal{B}$ and $\mu(T^{-1}A) = \mu(A)$). Again, B is a measurable closed T-invariant subgroup.

Lemma 3. *$B \cdot x$ is a measurable closed set for all $x \in X$.*

Let $E(\phi \mid \Im)$ be the conditional expectation of ϕ given a sub-σ-algebra $\Im$ and define the conditional information function of a countable partition ζ given a $\Im \subset \mathcal{B}$ to be

$$I_{\zeta \mid \Im}(x) = -\sum_{A \in \zeta} \log E(\chi_A \mid \Im)\chi_A(x),$$

where χ_A is the characteristic function of A. The conditional entropy of ζ given $\Im$ is defined as

$$H(\zeta \mid \Im) = \int_X I_{\zeta \mid \Im}(x) d\mu.$$

For these general entropy concepts, we recommend Petersen's book, see [8]. Then we consider the conditional entropy of any finite partition α w.r.t. this special right coset partition.

Definition 1. The *conditional entropy* of any partition α w.r.t. the right coset partition $[B] = \{B \cdot x : x \in X\}$ is defined as

$$H_\mu(\alpha \mid [B]) = H_\mu(\alpha \mid \Im([B])) = \int_X I_{\alpha \mid \Im([B])}(x) d\mu,$$

where $\Im([B])$ is the σ-algebra generated by this partition $[B]$.

Suppose that x and y are two different points of X such that the B-cosets of Tx and Ty are the same, i.e., $B \cdot Tx = B \cdot Ty$. Then $T(xy^{-1}) \in B$, and because B is T-invariant, which implies $xy^{-1} \in B$, so that $B \cdot x = B \cdot y$. This shows that the induced map on cosets is one-to-one.

Lemma 4. *$T^{-n}[B] = [B]$ for all positive integers n.*

Given a finite partition α, let $\alpha^n = \vee_{i=0}^{n-1} T^{-i}\alpha$ and $\alpha_i^j = \vee_{k=i}^{j} T^{-k}\alpha$.

Lemma 5. *The function $a_n = H_\mu(\alpha^n \mid [B])$ is subadditive.*

Definition 2 (Conditional Metric Entropy on $[B]$)**.** The *conditional entropy* of α given $[B]$ is the number

$$h_\mu(T \mid [B], \alpha) = \lim_{n\to\infty} \frac{1}{n} H_\mu(\alpha^n \mid [B]) = \inf_{n\to\infty} \frac{1}{n} H_\mu(\alpha^n \mid [B])$$

and we define the conditional entropy of T with respect to μ and $[B]$ to be

$$h_\mu(T \mid [B]) = \sup_{\alpha} h_\mu(T \mid [B], \alpha),$$

where α ranges over all finite partitions of X.

Next, we only describe those propositions of this conditional metric entropy on $[B]$. The reader can see [2] for more detail.

Lemma 6. *The conditional entropy $h_\mu(T \mid [B])$ is a measure-theoretic conjugacy invariant.*

Lemma 7. *For each positive integer r, $h_\mu(T^r \mid [B]) = r \cdot h_\mu(T \mid [B])$.*

Let β be a sub-algebra of $\mathcal{B}$. Then we define $\mathcal{B}(\beta)$ to be the σ-algebra generated by β. Let C be a finite sub-σ-algebra of $\mathcal{B}$, say $C = \{C_i : i = 1, 2, \ldots, n\}$. Then the non-empty sets of the form $B_1 \cap B_2 \cap \cdots \cap B_n$, where $B_i = C_i$ or $X \setminus C_i$, form a finite partition of X. We denote it by $\alpha(C)$ and define $h_\mu(T \mid [B], C) = h_\mu(T \mid [B], \alpha(C))$.

Lemma 8. *Let $(X, \mathcal{B}, \mu)$ be a probability space. If $\mathcal{B}_0$ is a sub-algebra of $\mathcal{B}$ with $\mathcal{B}(\mathcal{B}_0) = \mathcal{B}$, then for m.p.t. $T : X \to X$, we have*

$$h_\mu(T \mid [B]) = \sup h_\mu(T \mid [B], A),$$

where the supremum is taken over all finite sub-algebras A of $\mathcal{B}_0$.

Lemma 9. *Let $(X, \mathcal{B}, \mu)$ be a probability space and let $\{A_n\}_1^\infty$ be an increasing sequence of finite sub-algebras of $\mathcal{B}$ such that $\bigvee_{n=1}^\infty A_n = \mathcal{B}$. If $T : X \to X$ is a m.p.t., then*

$$h_\mu(T \mid [B]) = \lim_{n\to\infty} h_\mu(T \mid [B], A_n).$$

Again, we let $E(\phi \mid \Im)$ be the conditional expectation of ϕ given $\Im$.

Lemma 10. *Let α_i be a finite partition and $\Im_i$ be a sub-σ-algebra of*

$(X_i, \mathcal{B}_i, m_i)$, for $i = 1, 2$. Then

$$E(\chi_{A\times B} \mid \mathfrak{I}) = E(\chi_A \mid \mathfrak{I}_1) \cdot E(\chi_B \mid \mathfrak{I}_2), \text{ a.e.},$$

where $\mathfrak{I} = \mathfrak{I}_1 \times \mathfrak{I}_2$, $A \in \alpha_1$, $B \in \alpha_2$.

Theorem 1. *Let $(X_1, \mathcal{B}_1, m_1)$ and $(X_2, \mathcal{B}_2, m_2)$ be probability spaces and let $T_1 : X_1 \to X_1$ and $T_2 : X_2 \to X_2$ be m.p.t. Then*

$$h_\mu(T_1 \times T_2 \mid [B_1 \times B_2]) = h_{m_1}(T_1 \mid [B_1]) + h_{m_2}(T_2 \mid [B_2]),$$

where $\mu = m_1 \times m_2$ and B_i is a measurable closed T_i-invariant subgroup of X_i, $i = 1, 2$.

Theorem 2. *Let T be a m.p.t. and endomorphism of the probability space $(X, \mathcal{B}, \mu)$. Then the map $\mu \to h_\mu(T \mid [B])$ is affine where α is any finite partition of X. Hence, so is the map $\mu \to h_\mu(T \mid [B])$, i.e., for all $0 < \lambda < 1$, and when μ_1 and μ_2 are both invariant measures, we have*

$$h_{\lambda\mu_1+(1-\lambda)\mu_2}(T \mid [B]) = \lambda \cdot h_{\mu_1}(T \mid [B]) + (1-\lambda) \cdot h_{\mu_2}(T \mid [B]).$$

3. Variational Principle

Lemma 11. *If ζ and η are two finite partitions of X, then*

$$h_\mu(\zeta \mid [B]) \le h_\mu(\eta \mid [B]) + H_\mu(\zeta \mid \eta).$$

Let ζ be an arbitrary decomposition of the Lebesgue space X and $X|_\zeta$ be the factor space. Let the factor map $\pi : X \to X|_\zeta$ be $\pi(x) = C$, where $x \in C \in \zeta$. Then

$$\mu(A) = \int_{X|\zeta} \mu_C(A \cap C) d\pi_* \mu,$$

where μ_C is the conditional measure on C.

Lemma 12. *Let α be a partition of $(X, \mathcal{B}, \mu)$, consider the factor map $\pi_x : X \to X|_{B\cdot x}$ and let μ_x be the conditional measure of μ on $B \cdot x$. Then*

$$H_{\mu}\left(\bigvee_{i=0}^{n-1} T^{-i}\alpha \,|\, [B]\right) = \int_{X|[B]} H_{\mu_x}\left(\bigvee_{i=0}^{n-1} T^{-i}\alpha\right) d\pi_{x*}\mu.$$

Lemma 13. *Let* $\eta = \{B_0, B_1, \ldots, B_k\}$ *be a partition of* X *such that* $\beta = (B_0 \cup B_1, \ldots, B_0 \cup B_k)$ *is an open cover of* X. *Then*

$$\aleph\left(\bigvee_{i=0}^{n-1} T^{-i}\eta\right)\Bigg|_Y \leq \aleph\left(\bigvee_{i=0}^{n-1} T^{-i}\beta\right)\Bigg|_Y \cdot 2^n$$

for any subset Y *of* X.

Lemma 14 [10]**.** *Assume that* $1 < q < n$, *for* $0 \leq j \leq q-1$, *and put* $a(j) = \left[\frac{(n-j)}{q}\right]$, *where* $[b]$ *denotes the integer part of* b. *We have*:

(1) *Fix* $0 \leq j \leq q-1$. *Then we have*

$$\{0, 1, 2, \ldots, n-1\} = \{j + rq + i \,|\, 0 \leq r \leq a(j) - 1,\ 0 \leq i \leq q-1\} \cup S,$$

where $S = \{0, 1, \ldots, j-1, j + a(j)q, j + a(j)q + 1, \ldots, n-1\}$ *and the cardinality of* S *is at most* $2q$.

(2) *The numbers* $\{j + rq \,|\, 0 \leq j \leq q-1,\ 0 \leq r \leq a(j) - 1\}$ *are all distinct and are all no greater than* $n - q$.

Now we are ready to show the relation between this conditional entropy and topological entropy for this measurable closed invariant subgroup. Here, the main technique used is the construction made by M. Misiurewicz.

Theorem 3 (Variational Principle)**.** *Let* $T : X \to X$ *be an endomorphism and continuous map of a compact metric topological group* X *and let* B *be a measurable closed* T-*invariant subgroup. Then*

$$\sup_{\mu \in M(X, T)} h_{\mu}(T \,|\, [B]) = \sup_{x \in X} h_{top}(T \,|\, B \cdot x),$$

where $M(X, T)$ *is the collection of all invariant measures* μ *under* T.

Proof. Part 1. Let $\mu \in M(X, T)$. Then we first show that

$$h_{\mu}(T \,|\, [B]) \leq \sup_{x \in X} h_{top}(T \,|\, B \cdot x).$$

Let $\zeta = \{A_1, \ldots, A_k\}$ be a finite partition of X. Choose $\varepsilon > 0$ so that $\varepsilon < \frac{1}{k \log k}$. Then we can choose compact sets $B_j \subset A_j$, $1 \le j \le k$, with $\mu(A_j \setminus B_j) < \varepsilon$ and $B_i \cap B_j = \varnothing$ if $i \ne j$. Let $\eta = \{B_0, B_1, \ldots, B_k\}$, where $B_0 = X \setminus \bigcup_{j=1}^k B_j$. We have $\mu(B_0) < k\varepsilon$, and

$$\begin{aligned}
& H_\mu(\zeta \mid \eta) \\
&= -\sum_{i=0}^{k} \mu(B_i) \sum_{j=1}^{k} \frac{\mu(B_i \cap A_j)}{\mu(B_i)} \log \frac{\mu(B_i \cap A_j)}{\mu(B_i)} \\
&= -\mu(B_0) \sum_{j=1}^{k} \frac{\mu(B_0 \cap A_j)}{\mu(B_0)} \log \frac{\mu(B_0 \cap A_j)}{\mu(B_0)} \quad \text{for } i \ne 0,\ \frac{\mu(B_i \cap A_j)}{\mu(B_i)} = 0 \text{ or } 1 \\
&\le \mu(B_0) \log k \\
&< k\varepsilon \log k < 1.
\end{aligned}$$

So we have $H_\mu(\zeta \mid \eta) < 1$.

Then $\beta = \{B_0 \cup B_1, \ldots, B_0 \cup B_k\}$ is an open cover of X. We have if $n \ge 1$, $H_{\mu_x}(\vee_{i=0}^{n-1} T^{-i}\eta) \le \log \aleph(\vee_{i=0}^{n-1} T^{-i}\eta \mid_{B \cdot x})$, where μ_x is the conditional measure of μ on $B \cdot x$ and $\aleph(\vee_{i=0}^{n-1} T^{-i}\eta \mid_{B \cdot x})$ denotes the number of nonempty set in the partition $\vee_{i=0}^{n-1} T^{-i}\eta$ under $B \cdot x$. Let $\pi_x : X \to X \mid_{B \cdot x}$ be the factor map, by Lemmas 13 and 14

$$\begin{aligned}
H_\mu\left(\bigvee_{i=0}^{n-1} T^{-i}\eta \mid [B]\right) &= \int_{X \mid [B]} H_{\mu_x}\left(\bigvee_{i=0}^{n-1} T^{-i}\eta\right) d\pi_{x*}\mu \\
&\le \sup_{x \in X} H_{\mu_x}\left(\bigvee_{i=0}^{n-1} T^{-i}\eta\right) \\
&\le \sup_{x \in X} \log\left(\aleph\left(\bigvee_{i=0}^{n-1} T^{-i}\eta \mid_{B \cdot x}\right)\right) \\
&\le \sup_{x \in X} \log\left(\aleph\left(\bigvee_{i=0}^{n-1} T^{-i}\beta \mid_{B \cdot x}\right) \cdot 2^n\right).
\end{aligned}$$

Let this inequality be divided by n and n approach to infinity, therefore

$$h_{\mu}(\eta \mid [B]) \leq \sup_{x \in X} h_{top}(T \mid B \cdot x, \beta) + \log 2$$

$$\leq \sup_{x \in X} h_{top}(T \mid B \cdot x) + \log 2.$$

So by Lemma 11

$$h_{\mu}(\zeta \mid [B]) \leq h_{\mu}(\eta \mid [B]) + H_{\mu}(\zeta \mid \eta)$$

$$\leq \sup_{x \in X} h_{top}(T \mid B \cdot x) + \log 2 + 1.$$

This gives

$$h_{\mu}(T \mid [B]) \leq \sup_{x \in X} h_{top}(T \mid B \cdot x) + \log 2 + 1$$

for all $\mu \in M(X, T)$.

This inequality holds for T^n, which implies that

$$n \cdot h_{\mu}(T \mid [B]) \leq n \cdot \sup_{x \in X} h_{top}(T \mid B \cdot x) + \log 2 + 1.$$

We divide by n and let n approach to infinity. Hence,

$$h_{\mu}(T \mid [B]) \leq \sup_{x \in X} h_{top}(T \mid B \cdot x)$$

which implies

$$\sup_{x \in X} h_{\mu}(T \mid [B]) \leq \sup_{x \in X} h_{top}(T \mid B \cdot x).$$

Part 2. Given $\varepsilon > 0$ and $x \in X$, we wish to produce a T-invariant probability measure μ_x such that

$$h_{\mu_x}(T \mid [B]) \geq h_{top}(T \mid B \cdot x, \varepsilon),$$

where

$$h_{top}(T \mid B \cdot x, \varepsilon) = \limsup_{n \to \infty} \frac{1}{n} \log r(n, \varepsilon, B \cdot x).$$

Here, $r(n, \varepsilon, B \cdot x)$ is the maximal cardinality of (n, ε) separating set for $B \cdot x$.

Choose sequences $n_i \to \infty$ such that

$$h_{top}(T \mid B \cdot x, \varepsilon) = \lim_{i \to \infty} \frac{1}{n_i} \log r(n_i, \varepsilon, B \cdot x).$$

Let $E_{i,x}$ denote a maximal (n_i, ε)-separated set for $B \cdot x$ such that $card\, E_{i,x} = r(n_i, \varepsilon, B \cdot x)$. Thus we have

$$h_{top}(T \mid B \cdot x, \varepsilon) = \lim_{i \to \infty} \frac{1}{n_i} \log card\, E_{i,x}.$$

Letting δ_z denote the point mass at point $z \in X$, let

$$\sigma_i = \frac{1}{card\, E_{i,x}} \sum_{z \in E_{i,x}} \delta_z$$

and

$$\mu_i = \frac{1}{n_i} \sum_{j=0}^{n_i - 1} \sigma_i \circ T^{-j}$$

and, let

$$\mu = \lim_{i \to \infty} \mu_i.$$

Now we choose a partition $\alpha = \{A_1, A_2, \ldots, A_k\}$ of X so that $\operatorname{diam}(A_i) < \varepsilon$ and $\mu(\partial A_i) = 0$ for $1 \le i \le k$. Since no member of $\bigvee_{i=0}^{n-1} T^{-i}\alpha$ can contain more than one member of $E_{i,x}$, as in Lemma 14

$$\begin{aligned}
\log r(n, \varepsilon, B \cdot x) &= H_{\sigma_n}\left(\bigvee_{i=0}^{n-1} T^{-i}\alpha\right) \\
&= H_{\sigma_n}\left(\bigvee_{i=0}^{n-1} T^{-i}\alpha \mid [B]\right) \\
&= H_{\sigma_n}\left(\bigvee_{i=0}^{n-1} T^{-i}\alpha \mid T^{-k}([B])\right) \quad \text{for any } k \\
&\le \sum_{r=0}^{a(j)-1} H_{\sigma_n}\left(T^{-(rq+j)}\left(\bigvee_{i=0}^{q-1} T^{-i}\alpha \mid [B]\right)\right) + \sum_{l \in S} H_{\sigma_n}(T^{-l}\alpha)
\end{aligned}$$

$$\leq \sum_{r=0}^{a(j)-1} H_{\sigma_n \circ T^{-(rq+j)}}\left(\bigvee_{i=0}^{q-1} T^{-i}\alpha \,|\, [B]\right) + 2q \log(l).$$

Sum this inequality over j from 0 to $q - 1$ and by Lemma 14, then

$$q \log r(n,\ \varepsilon,\ B \cdot x) \leq \sum_{p=0}^{n-1} H_{\sigma_n \circ T^{-p}}\left(\bigvee_{i=0}^{q-1} T^{-i}\alpha \,|\, [B]\right) + 2q^2 \log(l).$$

If we divide by n and use Lemma 13 and the concavity of $-x \log x$, then we can get

$$\frac{q}{n} \log r(n,\ \varepsilon,\ B \cdot x) \leq H_{\mu_n}\left(\bigvee_{i=0}^{q-1} T^{-i}\alpha \,|\, [B]\right) + \frac{2q^2}{n} \log(l). \qquad (1)$$

Since the members of $\bigvee_{i=0}^{q-1} T^{-i}\alpha$ have boundaries of μ-measure zero, by Lemma 15 we can claim that

$$\lim_{j\to\infty} H_{\mu_{n_j}}\left(\bigvee_{i=0}^{q-1} T^{-i}\alpha \,|\, [B]\right) = H_{\mu}\left(\bigvee_{i=0}^{q-1} T^{-i}\alpha \,|\, [B]\right).$$

Therefore replacing n by n_j in (1) and letting j go to infinity we have

$$q \cdot r(\varepsilon,\ B \cdot x) \leq H_{\mu}\left(\bigvee_{i=0}^{q-1} T^{-i}\alpha \,|\, [B]\right) = H_{\mu}(\alpha_0^{q-1} \,|\, [B]),$$

where $r(\varepsilon,\ B \cdot x) = \lim_{j\to\infty} \frac{1}{n_j} \log r(n_j,\ \varepsilon,\ B \cdot x)$. We can divide by q and let q go to infinity to get

$$r(\varepsilon,\ B \cdot x) \leq h_{\mu}(T \,|\, [B],\ \alpha) \leq h_{\mu}(T \,|\, [B])$$

which implies that for each $x \in X$, we can find a dependent invariant measure μ_x such that

$$h_{top}(T \,|\, B \cdot x,\ \varepsilon) \leq h_{\mu_x}(T \,|\, [B]).$$

Then for all $x \in X$,

$$h_{top}(T \,|\, B \cdot x) \leq \sup_{\mu \in M(X, T)} h_{\mu}(T \,|\, [B])$$

which means

$$\sup_{x \in X} h_{top}(T \mid B \cdot x) \le \sup_{\mu \in M(X, T)} h_\mu(T \mid [B]).$$

Lemma 15. *For any finite partition* α, *if* $\mu_{n_j} \to \mu$, *then*

$$\lim_{j \to \infty} H_{\mu_{n_j}}\left(\bigvee_{i=0}^{q-1} T^{-i}\alpha \mid [B]\right) = H_\mu\left(\bigvee_{i=0}^{q-1} T^{-i}\alpha \mid [B]\right).$$

Proof. Since $B \cdot x$ is a closed set, $\limsup_{j \to \infty} \mu_{n_j}(B \cdot x) \le \mu(B \cdot x)$, and μ_{n_j} is supported on $B \cdot x$, $1 \le \mu(B \cdot x)$, this implies that $\mu(B \cdot x) = 1$, i.e., μ is supported on $B \cdot x$.

$$\begin{aligned}\lim_{j \to \infty} H_{\mu_{n_j}}\left(\bigvee_{i=0}^{q-1} T^{-i}\alpha \mid [B]\right) &= \lim_{j \to \infty} H_{\mu_{n_j}}\left(\bigvee_{i=0}^{q-1} T^{-i}\alpha\right) \\ &= H_\mu\left(\bigvee_{i=0}^{q-1} T^{-i}\alpha\right) \\ &= H_\mu\left(\bigvee_{i=0}^{q-1} T^{-i}\alpha \mid [B]\right).\end{aligned}$$

When this measurable closed invariant subgroup B is the whole space X, then $B \cdot x = B$. Then we can get the following lemma which is the usual variational principle.

Lemma 16. *If* $T : X \to X$ *is an endomorphism and continuous map of a compact metric topological group, then*

$$h_{top}(T) = \sup_{\mu \in M(X, T)} h_\mu(T),$$

where $h_{top}(T)$ *is the topological entropy of* T, $h_\mu(T)$ *is the measure-theoretic entropy of* T *and* $M(X, T)$ *is the collection of all invariant measures* μ *under* T.

Again, let $T : (X, \mathcal{B}, \mu) \to (X, \mathcal{B}, \mu)$ be a measure preserving transformation (m.p.t.) of probability space $(X, \mathcal{B}, \mu)$. B is a measurable closed invariant subgroup. Assume that $\{B_n\}$ is a sequence of sub-σ-

algebra of $\mathcal{B}$ and $\{X_n\}$ is a sequence of random variables. Then $\{(X_n, B_n) : n = 1, 2, 3, ...\}$ is a martingale if

(1) $B_n \subset B_{n+1}$,

(2) X_n is measurable w.r.t. B_n,

(3) $E[| X_n |] < \infty$,

(4) $E[X_{n+1} | B_n] = X_n$ a.e.

Lemma 17 (Martingale Convergence Theorem)**.** *Every* L^1 *bounded martingale converges a.e.*

Let α be a partition of X. Then we consider the conditional information function

$$I_{\alpha | \Im([B])}(x) = -\sum_{A \in \alpha} \log E(\chi_A | \Im([B]))\chi_A(x).$$

The Martingale Convergence Theorem can be applied to produce the following lemma.

Lemma 18. *If* $g_n = I_{\alpha | \Im(\alpha_1^n \vee [B])}$, *then*

$$g = \lim_{n \to \infty} I_{\alpha | \Im(\alpha_1^n \vee [B])} \text{ exists a.e. in } L^1.$$

Lemma 19. *Let* $B_\infty = \vee_{n=1}^{\infty} B_n$ *if* $\{B_n\}$ *is an increasing sequence of sub-σ-algebras of X and let* $B_\infty = \cap B_n$ *if* $\{B_n\}$ *is a decreasing sequence. Then*

$$\lim_{n \to \infty} H_\mu(\alpha | B_n) = H_\mu(\alpha | B_\infty).$$

Lemma 20. *If* α *be a finite partition, then*

$$\begin{aligned} h_\mu(T | [B], \alpha) &= \lim_{n \to \infty} H_\mu\left(\alpha | \bigvee_{l=1}^{n-1} T^{-l}\alpha \vee [B]\right) \\ &= \lim_{n \to \infty} \int I_{\alpha | \Im(\alpha_1^n \vee [B])} d\mu \\ &= \int \lim_{n \to \infty} I_{\alpha | \Im(\alpha_1^n \vee [B])} d\mu. \end{aligned}$$

Lemma 21. *If* α *and* β *are countable measurable partitions of* X *and* ς *is a sub-*σ*-algebra of* $\mathcal{B}$, *then*

$$I_{\alpha \vee \beta | \varsigma} = I_{\alpha | \varsigma} + I_{\beta | \Im(\alpha) \vee \varsigma}.$$

Next we are ready to show the similar Shannon-McMillan-Breiman theory when $[B]$ is a closed T-invariant subgroup. We recall a m.p.t. T of $(X, \mathcal{B}, \mu)$ is called *ergodic* if the only measurable set $B \in \mathcal{B}$ with $T^{-1}B = B$ satisfies $\mu(B) = 0$ or $\mu(B) = 1$.

Theorem 4 (Birkhoff Ergodic Theorem). *Suppose* $T : (X, \mathcal{B}, \mu) \to (X, \mathcal{B}, \mu)$ *is ergodic m.p.t., we have for* $\forall f \in L^1(\mu)$,

$$\lim_{n \to \infty} \frac{1}{n} \sum_{i=0}^{n-1} f(T^i x) = \int f \, d\mu \;\; a.e.$$

Theorem 5 (Conditional Shannon-McMillan-Breiman Theorem)**.** *If* $T : X \to X$ *is an ergodic m.p.t. of the probability space* $(X, \mathcal{B}, \mu)$, α *is a finite partition of* X *and* B *is a measurable closed* T*-invariant subgroup, then* $\mu(B) = 0$ *or* $\mu(B) = 1$ *and*

$$\lim_{n \to \infty} \frac{1}{n+1} I_{\alpha | \Im(\vee_{l=1}^{n} T^{-l}(\alpha) \vee [B])}(x) = h_\mu(T | [B], \alpha) \;\; a.e.$$

Proof. From ergodic definition, we have $\mu(B) = 0$ or $\mu(B) = 1$.

Let $g_n = I_{\alpha | \Im(\vee_{l=1}^{n} T^{-l}(\alpha) \vee [B])}$ for $n = 1, 2, 3, \ldots$.

From Lemma 4, we have $T^{-1}B = B$, which implies

$$\begin{aligned} I_{\alpha_0^n | \Im([B])} &= I_{T^{-1}(\alpha) \vee \cdots \vee T^{-n}(\alpha) | \Im([B])} + I_{\alpha | \Im(T^{-1}(\alpha) \vee \cdots \vee T^{-n}(\alpha) \vee [B])} \\ &= I_{\alpha_0^{n-1} | \Im([B])} \circ T + I_{\alpha | \Im(T^{-1}(\alpha) \vee \cdots \vee T^{-n}(\alpha) \vee [B])} \\ &= g_n + g_{n-1} \circ T + \cdots + g_1 \circ T^{n-1} + g_0 \circ T^n \\ &= \sum_{s=0}^{n} g_{n-s} \circ T^s, \end{aligned}$$

where

$$g_0 = \lim_{k\to\infty} I_{\alpha\,|\,\Im([B])}.$$

Let $g = \lim_{n\to\infty} I_{\alpha\,|\,\Im(\vee_{l=1}^{n-1} T^{-l}\alpha\vee[B])}$. By Lemma 18, g exists a.e. in L^1. Then we can write

$$\frac{1}{n+1} I_{\alpha_0^n\,|\,\Im([B])} = \frac{1}{n+1}\sum_{s=0}^{n} g_{n-s}\circ T^s$$

$$= \frac{1}{n+1}\sum_{s=0}^{n} g\circ T^s + \frac{1}{n+1}\sum_{s=0}^{n}(g_{n-s} - g)\circ T^s.$$

According to Theorem 4 and Lemma 20, this implies

$$\lim_{n\to\infty}\frac{1}{n+1}\sum_{s=0}^{n} g\circ T^s = \int_X g\,d\mu$$

$$= \int \lim_{n\to\infty} I_{\alpha\,|\,\Im(\alpha_1^n\vee[B])} d\mu$$

$$= h_\mu(T\,|\,[B], \alpha).$$

Therefore, we must show the following to prove this claim

$$\lim_{n\to\infty}\frac{1}{n+1}\sum_{s=0}^{n}|\, g_{n-s} - g\,|\circ T^s = 0 \text{ a.e.}$$

For each $N = 1, 2, 3, \ldots,$ let $G_N = \sup_{s\geq N}|\, g_s - g\,|$. Thus

$$\frac{1}{n+1}\sum_{s=0}^{n}|\, g_{n-s} - g\,|\circ T^s$$

$$= \frac{1}{n+1}\sum_{s=0}^{n-N}|\, g_{n-s} - g\,|\circ T^s + \frac{1}{n+1}\sum_{s=n-N+1}^{n}|\, g_{n-s} - g\,|\circ T^s$$

$$\leq \frac{1}{n+1}\sum_{s=0}^{n-N}|\, G_N\,|\circ T^s + \frac{1}{n+1}\sum_{s=n-N+1}^{n}|\, g_{n-s} - g\,|\circ T^s.$$

We fix N and let n go to infinity. Since $|g_{n-s} - g| \leq g^* + g \in L^1$, the second term above tends to 0 a.e. Similarly, $G_N \leq g^* + g \in L^1$, so we may apply Birkhoff Ergodic Theorem to the first term:

$$\lim_{n\to\infty} \frac{1}{n+1} \sum_{s=0}^{n-N} |g_{n-s} - g| \circ T^s \leq \lim_{n\to\infty} \frac{1}{n+1} \sum_{s=0}^{n-N} G_N T^s$$

$$= \int_X G_N d\mu \to 0$$

by the dominated convergence theorem and that $G_N \to 0$ a.e. which implies claim.

References

[1] P. Billingsley, Probability and Measure, John Wiley & Sons, Inc., New York, 1995.

[2] Wen-Chiao Cheng, Variational inequality relative to a measurable closed invariant set, preprint.

[3] Wen-Chiao Cheng and Sheldon Newhouse, Pre-image entropy, Ergodic Theory Dynam. Systems 25 (2005), 1091-1113.

[4] M. Hurley, On topological entropy of maps, Ergodic Theory Dynam. Systems 15 (1995), 557-568.

[5] R. Langevin and F. Przytycki, Entropie de l'image in inverse d'une application, Bull. Soc. Math. France 120 (1992), 237-250.

[6] R. Langevin and P. Walczak, Entropie d'une dynamique, C. R. Acad. Sci. Paris 312 (1991), 141-144.

[7] Z. Nitecki and F. Przytycki, Preimage entropy of mappings, Internat. J. Bifur. Chaos 9(9) (1999), 1815-1843.

[8] K. Petersen, Ergodic Theory, Cambridge University Press, 1981.

[9] V. A. Rohlin, On the fundamental ideals of measure theory, Amer. Math. Soc. Trans. 71 (1952).

[10] P. Walters, An Introduction to Ergodic Theory, Springer Lecture Notes, Vol. 458, 1982.

■