TAIWANESE JOURNAL OF MATHEMATICS
Vol. 12, No. 7, pp. 1791-1803, October 2008
This paper is available online at http://www.tjm.nsysu.edu.tw/

ESTIMATE FOR SUPREMUM OF CONDITIONAL ENTROPY ON A CLOSED SUBSET

Wen-Chiao Cheng

Abstract. This paper compares the conditional metric entropy $h_\mu(T \mid G)$, with the topological entropy, $h_{top}(T \mid G)$, of a continuous map T, where G is a closed fully T-invariant subset. The following Variational Inequality is proven,

$$h_{top}(T \mid G) \leq \sup_{\mu \in M(X,T)} h_\mu(T \mid \langle G \rangle) \leq h_{top}(T \mid G) + h_{top}(T \mid cl(X \backslash G))$$

where $M(X, T)$ is the collection of all invariant measures of X, which is an extension of the usual variational principle when $G = X$.

1. Introduction

Let (X, T) be a topological dynamical system, with X a compact metric space and metric d, and a surjective continuous map from X to itself. It is well known that invariant Borel probability measures are associated to (X, T). Measure-theoretic and topological entropies are among the most important global invariants of dynamical systems. Measure-theoretical entropy regarding an invariant measure gives the exponential growth rate of the statistically significant orbits. Adler, Konheim and McAndrew introduced the concept of topological entropy in 1965, with some indications of the analogy to measure-theoretic entropy. Topological entropy characterizes the total exponential complexity of the orbit structure with a single number.

Entropy knowledge provides a wealth of quantitative structural information about a system. The Variational Principle shows the relationship between these two kinds of entropy and has gained a lot of attention. Conditional topological and measure-theoretical entropies were recently established with a factor map and the local variational principle shows their relationship. (See [5] and [7])

Received September 18, 2006, accepted May 6, 2007.
Communicated by Song-Sun Lin.
2000 *Mathematics Subject Classification*: Primary 37A35; Secondary 37B40.
Key words and phrases: Conditional metric entropy, Topological entropy, Variational inequality.

The entropy concept can be localized by restricting in a closed subset in topological and in measure-theoretical situations. During this paper, we estimate the supremum of a special kind of conditional metric entropy by calculating topological entropy. Assume that (X, d) is a compact metric space and we denote the closed fully T-invariant subspaces of X as G. First we review the definition of conditional metric entropy giving the σ-algebra generated by G and discuss some basic propositions. Then we present an estimate of an upper bound and lower bound for this conditional entropy by the topological entropy restricted on G and the closure of the complement of G. The variational inequality derived is the consequence of the methods of M. Misiurewicz, see [4],[6], or [10].

2. Preliminaries

The general conditional entropy of an ergodic theory is usually defined as follows. Let $(X, \mathcal{B}, \mu)$ be a probability space and $E(\phi \mid \Im)$ 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

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

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

$$H_\mu(\zeta \mid \Im) = \int_X I_{\zeta|\Im}(x)\, d\mu$$

Petersen's book,[8], gives a general background in entropy theory. For any finite partition $\alpha = \{A_1, A_2, ..., A_k\}$ and $\beta = \{B_1, B_2, ..., B_p\}$ of X, let $\xi(\beta)$ be the σ-algebras generated by β. We define $H_\mu(\alpha \mid \beta) = H_\mu(\alpha \mid \xi(\beta))$. The discussion in [10] suggests the following equality.

$$H_\mu(\alpha \mid \beta) = H_\mu(\alpha \mid \xi(\beta)) = -\sum_{i=1}^{p} \mu(B_i) \sum_{j=1}^{k} \frac{\mu(B_i \cap A_j)}{\mu(B_i)} \log \frac{\mu(B_i \cap A_j)}{\mu(B_i)}$$

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 $\mu(T^{-1}A) = \mu(A)$). We also define

$$\alpha^n = \bigvee_{i=0}^{n-1} T^{-i}\alpha = \{A_{i_0} \cap T^{-1}A_{i_1} \cap ... \cap T^{-(n-1)}A_{i_{n-1}} : A_{i_j} \in \alpha, 0 \le j \le n-1\}$$

Then consider the conditional entropy of any finite partition α w.r.t. the special sub-σ-algebra generated by G, where G is a closed fully T-invariant subset of X, i.e. G is closed and $T^{-1}G = TG = G$. A simple example is the subshift of finite type on symbolic dynamics with two-sided shift.

Without loss of generality, we assume that $\langle G\rangle = \{\phi, G, X\backslash G, X\}$ and investigate $H_\mu(\alpha \mid \langle G\rangle)$, the conditional entropy of α given the sub-σ-algebra generated by G.

Lemma 2.1. *The function* $a_n = H_\mu(\alpha^n \mid \langle G\rangle)$ *is sub-additive.*

Proof. We have

$$\begin{aligned} a_{n+m} &= H_\mu(\alpha^{n+m} \mid \langle G\rangle) \\ &= H_\mu(\alpha^n \vee T^{-n}\alpha^m \mid \langle G\rangle) \\ &= H_\mu(\alpha^n \mid \langle G\rangle) + H_\mu(T^{-n}\alpha^m \mid \alpha^n \vee \langle G\rangle) \\ &\le H_\mu(\alpha^n \mid \langle G\rangle) + H_\mu(T^{-n}\alpha^m \mid \langle G\rangle) \\ &= H_\mu(\alpha^n \mid \langle G\rangle) + H_\mu(T^{-n}\alpha^m \mid T^{-n}\langle G\rangle) \\ &= H_\mu(\alpha^n \mid \langle G\rangle) + H_\mu(\alpha^m \mid \langle G\rangle) \\ &= a_n + a_m \end{aligned}$$

■

Since $H_\mu(\alpha^n \mid \langle G\rangle)$ is sub-additive, $\lim_{n\to\infty} \frac{1}{n} H_\mu(\alpha^n \mid \langle G\rangle)$ exists, then the following definition is clearly shown.

Definition 2.2. (Conditional Metric Entropy on $\langle G\rangle$) The conditional entropy of α given $\langle G\rangle$ is the number

$$h_\mu(T \mid \langle G\rangle, \alpha) = \lim_{n\to\infty} \frac{1}{n} H_\mu(\alpha^n \mid \langle G\rangle) = \inf_{n\to\infty} \frac{1}{n} H_\mu(\alpha^n \mid \langle G\rangle)$$

and we define the conditional entropy of T with respect to μ and $\langle G\rangle$ to be

$$h_\mu(T \mid \langle G\rangle) = \sup_\alpha h_\mu(T \mid \langle G\rangle, \alpha)$$

where α is any finite partition of X.

Lemma 2.3. *This conditional entropy* $h_\mu(T \mid \langle G\rangle)$ *is a measure-theoretic conjugacy invariant.*

The next lemma shows that the power rule holds for this conditional metric entropy.

Lemma 2.4. *For each positive integer* r, $h_\mu(T^r \mid \langle G\rangle) = r \cdot h_\mu(T \mid \langle G\rangle)$.

Proof. For any finite partition α, we first show that

$$h_\mu(T^r \mid \langle G\rangle, \bigvee_{i=0}^{r-1} T^{-i}\alpha) = r \cdot h_\mu(T \mid \langle G\rangle, \alpha)$$

Because

$$\begin{aligned} r \cdot h_\mu(T \mid \langle G\rangle, \alpha) &= \lim_{n\to\infty} \frac{r}{nr} H_\mu(\bigvee_{i=0}^{nr-1} T^{-i}\alpha \mid \langle G\rangle) \\ &= \lim_{n\to\infty} \frac{1}{n} H_\mu(\bigvee_{j=0}^{n-1} T^{-rj}(\bigvee_{i=0}^{r-1} T^{-i}\alpha) \mid \langle G\rangle) \\ &= h_\mu(T^r \mid \langle G\rangle, \bigvee_{i=0}^{r-1} T^{-i}\alpha) \end{aligned}$$

It follows that

$$\begin{aligned} r \cdot h_\mu(T \mid \langle G\rangle) &= r \cdot \sup_\alpha h_\mu(T \mid \langle G\rangle, \alpha) \\ &= \sup_\alpha h_\mu(T^r \mid \langle G\rangle, \bigvee_{i=0}^{r-1} T^{-i}\alpha) \\ &\le \sup_c h_\mu(T^r \mid \langle G\rangle, c) \text{ where } c \text{ is any finite partition.} \\ &= h_\mu(T^r \mid \langle G\rangle) \end{aligned}$$

On the other hand, we prove the reverse inequality. Since

$$h_\mu(T^r \mid \langle G\rangle, \alpha) \le h_\mu(T^r \mid \langle G\rangle, \bigvee_{i=0}^{r-1} T^{-i}\alpha) = r \cdot h_\mu(T^r \mid \langle G\rangle, \alpha)$$

This implies that $h_\mu(T^r \mid \langle G\rangle) = r \cdot h_\mu(T \mid \langle G\rangle)$. ■

Well-known results for this conditional metric entropy also hold, such as, product rule and affinity, see next two theorems.

Lemma 2.5. *Let α_i be a finite partition and $\Im_i$ be a sub-σ-algebra of the probability spaces $(X_i, \mathcal{B}_i, m_i)$, for $i = 1, 2$, then*

$$E(\chi_{A\times B} \mid \Im) = E(\chi_A \mid \Im_1) \cdot E(\chi_B \mid \Im_2), a.e.$$

where $\Im = \Im_1 \times \Im_2$, $A \in \alpha_1$, $B \in \alpha_2$.

Theorem 2.6. *Let* $(X_1,\mathcal{B}_1,m_1)$ *and* $(X_2,\mathcal{B}_2,m_2)$ *be probability spaces and let* $T_1 : X_1 \to X_1, T_2 : X_2 \to X_2$ *be m.p.t. Then*

$$h_\mu(T_1 \times T_2 \mid \langle G_1 \times G_2 \rangle) = h_{m_1}(T_1 \mid \langle G_1 \rangle) + h_{m_2}(T_2 \mid \langle G_2 \rangle)$$

where $\mu = m_1 \times m_2, G_i$ *is a closed fully* T_i*-invariant subspace of* $X_i, i = 1, 2$.

For the proof, please see [2].

Theorem 2.7. *Let* T *be a m.p.t. of the probability space* $(X, \mathcal{B},\mu)$ *and* G *be a closed fully* T*-invariant subset of* X*. Then the map* $\mu \to h_\mu(T \mid \langle G \rangle, \alpha)$ *is affine where* α *is any finite partition of* X*. Hence, so is the map* $\mu \to h_\mu(T \mid \langle G \rangle)$*, i.e., for all* $0 < \lambda < 1$*, and when* μ_1, μ_2 *are both invariant measures, we have*

$$h_{\lambda\mu_1+(1-\lambda)\mu_2}(T \mid \langle G \rangle) = \lambda \cdot h_{\mu_1}(T \mid \langle G \rangle) + (1 - \lambda) \cdot h_{\mu_2}(T \mid \langle G \rangle)$$

The proof follows completely from [3] and omitted.

Next, the topological entropy on each closed subspace G is as follows:

Assume that (X, d) is a compact metric space with metric d and $T : X \to X$ is a continuous function. From Bowen's paper [1], we define a new metric $d_n(x, y) = \max_{0\leq i\leq n-1} d(T^i x, T^i y), \forall x, y \in X$. Let n be a natural number, $\epsilon > 0$ and let K be a compact subset of X. A subset F of X is said to be a (n, ϵ) span K with respect to T if $\forall x \in K, \exists y \in F$ with $d_n(x, y) \leq \epsilon$, i.e.,

$$K \subset \bigcup_{y\in F} \bigcap_{i=0}^{n-1} T^{-i}\bar{B}(T^i y; \epsilon).$$

If n is a natural number, $\epsilon > 0$ and K is a compact subset of X, let $r(n, \epsilon, K)$ denote the smallest cardinality of any $(n, \epsilon)-$spanning set for K with respect to T. A subset E of K is said to be (n, ϵ) separated with respect to T if $x, y \in E, x \neq y$, implies $d_n(x, y) > \epsilon$, i.e., for $x \in E$ the set $\bigcap_{i=0}^{n-1} T^{-i}\bar{B}(T^i x; \epsilon)$ contains no other point of E. Let $s(n, \epsilon, K)$ be the largest cardinality of any (n, ϵ) separated subset of K with respect to T.

If α is an open cover of X, let $\aleph(\alpha|_G)$ denote the number of sets in a finite subcover of α with the smallest cardinality for G. The entropy of α on G is defined by $H(\alpha|_G) = \log \aleph(\alpha|_G)$ and the topological entropy w.r.t. a closed subset G is as follows:

$$\begin{aligned} h_{top}(T \mid G) &= \lim_{\epsilon\to 0} \limsup_{n\to\infty} \frac{1}{n} \log r(n, \epsilon, G) \\ &= \lim_{\epsilon\to 0} \limsup_{n\to\infty} \frac{1}{n} \log s(n, \epsilon, G) \\ &= \sup_{\text{open cover } \beta} \limsup_{n\to\infty} \frac{1}{n} \log \aleph(\bigvee_{i=0}^{n-1} T^{-i}\beta|_G) \end{aligned}$$

where $\bigvee_{i=0}^{n-1} T^{-i}\beta = \{A_{i_0} \cap T^{-1}A_{i_1} \cap ... \cap T^{-(n-1)}A_{i_{n-1}} : A_{i_j} \in \beta\}$

The topological entropy is an important invariant that characterized by a value, $h_{top}(T \mid G) \geq 0$, which is the complexity of the orbits of T. There is a lot of information on $h_{top}(T \mid G)$ in the literature. Next section shows that the supremum of conditional entropy is bounded by the topological entropy.

3. Variational Inequality

Lemma 3.1. *If ζ and η are two finite partitions of X, then*

$$h_\mu(T \mid \langle G\rangle, \zeta) \leq h_\mu(T \mid \langle G\rangle, \eta) + H_\mu(\zeta \mid \eta)$$

Measurable decompositions are necessary to show the variational inequality. See [9]. *Let ζ be an arbitrary decomposition of the Lebesgue space X and $X \mid_\zeta$ be the factor space. Let the factor map $\pi : X \to X \mid_\zeta$ 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.

Assuming that $G_1 = G$, $G_2 = X \setminus G$ and the decomposition $\zeta = \{G_1, G_2\}$ produce the following lemma.

Lemma 3.2. *Let α be a partition of $(X, \mathcal{B}, \mu)$, consider the factor map $\pi : X \to X \mid_\zeta$ and let μ_{G_j} be the conditional measure of μ on $G_j, j = 1, 2$. Then*

$$H_\mu(\bigvee_{i=0}^{n-1} T^{-i}\alpha \mid \langle G\rangle) = \int_{X|\zeta} H_{\mu_{G_j}}(\bigvee_{i=0}^{n-1} T^{-i}\alpha)\, d\pi_*\mu$$

Next two lemmas will be used in the main theorem, we refer Walter's book for the proof. See [10].

Lemma 3.3. [10] *Let $\eta = \{B_0, B_1, ..., B_k\}$ be a partition of X such that $\beta = \{B_0 \cup B_1, ..., B_0 \cup B_k\}$ is an open cover of X. Then*

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

for any closed subset Y *of* X.

Lemma 3.4. [10] *Assume that* $1 < q < n$, *for* $0 \leq j \leq q-1$, *and put* $a(j) = [\frac{(n-j)}{q}]$, *where* $[b]$ *denotes the integer part of* b. *It follows that:*

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

$$\{0,1,2,...,n-1\} = \{j + rq + i \mid 0 \leq r \leq a(j)-1, 0 \leq i \leq q-1\} \bigcup S$$

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

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

Now it is possible to show the relation between this conditional entropy and topological entropy for this closed fully T-invariant subset, G. Here, the proof of the variational principle below follows the argument of M. Misiurewicz.

Theorem 3.5. (Variational Inequality) *Let* $T : X \to X$ *be a continuous map of a compact metric space* X *and let* G *be a closed fully* T*-invariant subspace, then*

$$h_{top}(T \mid G) \leq \sup_{\mu \in M(X,T)} h_\mu(T \mid \langle G \rangle) \leq h_{top}(T \mid G) + h_{top}(T \mid cl(X \backslash G))$$

where $M(X,T)$ *is the collection of all invariant measures* μ *under* T *and* $cl(X\backslash G)$ *is the closure of* $X\backslash G$.

Proof.

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

$$h_\mu(T \mid \langle G \rangle) \leq h_{top}(T \mid G) + h_{top}(T \mid cl(X \backslash G)).$$

Let $\zeta = \{A_1,...,A_k\}$ be a finite partition of X. Choose $\epsilon > 0$ so that $\epsilon < \frac{1}{k \log k}$. Then we can choose compact sets $B_j \subset A_j, 1 \leq j \leq k$, with $\mu(A_j \setminus B_j) < \epsilon$ and $B_i \cap B_j = \phi$ if $i \neq j$. Let $\eta = \{B_0, B_1, ..., B_k\}$ where $B_0 = X \backslash \bigcup_{j=1}^k B_j$. Thus $\mu(B_0) < k\epsilon$, and we obtain

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

Therefore, $H_\mu(\zeta \mid \eta) < 1$.

Then $\beta = \{B_0 \cup B_1, ..., B_0 \cup B_k\}$ is an open cover of X. Assume $G_1 = G$, $G_2 = X \setminus G$. We have if $n \geq 1$, $H_{\mu_{G_i}}(\bigvee_{i=0}^{n-1} T^{-i}\eta) \leq \log \aleph(\bigvee_{i=0}^{n-1} T^{-i}\eta|_{G_i})$ where μ_{G_i} is the conditional measure of μ on G_i and $\aleph(\bigvee_{i=0}^{n-1} T^{-i}\eta|_{G_i})$ denotes the number of nonempty set in the partition $\bigvee_{i=0}^{n-1} T^{-i}\eta$ under $G_i, i = 1, 2$. Let $\pi : X \to X \mid_\zeta$ be the factor map where $\zeta = \{G_1, G_2\}$, by Lemma 3.2 and Lemma 3.3

$$\begin{aligned}
H_\mu(\bigvee_{i=0}^{n-1} T^{-i}\eta \mid \langle G\rangle) &= \int_{X|\zeta} H_{\mu_{G_i}}(\bigvee_{i=0}^{n-1} T^{-i}\eta)\, d\pi_*\mu \\
&\leq \int_G H_{\mu_{G_i}}(\bigvee_{i=0}^{n-1} T^{-i}\eta)\, d\pi_*\mu + \int_{cl(X\setminus G)} H_{\mu_{G_i}}(\bigvee_{i=0}^{n-1} T^{-i}\eta)\, d\pi_*\mu \\
&\leq \log(\aleph(\bigvee_{i=0}^{n-1} T^{-i}\eta|_G) + \log(\aleph(\bigvee_{i=0}^{n-1} T^{-i}\eta|_{cl(X\setminus G)}) \\
&\leq \log \aleph(\bigvee_{i=0}^{n-1} T^{-i}\beta|_G) \cdot 2^n) + \log \aleph(\bigvee_{i=0}^{n-1} T^{-i}\beta|_{cl(X\setminus G)} \cdot 2^n)
\end{aligned}$$

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

$$\begin{aligned}
h_\mu(T \mid \langle G\rangle, \eta) &\leq h_{top}(T \mid G, \beta) + \log 2 + h_{top}(T \mid cl(X\setminus G), \beta) + \log 2 \\
&\leq h_{top}(T \mid G) + h_{top}(T \mid cl(X\setminus G)) + 2\log 2
\end{aligned}$$

Then $\beta = \{B_0 \cup B_1, ..., B_0 \cup B_k\}$ is an open cover of X. Assume $G_1 = G$, $G_2 = X \setminus G$. We have if $n \geq 1$, $H_{\mu_{G_i}}(\bigvee_{i=0}^{n-1} T^{-i}\eta) \leq \log \aleph(\bigvee_{i=0}^{n-1} T^{-i}\eta|_{G_i})$ where μ_{G_i} is the conditional measure of μ on G_i and $\aleph(\bigvee_{i=0}^{n-1} T^{-i}\eta|_{G_i})$ denotes the number of nonempty set in the partition $\bigvee_{i=0}^{n-1} T^{-i}\eta$ under $G_i, i = 1, 2$. Let $\pi : X \to X \mid_\zeta$ be the factor map where $\zeta = \{G_1, G_2\}$, by Lemma 3.2 and Lemma 3.3

$$\begin{aligned}
H_\mu(\bigvee_{i=0}^{n-1} T^{-i}\eta \mid \langle G\rangle) &= \int_{X|\zeta} H_{\mu_{G_i}}(\bigvee_{i=0}^{n-1} T^{-i}\eta)\, d\pi_*\mu \\
&\leq \int_G H_{\mu_{G_i}}(\bigvee_{i=0}^{n-1} T^{-i}\eta)\, d\pi_*\mu + \int_{cl(X\setminus G)} H_{\mu_{G_i}}(\bigvee_{i=0}^{n-1} T^{-i}\eta)\, d\pi_*\mu \\
&\leq \log(\aleph(\bigvee_{i=0}^{n-1} T^{-i}\eta|_G) + \log(\aleph(\bigvee_{i=0}^{n-1} T^{-i}\eta|_{cl(X\setminus G)}) \\
&\leq \log \aleph(\bigvee_{i=0}^{n-1} T^{-i}\beta|_G) \cdot 2^n) + \log \aleph(\bigvee_{i=0}^{n-1} T^{-i}\beta|_{cl(X\setminus G)} \cdot 2^n)
\end{aligned}$$

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

$$\begin{aligned} h_\mu(T \mid \langle G\rangle, \eta) &\leq h_{top}(T \mid G, \beta) + \log 2 + h_{top}(T \mid cl(X\backslash G), \beta) + \log 2 \\ &\leq h_{top}(T \mid G) + h_{top}(T \mid cl(X\backslash G)) + 2\log 2 \end{aligned}$$

By Lemma 3.1

$$\begin{aligned} h_\mu(\zeta \mid \langle G\rangle) &\leq h_\mu(\eta \mid \langle G\rangle) + H_\mu(\zeta \mid \eta) \\ &\leq h_{top}(T \mid G) + h_{top}(T \mid cl(X\backslash G)) + 2\log 2 + 1 \end{aligned}$$

This gives $h_\mu(T \mid \langle G\rangle) \leq h_{top}(T \mid G) + h_{top}(T \mid cl(X\backslash G)) + 2\log 2 + 1$ for all $\mu \in M(X, T)$.

This inequality holds for T^n so that

$$n \cdot h_\mu(T \mid \langle G\rangle) \leq n \cdot h_{top}(T \mid G) + n \cdot h_{top}(T \mid cl(X\backslash G)) + 2\log 2 + 1$$

Divide by n and let n approach to infinity. Hence

$$h_\mu(T \mid \langle G\rangle) \leq h_{top}(T \mid G) + h_{top}(T \mid cl(X\backslash G)).$$

Part 2. Part 1 shows that

$$\sup_{\mu\in M(X,T)} h_\mu(T \mid \langle G\rangle) \leq h_{top}(T \mid G) + h_{top}(T \mid cl(X\backslash G))$$

so it is only necessary to demonstrate the opposite inequality.

Given $\epsilon > 0$, we wish to produce a T-invariant probability measure μ such that

$$h_\mu(T \mid \langle G\rangle) \geq h_{top}(T \mid G, \epsilon)$$

where

$$h_{top}(T \mid G, \epsilon) = \limsup_{n\to\infty} \frac{1}{n} \log s(n, \epsilon, G)$$

Here, $s(n, \epsilon, G)$ is the maximal cardinality of (n, ϵ) separating set for G.

Choose sequences $n_i \to \infty$ such that

$$h_{top}(T \mid G, \epsilon) = \lim_{i\to\infty} \frac{1}{n_i} \log s(n_i, \epsilon, G)$$

Let E_i denote a maximal (n_i, ϵ)-separated set in G such that $card\, E_i = s(n_i, \epsilon, G)$. Thus we have

$$h_{top}(T \mid G, \epsilon) = \lim_{i\to\infty} \frac{1}{n_i} \log card\, E_i$$

Letting δ_x denote the point mass at point $x \in X$, let

$$\sigma_i = \frac{1}{card\, E_i} \sum_{x \in E_i} \delta_x,$$

and

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

Without loss of generality, assume that

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

Next we choose a measurable partition $\alpha = \{A_1, A_2, ..., A_k\}$ of X so that diam$(A_i) < \epsilon$ and $\mu(\partial A_i) = 0$ for $1 \leq i \leq k$. Since no member of $\bigvee_{i-0}^{n-1} T^{-i}\alpha$ can contain more than one member of E_n, it follows that

$$\begin{aligned} \log s(n, \epsilon, G) &= H_{\sigma_n}(\bigvee_{i=0}^{n-1} T^{-i}\alpha) \\ &= H_{\sigma_n}(\bigvee_{i=0}^{n-1} T^{-i}\alpha \mid \langle G \rangle) \\ &\leq \sum_{r=0}^{a(j)-1} H_{\sigma_n}(T^{-(rq+j)}(\bigvee_{i=0}^{q-1} T^{-i}\alpha \mid \langle G \rangle)) + \sum_{l \in S} H_{\sigma_n}(T^{-l}\alpha) \\ &\leq \sum_{r=0}^{a(j)-1} H_{\sigma_n \circ T^{-(rq+j)}}(\bigvee_{i=0}^{q-1} T^{-i}\alpha \mid \langle G \rangle) + 2q \log(l) \end{aligned}$$

Summing over j from 0 to $q-1$, we obtain

$$q \log s(n, \epsilon, G) \leq \sum_{p=0}^{n-1} H_{\sigma_n \circ T^{-p}}(\bigvee_{i=0}^{q-1} T^{-i}\alpha \mid \langle G \rangle) + 2q^2 \log(l).$$

Dividing by n and using the concavity of $-x \log x$, we can get

$$\frac{q}{n} \log s(n, \epsilon) \leq H_{\mu_n}(\bigvee_{i=0}^{q-1} T^{-i}\alpha \mid \langle G \rangle) + \frac{2q^2}{n} \log(l)$$

Since the members of $\bigvee_{i=0}^{q-1} T^{-i}\alpha$ have boundaries of μ-measure zero, by Lemma 3.6 we can claim that

$$\lim_{j \to \infty} H_{\mu_{n_j}}(\bigvee_{i=0}^{q-1} T^{-i}\alpha \mid \langle G \rangle) = H_{\mu}(\bigvee_{i=0}^{q-1} T^{-i}\alpha \mid \langle G \rangle)$$

Therefore replacing n by n_j in last inequality and letting j go to infinity produces

$$qs(\epsilon, G) \le H_\mu(\bigvee_{i=0}^{q-1} T^{-i}\alpha \mid \langle G\rangle) = H_\mu(\alpha_0^{q-1} \mid \langle G\rangle).$$

where $s(\epsilon, G) = \lim_{j\to\infty} \frac{1}{n_j} \log s(n_j, \epsilon, G)$. Divide by q and let q go to infinity. Thus,

$$s(\epsilon, G) \le h_\mu(T \mid \langle G\rangle, \alpha) \le h_\mu(T \mid \langle G\rangle)$$

which proves the theorem. ■

Lemma 3.6. *For any finite partition* α,

$$\lim_{j\to\infty} H_{\mu_{n_j}}(\bigvee_{i=0}^{q-1} T^{-i}\alpha \mid \langle G\rangle) = H_\mu(\bigvee_{i=0}^{q-1} T^{-i}\alpha \mid \langle G\rangle)$$

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

$$\begin{aligned}
\lim_{j\to\infty} H_{\mu_{n_j}}(\bigvee_{i=0}^{q-1} T^{-i}\alpha \mid \langle G\rangle) &= \lim_{j\to\infty} H_{\mu_{n_j}}(\bigvee_{i=0}^{q-1} T^{-i}\alpha) \\
&= H_\mu(\bigvee_{i=0}^{q-1} T^{-i}\alpha) \\
&= H_\mu(\bigvee_{i=0}^{q-1} T^{-i}\alpha \mid \langle G\rangle)
\end{aligned}$$

■

In the process of proving part 1 in Theorem 3.5, we also can show this way

$$\begin{aligned}
&H_\mu(\bigvee_{i=0}^{n-1} T^{-i}\eta \mid \langle G\rangle) \\
&= \int_{X|\zeta} H_{\mu_{G_i}}(\bigvee_{i=0}^{n-1} T^{-i}\eta)\, d\pi_*\mu \\
&\le \int_G H_{\mu_{G_i}}(\bigvee_{i=0}^{n-1} T^{-i}\eta)\, d\pi_*\mu + \int_{X\setminus G} H_{\mu_{G_i}}(\bigvee_{i=0}^{n-1} T^{-i}\eta)\, d\pi_*\mu \\
&\le \mu(G)\log(\aleph(\bigvee_{i=0}^{n-1} T^{-i}\eta|_G) + (1-\mu(G))\log(\aleph(\bigvee_{i=0}^{n-1} T^{-i}\eta|_{cl(X\setminus G)}) \\
&\le \mu(G)\log\aleph(\bigvee_{i=0}^{n-1} T^{-i}\beta|_G)\cdot 2^n) + (1-\mu(G))\log\aleph(\bigvee_{i=0}^{n-1} T^{-i}\beta|_{cl(X\setminus G)}\cdot 2^n)
\end{aligned}$$

Thus, more accurate variational inequality is obtained.

Lemma 3.7. *Let $T : X \to X$ be a continuous map of a compact metric space X and let G be a closed fully T-invariant subspace, then*

$$\begin{aligned} h_{top}(T \mid G) &\leq \sup_{\mu \in M(X,T)} h_\mu(T \mid \langle G \rangle) \\ &\leq \sup_{\mu \in M(X,T)} \{\mu(G) h_{top}(T \mid G) + (1 - \mu(G)) h_{top}(T \mid cl(X \backslash G))\} \end{aligned}$$

where $M(X,T)$ is the collection of all invariant measures μ under T and $cl(X\backslash G)$ is the closure of $X\backslash G$.

If this closed T-fully invariant subset G is the whole space X, then $cl(X\backslash B) = \phi$. Then allows the following Lemma, which is the usual variational principle.

Lemma 3.8. *Let $T : X \to X$ be a continuous function of a compact metric space, 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.

ACKNOWLEDGMENT

The author would like to thank the referee for those useful comments.

REFERENCES

1. R. Bowen, Entropy for group endomorphisms and homogeneous spaces, *Trans. Amer. Math. Soc.*, **153** (1971), 401-414.
2. Wen-Chiao Cheng, Variational Inequality Relative to a Closed Invariant Subgroup, *Far East Journal of Dynamical Systems*, **9(1)** (2007), 1-16.
3. Wen-Chiao Cheng and Sheldon Newhouse, Pre-image entropy, *Ergodic Theory and Dynamical Systems*, **25** (2005), 1091-1113.
4. M. Denker, C. Grillenberger and K. Sigmund, *Ergodic Theory on Compact Spaces*, Spring Lecture Notes in Math. 527, Spring: New York, 1976.
5. T. Downarowicz and J. Serafin, Fiber entropy and conditional variational principles in compact non-metrizable spaces, *Fund. Math.*, **172** (2002), 217-247.
6. Anatole Katok and Boris Hasselblatt, *An introduction to the modern theory of dynamical systems. Encyclopedia of Mathematics and Its Applications, Vol. 54,* Cambridge University Press, Cambridge, 1995.

7. Wen Huang, Xiangdong Ye and Guohua Zhang, A local variational principle for conditional entropy, *Ergodic Theory and Dynamical Systems*, **26** (2006), 219-245.

8. K. Petersen, *Ergodic Theory*, Cambridge University Press, 1981.

9. V. A. Rohlin, *On the Fundamental Ideals of Measure Theory*, American Mathematical Society, Translation Number 71, 1952.

10. P. Walters, An Introduction to ergodic theory, *Springer Lecture Notes*, **458** (1982).

Wen-Chiao Cheng
Department of Applied Mathematics,
Chinese Culture University,
55, Hwa-Kang Road, Yang-Ming-Shan
Taipei, Taiwan 11114, R.O.C.
E-mail: zwq2@faculty.pccu.edu.tw