Pacific Journal of Mathematics

VARIATIONAL INEQUALITY FOR CONDITIONAL PRESSURE ON A BOREL SUBSET

YUAN LI, ERCAI CHEN AND WEN-CHIAO CHENG

Volume 256 No. 1 March 2012

PACIFIC JOURNAL OF MATHEMATICS
Vol. 256, No. 1, 2012

VARIATIONAL INEQUALITY FOR CONDITIONAL PRESSURE ON A BOREL SUBSET

YUAN LI, ERCAI CHEN AND WEN-CHIAO CHENG

We define topological conditional pressure on a Borel subset and investigate its properties. We then estimate the supremum of metric conditional entropy with potential energy. The usual basic properties hold for this topological conditional pressure. In particular, we find a variational inequality that is an extension of the variational principle for topological pressure.

1. Introduction

A dynamical system is a pair (X, T) in which X is a set (called a phase space) and T is a group or semigroup of self transformations of X. In the most classical case, these are the iterations of a single transformation. Usually, the space X is endowed with a structure that the acting transformations must respect.

Entropy is an important notion in dynamical systems. Kolomogorov and Sinai developed the metric or measure-theoretic entropy of a transformation based on Shannon's information theory in 1959. Topological entropy was first introduced in 1965 by Adler, Konheim and McAndrew and defined by Bowen later on a metric space. Measure-theoretic entropy measures the maximal loss of information in the iteration of finite partitions in a measure-preserving transformation. Topological entropy, on the other hand, measures the maximal exponential growth rate of orbits for an arbitrary topological dynamical system. These two notions are connected by a variational principle. This relation, which states that the topological entropy is the supremum of the metric entropies for all invariant probability measures of a given topological system, has gained a lot of attention. Good references for those entropy invariants are [Katok and Hasselblatt 1995; Walters 1982], which contain many of the early references.

Researchers have recently characterized the local structure of maps by defining entropy pairs, entropy tuples, entropy sets, or entropy points in both topological and measure-theoretical situations. Several studies examine the connection between measure-theoretic entropy notions and topological entropy ones. These studies

Wen-Chiao Cheng is the corresponding author.
MSC2010: primary 37A35, 37B40; secondary 37D35.
Keywords: variational principle, topological conditional pressure, conditional entropy.

investigate their local behavior, and in many cases, establish new variational principles. See [Downarowicz and Serafin 2002; Huang and Ye 2006; Huang et al. 2006; Romagnoli 2003] for related propositions and results.

Topological pressure is a natural generalization of topological entropy. Ruelle [1973] first introduced the concept of topological pressure for additive potentials, and applied this concept to expansive dynamical systems and formulated a variational principle for topological pressure. Walters [1975] later generalized these results to continuous maps on compact metric spaces. For an arbitrary set, we emphasize that it need not be invariant or compact, as it generalizes the notion of topological pressure proposed by Pesin and Pitskel' [1984], and the notions of lower and upper capacity topological pressures introduced by Pesin [1988]. Propositions and applications to various types of entropy and pressure are discussed in [Ledrappier 1979; Mihailescu and Urbański 2004; Misiurewicz 1976; Mummert 2007]. Theories of topological pressure, variational principle, and equilibrium states play a fundamental role in statistical mechanics, ergodic theory, and dynamical systems. See [Bowen 1975; Ruelle 1978; Walters 1982].

Assume that G is a closed T-invariant subset of T, that is, $T^{-1}G = G$ and consider the partition $\langle G\rangle = \{G, X \setminus G\}$, the supremum of the conditional entropy $h_\mu(T \mid \langle G\rangle)$ as estimated in [Cheng 2008]. This paper extends the conditional entropy to topological conditional pressure and estimates the supremum of a special kind of conditional entropy with a potential function by calculating this new pressure. Assume that (X, d) is a compact metric space and denote the closed T-invariant subspaces of X by G. We first review the definition of conditional metric entropy given the partition $\langle G\rangle = \{G, X \setminus G\}$ and discuss some basic propositions. We then present an estimate of an upper bound and lower bound for this conditional entropy with potential energy using the topological conditional pressure restricted on G and the closure of the complement of G. The resulting variational inequality derived is based on the methods of P. Walters [1982; 1975]. Then, using the results of [Pesin and Pitskel' 1984; Pesin 1988] and one more condition, the main variational inequality also holds for any T-invariant Borel subset G.

2. Basic notation and statement of results

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

$$\alpha = \{A_1, A_2, \dots, A_m\} \quad \text{and} \quad \beta = \{B_1, B_2, \dots, B_n\}$$

be finite partitions of X. Then Walters [1982] gives the equality

$$H_\mu(\alpha \mid \beta) = -\sum_{i=1}^{n} \mu(B_i) \sum_{j=1}^{m} \frac{\mu(B_i \cap A_j)}{\mu(B_i)} \log \frac{\mu(B_i \cap A_j)}{\mu(B_i)}.$$

Next, let $T:(X,\mathcal{B},\mu)\to(X,\mathcal{B},\mu)$ be a measure-preserving map of probability space $(X,\mathcal{B},\mu)$ (that is, if $A\in\mathcal{B}$, then $T^{-1}A\in\mathcal{B}$ and $\mu(T^{-1}A)=\mu(A)$), and define $\alpha^n=\bigvee_{i=0}^{n-1}T^{-i}\alpha$. Then, investigate the conditional entropy of any finite partition α with respect to the partition $\langle G\rangle=\{G, X\setminus G\}$, where G is a subset of X and $X\setminus G$ is the complement of G.

If G is a T-invariant subset of X, that is, if $T^{-1}G=G$, then the sequence $a_n=H_\mu(\alpha^n\mid\langle G\rangle)$ is subadditive. Thus, the conditional entropy of α given $\langle G\rangle$ is the value

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

and the conditional entropy of T with respect to μ and $\langle G\rangle$ is

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

where α is any finite partition of X. If $\langle G\rangle=\{X,\phi\}$, then $h_\mu(T\mid X)=h_\mu(T)$ is the usual measure-theoretic entropy.

The basic properties of conditional entropy $h_\mu(T\mid\langle G\rangle)$, such as power rule, product rule and affinity, are stated as follows; see [Cheng 2008]. A simple example is the subshift of finite type on symbolic dynamics with two-sided shift.

Lemma 2.1. *$h_\mu(T\mid\langle G\rangle)$ is a measure-theoretic conjugacy invariant.*

Lemma 2.2. *$h_\mu(T^r\mid\langle G\rangle)=r\cdot h_\mu(T\mid\langle G\rangle)$ for each positive integer r.*

Lemma 2.3. *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 measure-preserving maps. 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 $\mu=m_1\times m_2$ and G_i is a T_i-invariant subspace of X_i for $i=1,2$.

Lemma 2.4. *Let T be a measure-preserving map of the probability space $(X,\mathcal{B},\mu)$ and G be a 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 $\mu\to h_\mu(T\mid\langle G\rangle)$. In other words, for all $0<\lambda<1$ and invariant measures μ_1 and μ_2, we have*

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

Next, recall the notion of topological entropy by using spanning sets and separated sets.

Let $\epsilon>0$ and $n\in\mathbb{N}\setminus\{0\}$. Let $B(x,n,\epsilon)$ denote Bowen's ball of order n, radius ϵ and center x:

$$B(x,n,\epsilon):=\{y\in X\mid d(T^k(x),T^k(y))<\epsilon\text{ for all }0\leqslant k<n\}.$$

A set $E \subset X$ is said to be a (n, δ)-spanning subset of K if $K \subset \bigcup_{x \in E} B(x, n, \delta)$, and a (n, δ)-separated subset of K if for all $x, y \in E$, there exists a nonnegative $k < n$ such that $d(T^k(x), T^k(y)) \geqslant \delta$. Use the notation $r(n, \delta, K)$ for the minimal cardinality of a (n, δ)-spanning subset of a set $K \subset X$ and the notation $s(n, \delta, K)$ for the maximal cardinality of a (n, δ)-separated subset of K.

Topological conditional pressure is defined as follows: Let (X, d) be a compact metric space, $C(X, R)$ be the space of real-valued continuous functions of X, and $T : X \to X$ be a continuous map. For $\varphi \in C(X, R)$, G a closed subset of X, and $n \geq 1$, let $\sum_{i=0}^{n-1} \varphi(T^i x) = (S_n\varphi)(x)$ Set

$$P_n(T, \varphi, \epsilon, G) = \sup\Big\{\sum_{z \in E} \epsilon^{(S_n\varphi)(z)} : E \text{ is an } (n, \epsilon)\text{-separated subset of } G\Big\},$$

$$P(T, \varphi, \epsilon, G) = \limsup_{n \to \infty} \frac{1}{n} \log P_n(T, \varphi, \epsilon, G).$$

It is easy to show that $P(T, \varphi, \epsilon, G)$ is an increasing function of ϵ. Therefore, the topological conditional pressure of T with respect to φ and G can be defined as follows:

$$P_G(T, \varphi) = \lim_{\epsilon \to 0} P(T, \varphi, \epsilon, G).$$

This real-valued function φ is called a potential function and $\int \varphi d\mu$ is a potential energy. If $\varphi = 0$, then $P_G(T, \varphi)$ is reduced to the topological entropy on G, that is, $h_{\text{top}}(T \mid G)$. If $G = \varnothing$, we define $P_\varnothing(T, \varphi) = 0$.

The basic proposition concerning topological conditional pressure is the power rule.

Lemma 2.5. *For any positive integer k, let G be a closed subset of X. Then*

$$P_G(T^k, S_k\varphi) = k P_G(T, \varphi),$$

where $(S_k\varphi)(x) = \sum_{i=0}^{k-1} \varphi(T^i x)$.

Proof. For any fixed positive integer k, if E is an (n, ϵ)-separated set of G with respect to T^k, then E should be an (nk, ϵ)-separated subset of G with respect to T. Thus, it follows that

$$\begin{aligned}
&P_n(T^k, S_k\varphi, \epsilon, G)\\
&\quad = \sup_E\Big\{\sum_{z \in E} e^{\sum_{i=0}^{n-1}\sum_{j=0}^{k-1} \varphi(T^j(T^{ki}(z))}:\\
&\qquad\qquad E \text{ is an } (n, \epsilon)\text{-separated subset of } G \text{ with respect to } T^k\Big\}\\
&\quad \leq \sup_E\Big\{\sum_{z \in E} e^{\sum_{j=0}^{nk-1} \varphi((T^{ki+j}(z))}:\\
&\qquad\qquad E \text{ is an } (nk, \epsilon)\text{-separated subset of } G \text{ with respect to } T\Big\}\\
&\quad = P_{nk}(T, \varphi, \epsilon, G).
\end{aligned}$$

This implies that

$$P(T^k, S_k\varphi, \epsilon, G) \le kP(T, \varphi, \epsilon, G), \tag{2-1}$$

which gives the inequality

$$P_G(T^k, S_k\varphi) \le kP_G(T, \varphi).$$

To show the reverse inequality, for any $\epsilon > 0$, we choose $\delta > 0$ small enough so that $d(x, y) \leqslant \delta$ implies $\max_{1\leqslant i\leqslant k-1} d(T^i x, T^i y) < \epsilon$. If E is a (nk, ϵ)-separated subset of X for T, then E is a (n, δ)-separated subset of X for T^k. Hence $P_{nk}(T, \varphi, \epsilon, G) \leqslant P_n(T^k, S_k\varphi, \delta, G)$. Therefore

$$P_G(T^k, S_k\varphi) \geqslant kP_G(T, \varphi). \qquad \square$$

If $\varphi = 0$, then $P_G(T, \varphi)$ is equal to the topological entropy $h_{\text{top}}(T \mid G)$, which measures the orbit structure complexity of the map. The conditional entropy and topological entropy concentrated on G are related as follows.

Theorem 2.1 (variational inequality for conditional entropy [Cheng 2008]). *Let $T : X \to X$ be a continuous map of a compact metric space X and let G be a closed T-invariant subspace. Then*

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

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

If this closed T-invariant subset G is the whole space X, then $X \setminus G = \varnothing$. This allows the classical variational principle to be stated as

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

where $h_{\text{top}}(T)$ 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.

3. The compact case

Let (X, d) be a compact metric space, let $C(X, R)$ be the space of real-valued continuous functions of X and let $T : X \to X$ be a continuous map. This section provides a proof of the variational inequality.

Lemma 3.1 [Walters 1982]. *Let $a_1, \dots, a_k$ be given real numbers. If $p_i \geqslant 0$ and $\sum_{i=1}^k p_i = 1$, then*

$$\sum_{i=1}^k p_i(a_i - \log p_i) \leqslant \log\Big(\sum_{i=1}^k e^{a_i}\Big)$$

and equality holds if and only if

$$p_i = \frac{e^{a_i}}{\sum_{i=1}^{k} e^{a_i}}.$$

Lemma 3.2 [Walters 1982]. *Assume that* $1 < q < n$ *for* $0 \leqslant j \leqslant q-1$, *and set* $a(j) = [(n-j)/q]$, *where* $[b]$ *denotes the integer part of* b.

(1) *Fix* $0 \leqslant j \leqslant q-1$. *Then*

$$\{0, 1, 2, \dots, n-1\} = \bigl\{ j + rq + i \mid 0 \leqslant r \leqslant a(j), 0 \leqslant i \leqslant q-1 \bigr\} \cup S,$$

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

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

For $(X, \mathcal{B})$ a measurable space, we denote by $M(X)$ the collection of all Borel probability measures on $(X, \mathcal{B})$.

Lemma 3.3 [Pesin 1988]. *Let* X *be a compact metric space and* $\mu \in M(X)$.

(1) *If* $x \in X$ *and* $\delta > 0$, *there exists* $\delta' < \delta$ *such that* $\mu(\partial B(x; \delta')) = 0$, *where* ∂B *denotes the boundary of set* B.

(2) *If* $\delta > 0$, *there is a finite partition* $\xi = \{A_1, A_2, \dots, A_k\}$ *of* $(X, \mathcal{B}, \mu)$ *such that* $\operatorname{diam}(A_j) < \delta$ *and* $\mu(\partial A_j) = 0$ *for each* $j = 1, 2, \dots, k$.

Lemma 3.4 [Pesin 1988]. *Let* (X, d) *be a compact metric space, let* $\mu_i \in M(X)$ *for* $1 \leqslant i \leqslant n$, *and suppose* $p_i \geq 0$ *satisfy* $\sum_{i=1}^{n} p_i = 1$. *Then*

$$H_{\sum_{i=1}^{n} p_i \mu_i}(\xi) \geqslant \sum_{i=1}^{n} p_i H_{\mu_i}(\xi).$$

Again, we follow Walter's proof of the standard variational principle and simply call them the SVP arguments (as in [Walters 1982, pages 218–221]), but we must make some modifications to obtain the variational inequality.

Theorem 3.1. *Let* $T : X \to X$ *be a continuous map of a compact metric space, let* $\varphi \in C(X, R)$, *and let* G *be a closed* T*-invariant subset of* X (*that is,* $T^{-1}G = G$). *Then*

$$P_G(T, \varphi) \leq \sup_{\mu \in M(X,T)} \{h_\mu(T \mid \langle G \rangle) + \textstyle\int \varphi \, d\mu\} \leq \max\{P_G(T, \varphi), P_{\overline{X \setminus G}}(T, \varphi)\}.$$

Proof. We follow the SVP argument.

Let $\epsilon > 0$. We want to find $\mu \in M(X, T)$ with $h_\mu(T|\langle G\rangle)+\int \varphi d\mu \geq P_G(T, \varphi, \varepsilon)$, which clearly implies

$$P_G(T, \varphi) \leq \sup_{\mu \in M(X,T)} \{h_\mu(T \mid \langle G\rangle) + \textstyle\int \varphi d\mu\}.$$

Let E_n be an (n, ϵ) separated set of G with

$$\log \sum_{y \in E_n} e^{(s_n\varphi)(y)} \geq \log P_n(T, \varphi, \varepsilon, G) - 1.$$

Let $\sigma_n \in M(G)$ be the atomic measure concentrated on E_n by the formula

$$\sigma_n = \frac{\sum_{y \in E_n} e^{(s_n\varphi)(y)} \cdot \delta_y}{\sum_{z \in E_n} e^{(s_n\varphi)(z)}}.$$

Let $\mu_n \in M(G)$ be defined by

$$\mu_n = \frac{1}{n} \sum \sigma_n \circ T^{-i}.$$

Since $M(G)$ is compact, it is possible to choose a subsequence $\{n_j\}$ of natural numbers such that

$$\lim_{j \to \infty} \frac{1}{n_j} \log P_{n_j}(T, \varphi, \epsilon, G) = P(T, \varphi, \epsilon, G)$$

and μ_{n_j} converges in $M(G)$ to some $\mu \in M(G)$. Thus, $\mu \in M(G, T) \subseteq M(X, T)$. We shall show $h_\mu(T \mid \langle G\rangle) + \int \varphi d\mu \geq P(T, \varphi, \epsilon, G)$.

By Lemma 3.3, we can choose a partition $\xi = \{A_1, A_2, \dots, A_k\}$ of X such that $\operatorname{diam}(A_j) < \epsilon$ and $\mu(\partial A_i) = 0$ for $1 \leqslant i \leqslant k$. Since each element of $\bigvee_{j=0}^{n-1} T^{-j}\xi$ contains at most one element of E_n,

$$\begin{aligned} H_{\sigma_n}\Big(\bigvee_{j=0}^{n-1} T^{-j}\xi \;\Big|\; \langle G\rangle\Big) + \int s_n\varphi d\sigma_n &= H_{\sigma_n}\Big(\bigvee_{j=0}^{n-1} T^{-j}\xi\Big) + \int s_n\varphi d\sigma_n \\ &= \sum_{y \in E_n} \sigma_n(y)((s_n\varphi)(y) - \log(\sigma_n(y))) \\ &= \log \sum_{y \in E_n} e^{(s_n\varphi)(y)}. \end{aligned}$$

Fix natural members q and, n with $1 \leqslant q \leqslant n$ and, using Lemma 3.2, define $a(j)$ for $0 \leqslant j \leqslant q-1$, by $a(j) = [(n-j)/q]$. Fix $0 \leqslant j \leqslant q-1$. Note that

$$\bigvee_{j=0}^{n-1} T^{-j}\xi = \bigvee_{r=0}^{a(j)-1} T^{-(rq+j)} \bigvee_{i=0}^{q-1} T^{-i} \vee \bigvee_{l \in S} T^{-l}\xi$$

and S has a cardinality of at most $2q$. Therefore

$$\begin{aligned}
&\log \sum_{y\in E_n} e^{(S_n\varphi)(y)}\\
&= H_{\sigma_n}\Big(\bigvee_{j=0}^{n-1} T^{-j}\xi \,\Big|\, \langle G\rangle\Big) + \int S_n\varphi d\sigma_n\\
&\le \sum_{r=0}^{a(j)-1} H_{\sigma_n}\Big(T^{-(rq+j)}\bigvee_{i=0}^{q-1} T^{-i}\xi \,\Big|\, \langle G\rangle\Big) + H_{\sigma_n}\Big(\bigvee_{k\in S} T^{-k}\xi \,\Big|\, \langle G\rangle\Big) + \int S_n\varphi d\sigma_n\\
&= \sum_{r=0}^{a(j)-1} H_{\sigma_n}\Big(T^{-(rq+j)}\bigvee_{i=0}^{q-1} T^{-i}\xi \,\Big|\, \langle G\rangle\Big) + H_{\sigma_n}\Big(\bigvee_{k\in S} T^{-k}\xi\Big) + \int S_n\varphi d\sigma_n\\
&\le \sum_{r=0}^{a(j)-1} H_{\sigma_n}\Big(T^{-(rq+j)}\bigvee_{i=0}^{q-1} T^{-i}\xi \,\Big|\, \langle G\rangle\Big) + 2q\log k + \int S_n\varphi d\sigma_n\\
&\le \sum_{r=0}^{a(j)-1} H_{\sigma_n\circ T^{-(rq+j)}}\Big(\bigvee_{i=0}^{q-1} T^{-i}\xi \,\Big|\, \langle G\rangle\Big) + 2q\log k + \int S_n\varphi d\sigma_n.
\end{aligned}$$

Summing this over j from 0 to $q-1$ leads to

$$q\log \sum_{y\in E_n} e^{(s_n\varphi)(y)} \le \sum_{p=0}^{n-1} H_{\sigma_n\circ T^{-p}}\Big(\bigvee_{i=0}^{q-1} T^{-i}\xi \,\Big|\, \langle G\rangle\Big) + 2q^2\log k + q\int S_n\varphi d\sigma_n.$$

Now, after dividing this by n and using Lemma 3.4, we have

$$\text{(3-1)}\qquad \frac{q}{n}\log \sum_{y\in E_n} e^{(s_n\varphi)(y)} \le H_{\mu_n}\Big(\bigvee_{i=0}^{q-1} T^{-i}\xi \,\Big|\, \langle G\rangle\Big) + \frac{2q^2}{n}\log k + q\int \varphi d\mu_n.$$

Because $\mu(\partial A_i)=0$ for all i, we have by [Cheng 2008, Lemma 3.6] that

$$\lim_{j\to\infty} H_{\mu_{n_j}}\Big(\bigvee_{i=0}^{q-1} T^{-i}\xi \,\Big|\, \langle G\rangle\Big) = H_{\mu}\Big(\bigvee_{i=0}^{q-1} T^{-i}\xi \,\Big|\, \langle G\rangle\Big).$$

Then, replacing n by n_j in (3-1) and letting j go to infinity, we obtain

$$qP(T,\varphi,\varepsilon,G) \le H_{\mu}\Big(\bigvee_{i=0}^{q-1} T^{-i}\xi \,\Big|\, \langle G\rangle\Big) + q\int \varphi d\mu.$$

Dividing by q and letting $q\to\infty$ yields

$$P(T,\varphi,\varepsilon,G) \le h_\mu(T\mid\langle G\rangle,\xi) + \textstyle\int\varphi d\mu \le h_\mu(T\mid\langle G\rangle) + \int\varphi d\mu,$$

which implies that

$$P_G(T,\varphi) \le h_\mu(T \mid \langle G\rangle) + \int \varphi d\mu$$

and completes the SVP argument.

Now, for any $\mu \in M(X,T)$, let η be a partition of $(X, \mathscr{B}, \mu)$ and let $G^c = X \setminus G$. Then

$$\begin{aligned} H_\mu\Big(\bigvee_{i=0}^{n-1} T^{-i}\eta \,\Big|\, \langle G\rangle\Big) &= -\mu(G) \sum_{C\in\bigvee_{i=0}^{n-1} T^{-i}\eta} \frac{\mu(C\bigcap G)}{\mu(G)} \log \frac{\mu(C\bigcap G)}{\mu(G)} \\ &\qquad - \mu(G^c) \sum_{C\in\bigvee_{i=0}^{n-1} T^{-i}\eta} \frac{\mu(C\bigcap G^c)}{\mu(G^c)} \log \frac{\mu(C\bigcap G^c)}{\mu(G^c)} \\ &= \mu(G) H_{\mu_G}\Big(\bigvee_{i=0}^{n-1} T^{-i}\eta\Big) + \mu(G^c) H_{\mu_{G^c}}\Big(\bigvee_{i=0}^{n-1} T^{-i}\eta\Big), \end{aligned}$$

where μ_G and μ_{G^c} denotes the conditional probability measures induced by μ on G and G^c, respectively. So, $\mu_G \in M(G,T)$, and $\mu_{G^c} \in M(G^c,T) \subset M(\overline{G^c},T)$. Thus, we have

$$\begin{aligned} h_\mu(T \mid \langle G\rangle, \eta) &= \mu(G) h_{\mu_G}(T,\eta) + \mu(G^c) h_{\mu_{G^c}}(T,\eta) \\ &\le \mu(G) h_{\mu_G}(T) + \mu(G^c) h_{\mu_{G^c}}(T), \end{aligned}$$

which implies

$$h_\mu(T \mid \langle G\rangle) \le \mu(G) h_{\mu_G}(T) + \mu(G^c) h_{\mu_{G^c}}(T).$$

On the other hand, for all $\epsilon > 0$ there is a partition η such that

$$h_{\mu_G}(T) \le h_{\mu_G}(T,\eta) + \epsilon/2.$$

Similarly, there is a partition ξ such that

$$h_{\mu_{G^c}}(T) \le h_{\mu_{G^c}}(T,\xi) + \epsilon/2.$$

Therefore, we can construct a partition $\zeta = \eta \vee \xi$ such that

$$h_{\mu_G}(T) \le h_{\mu_G}(T,\zeta) + \epsilon/2,$$
$$h_{\mu_{G^c}}(T) \le h_{\mu_{G^c}}(T,\zeta) + \epsilon/2.$$

Then

$$\mu(G) h_{\mu_G}(T) \le \mu(G)\{h_{\mu_G}(T,\zeta) + \epsilon/2\},$$
$$\mu(G^c) h_{\mu_{G^c}}(T) \le \mu(G^c)\{h_{\mu_{G^c}}(T,\zeta) + \epsilon/2\},$$

which implies that

$$\mu(G)h_{\mu_G}(T)+\mu(G^c)h_{\mu_{G^c}}(T)\le h_\mu(T\mid\langle G\rangle).$$

Thus,

$$h_\mu(T\mid\langle G\rangle)=\mu(G)h_{\mu_G}(T)+\mu(G^c)h_{\mu_{G^c}}(T).$$

On the other hand, for G is a closed T-invariant subset of X, it is not hard to show that $T\overline{G^c}\subseteq\overline{G^c}$. Thus, the topological pressure on the compact subset $\overline{G^c}$ is well defined. Since

$$\int\varphi d\mu=\int_G\varphi d\mu+\int_{G^c}\varphi d\mu=\mu(G)\int_G\varphi d\mu_G+\mu(G^c)\int_{G^c}\varphi d\mu_{G^c}.$$

by the SVP argument and Lemma 2.5, we obtain

$$\mu(G)h_{\mu_G}(T)+\mu(G)\int_G\varphi d\mu_G\le\mu(G)P_G(T,\varphi),$$
$$\mu(G^c)h_{\mu_{G^c}}(T)+\mu(G^c)\int_{G^c}\varphi d\mu_{G^c}\le\mu(G^c)P_{\overline{G^c}}(T,\varphi).$$

Combining these two, we obtain

$$\mu(G)h_{\mu_G}(T)+\mu(G^c)h_{\mu_{G^c}}(T)+\int\varphi d\mu\le\mu(G)P_G(T,\varphi)+\mu(G^c)P_{\overline{G^c}}(T,\varphi),$$

which implies, as desired,

$$h_\mu(T\mid\langle G\rangle)+\int\varphi d\mu\le\max\{P_G(T,\varphi),\,P_{\overline{X\setminus G}}(T,\varphi)\}\quad\text{for any }\ \mu\in M(X,T).\ \square$$

If this closed T-invariant subset G represents the whole space X, then $X\setminus G$ is empty. This allows the following lemma, which is the usual variational principle of topological pressure.

Lemma 3.5. *Let $T:X\to X$ be a continuous map of a compact metric space X and let φ be any real-valued continuous function of X. Then*

$$P(T,\varphi)=\sup_{\mu\in M(X,T)}\Big\{h_\mu(T)+\int\varphi d\,\mu\Big\},$$

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

4. The noncompact case

Using different definitions of topological pressure, an even more general variational inequality can be obtained, one that is also true for any T-invariant Borel subset G of X (that is, $T^{-1}G=G$). First, we define the topological pressure of any subset. The following notations come from [Pesin and Pitskel' 1984; Pesin 1988].

Let X be a compact metric space, let $Y \subset X$, and let $T : Y \to Y$ be a continuous mapping. Let $\mathcal{U}$ be a finite open cover of X. Denote by $\mathcal{W}_m(\mathcal{U})$ the set of collections of length m of elements of cover $\mathcal{U} : \underline{U} = U_{i_0} U_{i_1} \cdots U_{i_{m-1}}$. For any real-valued continuous function φ of X, set

$$Z(\underline{U}) = \{x \in Z : T^k(x) \in U_{i_k}, k = 0, \dots, m-1\},$$

and

$$S_m\varphi(\underline{U}) = \sup\Big\{\sum_{k=0}^{m-1} \varphi(T^k(x)) : x \in Z(\underline{U})\Big\}.$$

If $Z(\underline{U}) = \varnothing$, we assume that $S_m\varphi(\underline{U}) = -\infty$. Set $\mathcal{W}(\mathcal{U}) = \bigcup_{m \geqslant 0} \mathcal{W}_m(\mathcal{U})$. We will say that $\Gamma \subset \mathcal{W}(\mathcal{U})$ covers Z if $Z \subset \bigcup_{\underline{U} \in \Gamma} Z(\underline{U})$. The number of elements of collection $\underline{U}$ will be denoted by $m(\underline{U})$. Put

$$M(\mathcal{U}, \lambda, Z, \varphi, N) = \inf_{\Gamma \subset \mathcal{W}(\mathcal{U})} \Big\{\sum_{\underline{U} \in \Gamma} \exp(-\lambda m(\underline{U}) + S_{m(\underline{U})}\varphi(\underline{U}))\Big\},$$

where Γ covers Z and for every $\underline{U} \in \Gamma$ and $m(\underline{U}) \geqslant N$. It is readily verified that the function $M(\mathcal{U}, \lambda, Z, \varphi, N)$ increases monotonically with the growth of N. Then define

$$m(\mathcal{U}, \lambda, Z, \varphi) = \lim_{N \to \infty} M(\mathcal{U}, \lambda, Z, \varphi, N).$$

For any fixed subset Z, let $P_Z(\mathcal{U}, \varphi) = \inf\{\lambda : m(\mathcal{U}, \lambda, Z, \varphi) = 0\}$. The topological pressure of any set Z with respect to T is defined by

$$P_Z(T, \varphi) = \lim_{\operatorname{diam} \mathcal{U} \to 0} P_Z(\mathcal{U}, \varphi).$$

Given a T-invariant subset Z, denote by $M(Z, T) \subset M(X, T)$ the set of measures μ for which $\mu(Z) = 1$ for all $\mu \in M(X, T)$. For $x \in X$ and $n > 0$, define the probability measure

$$\mathcal{E}_n(x) := \frac{1}{n} \sum_{k=0}^{n-1} \delta_{T^k(x)},$$

where δ_x is the δ-measure supported at the point x. Denote by $V(x)$ the set of limit measures of the sequence of measures $\mathcal{E}_n(x)$. It is easy to see that $V(x) \subset M(X, T)$ for any $x \in X$.

Lemma 4.1. *Assume that (X, d) is a compact metric space, $T : X \to X$ is a continuous map, G is a T-invariant Borel subset of X (that is, $T^{-1}G = G$), and $\mu \in M(G, T)$. If we consider $T|_G : G \to G$, then $h_\mu(T|_G) = h_\mu(T \mid \langle G \rangle)$.*

Proof. Let η be a partition of X and let $G^c = X \setminus G$. Then

$$\begin{aligned} H_\mu\Big(\bigvee_{i=0}^{n-1} T^{-i}\eta \mid \langle G\rangle\Big) &= -\mu(G) \sum_{C\in\bigvee_{i=0}^{n-1} T^{-i}\eta} \frac{\mu(C\cap G)}{\mu(G)} \log \frac{\mu(C\cap G)}{\mu(G)} \\ &\qquad - \mu(G^c) \sum_{C\in\bigvee_{i=0}^{n-1} T^{-i}\eta} \frac{\mu(C\cap G^c)}{\mu(G^c)} \log \frac{\mu(C\cap G^c)}{\mu(G^c)} \\ &= -\mu(G) \sum_{C\in\bigvee_{i=0}^{n-1} T^{-i}\eta} \frac{\mu(C\cap G)}{\mu(G)} \log \frac{\mu(C\cap G)}{\mu(G)} \\ &= - \sum_{C\in\bigvee_{i=0}^{n-1} T^{-i}\eta} \mu(C\cap G) \log \mu(C\cap G) \\ &= H_\mu\Big(\bigvee_{i=0}^{n-1} T^{-i}\eta \,\Big|\, G\Big), \end{aligned}$$

which implies that $h_\mu(T|_G) = h_\mu(T \mid \langle G\rangle)$. □

Theorem 4.1. *Let $T : X \to X$ be a continuous map of a compact metric space, let $\varphi \in C(X, R)$, and let G be a T-invariant Borel subset of X (that is, $T^{-1}G = G$). Suppose that for each $x \in G$, the intersection $V(x) \cap M(G, T)$ is nonempty. Then*

$$P_G(T,\varphi) \le \sup_{\mu\in M(X,T)} \{h_\mu(T \mid \langle G\rangle) + \textstyle\int \varphi d\mu\} \le \max\{P_G(T,\varphi), P_{G^c}(T,\varphi)\},$$

where G^c is the complement of G, that is, $G^c = X \setminus G$.

Proof. On the one hand, the combination of [Pesin and Pitskel' 1984, Lemma 4.1 and Theorem 2] indicates that

$$P_G(T,\varphi) \le \sup_{\mu\in M(X,T)} \{h_\mu(T \mid \langle G\rangle) + \textstyle\int \varphi d\mu\}.$$

On the other hand, by an argument similar to that of Theorem 3.1, we have

$$h_\mu(T \mid \langle G\rangle) = \mu(G) h_{\mu_G}(T) + \mu(G^c) h_{\mu_{G^c}}(T)$$

and

$$\int \varphi d\mu = \mu(G) \int_G \varphi d\mu_G + \mu(G^c) \int_{G^c} \varphi d\mu_{G^c}.$$

Therefore, from [Pesin and Pitskel' 1984, Theorem 1], we can obtain that

$$h_{\mu_G}(T) + \int_G \varphi d\mu_G \le P_G(T,\varphi)$$

and

$$h_{\mu_{G^c}}(T) + \int_{G^c} \varphi \, d\mu_{G^c} \le P_{G^c}(T, \varphi). \qquad \square$$

Acknowledgements

This work was carried out when Wen-Chiao Cheng visited Institute of Mathematics at Nanjing Normal University. Wen-Chiao Cheng sincerely appreciates the warm hospitality. During this visit in 2011, Wen-Chiao Cheng was partially supported by NSC Grant NSC 99-2115-M-034-001. The author Ercai Chen was supported by the National Natural Science Foundation of China 10971100 and National Basic Research Program of China (973 Program) 2007CB814800.

References

[Bowen 1975] R. Bowen, *Equilibrium states and the ergodic theory of Anosov diffeomorphisms*, Lecture Notes in Mathematics **470**, Springer, Berlin, 1975. MR 56 #1364 Zbl 0308.28010

[Cheng 2008] W.-C. Cheng, "Estimate for supremum of conditional entropy on a closed subset", *Taiwanese J. Math.* **12**:7 (2008), 1791–1803. MR 2010b:37007 Zbl 1221.37027

[Downarowicz and Serafin 2002] T. Downarowicz and J. Serafin, "Fiber entropy and conditional variational principles in compact non-metrizable spaces", *Fund. Math.* **172**:3 (2002), 217–247. MR 2003b:37027 Zbl 1115.37308

[Huang and Ye 2006] W. Huang and X. Ye, "A local variational relation and applications", *Israel J. Math.* **151** (2006), 237–279. MR 2006k:37033 Zbl 1122.37013

[Huang et al. 2006] W. Huang, X. Ye, and G. Zhang, "A local variational principle for conditional entropy", *Ergodic Theory Dynam. Systems* **26**:1 (2006), 219–245. MR 2006j:37015 Zbl 1176.37011

[Katok and Hasselblatt 1995] A. Katok and B. Hasselblatt, *Introduction to the modern theory of dynamical systems*, Encyclopedia of Mathematics and its Applications **54**, Cambridge University Press, 1995. MR 96c:58055 Zbl 0878.58020

[Ledrappier 1979] F. Ledrappier, "A variational principle for the topological conditional entropy", pp. 78–88 in *Ergodic theory* (Oberwolfach, 1978), edited by M. Denker and K. Jacobs, Lecture Notes in Math. **729**, Springer, Berlin, 1979. MR 80j:54011 Zbl 0415.54025

[Mihailescu and Urbański 2004] E. Mihailescu and M. Urbański, "Inverse topological pressure with applications to holomorphic dynamics of several complex variables", *Commun. Contemp. Math.* **6**:4 (2004), 653–679. MR 2005i:37033 Zbl 1091.37011

[Misiurewicz 1976] M. Misiurewicz, "Topological conditional entropy", *Studia Math.* **55**:2 (1976), 175–200. MR 54 #3672 Zbl 0355.54035

[Mummert 2007] A. Mummert, "A variational principle for discontinuous potentials", *Ergodic Theory Dynam. Systems* **27**:2 (2007), 583–594. MR 2008b:37047 Zbl 1109.37017

[Pesin 1988] Y. B. Pesin, "Dimension-like characteristics for invariant sets of dynamical systems", *Uspekhi Mat. Nauk* **43**:4(262) (1988), 95–128, 255. In Russian; translated in *Russian Math. Surveys* **43**:4 (1988), 111–151. MR 90e:28011 Zbl 0684.58024

[Pesin and Pitskel' 1984] Y. B. Pesin and B. S. Pitskel', "Topological pressure and the variational principle for noncompact sets", *Funkt. Anal. i Prilozhen.* **18**:4 (1984), 50–63, 96. In Russian. MR 86i:28031

[Romagnoli 2003] P.-P. Romagnoli, "A local variational principle for the topological entropy", *Ergodic Theory Dynam. Systems* **23**:5 (2003), 1601–1610. MR 2004i:37030 Zbl 1056.37017

[Ruelle 1973] D. Ruelle, "Statistical mechanics on a compact set with Z^v action satisfying expansiveness and specification", *Trans. Amer. Math. Soc.* **187** (1973), 237–251. MR 54 #5441 Zbl 0278.28012

[Ruelle 1978] D. Ruelle, *Thermodynamic formalism*, Encyclopedia of Mathematics and its Applications **5**, Addison-Wesley, Reading, MA, 1978. MR 80g:82017 Zbl 0401.28016

[Walters 1975] P. Walters, "A variational principle for the pressure of continuous transformations", *Amer. J. Math.* **97**:4 (1975), 937–971. MR 52 #11006 Zbl 0318.28007

[Walters 1982] P. Walters, *An introduction to ergodic theory*, Graduate Texts in Mathematics **79**, Springer, New York, 1982. MR 84e:28017 Zbl 0475.28009

Received March 3, 2011. Revised November 9, 2011.

YUAN LI
KEY LABORATORY OF VIRTUAL GEOGRAPHIC ENVIRONMENT
NANJING NORMAL UNIVERSITY
NANJING, 210046
CHINA
liyuannjnu@163.com

ERCAI CHEN
SCHOOL OF MATHEMATICAL SCIENCE
NANJING NORMAL UNIVERSITY
NANJING, 210093
CHINA
ecchen@njnu.edu.cn

WEN-CHIAO CHENG
DEPARTMENT OF APPLIED MATHEMATICS
CHINESE CULTURE UNIVERSITY
YANGMINGSHAN, TAIPEI 11114
TAIWAN
zwq2@faculty.pccu.edu.tw

PACIFIC JOURNAL OF MATHEMATICS

http://pacificmath.org

Founded in 1951 by
E. F. Beckenbach (1906–1982) and F. Wolf (1904–1989)

See inside back cover or pacificmath.org for submission instructions.

The subscription price for 2012 is US $420/year for the electronic version, and $485/year for print and electronic.
Subscriptions, requests for back issues from the last three years and changes of subscribers address should be sent to Pacific Journal of Mathematics, P.O. Box 4163, Berkeley, CA 94704-0163, U.S.A. Prior back issues are obtainable from Periodicals Service Company, 11 Main Street, Germantown, NY 12526-5635. The Pacific Journal of Mathematics is indexed by Mathematical Reviews, Zentralblatt MATH, PASCAL CNRS Index, Referativnyi Zhurnal, Current Mathematical Publications and the Science Citation Index.

The Pacific Journal of Mathematics (ISSN 0030-8730) at the University of California, c/o Department of Mathematics, 969 Evans Hall, Berkeley, CA 94720-3840, is published monthly except July and August. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices. POSTMASTER: send address changes to Pacific Journal of Mathematics, P.O. Box 4163, Berkeley, CA 94704-0163.

PJM peer review and production are managed by EditFLOW™ from Mathematical Sciences Publishers.

PUBLISHED BY PACIFIC JOURNAL OF MATHEMATICS
at the University of California, Berkeley 94720-3840
A NON-PROFIT CORPORATION
Typeset in LaTeX

PACIFIC JOURNAL OF MATHEMATICS

Volume 256 No. 1 March 2012