LOCALIZED VARIATIONAL PRINCIPLE FOR NON-BESICOVITCH METRIC SPACES

TAMARA KUCHERENKO

Abstract. We consider the localized entropy of a point $w \in \mathbb{R}^m$ which is computed by considering only those $(n, \varepsilon)$-separated sets whose statistical sums with respect to an $m$-dimensional potential $\Phi$ are "close" to a given value $w$. Previously, a local version of the variational principle was established for systems on non-Besicovitch compact metric spaces. We extend this result to all compact metric spaces.

1. Introduction

1.1. Motivation. We say that a metric space $(X, d)$ satisfies the Besicovitch Covering Property if there exists an integer $N$ so that for each family $\mathcal{B}$ of closed balls, whose centers form a bounded subset of $X$, there is a subfamily $\mathcal{F}$ covering the set of centers of the balls in $\mathcal{B}$, and such that each point of $X$ is contained in at most $N$ balls from $\mathcal{F}$. A large variety of dynamical systems satisfies this property. Some examples are subshifts of finite type, hyperbolic systems and continuous maps on compact smooth Riemannian manifolds (see [1, 7]).

Whether or not a metric space satisfies the Besicovitch Covering Property hinges on the metric it is endowed with. For example, any uncountable complete separable metric space can be re-metricised with a bilipschitz equivalent metric so that the Besicovitch Covering Property is not satisfied [9]. One source of examples of spaces without the Besicovitch Covering Property property are Heisenberg groups with Korányi distance [4, 12] or Carnot-Carathéodory distance [11], which are geometrically sub-Riemannian manifolds. It is still not clear whether for these spaces there exist equivalent metrics for which the Besicovitch Covering Property holds.

In [5] we prove a local version of the variational principle for compact metric spaces which satisfy the Besicovitch Covering Property. The aim of this note is to extend this result to all compact metric spaces.

1.2. Basic definitions and statement of the results. Let $f : X \to X$ be a continuous map on a compact metric space $(X, d)$. We consider continuous potential $\Phi = (\phi_1, \cdots, \phi_m) : X \to \mathbb{R}^m$.

Key words and phrases. topological pressure, generalized rotation sets, variational principle, equilibrium states, thermodynamic formalism.
For $n \in \mathbb{N}$ and $\varepsilon > 0$, we say that a set $F \subset X$ is $(n, \varepsilon)$-separated if for all $x, y \in F$ with $x \neq y$ we have $d_n(x, y) \overset{\text{def}}{=} \max_{k=0, \ldots, n-1} d(f^k(x), f^k(y)) \geq \varepsilon$. Note that $d_n$ is a metric (called Bowen metric) that induces the same topology on $X$ as $d$. For $x \in X$ and $n \in \mathbb{N}$, we denote by $\frac{1}{n} S_n(\Phi, f)(x)$ the $m$-dimensional Birkhoff average at $x$ of length $n$ with respect to $\Phi$ and $f$, where $S_n(\Phi, f)(x) = (S_n(\phi_1, f)(x), \ldots, S_n(\phi_m, f)(x))$ and

$$S_n(\phi_i, f)(x) = \sum_{k=0}^{n-1} \phi_i(f^k(x)).$$

Given $w \in \mathbb{R}^m$ and $r > 0$ we say a set $F \subset X$ is a $(n, \varepsilon, w, r)$-set for $\Phi$ and $f$ if $F$ is $(n, \varepsilon)$-separated set and for all $x \in F$ the Birkhoff average $\frac{1}{n} S_n(\Phi, f)(x)$ is contained in the Euclidean ball $B(w, r)$ with center $w$ and radius $r$. For all $n \in \mathbb{N}$ and $\varepsilon, r > 0$ we pick a maximal (with respect to the inclusion) $(n, \varepsilon, w, r)$-set $F_n(\varepsilon, w, r)$.

Then the localized topological entropy at $w \in \mathbb{R}^m$ (with respect to $\Phi$ and $f$) is defined by

$$h_{\text{top}}(w, \Phi, f) = \lim_{r \to 0} \lim_{\varepsilon \to 0} \limsup_{n \to \infty} \frac{1}{n} \log \text{card } F_n(\varepsilon, w, r) \quad (1)$$

This definition is analogous to that of the classical topological entropy with the exception that we here only consider orbits with Birkhoff averages close to $w$.

Alternatively, there we can use a measure-theoretic approach to to define a localized entropy. We denote by $\mathcal{M}(f)$ the set of all Borel $f$-invariant probability measures on $X$ endowed with the weak$^*$ topology. Following [3], we define the generalized rotation set of $\Phi$ by

$$\text{Rot}(\Phi) = \{rv_{\Phi}(\mu) : \mu \in \mathcal{M}\},$$

where $rv_{\Phi}(\mu) = (\int \phi_1 d\mu, \ldots, \int \phi_m d\mu)$ denotes the rotation vector of the measure $\mu$. We call $\mathcal{N}_w(\Phi) = \{\mu \in \mathcal{M} : rv_{\Phi}(\mu) = w\}$ the rotation class of $w$. We refer to [3, 5, 14] for further details about rotation sets.

For $w \in \text{Rot}(\Phi)$, we define the localized measure-theoretic entropy at $w$ (with respect to $\Phi$ and $f$) by

$$h_{\mu}(f, \Phi, w) = \sup \{h_{\mu}(f) : \mu \in \mathcal{N}_w(\Phi)\}.$$

The classical variational principle (without localization) states that the topological and the measure-theoretic versions of the entropy coincide. However, it turns out that in the case of localized entropy the measure-theoretic and topological entropies may differ, and strict inequalities can occur in both directions [6]. On the other hand, the following result gives a fairly complete description of the assumptions needed to still have a variational principle.

**Theorem 1.** [5] Let $f : X \to X$ be a continuous map on a compact metric space $X$ which satisfies the Besicovitch Covering Property. Let $\Phi : X \to \mathbb{R}^m$
be continuous and let \( w \in \text{Rot}(\Phi) \) be such that the map \( v \mapsto h_m(v, f, \Phi) \) is continuous at \( w \) and \( h_m(w, f, \Phi) \) is approximated by ergodic measures. Then
\[
h_{\text{top}}(w, f, \Phi) = h_m(w, f, \Phi).
\]

Here, we say that \( h_m(w, f, \Phi) \) is approximated by ergodic measures if there exists a sequence of ergodic measures \((\mu_n)_{n \in \mathbb{N}}\) such that \( \nu\Phi(\mu_n) \to w \) and \( h_{\mu_n}(f) \to h_m(w, f, \Phi) \) as \( n \to \infty \). The assumption that \( h_m(f, \Phi, w) \) is approximated by ergodic measures cannot be dropped in Theorem 1. Indeed, there are examples which do not satisfy this assumption and \( h_{\text{top}}(f, \Phi, w) < h_m(f, \Phi, w) \) holds. On the other hand, without the assumption that \( v \mapsto h_m(f, \Phi, v) \) is continuous at \( w \), we arrive at the opposite inequality [6].

All conditions except the Besicovitch property are necessary for the conclusion of this theorem. However, the Besicovitch property appears to be nonessential. Moreover, the inequality \( h_{\text{top}}(f, \Phi, w) \leq h_m(f, \Phi, w) \) is proven without any additional assumptions on the metric space \( X \). Here we present an alternate proof of the opposite inequality that does not rely on the Besicovitch property.

2. Localized Variational Principle for Entropy

This section is devoted to the proof of the following theorem.

**Theorem 2.** Let \( f : X \to X \) be a continuous map on a compact metric space \( X \). Let \( \Phi : X \to \mathbb{R}^m \) be continuous and let \( w \in \text{Rot}(\Phi) \) be such that the map \( v \mapsto h_m(f, \Phi, v) \) is continuous at \( w \) and \( h_m(f, \Phi, w) \) is approximated by ergodic measures. Then \( h_{\text{top}}(w, f, \Phi) = h_m(w, f, \Phi) \).

Note that the inequality \( h_{\text{top}}(f, \Phi, w) \leq h_m(f, \Phi, w) \) was proven in [5]. The proof of the opposite inequality \( h_m(f, \Phi, w) \leq h_{\text{top}}(f, \Phi, w) \) relies on the following three lemmas.

We fix \( w \in \text{Rot}(\Phi) \) and \( r > 0 \). We denote by
\[
h(r, w, \Phi, f) = \lim_{\varepsilon \to 0} \limsup_{n \to \infty} \frac{1}{n} \log \text{card } F_n(\varepsilon, w, r) \tag{2}
\]
Here \( F_n(\varepsilon, w, r) \) stands for a maximal \((\varepsilon, n, w, r)\)-set. Then the localized topological entropy at \( w \) (with respect to \( \Phi \) and \( f \)) is
\[
h_{\text{top}}(w, \Phi, f) = \lim_{r \to 0} h(r, w, \Phi, f) \tag{3}
\]

**Lemma 1.** Let \( X \) be a metric space, \( f : X \to X \) and \( \Phi : X \to \mathbb{R}^m \). For \( k \in \mathbb{N} \) denote by \( \Phi_k = \frac{1}{k}S_k(\Phi, f) \). Then for any \( n \in \mathbb{N} \) we have
\[
\frac{1}{n}S_n(\Phi_k, f^k) = \frac{1}{kn}S_{kn}(\Phi, f)
\]

The equality is this lemma is proved by standard algebraic manipulation and thus we omit it here.
**Lemma 2.** Let $f : X \to X$ be a continuous map on a compact metric space, $\Phi : X \to \mathbb{R}^m$ be a continuous potential and $w \in \text{Rot}(\Phi)$. For any $r > 0$, $\varepsilon > 0$ and $k \in \mathbb{N}$ we have

$$h(r, w, f^k, \Phi_k) = k \cdot h(r, w, f, \Phi),$$

where $\Phi_k = \frac{1}{k} S_k(\Phi, f)$.

**Proof.** Let $F$ be any $(n, \varepsilon, r, w)$-set with respect to $\Phi_k$ and $f^k$. We will show that $F$ is also a $(kn, \varepsilon, r, w)$-set with respect to $\Phi$ and $f$. For any $x, y \in F$ we have

$$\max_{0 \leq i \leq kn-1} d(f^i(x), f^i(y)) \geq \max_{0 \leq j \leq n-1} d(f^{jk}(x), f^{jk}(y)) > \varepsilon.$$ 

Moreover,

$$\frac{1}{kn} S_{kn}(\Phi, f)(x) = \frac{1}{n} S_n(\Phi_k, f^k)(x) \in B(r, w).$$

Since every $(n, \varepsilon, r, w)$-set with respect to $\Phi_k$ and $f^k$ is a $(kn, \varepsilon, r, w)$-set with respect to $\Phi$ and $f$, we obtain

$$\text{card } F_n(\varepsilon, r, w, f^k, \Phi_k) \leq \text{card } F_{kn}(\varepsilon, r, w, f, \Phi).$$

Therefore,

$$\frac{1}{n} \log \text{card } F_n(\varepsilon, r, w, f^k, \Phi_k) \leq k \cdot \frac{1}{kn} \log \text{card } F_{kn}(\varepsilon, r, w, f, \Phi).$$

Passing to the upper limit as $n \to \infty$ and to the limit as $\varepsilon \to 0$ we obtain $h(r, w, f^k, \Phi_k) \leq k h(r, w, f, \Phi)$.

To prove the opposite inequality we fix an $\varepsilon > 0$ and $n \in \mathbb{N}$ and let $F$ be any $(\varepsilon, kn, r, w)$-set with respect to $\Phi$ and $f$. We use the uniform continuity of $f$ on $X$ to find a $0 < \delta < \varepsilon$ such that for $i = 0, \ldots, k$ we have $d(f^i(x), f^i(y)) < \frac{\varepsilon}{2}$ whenever $d(x, y) < \delta$. Denote be $F_n$ any maximal $(\delta, n, r, w)$-set with respect to $\Phi_k$ and $f^k$. If $x \in F$ then $\frac{1}{n} S_n(f^k, \Phi_k)(x) \in B(r, w)$ by Lemma 1. The maximality of $F_n$ implies the existence of $y_x \in F_n$ such that

$$\max_{0 \leq j \leq n-1} d(f^{jk}(x), f^{jk}(y_x)) < \delta.$$ 

Then $d(f^i(x), f^i(y_x)) \leq \frac{\varepsilon}{2}$ for all $0 \leq i \leq kn-1$. The map $x \in F \mapsto y_x \in F_n$ is injective. Indeed, if for some $x_1, x_2 \in F$ we have $y_{x_1} = y_{x_2}$ then the triangle inequality yields

$$\max_{0 \leq i \leq kn-1} d(f^i(x_1), f^i(y_2)) < \varepsilon.$$ 

This contradicts the fact that $F$ is $(kn, \varepsilon)$-separated. Therefore, $\text{card } F \leq \text{card } F_n$. Since $F$ was arbitrary $(\varepsilon, kn, r, w)$-set with respect to $\Phi$ and $f$, we obtain

$$k \cdot \frac{1}{kn} \log \text{card } F_{kn}(\varepsilon, w, r, f, \Phi) \leq \frac{1}{n} \log \text{card } F_n(\delta, w, r, f^k, \Phi_k).$$

Letting $n \to \infty$ and $\varepsilon \to 0$ we obtain the desired inequality. □
Remark. The consequence of this lemma is an analog of the power rule for classical entropy $h_{\text{top}}(w, f^k, \Phi_k) = k \cdot h_{\text{top}}(w, f, \Phi)$.

Before formulating the next lemma we recall the standard definition of the entropy of the measure $\mu \in \mathcal{M}(f)$. Let $A = \{A_1, A_2, \ldots, A_k\}$ be a finite partition of $X$. Then the entropy of the partition $A$ is

$$H_\mu(A) = -\sum_{i=1}^{k} \mu(A_i) \log \mu(A_i).$$

Note that the convexity of the function $x \mapsto x \log x$ implies $H_\mu(A) \leq \log \text{card } (\mathcal{A})$.

The join of the partitions $f^{-j}(A) = \{f^{-j}(A_1), \ldots, f^{-j}(A_k)\}$ is the partition $A^n = \bigvee_{j=0}^{n-1} f^{-j}(A)$, which consists of all sets of the form $\cap_{j=0}^{n-1} f^{-j}(A_i)$ with $A_i \in A$. The entropy of the partition $A$ is

$$h_\mu(f, A) = \lim_{n \to \infty} \frac{1}{n} H_\mu(A^n).$$

Finally, the entropy of the measure $\mu$ with respect to $f$ is

$$h_\mu(f) = \sup \{h_\mu(f, A) : A \text{ is a finite partition of } X\}$$

Lemma 3. Let $f : X \to X$ be a continuous map on a compact metric space, $\Phi : X \to \mathbb{R}^m$ be a continuous potential, $w \in \text{Rot}(\Phi)$ and $r > 0$. Suppose that $\mu \in \mathcal{M}(f)$ is such that $rv_\Phi(\mu) \in B(r, w)$ and for $\mu$-almost all $x \in X$

$$\lim_{n \to \infty} \frac{1}{n} S_n(f, \Phi)(x) = rv_\Phi(\mu).$$

Then

$$h_\mu(f) \leq h(r, w, f, \Phi) + \log 2 + 1.$$

Proof. Let $\mathcal{A} = \{A_1, A_2, \ldots, A_k\}$ be any Borel partition of $X$. Choose $\varepsilon > 0$ such that $\varepsilon < \frac{1}{k \log k}$. Since the sequence $\frac{1}{n} S_n(f, \Phi)$ converges $\mu$-almost everywhere to $rv_\Phi(\mu)$, we use Egoroff’s theorem to find subsets $D_i \subset A_i$, $i = 1, \ldots, k$ with the following properties

- $D_i$ is compact
- $\mu(A_i \setminus D_i) < \varepsilon$
- $\frac{1}{n} S_n(f, \Phi) \to rv_\Phi(\mu)$ uniformly on $D_i$

Let $D_0 = X \setminus \bigcup_{i=1}^{k} D_i$. Then $\mathcal{D} = \{D_0, D_1, \ldots, D_k\}$ is also a partition of $X$. Denote by $r_\mu = \min\{\|rv_\Phi(\mu) - w\|, r - \|rv_\Phi(\mu) - w\|\}$ and by

$$l = 4 \left[ \sup \{\|\Phi(x)\| : x \in X\} \right].$$

We consider the join of the partitions

$$\mathcal{D}^{ln} = \bigvee_{j=0}^{ln-1} f^{-j}(\mathcal{D}).$$
We split $D_n$ into "good sets" $G^n$ and "bad sets" $B^n_i$, $(i = 0, ..., l)$ in the following way.

\[ B^n_0 = D_0 \cap f^{-1}(D_0) \cap ... \cap f^{-ln+1}(D_0) \]

\[ B^n_i = \{ D \in D^n \setminus \cup_{j=0}^{i-1} B^n_j : D \subset D_0 \cap f^{-1}(D_0) \cap ... \cap f^{-(l-i)n+1}(D_0) \} \]

Then $G^n = D^n \setminus \cup_{i=0}^{l} B^n_i$. We also denote by

\[ E^n = \{ x \in X : \frac{1}{m} S_m(f, \Phi) \in B(r, w) \text{ for any } m > ln \} \]

Since $\frac{1}{n} S_n(f, \Phi)$ converges uniformly on $D_i (i \neq 0)$ to the rotation vector of $\mu$, there is $N_\mu > 0$ such that for any $n > N_\mu$ and any $x \in \cup_{i=1}^k D_i$ we have

\[ \left\| \frac{1}{n} S_n(f, \Phi)(x) - rv\Phi(\mu) \right\| < \frac{r_\mu}{2}. \]

From now on we will consider $n > N_\mu$.

First we will show that for such $n$ any set $G \in G^n$ is a subset of $E_n$. Pick any $x \in G$. Then there is $s < n$ such that $f^s(x) \in D_i$ for some $i \neq 0$. Then for any $m > ln$ we have

\[ \left\| \frac{1}{m} S_m(f, \Phi)(x) - w \right\| \leq \left\| \frac{1}{m} S_m(f, \Phi)(x) - \frac{1}{m} S_m(f, \Phi)(f^s(x)) \right\| \]

\[ + \left\| \frac{1}{m} S_m(f, \Phi)(f^s(x)) - w \right\| \]

To estimate the first term we note that

\[ \| S_m(f, \Phi)(x) - S_m(f, \Phi)(f^s(x)) \| \leq \Phi(x) + \| \Phi(f(x)) \| + \cdots + \| \Phi(f^{s-1}(x)) \| \]

\[ + \| \Phi(f^m(x)) \| + \cdots + \| \Phi(f^{m+s-1}(x)) \| \]

\[ \leq 2s \cdot \sup \{ \| \Phi(x) \| : x \in X \} \]

\[ \leq 2s \cdot \frac{lr_\mu}{4} \]

\[ \leq \frac{lnr_\mu}{2} \]

To estimate the second term we use the fact that $f^s(x) \in \cup_{i=1}^k D_i$ and $m \geq N_\mu$ implies $\frac{1}{m} S_m(f, \Phi)(f^s(x)) \in B\left(\frac{r_\mu}{2}, rv\Phi(\mu)\right)$. Therefore,

\[ \left\| \frac{1}{m} S_m(f, \Phi)(f^s(x)) - w \right\| \leq \frac{r_\mu}{2} + \| w - rv\Phi(\mu) \| \]

Combining these two estimates we obtain

\[ \left\| \frac{1}{m} S_m(f, \Phi)(x) - w \right\| \leq \frac{1}{m} \cdot \frac{lnr_\mu}{2} + \frac{r_\mu}{2} + \| w - rv\Phi(\mu) \| \]

\[ < r_\mu + \| w - rv\Phi(\mu) \| \]

\[ \leq r \]

Therefore, $x \in E_n$. 
Now we will show that the cardinality of $G^n$ is comparable to the cardinality of $\mathcal{D}^n$. If a set $B \in \mathcal{B}_i^n$, $(0 < i \leq l)$ then

$$B = D_0 \cap f^{-i}(D_0) \cap ... \cap f^{-(l-i)n+1}(D_0) \cap f^{-(l-i)n}(D_j) \cap ... \cap f^{-l+1}(D_{jn})$$

If $B$ is not empty then the set $D_0 \cap ... \cap f^{-l+1}(D_{jn})$ is also not empty. Moreover, different sets $B \in \mathcal{B}_i^n$ correspond to different sets of the form above. By construction of $\mathcal{B}_i^n$ there is a set $G$ in $\mathcal{G}^n$ such that $G \subset D_{j0} \cap ... \cap f^{-l+1}(D_{jn})$. Therefore,

$$\text{card } G^n \geq \max_{0 \leq i \leq l} \{ \mathcal{B}_i^n \}$$

Since the families $\mathcal{B}_i$ are disjoint and $\mathcal{D}^n = \cup_{i=0}^l \mathcal{B}_i^n \cup \mathcal{G}^n$, we obtain

$$\text{card } \mathcal{D}^n \leq (l + 2) \text{card } \mathcal{G}^n \quad (4)$$

Consider

$$C = \{ D_0 \cup D_1, D_0 \cup D_2, ..., D_0 \cup D_k \}.$$ 

Since the sets $D_i$ $(1 \leq i \leq k)$ are compact, $C$ is an open cover of $X$. We denote by $\mathcal{C}_n$ a subfamily of the join $\bigvee_{j=0}^{l-1} f^{-j}(C)$ which covers $E_n$ and has minimal cardinality. Next we will show that

$$\text{card } \mathcal{G}^n \leq 2^{ln} \text{card } \mathcal{C}_n \quad (5)$$

Let $G \in \mathcal{G}^n$,

$$G = D_{s1} \cap f^{-1}(D_{s2}) \cap ... \cap f^{-ln+1}(D_{st})$$

Then $G \subset E_n$ and thus there is a set $C \in \mathcal{C}_n$ such that $G \cap C \neq \emptyset$. The set $C$ is of the form

$$C = (D_0 \cup D_{j1}) \cap f^{-1}(D_0 \cup D_{j2}) \cap ... \cap f^{-ln+1}(D_0 \cup D_{jn}).$$

Since $C \cap G$ is not empty, for $s = 1, ..., ln$ we must have $D_{is} \cap (D_0 \cup D_{js}) \neq \emptyset$. Since the sets $\{ D_i \}_{i=0}^k$ form a partition of $X$, either $i_s = 0$ or $i_s = j_s$. Also, in this case $G \subset C$. This implies that any set in $\mathcal{G}^n$ is a subset of some set $C \in \mathcal{C}_n$. Moreover, each set in $\mathcal{C}_n$ can contain at most $2^{ln}$ sets from $\mathcal{G}^n$. We conclude that $\text{card } \mathcal{G}^n \leq 2^{ln} \text{card } \mathcal{C}_n$.

Let $\delta$ be a Lebesgue number of the cover $C$, that is any subset of $X$ of diameter less than or equal to $\delta$ lies in some member of $C$. Then $\delta$ is also a Lebesgue number of the cover $\bigvee_{j=0}^{ln-1} f^{-j}(C)$ in the $d_n$-metric. Since $\mathcal{C}_n$ is a minimal cover of $E_n$, every set $C \in \mathcal{C}_n$ contains a point $x_C \in E_n$ which is not in any other element of $\mathcal{C}_n$. Then the ball in the $d_n$-metric centered at $x_C$ of diameter $\delta$ is contained in $C$. Therefore, points $\{ x_C : C \in \mathcal{C}_n \}$ form a $(\delta/2, ln)$-separated set of $E_n$.

Recall that $F_{ln}(\delta/2, w, r)$ denotes the maximal $(\delta/2, ln)$-separated set with the property that $\frac{1}{ln} S_{ln}(f, \Phi)(x) \in B(r, w)$ for any $x \in F_{ln}(\delta/2, w, r)$. We see that

$$\text{card } \mathcal{C}_n = \text{card } \{ x_C : C \in \mathcal{C}_n \} \leq \text{card } F_{ln}(\delta/2, w, r) \quad (6)$$
Combining this last inequality with (4) and (5) we obtain
\[
\text{card } \mathcal{D}^n \leq (l + 2)\text{card } \mathcal{G}^n \\
\leq 2^n(l + 2)\text{card } \mathcal{C}_n \\
\leq 2^n(l + 2)\text{card } F_{in}(\delta/2, w, r)
\]

Using the fact that \(H(\mathcal{D}^n) \leq \log \text{card } \mathcal{D}^n\) we estimate

\[
h(\mathcal{D}^n) = \lim_{n \to \infty} \frac{1}{ln} H(\mathcal{D}^n) \\
\leq \lim_{n \to \infty} \frac{1}{ln} (ln \log 2 + \log(l + 2) + \log \text{card } F_{in}(\delta/2, w, r)) \\
\leq \log 2 + \lim_{n \to \infty} \text{card } F_{in}(\delta/2, w, r) \\
\leq \log 2 + h(r, w, f, \Phi)
\]

Now it is left to compare \(h(\mathcal{D}^n)\) and \(h(\mathcal{A})\). It is a standard argument to show that \(h(\mathcal{D}^n) \leq h(\mathcal{A}) + 1\) (see [13] or [10]). We will outline it here for the sake of completeness. We have ([13, ?])

\[
h(\mathcal{D}^n) \leq h(\mathcal{A}) + H(\mathcal{A}|\mathcal{D}),
\]

where \(H(\mathcal{A}|\mathcal{D})\) is the conditional entropy of \(\mathcal{A}\) given \(\mathcal{D}\) defined by

\[
H(\mathcal{A}|\mathcal{D}) = -\sum_{i=0}^{k} \sum_{j=1}^{k} \mu(D_i \cap A_j) \log \frac{\mu(D_i \cap A_j)}{\mu(D_i)}
\]

The expression in the brackets above is the entropy of the cover \(\mathcal{A}\) restricted to the set \(D_0\), and thus it is bounded by the log card \(\mathcal{A} = \log k\). Therefore, \(H(\mathcal{A}|\mathcal{D}) \leq \mu(D_0) \log k \leq k\varepsilon \log k \leq 1\).

We arrive at the inequality \(h(\mathcal{D}^n) \leq 1 + \log 2 + h(r, w, f, \Phi)\). Since \(\mathcal{A}\) was an arbitrarily chosen Borel partition we obtain \(h(\mathcal{A}) \leq 1 + \log 2 + h(r, w, f, \Phi)\).

Now we are ready to prove the main theorem.

**Proof of Theorem 2.** Since \(h_m(w, f, \Phi)\) is approximated by ergodic measures, for any \(\varepsilon > 0\) and \(r > 0\) there is an ergodic measure \(\mu = \mu(\varepsilon, r)\) such that \(rv(\mu) \in B(r, w)\) and \(|h(\mu) - h_m(w, f, \Phi)| < \varepsilon\). Since \(\mu\) is ergodic, we can apply Lemma 3 and obtain

\[
h(\mu) \leq 1 + \log 2 + h(r, w, f, \Phi).
\]
Fix any $k \in \mathbb{N}$. As before, denote $\Phi_k = \frac{1}{k} S_k(f, \Phi)$. Then $rv_{\Phi_k}(\mu) = rv_{\Phi}(\mu)$ and by Lemma 1 we have
\[ \frac{1}{n} S_n(f^k, \Phi_k)(x) = \frac{1}{kn} S_{kn}(f, \Phi)(x) \to rv_{\Phi_k}(\mu) \quad \text{for } \mu\text{-almost all } x. \]
Therefore, measure $\mu$ satisfies the assumptions of Lemma 3 for the maps $f^k$ and $\Phi_k$. We obtain
\[ h_{\mu}(f^k) \leq 1 + \log 2 + h(r, w, f^k, \Phi_k) \]
Application of the power rule for measure-theoretic entropy of $\mu$ on the left-hand side and Lemma 2 on the right gives
\[ kh_{\mu}(f) \leq 1 + \log 2 + kh(r, w, f, \Phi) \]
Since $k \in \mathbb{N}$ was arbitrary, we obtain $h_{\mu}(f) \leq h(r, w, f, \Phi)$. By the choice of measure $\mu$ we have $h_m(w, f, \Phi) - \varepsilon \leq h(r, w, f, \Phi)$. Finally, letting $\varepsilon$ and $r$ approach 0 we obtain the desired inequality
\[ h_m(w, f, \Phi) \leq h_{\text{top}}(w, f, \Phi). \]
This completes the proof of the theorem. \qed

References


Department of Mathematics, The City College of New York, New York, NY, 10031, USA

\textit{E-mail address: tkucherenko@ccny.cuny.edu}