*Pacific Journal of Mathematics*

# FORWARD GENERATOR FOR PREIMAGE ENTROPY

WEN-CHIAO CHENG

Volume 223 No. 1 January 2006

PACIFIC JOURNAL OF MATHEMATICS
Vol. 223, No. 1, 2006

# FORWARD GENERATOR FOR PREIMAGE ENTROPY

WEN-CHIAO CHENG

**We investigate preimage entropy and show how to calculate it under the assumption of forward expansiveness. We then define new invariants of noninvertible maps, called the upper preimage entropy and the metric preimage entropy, and obtain variational principles for them. Last, we prove a similar result for the Kolmogorov–Sinai generator.**

## 1. Introduction

The topological entropy $h(T)$ of a continuous map $T$ of a compact metric space to itself is a measure of its dynamical complexity. It was first defined by Adler, Konheim and McAndrew, and later was given several equivalent definitions by Bowen and others (see [Bowen 1972] for an exposition). These definitions have led to results connecting topological and measure-theoretic entropies.

More recently, the preimage relation entropy $h_r(T)$ of a compact metric space was introduced in [Langevin and Walczak 1991] and used as a new tool for studying the topology and dynamics of endomorphisms of compact metric spaces. Hurley [1995] and Nitecki and Przytycki [1999] then introduced several other entropy-like invariants for noninvertible maps. One, the preimage branch entropy $h_i(T)$, is closely related to $h_r(T)$. Two others are based on how many branches of the inverse of the iterated map $T^{-n}$ at a point $x$ can be distinguished by measurements of finite accuracy; they are called pointwise preimage entropies and are denoted by $h_p(T)$ and $h_m(T)$. The following inequalities hold for any continuous map $T$ on a compact metric space [Hurley 1995; Nitecki and Przytycki 1999]:

$$h_p(T) \leq h_m(T) \leq h(T) \leq h_i(T) + h_m(T) \leq h_r(T) + h_m(T).$$

In this paper we first concentrate on the pointwise preimage entropies $h_p(T)$ and $h_m(T)$, whose definitions are in some sense analogous to (and were motivated by) Bowen's notion of "local entropy" (see [Bowen 1972]). The definitions and some basic properties of these invariants are reviewed in Section 2.

*MSC2000:* 37B40.
*Keywords:* preimage entropy, forward generator, expansiveness.

In Section 3 we consider forward expansive maps. We show that, as in the case of the standard topological entropy, both $h_p(T)$ and $h_m(T)$ can be computed using the forward generator.

In Section 4 we introduce modified preimage entropies in the topological and measure-theoretic contexts. We obtain variational principles and a similar result for generators in the case of a measure-preserving transformation. We also show that the construction of the metric preimage entropy equals the conditional metric preimage entropy with infinite past $\sigma$-algebra.

## 2. Pointwise preimage entropy

As in Adler, Konheim and McAndrew's definition of topological entropy, we take an open cover $U$ of a compact space $X$ and a continuous map $T : X \to X$, and set

$$\bigvee_{i=0}^{n} T^{-i}(U) = \{U_{j_0} \cap T^{-1}U_{j_1} \cap \cdots \cap T^{-n}U_{j_n} \mid U_{j_k} \in U, 0 \leq k \leq n\}.$$

This is an open cover and a refinement of $U$. If $x \in U_{j_0} \cap T^{-1}U_{j_1} \cap \cdots \cap T^{n}U_{j_n}$, then $x \in U_{j_0}$, $Tx \in U_{j_1}, \ldots, T^n x \in U_{j_n}$.

**Definition 2.1.** Let $T : X \to X$ be a continuous map of a compact space $X$ and take $x \in X$. For $N = 1, 2, \ldots$, the *$N$-th preimage set* of $x$ under $T$ is the set $T^{-N}(x) := \{z \in X \mid T^N(z) = x\}$, and the *$N$-th branch* at $x$ is the set

$$B_N(x, T) = \{(z_N, z_{N-1}, \ldots, z_0) \mid T(z_{i+1}) = z_i,\ 0 \leq i \leq N-1 \text{ and } z_0 = x\}.$$

Let $O(X)$ be the collection of all (finite or infinite) open covers of $X$. Given $U \in O(X)$, let $U^N$ be the open cover of $X^N$ by product sets $U_1 \times U_2 \times \cdots \times U_N$, where $U_i \in U$. For a subset $S_N \subset X^N$, define $\aleph(U, N, S_N)$ to be the least cardinality of a subset of $U^N$ that covers $S_N$.

**Remark 2.2.** The continuity of $T$ and the compactness of $X$ insure that $B_N(x, T)$ is compact, and hence that the numbers $\aleph(U, N, B_N(x, T))$ are all finite and bounded for fixed $N$ over $x \in X$.

**Definition 2.3** (Pointwise preimage entropies [Nitecki and Przytycki 1999]). For $T : X \to X$ be a continuous mapping from a compact space $X$ to itself, define

$$h_p(T) = \sup_{x \in X} \sup_{U \in O(X)} \limsup_{N \to \infty} \frac{1}{N} \log \aleph(U, N, B_N(x, T)),$$

$$h_m(T) = \sup_{U \in O(X)} \limsup_{N \to \infty} \frac{1}{N} \log \sup_{x \in X} \aleph(U, N, B_N(x, T)).$$

**Remark 2.4.** If $T$ is a homeomorphism, then $h_p(T) = h_m(T) = 0$.

As in the case of the topological entropy, we can give metric definitions for these invariants by reinterpreting the numbers $\aleph(U, N, S_N)$ in terms of $\varepsilon$-spanning and

$\varepsilon$-separated sets. Given a metric space $(X, d)$ and some $\varepsilon > 0$, we say that a subset $S \subset X$ is $\varepsilon$-separated if any two distinct points of $S$ are at least $\varepsilon$ apart; and we say that $A \subset X$ is $\varepsilon$-spanned by $R \subset X$ (or that $R$ $\varepsilon$-spans $A$) if any point of $A$ is at a distance of at most $\varepsilon$ from $R$. Set

$$r(\varepsilon, d, A) = \min\{\text{card}\, R \mid R \text{ is } \varepsilon\text{-spans } A\},$$
$$s(\varepsilon, d, A) = \max\{\text{card}\, S \mid S \subset A \text{ is } \varepsilon\text{-separated}\}.$$

**Theorem 2.5** [Nitecki and Przytycki 1999]. *Let $(X, d)$ be a compact metric space. For any positive integer $N$, let $d^N$ be the metric on $X^N$ given by*

$$d^N((x_1, \ldots, x_N), (y_1, \ldots, y_N)) = \max_{1\le i\le N} d(x_i, y_i).$$

*Then for $T : X \to X$ continuous, we have*

$$h_p(T) = \sup_{x\in X} \lim_{\varepsilon\to 0} \limsup_{N\to\infty} \frac{1}{N} \log s(\varepsilon, d^N, B_N(x, T)),$$
$$h_m(T) = \lim_{\varepsilon\to 0} \limsup_{N\to\infty} \frac{1}{N} \log \sup_{x\in X} s(\varepsilon, d^N, B_N(x, T)).$$

*In both formulas, $s(\varepsilon, d^N, B_N(x, T))$ can be replaced by $r(\varepsilon, d^N, B_N(x, T))$.*

We define a new metric $d_n$ on $X$ by setting

$$d_n(x, y) = \max_{0\le i\le n-1} d(T^i(x), T^i(y)).$$

A subset $F$ of $X$ is said to *$(n, \varepsilon)$-span* $K$ if for all $x \in K$ there exists $y \in F$ with $d_n(x, y) \le \varepsilon$. Let $r_n(\varepsilon, K)$ denote the smallest cardinality of any $(n, \varepsilon)$-spanning set for $K$. A similar definition holds for $(n, \varepsilon)$ separated sets and $s_n(\varepsilon, K)$.

**Definition 2.6** [Walters 1982]. Let $U$ be a cover of $X$ and $Y$ a subset of $X$. We denote by $\aleph(U)|_Y$ the smallest cardinality of a subcover of $U$ which can cover $Y$.

For $U \in O(X)$, we easily see that $r(\varepsilon, d^N, B_N(x, T)) = r_N(\varepsilon, T^{-N}(x))$, that $s(\varepsilon, d^N, B_N(x, T)) = s_N(\varepsilon, T^{-N}(x))$, and that

$$\aleph(U, N, B_N(x, T)) = \aleph\big(\textstyle\bigvee_{n=0}^{N} T^{-n}U\big)|_{T^{-N}(x)}.$$

**Remark 2.7.** The pointwise preimage entropies are also given by

$$\begin{aligned} h_p(T) &= \sup_{x\in X} \lim_{\varepsilon\to 0} \limsup_{N\to\infty} \frac{1}{N} \log s_N(\varepsilon, T^{-N}(x)) \\ &= \sup_{x\in X} \lim_{\varepsilon\to 0} \limsup_{N\to\infty} \frac{1}{N} \log r_N(\varepsilon, T^{-N}(x)) \\ &= \sup_{x\in X} \sup_{U\in O(X)} \limsup_{N\to\infty} \frac{1}{N} \log \aleph\big(\textstyle\bigvee_{n=0}^{N} T^{-n}U\big)|_{T^{-N}(x)} \end{aligned}$$

and

$$\begin{aligned} h_m(T) &= \lim_{\varepsilon\to 0} \limsup_{N\to\infty} \frac{1}{N} \log \sup_{x\in X} s_N(\varepsilon, T^{-N}(x)) \\ &= \lim_{\varepsilon\to 0} \limsup_{N\to\infty} \frac{1}{N} \log \sup_{x\in X} r_N(\varepsilon, T^{-N}(x)) \\ &= \sup_{U\in O(X)} \limsup_{N\to\infty} \frac{1}{N} \log \sup_{x\in X} \aleph\big(\textstyle\bigvee_{n=0}^{N} T^{-n}U\big)|_{T^{-N}(x)}. \end{aligned}$$

If $X$ is the circle or a closed interval, $h_p(T)$ and $h_m(T)$ always coincide with the topological entropy of $T$ [Langevin and Przytycki 1992; Langevin and Walczak 1991]. But there exist $X$ and $T$ such that $h_p(T)=0$ and $h_m(T)>0$; in fact, $X$ can be taken to be a zero-dimensional compact metric space [Nitecki and Przytycki 1999].

**Theorem 2.8** [Nitecki and Przytycki 1999]. *If $T_1 : X \to X$ and $T_2 : Y \to Y$ are topologically conjugate, then $h_p(T_1) = h_p(T_2)$ and $h_m(T_1) = h_m(T_2)$.*

**Theorem 2.9.** *If $d, d'$ are metrics on a compact set $X$ defining the same topology, the pointwise preimage entropies with respect to $d$ and $d'$ coincide.*

If $T_2$ is a factor of $T_1$ then $h(T_2) \le h(T_1)$, where $h$ is the topological entropy. This inequality need not hold for pointwise preimage entropies, which can increase when we pass to factors. As an example, consider for a map $f : X \to X$ the shift $\sigma_f$ defined on the sequence space $\Sigma_f = \big\{\{x_i\}_{i=0}^{\infty} \mid f(x_i) = x_{i-1}, i = 1, 2, \dots\big\}$ by

$$\sigma_f(x_0, x_1, \dots) = (f(x_0), f(x_1), \dots) = (f(x_0), x_0, x_1, \dots).$$

The product topology on $\Sigma_f \subset X^N$ makes $\Sigma_f$ compact and $\sigma_f$ a homeomorphism. Furthermore, if $f$ is surjective, $f$ is a factor of $\sigma_f$ via the projection $\varphi(\{x_i\}_{i=0}^{\infty}) = x_0$. By Remark 2.4, we have

$$h_p(\sigma_f) = h_m(\sigma_f) = 0.$$

Taking $f$ such that $h_p(f) = h_m(f) > 0$ (for instance, the standard expanding map $z \mapsto 2z \bmod 1$ of the circle $S^1 = \mathbb{R}/\mathbb{Z}$, for which $h_p(f) = h_m(f) = \log 2$) and setting $g = \sigma_f$, we see that there exist maps $f : X \to X$, $g : Y \to Y$ with $f$ a factor of $g$ and

$$h_m(f) = h_p(f) > h_m(g) = h_p(g).$$

The pointwise preimage entropies are subadditive under Cartesian products and multiplicative under iteration:

**Lemma 2.10** [Nitecki and Przytycki 1999]. *Let $X, X'$ be compact metric spaces and $T : X \to X$, $T' : X' \to X'$ continuous maps. Then*

$$h_p(T \times T') \le h_p(T) + h_p(T'), \qquad h_p(T^k) = k h_p(T),$$
$$h_m(T \times T') \le h_m(T) + h_m(T'), \qquad h_m(T^k) = k h_m(T).$$

## 3. Topological forward generator

Let $X$ be a compact metric space and $T : X \to X$ a continuous map. A finite open cover $\alpha$ of $X$ is a *forward generator* for $T$ if for every sequence $(A_n)_0^\infty$ of members of $\alpha$ the set $\bigcap_{n=0}^\infty T^{-n}\overline{A}_n$ contains at most one point of $X$.

**Lemma 3.1.** *Let $T : X \to X$ be a continuous map of a compact metric space $(X, d)$. Let $\alpha$ be a forward generator for $T$. For any $\varepsilon > 0$, there exists $N > 0$ such that each set in $\bigvee_{n=0}^N T^{-n}\alpha$ has diameter less than $\varepsilon$.*

*Proof.* Suppose the lemma does not hold. Then we can find a positive $\varepsilon$ such that, for all $j > 0$, there exist $x_j, y_j$ satisfying the conditions $d(x_j, y_j) > \varepsilon$ and

$$x_j, y_j \in \bigcap_{i=0}^{j} T^{-i} A_{j,i},$$

for some sequence $\{A_{j,i}\}_{i=0}^{j}$ of sets $A_{j,i} \in \alpha$. Using the compactness of $X$, we can assume (passing to a subsequence if necessary) that $x_j \to x$ and $y_j \to y$. We have $x \neq y$. Consider the sets $A_{j,0}$. Infinitely many of them coincide, since $\alpha$ is finite. Thus $x_j, y_j \in A_0$, say, for infinitely many $j$, and hence $x, y \in \overline{A}_0$. Similarly, for each $n$, infinitely many $A_{j,n}$ coincide and we obtain $A_n \in \alpha$ with $x, y \in T^{-n}\overline{A}_n$. Thus

$$x, y \in \bigcap_{0}^{\infty} T^{-n}\overline{A}_n,$$

contradicting the assumption that $\alpha$ is a forward generator. □

**Definition 3.2** [Nitecki and Przytycki 1999]. A continuous map $T$ from a compact metric space $(X, d)$ to itself is said to be *forward expansive* if there exists $\delta > 0$ such that, for any distinct $x \neq y \in X$, the forward images $T^n x$ and $T^n y$ are more than $\delta$ apart, for some $n$.

**Lemma 3.3.** *A continuous map $T$ from a compact metric space $(X, d)$ to itself is forward expansive if and only if is has a forward generator.*

*Proof.* Suppose $T$ is forward expansive. Let $\delta$ be as in the definition and let $\alpha$ be a finite cover of $X$ by open balls of radius $\delta/2$. Suppose that $x, y \in \bigcap_0^\infty T^{-n}\overline{A}_n$, where $A_n \in \alpha$. Then $d(T^n(x), T^n(y)) \leq \delta$ for all $n \in \mathbb{N} \cup \{0\}$ so, by assumption $x = y$. Then $\alpha$ is a forward generator.

Conversely, suppose $\alpha$ is a forward generator. Let $\delta$ be a Lebesgue number for $\alpha$, If $d(T^n(x), T^n(y)) \leq \delta$ for all $n \in \mathbb{N}\bigcup\{0\}$, then for all $n \in N$ exists $A_n \in \alpha$ with $T^n(x), T^n(y) \in A_n$ and so, $x, y \in \bigcap_0^\infty T^{-n} A_n$. Since this intersection contains at most one point we have $x = y$. Hence $T$ is forward expansive. □

**Example 3.4.** Take $X = \{1, 2, \ldots m\}^N$ and $T =$ left shift. Then $\{[k] : 1 \leq k \leq m\}$ is a forward generator, where $[k] = \big\{(k x_1 x_2 x_3 \ldots) : x_i \in \{1, 2, \ldots, m\}\big\}$.

Recall from Definitions 2.1 and 2.6 the notations $\aleph(U, N, Y)$ and $\aleph(U|_Y)$. Also recall that $O(X)$ is the collection of all open covers of $X$. Take $U \in O(X)$ and $x \in X$, and set

$$h_p(T, U) = \sup_{x \in X} \limsup_{N \to \infty} \frac{1}{N} \log \aleph\big(\textstyle\bigvee_{n=0}^{N} T^{-n} U\big)|_{T^{-N}(x)},$$
$$h_m(T, U) = \limsup_{N \to \infty} \frac{1}{N} \log \sup_{x \in X} \aleph\big(\textstyle\bigvee_{n=0}^{N} T^{-n} U\big)|_{T^{-N}(x)}.$$

**Theorem 3.5.** *Let $T : X \to X$ be a forward expansive continuous map of the compact metric space $(X, d)$. If $\alpha$ is a forward generator for $T$, then*

$$h_p(T) = h_p(T, \alpha) \quad \textit{and} \quad h_m(T) = h_m(T, \alpha).$$

*Proof.* Since $\alpha$ is a forward generator, for any $U \in O(X)$, we can choose $N$ large enough such that $\bigvee_{n=0}^{N} T^{-n}\alpha$ is a refinement of $U$. This implies that

$$\log \aleph\big(\textstyle\bigvee_{n=0}^{k} T^{-n} U\big)|_{T^{-k}(x)} \le \log \aleph\big(\bigvee_{n=0}^{k} T^{-n} \bigvee_{n=0}^{N} T^{-n}\alpha\big)|_{T^{-k}(x)} \quad \text{for any } k.$$

Then

$$\begin{aligned}
&\limsup_{k \to \infty} \frac{1}{k} \log \aleph\big(\textstyle\bigvee_{n=0}^{k} T^{-n} U\big)|_{T^{-k}(x)} \\
&\qquad \le \limsup_{k \to \infty} \frac{1}{k} \log \aleph\big(\textstyle\bigvee_{n=0}^{k} T^{-n} \bigvee_{n=0}^{N} T^{-n}\alpha\big)|_{T^{-k}(x)} \\
&\qquad = \limsup_{k \to \infty} \frac{1}{k} \log \aleph\big(\textstyle\bigvee_{n=0}^{k+N} T^{-n}\alpha\big)|_{T^{-k}(x)} \\
&\qquad \le \limsup_{k \to \infty} \frac{1}{k} \log \aleph\big(\textstyle\bigvee_{n=0}^{k+N} T^{-n}\alpha\big)|_{T^{-(k+N)}(x)} \\
&\qquad = \limsup_{k \to \infty} \frac{k+n}{k} \frac{1}{k+N} \log \aleph\big(\textstyle\bigvee_{n=0}^{k+N} T^{-n}\alpha\big)|_{T^{-(k+N)}(x)} \\
&\qquad \le \limsup_{k \to \infty} \frac{k+N}{k} \limsup_{k \to \infty} \frac{1}{k+N} \log \aleph\big(\textstyle\bigvee_{n=0}^{k+N} T^{-n}\alpha\big)|_{T^{-(k+N)}(x)} \\
&\qquad = \limsup_{k \to \infty} \frac{1}{k+N} \log \aleph\big(\textstyle\bigvee_{n=0}^{k+N} T^{-n}\alpha\big)|_{T^{-(k+N)}(x)}.
\end{aligned}$$

Let

$$h_p(T, U, x) = \limsup_{k \to \infty} \frac{1}{k} \log \aleph(\textstyle\bigvee_{n=0}^{k} T^{-n} U)|_{T^{-k}(x)}.$$

Then $h_p(T, U, x) \le h_p(T, \alpha, x)$ for all open covers $U$ and any fixed $x$ in $X$, which implies

$$h_p(T) = \sup_{x \in X} \sup_{U \in O(X)} h_p(T, U, x) \le h_p(T, \alpha) = \sup_{x \in X} h_p(T, \alpha, x).$$

Then $h_p(T) = h_p(T, \alpha)$.

Similarly, since $U$ is refined by $\bigvee_{n=0}^{N} T^{-N}\alpha$, we have

$$\aleph\big(\bigvee_{n=0}^{k} U\big)|_{T^{-k}(x)} \leq \aleph\big(\bigvee_{n=0}^{k}\bigvee_{n=0}^{N} T^{-n}\alpha\big)|_{T^{-k}(x)}.$$

Thus

$$\begin{aligned} h_m(T, U) &= \limsup_{k\to\infty} \frac{1}{k} \log \sup_{x\in X} \aleph\big(\bigvee_{n=0}^{k} T^{-n}U\big)|_{T^{-k}(x)} \\ &\leq \limsup_{k\to\infty} \frac{1}{k} \log \sup_{x\in X} \aleph\big(\bigvee_{n=0}^{k} T^{-n}\bigvee_{n=0}^{N} T^{-n}\alpha\big)|_{T^{-k}(x)}. \end{aligned}$$

A similar calculation yields $h_m(T, U) \leq h_m(T, \alpha)$ for all open covers $U$. Thus $h_m(T) = h_m(T, \alpha)$. □

## 4. Measure-theoretic forward generator

We continue to consider a continuous self-map $T$ of a compact metric space $(X, d)$. Given a subset $K \subset X$, a $\delta > 0$, and a positive integer $n$, we set

$$r(n, \delta, K) = r(n, \delta, K, T) = \max\{\operatorname{card} E : E \subseteq K \text{ is } (n, \delta)\text{-separated}\}.$$

**Definition 4.1.** The *upper preimage entropy* of a continuous self-map $T : X \to X$ of a compact metric space $X$ is the number

$$\begin{aligned} h_{\mathrm{top}}(T \,|\, \xi^-) &= \lim_{\delta\to 0} \limsup_{n\to\infty} \frac{1}{n} \log \sup_{k\geq 0, x\in X} r(n, \delta, T^{-k}x) \\ &= \sup_{\alpha\in O(X)} \limsup_{n\to\infty} \frac{1}{n} \log \sup_{\substack{k\geq 0\\ x\in X}} \aleph\big(\bigvee_{i=0}^{n-1} T^{-i}\alpha\big)|_{T^{-k}(x)}. \end{aligned}$$

One can check that $h_p(T) \leq h_m(T) \leq h_{\mathrm{top}}(T \,|\, \xi^-) \leq h(T)$.

**Example 4.2.** Consider $S : \{1, 2\}^{\mathbb{N}} \to \{1, 2\}^{\mathbb{N}}$ and $T : \{1, 2\}^{\mathbb{Z}} \to \{1, 2\}^{\mathbb{Z}}$, where $S$ and $T$ are left shifts. We know that $h_p(S) = h_m(S) = h_{\mathrm{top}}(S) = \log 2$, that $h_p(T) = h_m(T) = h_{\mathrm{top}}(T) = 0$, and that the product rule holds, i.e.,

$$h_{\mathrm{top}}(S \times T \,|\, \xi^-) = h_{\mathrm{top}}(S \,|\, \xi^-) + h_{\mathrm{top}}(T \,|\, \xi^-)$$

(see [Cheng and Newhouse 2005]) and

$$h(S \times T) = h(S) + h(T).$$

It follows that

$$h_p(S \times T) = h_m(S \times T) = h_{\mathrm{top}}(S \times T \,|\, \xi^-) = \log 2 < h(S \times T) = 2 \log 2.$$

Now we introduce conditional entropy. Let $\zeta = \{A_1, A_2, \dots\}$ be a countable partition of $X$ into measurable sets. For each $x \in X$, denote by $\zeta(x)$ the element of $\zeta$ to which $x$ belongs. The information function associated to $\zeta$ is defined to be

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

so that $I_\zeta(x)$ takes the constant value $-\log m(A)$ on the cell $A$ of $\zeta$. Clearly

$$H(\zeta) = \int_X I_\zeta(x)\, dm(x).$$

It is useful to consider conditional information and entropy, which take into account information that may already be in hand. Let $\mathcal{I}$ be a sub-$\sigma$-algebra of $\mathcal{B}$. Recall that for $\phi \in L^1(X)$, the conditional expectation $E(\phi \mid \mathcal{I})$ of $\phi$ given $\mathcal{I}$ is an $\mathcal{I}$-measurable function on $X$ satisfying

$$\int_F E(\phi \mid \mathcal{I})\, dm = \int_F \phi\, dm$$

for all $F \in \mathcal{I}$; the name comes from the fact that $E(\phi \mid \mathcal{I})(x)$ represents our expected value for $\phi$ if we are given the foreknowledge $\mathcal{I}$. Thus we let $m(A \mid \mathcal{I}) = E(\chi_A \mid \mathcal{I})$ and define the conditional information function of a countable partition $\zeta$ given a $\sigma$-algebra $\mathcal{I} \subset \mathcal{B}$ to be

$$I_{\zeta|\mathcal{I}}(x) = -\sum_{A\in\zeta} \log m(A \mid \mathcal{I})\chi_A(x).$$

The conditional entropy of $\zeta$ given $\mathcal{I}$ is defined by

$$H(\zeta \mid \mathcal{I}) = \int_X I_{\zeta|\mathcal{I}}(x)\, dm.$$

Next, let $\xi$ denote the point partition of $X$, we also identify with the $\sigma$-algebra $\mathcal{B}$ of Borel measurable sets. For $n > 0$, set

$$\xi^{-n} = T^{-n}\xi.$$

Given a finite partition $\alpha$, let $\alpha^n = \bigvee_{i=0}^{n-1} T^{-i}\alpha$. For a $T$-invariant probability $\mu$, let

$$H_\mu(\alpha^n \,|\, \xi^{-k})$$

denote the conditional entropy of $\alpha^n$ given the $\sigma$-algebra $T^{-k}\mathcal{B}$. We call this the *entropy of $\alpha^n$ given the preimage partition $\xi^{-k}$.*

Since $H_\mu(\cdot \mid \cdot)$ is increasing in the first variable and decreasing in the second, the inequalities $n \geq m, l \geq k$ imply

$$H_\mu(\alpha^n \,|\, \xi^{-l}) \geq H_\mu(\alpha^m \,|\, \xi^{-k}).$$

Set

$$H_\mu(\alpha^n \mid \xi^-) = H_\mu(\alpha^n \mid \xi^{-\infty}) = \sup_{k\geq 0} H_\mu(\alpha^n \mid \xi^{-k}) = \lim_{k\to\infty} H_\mu(\alpha^n \mid \xi^{-k}).$$

We can also define

$$\xi^- = \bigcap_{k=1}^{\infty} \xi^{-k},$$

and we call this eventual range $\xi^-$ the infinity past $\sigma$-algebra.

**Lemma 4.3** [Bowen 1972]. *The quantity $a_n = H_\mu(\alpha^n \mid \xi^-)$ is subadditive.*

**Definition 4.4.** For any finite partition $\alpha$, the *entropy of $\alpha$ given $\xi^-$* is the number

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

and we define the *metric preimage entropy of $T$ given $\xi^-$ with respect to $\mu$* to be

$$h_\mu(T \mid \xi^-) = \sup_\alpha h_\mu(\alpha \mid \xi^-) = \sup_\alpha h_\mu(T \mid \xi^-, \alpha).$$

**Lemma 4.5.** *The metric preimage entropy $h_\mu(T \mid \xi^-)$ is a measure-theoretic conjugacy invariant. The upper preimage entropy $h_{\text{top}}(T \mid \xi^-)$ is a topological conjugacy invariant.*

**Theorem 4.6** (Variational principle [Cheng and Newhouse 2005]). *Let $T : X \to X$ be a continuous map of a compact metric space $X$. Then*

$$h_{\text{top}}(T \mid \xi^-) = \sup_\mu h_\mu(T \mid \xi^-),$$

*where $\mu$ runs over all $T$-invariant Borel probability measures on $X$.*

**Theorem 4.7.** *Let $T : X \to X$ be a forward expansive continuous function of a compact metric space $(X, d)$. If $\alpha$ is a forward generator for $T$, then*

$$h_{\text{top}}(T \mid \xi^-) = h_{\text{top}}(T \mid \xi^-, \alpha).$$

The proof is similar to that of Theorem 3.5.

**Lemma 4.8.** *Let $\zeta$ and $\eta$ be two finite partitions of $X$. Then*

$$h_\mu(\zeta \mid \xi^-) \leq h_\mu(\eta \mid \xi^-) + H_\mu(\zeta \mid \eta).$$

*Proof.* We have

$$\begin{aligned} H_\mu\big(\textstyle\bigvee_{i=0}^{n-1} T^{-i}\zeta \mid \xi^{-k}\big) &\leq H_\mu\big(\big(\textstyle\bigvee_{i=0}^{n-1} T^{-i}\zeta \vee \bigvee_{i=0}^{n-1} T^{-i}\eta\big) \mid \xi^{-k}\big) \\ &= H_\mu\big(\textstyle\bigvee_{i=0}^{n-1} T^{-i}\eta \mid \xi^{-k}\big) + H_\mu\big(\textstyle\bigvee_{i=0}^{n-1} T^{-i}\zeta \mid \bigvee_{i=0}^{n-1} T^{-i}\eta \vee \xi^{-k}\big) \\ &\leq H_\mu\big(\textstyle\bigvee_{i=0}^{n-1} T^{-i}\eta \mid \xi^{-k}\big) + H_\mu\big(\textstyle\bigvee_{i=0}^{n-1} T^{-i}\zeta \mid \bigvee_{i=0}^{n-1} T^{-i}\eta\big). \end{aligned}$$

Let $k \to \infty$; since $H_\mu\big(\bigvee_{i=0}^{n-1} T^{-i}\zeta \mid \bigvee_{i=0}^{n-1} T^{-i}\eta\big) \le n \cdot H_\mu(\zeta \mid \eta)$, this implies that $H_\mu(\zeta_0^n \,|\, \xi^-) \le H_\mu(\eta_0^n \,|\, \xi^-) + n \cdot H_\mu(\zeta \mid \eta)$. Divide by $n$ and let $n$ go to infinity; then $h_\mu(\zeta \,|\, \xi^-) \le h_\mu(\eta \,|\, \xi^-) + H_\mu(\zeta \mid \eta)$. □

**Lemma 4.9.** *For any fixed k,*

$$h_\mu(T \,|\, \xi^-, \alpha) = h_\mu\big(T \,|\, \xi^-, \textstyle\bigvee_{i=0}^{k} T^{-i}\alpha\big).$$

*Proof.*
$$\begin{aligned} h_\mu\big(T \,|\, \xi^-, \textstyle\bigvee_{i=0}^{k} T^{-i}\alpha\big) &= \lim_{n\to\infty} \frac{1}{n} H_\mu\big(\textstyle\bigvee_{i=0}^{n-1} T^{-i}\big(\bigvee_{i=0}^{k} T^{-i}\alpha\big) \,|\, \xi^-\big) \\ &= \lim_{n\to\infty} \frac{1}{n} H_\mu\big(\textstyle\bigvee_{i=0}^{k+n-1} T^{-i}\alpha \,|\, \xi^-\big) \\ &= \lim_{n\to\infty} \frac{k+n-1}{n} \frac{1}{k+n-1} H_\mu\big(\textstyle\bigvee_{i=0}^{k+n-1} T^{-i}\alpha \,|\, \xi^-\big) \\ &= h_\mu(T \,|\, \xi^-, \alpha). \end{aligned}$$
□

**Lemma 4.10.** *If $\{A_n\}$ is an increasing sequence of finite partitions of $X$ and $C$ is a partition with $C \le \bigvee_{i=0}^{\infty} A_i$, then $H_\mu(C \mid A_n) \to 0$ as $n \to \infty$.*

Let $C = \{C_i : i = 1, 2, \ldots, n\}$ be a finite sub-$\sigma$-algebra of $\mathcal{B}$. The nonempty sets of the form $B_1 \cap B_2 \cdots \cap B_n$, where $B_i = C_i$ or $X \backslash C_i$, form a finite partition of $X$. We denote it by $\alpha(C)$ and we define $h_\mu(T \,|\, \xi^-, C) = h_\mu(T \,|\, \xi^-, \alpha(C))$.

As in the case of measure-theoretic entropy, the main method for calculating $h_\mu(T \,|\, \xi^-)$ is supplied by the next theorem.

**Theorem 4.11** (Kolmogorov–Sinai forward generator). *Let $T$ be a measure-preserving transformation of $(X, \mathcal{B}, \mu)$ and $\mathcal{R}$ a finite sub-$\sigma$-algebra such that*

$$\textstyle\bigvee_{n=0}^{\infty} T^{-n}(\alpha(\mathcal{R})) = \mathcal{B}.$$

*Then*

$$h_\mu(T \,|\, \xi^-) = h_\mu(T \,|\, \xi^-, \mathcal{R}).$$

*Proof.* Let $C$ be any partition. We show that $h_\mu(T \mid \xi^-, C) \le h_\mu(T \,|\, \xi^-, \alpha(\mathcal{R}))$. For $n \ge 1$, by Lemmas 4.8 and 4.9,

$$\begin{aligned} h_\mu(T \,|\, \xi^-, C) &\le h_\mu\big(T \,|\, \xi^-, \textstyle\bigvee_{i=0}^{n} T^{-i}\alpha(\mathcal{R})\big) + H_\mu(C \mid \bigvee_{i=0}^{n} T^{-i}\alpha(\mathcal{R})) \\ &= h_\mu\big(T \,|\, \xi^-, \alpha(\mathcal{R})\big) + H_\mu\big(C \mid \textstyle\bigvee_{i=0}^{n} T^{-i}\alpha(\mathcal{R})\big). \end{aligned}$$

Let $A_n = \bigvee_{i=0}^{n} T^{-i}\alpha(A)$ be as in Lemma 4.10. Then $H_\mu(C \mid A_n) \to 0$ as $n \to \infty$. This implies $h_\mu(T \mid \xi^-, C) \le h_\mu(T \mid \xi^-, \alpha(\mathcal{R}))$. Therefore

$$h_\mu(T \mid \xi^-, C) \le h_\mu(T \mid \xi^-, \mathcal{R}).$$
□

We end this section with some propositions about $h_\mu(T \,|\, \xi^-)$ and from those results we conclude that $h_\mu(T \,|\, \xi^-) \le h_\mu(T)$, where $h_\mu(T)$ is the measure-theoretic entropy.

**Lemma 4.12.** *We let $\mathcal{B}_\infty = \bigvee_{n=1}^{\infty} \mathcal{B}_n$ if $\{\mathcal{B}_n\}$ is an increasing sequence of sub-$\sigma$-algebras of $X$ and let $\mathcal{B}_\infty = \bigcap \mathcal{B}_n$ if $\{\mathcal{B}_n\}$ is a decreasing sequence. If $\alpha$ is a finite partition, then*

$$\lim_{n\to\infty} H_\mu(\alpha \mid \mathcal{B}_n) = H_\mu(\alpha \mid \mathcal{B}_\infty).$$

*Proof.* We show the decreasing case; the discussion for the increasing sequence is similar — see [Petersen 1983, Proposition 5.2.11].

Take $A \in \alpha$. Because $E(E(\chi_A \mid \mathcal{B}_{n-1}) \mid \mathcal{B}_n) = E(\chi_A \mid \mathcal{B}_n)$, by the reverse martingale theorem and [Billingsley 1995, Theorem 35.9], we have

$$\lim_{n\to\infty} E(\chi_A \mid \mathcal{B}_n) = E(\chi_A \mid \mathcal{B}_\infty).$$

Also $I_{\alpha|B_n} = -\sum_{A\in\alpha} \log E(\chi_A \mid \mathcal{B}_n)\cdot E(\chi_A \mid \mathcal{B}_n)$ is a bounded continuous function; thus, by the bounded convergence theorem, we get

$$\lim_{n\to\infty} H(\alpha \mid \mathcal{B}_n) = \lim_{n\to\infty} \int I_{\alpha|\mathcal{B}_n}\, d\mu = \int \lim_{n\to\infty} I_{\alpha|\mathcal{B}_n}\, d\mu = H(\alpha \mid \mathcal{B}_\infty). \qquad \square$$

**Lemma 4.13.** *Let $\alpha$ be a finite partition. Then*

$$\begin{aligned} h_\mu(\alpha \mid \xi^-) = h_\mu(T \mid \xi^-, \alpha) &= \lim_{n\to\infty}\lim_{k\to\infty} H_\mu\big(\alpha \mid \textstyle\bigvee_{l=1}^{n-1} T^{-l}\alpha \vee T^{-k}(\xi)\big) \\ &= \lim_{n\to\infty} H_\mu\big(\alpha \mid \lim_{k\to\infty} \textstyle\bigvee_{l=1}^{n-1} T^{-l}\alpha \vee T^{-k}(\xi)\big) \\ &= \lim_{n\to\infty} H_\mu\big(\alpha \mid \textstyle\bigvee_{l=1}^{n-1} T^{-l}\alpha \vee \xi^-\big). \end{aligned}$$

*Proof.* We have

$$\begin{aligned} &\lim_{k\to\infty} H_\mu(\alpha \vee T^{-1}\alpha \vee \cdots \vee T^{-(j-1)}\alpha \mid T^{-k}(\xi)) \\ &\qquad = \lim_{k\to\infty} H_\mu\big(\alpha \mid \textstyle\bigvee_{l=1}^{j-1} T^{-l}\alpha \vee T^{-k}(\xi)\big) + \lim_{k\to\infty} H_\mu\big(\textstyle\bigvee_{l=1}^{j-1} T^{-l}\alpha \mid T^{-k}(\xi)\big), \end{aligned}$$

which implies

$$\begin{aligned} &\lim_{k\to\infty} H_\mu\big(\alpha \mid \textstyle\bigvee_{l=1}^{j-1} T^{-l}\alpha \vee T^{-k}(\xi)\big) \\ &\quad = \lim_{k\to\infty} H_\mu(\alpha \vee T^{-1}\alpha \vee \cdots \vee T^{-(j-1)}\alpha \mid T^{-k}(\xi)) - \lim_{k\to\infty} H_\mu\big(\textstyle\bigvee_{l=1}^{j-1} T^{-l}\alpha \mid T^{-k}(\xi)\big) \\ &\quad = \lim_{k\to\infty} H_\mu\big(\textstyle\bigvee_{l=0}^{j-1} T^{-l}\alpha \mid T^{-k}(\xi)\big) - \lim_{k\to\infty} H_\mu\big(\textstyle\bigvee_{l=0}^{j-2} T^{-l}\alpha \mid T^{-(k-1)}(\xi)\big). \end{aligned}$$

We thus get

$$\begin{aligned} &\sum_{j=2}^{n} \lim_{k\to\infty} H_\mu\big(\alpha \mid \textstyle\bigvee_{l=1}^{j-1} T^{-l}\alpha \vee T^{-k}(\xi)\big) \\ &\qquad\qquad = \lim_{k\to\infty} H_\mu\big(\textstyle\bigvee_{l=0}^{n-1} T^{-l}\alpha \mid T^{-k}(\xi)\big) - \lim_{k\to\infty} H_\mu(\alpha \mid T^{-k}(\xi)) \end{aligned}$$

By Cesàro's Theorem and Lemma 4.12,

$$\begin{aligned} h_\mu(\alpha \,|\, \xi^-) &= \lim_{n\to\infty} \lim_{k\to\infty} H_\mu\big(\alpha \,|\, \textstyle\bigvee_{l=1}^{n-1} T^{-l}\alpha \vee T^{-k}(\xi)\big) \\ &= \lim_{n\to\infty} H_\mu\big(\alpha \,|\, \lim_{k\to\infty} \textstyle\bigvee_{l=1}^{n-1} T^{-l}\alpha \vee T^{-k}(\xi)\big) \\ &= \lim_{n\to\infty} H_\mu\big(\alpha \,|\, \textstyle\bigvee_{l=1}^{n-1} T^{-l}\alpha \vee \xi^-\big). \end{aligned}$$ □

**Lemma 4.14.** $h_\mu(T \,|\, \xi^-) \le h_\mu(T)$.

*Proof.* For any finite partition $\alpha$, $H_\mu\big(\alpha \,|\, \bigvee_{i=1}^{n-1} T^{-i}\alpha \vee \xi^-\big) \le H_\mu\big(\alpha \,|\, \bigvee_{i=1}^{n-1} T^{-i}\alpha\big)$ and $h_\mu(T, \alpha) = \lim_{n\to\infty} H_\mu\big(\alpha \,|\, \bigvee_{i=1}^{n-1} T^{-i}\alpha\big)$. By Lemma 4.13, $h_\mu(T \,|\, \xi^-) \le h_\mu(T)$. □

## References


[Billingsley 1995] P. Billingsley, *Probability and measure*, 3rd ed., John Wiley & Sons, New York, 1995. MR 95k:60001 Zbl 0822.60002

[Bowen 1972] R. Bowen, "Entropy-expansive maps", *Trans. Amer. Math. Soc.* **164** (1972), 323–331. MR 44 #2907 Zbl 0229.28011

[Cheng and Newhouse 2005] W.-C. Cheng and S. E. Newhouse, "Pre-image entropy", *Ergodic Theory Dynam. Systems* **25**:4 (2005), 1091–1113. MR 2158398

[Hurley 1995] M. Hurley, "On topological entropy of maps", *Ergodic Theory Dynam. Systems* **15**:3 (1995), 557–568. MR 96h:58097 Zbl 0833.54021

[Langevin and Przytycki 1992] R. Langevin and F. Przytycki, "Entropie de l'image inverse d'une application", *Bull. Soc. Math. France* **120**:2 (1992), 237–250. MR 94b:58085 Zbl 0757.58023

[Langevin and Walczak 1991] R. Langevin and P. Walczak, "Entropie d'une dynamique", *C. R. Acad. Sci. Paris Sér. I Math.* **312**:1 (1991), 141–144. MR 92e:58159 Zbl 0723.54020

[Nitecki and Przytycki 1999] Z. Nitecki and F. Przytycki, "Preimage entropy for mappings", *Internat. J. Bifur. Chaos Appl. Sci. Engrg.* **9**:9 (1999), 1815–1843. MR 2000j:37025

[Petersen 1983] K. Petersen, *Ergodic theory*, Cambridge Studies in Advanced Mathematics **2**, Cambridge University Press, Cambridge, 1983. MR 87i:28002 Zbl 0507.28010

[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 April 25, 2004. Revised February 2, 2005.

WEN-CHIAO CHENG
DEPARTMENT OF MATHEMATICS
NATIONAL CHUNG CHENG UNIVERSITY
MIN-HSIUNG, CHIA-YI
TAIWAN, 621

wccheng@math.ccu.edu.tw