The ergodic theory of free group actions: entropy and the $f$-invariant

Lewis Bowen

Abstract. Previous work introduced two measure-conjugacy invariants: the $f$-invariant (for actions of free groups) and $\Sigma$-entropy (for actions of sofic groups). The purpose of this paper is to show that the $f$-invariant is essentially a special case of $\Sigma$-entropy. There are two applications: the $f$-invariant is invariant under group automorphisms and there is a uniform lower bound on the $f$-invariant of a factor in terms of the original system.

Mathematics Subject Classification (2010). 37A35.

Keywords. Free groups, entropy, $f$-invariant.

1. Introduction

The paper [Bo08b] introduced a measure-conjugacy invariant, called $\Sigma$-entropy, for measure-preserving actions of a sofic group. This was applied, for example, to classify Bernoulli shifts over an arbitrary countable linear group. Previously, [Bo08a] introduced the $f$-invariant for measure-preserving actions of free groups. The invariants of both papers have strong analogies with classical Kolmogorov–Sinai entropy. The purpose of this paper is to show that the $f$-invariant is essentially a special case of $\Sigma$-entropy. We apply this result to show the $f$-invariant does not change under group automorphisms and that there is a lower bound on the $f$-invariant of a factor in terms of the $f$-invariant of the system. The introductions to [Bo08a]–[Bo08b] provide further background and motivation for $\Sigma$-entropy and the $f$-invariant.

To define $\Sigma$-entropy precisely, let $G$ be a countable group and let $\Sigma = \{\sigma_i\}_{i=1}^{\infty}$ be a sequence of homomorphisms $\sigma_i : G \to \text{Sym}(m_i)$ where $\text{Sym}(m_i)$ denotes the full symmetric group of the set $\{1, \ldots, m_i\}$. $\Sigma$ is asymptotically free if

$$\lim_{i \to \infty} \frac{\left| \{1 \leq j \leq m_i \mid \sigma_i(g_1)j = \sigma_i(g_2)j \} \right|}{m_i} = 0,$$

for every pair $g_1, g_2 \in G$ with $g_1 \neq g_2$. The treatment of $\Sigma$-entropy given next differs from [Bo08b] in two respects: for simplicity, we assume that each $\sigma_i$ is a homomorphism and we use observables rather than partitions to define it.
We will write $G \curvearrowright X$ to mean $(X, \mu)$ is a standard probability measure space and $T = (T_g)_{g \in G}$ is an action of $G$ on $(X, \mu)$ by measure-preserving transformations. This means that for each $g \in G$, $T_g : X \to X$ is a measure-preserving transformation and $T_{g_1} T_{g_2} = T_{g_1 g_2}$. An observable of $(X, \mu)$ is a measurable map $\phi : X \to A$ where $A$ is a finite or countably infinite set. We will say that $\phi$ is finite if $A$ is finite. Roughly speaking, the $\Sigma$-entropy rate of $\phi$ is the exponential rate of growth of the number of observables $\psi : \{1, \ldots, m_i \} \to A$ that approximate $\phi$. In order to make precise what it means to approximate, we need to introduce some definitions.

If $\phi : X \to A$ and $\psi : X \to B$ are two observables, then the join of $\phi$ and $\psi$ is the observable $\phi \lor \psi : X \to A \times B$ defined by $\phi \lor \psi(x) = (\phi(x), \psi(x))$. If $g \in G$ then $T_g : X \to A$ is defined by $T_g \phi(x) = \phi(T_g x)$. If $H \subset G$ is finite, then let $\phi^H := \bigvee_{h \in H} T_h \phi$ maps $X$ into $A^H$, the direct product of $|H|$ copies of $A$. Let $\phi^H \mu$ denote the pushforward of $\mu$ on $A^H$. In other words, $\phi^H \mu(S) = \mu((\phi^H)^{-1}(S))$ for $S \subset A^H$.

For each $i$, let $\zeta_i$ denote the uniform probability measure on $\{1, \ldots, m_i \}$. If $h : \{1, \ldots, m_i \} \to A$ is an observable and $H \subset G$ then let $\psi^H := \bigvee_{h \in H} \sigma_i(h) \psi$, where $\sigma_i(h) \psi : \{1, \ldots, m_i \} \to A$ is defined by $\sigma_i(h) \psi(j) = \psi(\sigma_i(h) j)$. Of course, $\psi^H$ depends on $\sigma_i$ but, to keep the notation simple, we will leave this dependence implicit. Let $\psi^H \zeta_i$ be the pushforward of $\zeta_i$ on $A^H$. Finally, let $d_{\sigma_i}^H(\phi, \psi)$ be the $l^1$-distance between $\phi^H \mu$ and $\psi^H \zeta_i$. In other words,

$$d_{\sigma_i}^H(\phi, \psi) = \sum_{a \in A^H} |\phi^H \mu(a) - \psi^H \zeta_i(a)|.$$

**Definition 1.** If $\phi : X \to A$ is an observable and $A$ is finite then define the $\Sigma$-entropy rate of $\phi$ by

$$h(\Sigma, T, \phi) := \inf_{H \subset G} \inf_{\varepsilon > 0} \limsup_{i \to \infty} \frac{1}{m_i} \log(|\{\psi : \{1, \ldots, m_i \} \to A \mid d_{\sigma_i}^H(\phi, \psi) \leq \varepsilon\}|).$$

The first infimum above is over all finite subsets $H \subset G$.

**Definition 2.** Define the entropy of $\phi$ by

$$H(\phi) := -\sum_{a \in A} \mu(\phi^{-1}(a)) \log(\mu(\phi^{-1}(a))).$$

**Definition 3.** If $\phi : X \to A$ is an observable and $A$ is countably infinite then let $\pi_n : A \to A_n$ be a sequence of maps such that

1. $A_n$ is a finite set for all $n$;
2. for each $i > j$ there is a map $\pi_{ij} : A_i \to A_j$ such that $\pi_j = \pi_{ij} \circ \pi_i$;
3. $\pi_n$ is asymptotically injective in the sense that for all $a, b \in A$ with $a \neq b$ there exists $N$ such that $n > N$ implies $\pi_n(a) \neq \pi_n(b)$.

Now define

$$h(\Sigma, T, \phi) := \lim_{n \to \infty} h(\Sigma, T, \pi_n \circ \phi).$$
In [Bo08b] it is proven that if $H(\phi) < \infty$ then this limit exists and is independent of the choice of sequence $\{\pi_n\}$.

An observable $\phi$ is generating if the smallest $G$-invariant $\sigma$-algebra on $X$ that contains $\{\phi^{-1}(a)\}_{a \in A}$ is equal to the $\sigma$-algebra of all measurable sets up to sets of measure zero. The next theorem is (part of) the main result of [Bo08b].

**Theorem 1.1.** Let $\Sigma = \{\sigma_i\}$ be an asymptotically free sequence of homomorphisms $\sigma_i : G \to \text{Sym}(m_i)$ for a group $G$. Let $G \curvearrowright^T (X, \mu)$. If $\phi_1$ and $\phi_2$ are two finite-entropy generating observables then $h(\Sigma, T, \phi_1) = h(\Sigma, T, \phi_2)$.

This motivates the following definition.

**Definition 4.** If $\Sigma$ and $T$ are as above then the $\Sigma$-entropy of the action $T$ is defined by $h(\Sigma, T) := h(\Sigma, \phi)$, where $\phi$ is any finite-entropy generating observable (if one exists).

Next let us discuss a slight variation on $\Sigma$-entropy. Let $\{m_i\}_{i=1}^{\infty}$ be a sequence of natural numbers. For each $i \in \mathbb{N}$, let $\mu_i$ be a probability measure on the set of homomorphisms from $G$ to $\text{Sym}(m_i)$. Let $\alpha_i : G \to \text{Sym}(m_i)$ be chosen at random according to $\mu_i$. The sequence $\Sigma = \{\mu_i\}_{i=1}^{\infty}$ is said to be asymptotically free if for every pair $g_1, g_2 \in G$ with $g_1 \neq g_2$,

$$\lim_{i \to \infty} \frac{\mathbb{E}[\{|1 \leq j \leq m_i | \alpha_i(g_1)j = \alpha_i(g_2)j\}|]}{m_i} = 0$$

where $\mathbb{E}[\cdot]$ denotes expected value. The $\Sigma$-entropy rate of an observable $\phi : X \to A$ with $A$ finite is defined by

$$h(\Sigma, T, \phi) := \inf_{H < G} \inf_{\varepsilon > 0} \limsup_{i \to \infty} \frac{1}{m_i} \log(\mathbb{E}[\{\psi : \{1, \ldots, m_i\} \to A | d_{\Sigma}^{H}(\phi, \psi) \leq \varepsilon\}]$$

With these definitions in mind, Theorem 1.1 is still true if “homomorphisms” is replaced with “probability measures on the set of homomorphisms”.

Let us note one more generalization. If $G$ is a semigroup with identity then the above definitions still make sense. Using results from [Bo08c] it can be shown that Theorem 1.1 remains true.

Now let us recall the $f$-invariant from [Bo08a]. Let $G = \langle s_1, \ldots, s_r \rangle$ be either a free group or free semigroup of rank $r$. Let $G \curvearrowright^T (X, \mu)$. Let $\alpha$ be a partition of $X$ into at most countably many measurable sets. The entropy of $\alpha$ is defined by

$$H(\alpha) := -\sum_{A \in \alpha} \mu(A) \log(\mu(A))$$

where, by convention, $0 \log(0) = 0$. If $\alpha$ and $\beta$ are partitions of $X$ then the join is the partition $\alpha \vee \beta := \{A \cap B | A \in \alpha, B \in \beta\}$. Let $B(e, n)$ denote the ball of radius
n in G with respect to the word metric induced by its generating set (which is either \{s_1, \ldots, s_r\} if G is a semigroup or \{s_1^{\pm 1}, \ldots, s_r^{\pm 1}\} is G is a group). Define

\[
F(T, \alpha) := (1 - 2r)H(\alpha) + \sum_{i=1}^{r} H(\alpha \cup T^{-1}_{s_i} \alpha),
\]

\[
\alpha^n := \bigvee_{a \in B(\epsilon, n)} T^{-1}_{g} \alpha,
\]

\[
f(T, \alpha) := \inf_n F(T, \alpha^n).
\]

The partition \(\alpha\) is generating if the smallest \(G\)-invariant \(\sigma\)-algebra containing \(\alpha\) equals the \(\sigma\)-algebra of all measurable sets up to sets of measure zero.

**Theorem 1.2.** Let \(G = \langle s_1, \ldots, s_r \rangle\) be a free group or free semigroup. Let \(G \simeq T (X, \mu)\). If \(\alpha_1\) and \(\alpha_2\) are two generating partitions with \(H(\alpha_1) + H(\alpha_2) < \infty\) then \(f(T, \alpha_1) = f(T, \alpha_2)\).

This theorem was proven in [Bo08c]. The special case in which \(G\) is a group and \(\alpha_1, \alpha_2\) are finite is the main result of [Bo08a]. Because of this theorem, we define the \(f\)-invariant of the action by \(f(T) := f(T, \alpha)\), where \(\alpha\) is any finite-entropy generating partition of \(X\) (if one exists).

In order to relate this result with \(\Sigma\)-entropy, let us make the following definitions. If \(\phi: X \to A\) is an observable, then let \(\tilde{\phi} = \{\phi^{-1}(a)\}_{a \in A}\) be the corresponding partition of \(X\). Define \(F(T, \phi) := F(T, \tilde{\phi})\) and \(f(T, \phi) := f(T, \tilde{\phi})\). The main result of this paper is:

**Theorem 1.3.** Let \(G = \langle s_1, \ldots, s_r \rangle\) be a free group or free semigroup of rank \(r \geq 1\). Let \(G \simeq T (X, \mu)\). Let \(\phi\) be a finite observable. For \(i \geq 1\), let \(\mu_i\) be the uniform probability measure on the set of all homomorphisms from \(G\) to \(\text{Sym}(i)\). Let \(\Sigma = \{\mu_i\}_{i=1}^{\infty}\). Then \(h(\Sigma, T, \phi) = f(T, \phi)\).

We will prove a refined version of this theorem as follows. Recall the definition of \(d^H_{\sigma_i}(\phi, \psi)\) given above. Define

\[
d^*_{\sigma_i}(\phi, \psi) := \sum_{i=1}^{r} d^e_{\sigma_i, s_i}(\phi, \psi).
\]

**Theorem 1.4.** Let \(G\) and \(T\) be as in the previous theorem. Let \(\phi: X \to A\) be a finite observable. Let \(\sigma_i: G \to \text{Sym}(i)\) be a homomorphism chosen uniformly at random. Then

\[
F(T, \phi) = \inf_{\epsilon > 0} \limsup_{i \to \infty} \frac{1}{i} \log(\mathbb{E}[\{\psi: \{1, \ldots, i\} \to A \mid d^*_{\sigma_i}(\phi, \psi) \leq \epsilon\}])
\]

This theorem is proven in Section 2. In Section 3 we deduce Theorem 1.3 from it.
1.1. Application I: automorphism invariance. Let $G$ be a countable group or semigroup. Let $G \sim \alpha (X, \mu)$. Let $\omega: G \to G$ be an automorphism. Let $T^\omega = (T^\omega_g)_{g \in G}$ where $T^\omega_g x := T_{\omega(g)} x$ for all $x \in X$. This new action of $G$ is not necessarily isomorphic to the original action. That is, there might not exist a map $\phi: X \to X$ such that $\phi(T^\omega_g x) = T^\omega_g \phi(x)$ for a.e. $x \in X$ and all $g \in G$.

Suppose that $\Sigma = \{\sigma_i\}$ is an asymptotically free sequence of homomorphisms $\sigma_i: G \to \text{Sym}(m_i)$. Let $\Sigma^\omega = \{\sigma_i \circ \omega\}$. A short exercise reveals that $h(\Sigma, T, \phi) = h(\Sigma^\omega, T^\omega, \phi)$ for any $\phi$.

If $\sigma_i: G \to \text{Sym}(i)$ is chosen uniformly at random, it follows that the law of $\sigma_i \circ \omega$ is the same as the law of $\sigma_i$. Therefore, if $\mu_i$ is the uniform probability measure on the set of homomorphisms from $G$ to $\text{Sym}(i)$ and $\Sigma = \{\mu_i\}$, then $h(\Sigma, T, \phi) = h(\Sigma, T^\omega, \phi)$. Theorem 1.3 now implies:

**Theorem 1.5.** Let $G$ and $T$ be as in Theorem 1.3. Let $\omega: G \to G$ be an automorphism. Then $f(T, \phi) = f(T^\omega, \phi)$ for any finite observable $\phi$.

This implies that $f(T, \phi)$ does not depend on the choice of free generator set $\{s_1, \ldots, s_r\}$ for $G$ since any two free generating sets are related by an automorphism.

1.2. Application II: lower bounds on the $f$-invariant of a factor

**Definition 5.** Let $G \sim \alpha (X, \mu)$ and $G \sim \beta (Y, \nu)$. Then $S$ is a factor of $T$ if there exists a measurable map $\phi: X \to Y$ such that $\phi_* \mu = \nu$ and $\phi(T_g x) = S_g \phi(x)$ for all $g \in G$ and a.e. $x \in X$.

To motivate this section, let us point out two curious facts.

First, Ornstein proved in [Or70] that every factor of a Bernoulli shift over $\mathbb{Z}$ is measurably conjugate to a Bernoulli shift. It is not known whether this holds when $\mathbb{Z}$ is replaced with a nonabelian free group. A counterexample due to Sorin Popa [Po08] (based on [PS07]) shows that if $G$ is an infinite property $T$ group then there exists a factor of a Bernoulli shift over $G$ that is not measurably conjugate to a Bernoulli shift.

Second, the $f$-invariant of an action can be negative. For example, if $X$ is a set with $n$ elements, $\mu$ is the uniform measure on $X$ and $T = (T_g)_{g \in G}$ is a measure-preserving action of $G = \{s_1, \ldots, s_r\}$ on $X$ then $f(T) = -(r-1) \log(n)$.

From these two facts a natural question arises: can the $f$-invariant of a factor of a Bernoulli shift over $G$ be negative? To answer this, let us recall the following result from [Bo08b], Corollary 8.3.

**Lemma 1.6.** Let $G$ be a countable group. Let $\Sigma = \{\sigma_i\}_{i=1}^\infty$ be an asymptotically free sequence of homomorphisms $\sigma_i: G \to \text{Sym}(m_i)$. Let $T$ be a measure-preserving action of $G$ and let $S$ be a factor of $T$. Assume that there exist finite-entropy
generating partitions for $T$ and $S$. Also let $\phi$ be a generating observable for $T$ with $H(\phi) < \infty$. Then

$$h(\Sigma, S) \geq h(\Sigma, T) - H(\phi).$$

So Theorem 1.3 implies:

**Theorem 1.7.** Let $G = \langle s_1, \ldots, s_r \rangle$ be a free group on $r$ generators. Let $T$ be a measure-preserving action of $G$ and let $S$ be a factor of $T$. Assume there exists finite generating partitions for $T$ and $S$. Let $\alpha$ be a finite generating partition for $T$. Then

$$f(S) \geq f(T) - H(\alpha).$$

In order to apply this to Bernoulli shifts, let us recall the definitions. Let $K$ be a finite or countable set and $\kappa$ a probability measure on $K$. Let $(K^G, \kappa^G)$ denote the product measure space. Define $T_g : K^G \to K^G$ by $T_g(x)(h) = x(hg)$. This defines a measure-preserving action of $G$ on $(K^G, \kappa^G)$. It is the *Bernoulli shift* over $G$ with base measure $\kappa$. In [Bo08a] it was shown that $f(T) = H(\kappa)$ where

$$H(\kappa) := -\sum_{k \in K} \mu(\{k\}) \log(\mu(\{k\})).$$

Let $\alpha$ be the canonical partition of $K^G$, i.e., $\alpha = \{A_k : k \in K\}$ where $A_k = \{x \in K^G \mid x(e) = k\}$. Note $H(\alpha) = H(\kappa) = f(T)$. So the theorem above implies the following result.

**Corollary 1.8.** If $S$ is a factor of the Bernoulli shift and if there exists a finite generating partition for $S$ then $f(S) \geq 0$.

It is unknown whether there exists a nontrivial factor $S$ of a Bernoulli shift over a free group $G$ such that $f(S) = 0$.

In [Bo08c], classical Markov chains are generalized to Markov chains over free groups. An explicit example was given of a Markov chain with finite negative $f$-invariant. It follows that this Markov chain cannot be measurably conjugate to a factor of a Bernoulli shift. It can be shown that this Markov chain is uniformly mixing. To contrast this with the classical case, recall that Friedman and Ornstein proved in [FO70] that every mixing Markov chain over the integers is isomorphic to a Bernoulli shift.

Now we can construct a mixing Markov chain with positive $f$-invariant that is not isomorphic to a Bernoulli shift as follows. Let $T$ denote a mixing Markov chain with negative $f$-invariant. Let $S$ denote a Bernoulli shift with $f(S) > -f(T)$. Consider the product action $T \times S$. A short computation reveals that, in general, $f(T \times S) = f(T) + f(S)$. Therefore $T \times S$ has positive $f$-invariant. It can be shown that $T \times S$ is a mixing Markov chain. However it cannot be isomorphic to a Bernoulli shift since it factors onto $T$ which has negative $f$-invariant.
2. Proof of Theorem 1.4

Let $G = \langle s_1, \ldots, s_r \rangle$ be a free group or free semigroup of rank $r$. Let $G \curvearrowright T (X, \mu)$. Let $\phi: X \to A$ be a finite observable.

We will need to consider certain perturbations of the measure $\mu$ with respect to the given observable $\phi: X \to A$. For this purpose we introduce the notion of weights on the graph $\mathcal{G} = (V, E)$ that is defined as follows. The vertex set $V$ equals $A$. For every $a, b \in A$ and every $i \in \{1, \ldots, r\}$ there is a directed edge from $a$ to $b$ labeled $i$. This edge is denoted $(a, b; i)$. We allow the possibility that $a = b$. A weight on $\mathcal{G}$ is a function $W: V \cup E \to [0, 1]$ satisfying

$$W(a) = \sum_{b \in A} W(a, b; i) = \sum_{b \in A} W(b, a; i)$$

for all $i = 1 \ldots r, a \in A$

$$1 = \sum_{a \in A} W(a).$$

For example,

$$W_\mu(a) := \mu(\phi^{-1}(a)),$$

$$W_\mu(a, b; i) := \mu(\{x \in X \mid \phi(x) = a, \phi(T_{s_i} x) = b\})$$

is the weight associated to $\mu$. For a homomorphism $\sigma: G \to \text{Sym}(n)$ and a function $\psi: \{1, \ldots, n\} \to A$ we define the weight $W_{\sigma, \psi}$ by

$$W_{\sigma, \psi}(a) := |\psi^{-1}(a)|/n,$$

$$W_{\sigma, \psi}(a, b; i) := \{|j \mid \psi(j) = a, \psi(\sigma(s_i) j) = b\|/n.$$ 

Note that

$$d_\phi^*(\phi, \psi) = \sum_{i=1}^r \sum_{a, b \in A} |W_\mu(a, b; i) - W_{\sigma, \psi}(a, b; i)|.$$

So given two weights $W_1, W_2$ define

$$d_*(W_1, W_2) := \sum_{i=1}^r \sum_{a, b \in A} |W_1(a, b; i) - W_2(a, b; i)|.$$

**Proposition 2.1.** Let $n$ be a positive integer. Let $W$ be a weight. Suppose that $W(a, b; i)n \in \mathbb{Z}$ for every $a, b \in A$ and every $i = 1 \ldots r$. If $\sigma: G \to \text{Sym}(n)$ is chosen uniformly at random then

$$\mathbb{E}[\{|\psi: \{1, \ldots, n\} \to A \mid d_*(W, W_{\sigma, \psi}) = 0\}|] = \frac{n^{1-r} \prod_{a \in A} (nW(a))^{2r-1}}{\prod_{i=1}^r \prod_{a, b \in A} (nW(a, b; i))!}.$$

**Proof.** Note that if $d_*(W, W_{\sigma, \psi}) = 0$ then $W_{\sigma, \psi}(a) = W(a)$ for all $a \in A$. Equivalently,

$$|\psi^{-1}(a)| = nW(a)$$

for all $a \in A$. (1)
The number of functions $\psi : \{1, \ldots, n\} \to A$ that satisfy this requirement is

$$\frac{n!}{\prod_{a \in A}(nW(a))!}.$$ 

If $\psi_1, \psi_2$ are two different functions that satisfy equation (1) then there is a permutation $\tau \in \text{Sym}(n)$ such that $\psi_1 = \psi_2 \circ \tau$. If $\sigma^\tau : G \to \text{Sym}(n)$ is the homomorphism defined by $\sigma^\tau(g) = \tau \sigma(g) \tau^{-1}$ then $W_{\sigma, \psi_1} = W_{\sigma^\tau, \psi_2}$. Since $\sigma : G \to \text{Sym}(n)$ is chosen uniformly at random, this implies that the probability that $d_*(W, W_{\sigma, \psi_1}) = 0$ is the same as the probability that $d_*(W, W_{\sigma, \psi_2}) = 0$. So fix a particular function $\psi_0$ satisfying equation (1). Then

$$\mathbb{E}[\{\psi : \{1, \ldots, n\} \to A \mid d_*(W, W_{\sigma, \psi}) = 0\}] = \frac{n! \text{Prob}[d_*(W, W_{\sigma, \psi_0}) = 0]}{\prod_{a \in A}(nW(a))!}. \quad (2)$$

For any two weights $W_1, W_2$ and $1 \leq i \leq r$, define

$$d_i(W_1, W_2) := \sum_{a, b \in A} |W_1(a, b; i) - W_2(a, b; i)|.$$

So $d_* = \sum_{i=1}^r d_i$.

The homomorphism $\sigma : G \to \text{Sym}(n)$ is determined by its values $\sigma(s_1), \ldots, \sigma(s_r)$. The event $d_i(W, W_{\sigma, \psi_0}) = 0$ is determined by $\sigma(s_i)$. So if $i \neq j$ then the events $d_i(W, W_{\sigma, \psi_0}) = 0$ and $d_j(W, W_{\sigma, \psi_0}) = 0$ are independent. Therefore,

$$\mathbb{E}[\{\psi : \{1, \ldots, n\} \to A \mid d_*(W, W_{\sigma, \psi}) = 0\}] = \frac{n! \prod_{i=1}^r \text{Prob}[d_i(W, W_{\sigma, \psi_0}) = 0]}{\prod_{a \in A}(nW(a))!}. \quad (2)$$

Fix $i \in \{1, \ldots, r\}$. We will compute $\text{Prob}[d_i(W, W_{\sigma, \psi_0}) = 0]$. The element $\sigma(s_i)$ induces a pair of partitions $\alpha, \beta$ of $\{1, \ldots, n\}$ as follows: $\alpha := \{P_{a, b} \mid a, b \in A\}$ and $\beta := \{Q_{a, b} \mid a, b \in A\}$, where

$$P_{a, b} = \{j \mid \psi_0(j) = a \text{ and } \psi_0(\sigma(s_i)j) = b\},$$

$$Q_{a, b} = \{j \mid \psi_0(j) = b \text{ and } \psi_0(\sigma(s_i)^{-1}j) = a\}.$$

Also there is a bijection from $M_{a, b} : P_{a, b} \to Q_{a, b}$ defined by $M_{a, b}(j) = \sigma(s_i)j$. Conversely, $\sigma(s_i)$ is uniquely determined by these partitions and bijections.

Note that $|P_{a, b}| = |Q_{a, b}| = nW_{\sigma, \psi_0}(a, b; i)$. Thus $d_i(W, W_{\sigma, \psi_0}) = 0$ if and only $|P_{a, b}| = |Q_{a, b}| = nW(a, b; i)$ for all $a, b \in A$. If this occurs then $|\bigcup_{b \in A} P_{a, b}| = nW(a)$ for all $a \in A$. So the number of pairs of partitions $\alpha, \beta$ that satisfy this requirement is

$$\frac{\prod_{a \in A}(nW(a))!^2}{\prod_{a, b \in A}(nW(a, b; i))!^2}.$$
Given such a pair of partitions, the number of collections of bijections

\[ M_{a,b} : P_{a,b} \rightarrow Q_{a,b} \]

(for \( a, b \in A \)) equals \( \prod_{a,b \in A} (nW(a, b; i))! \). Since there are \( n! \) elements in Sym\((n)\) it follows that

\[
\text{Prob}[d_i(W, W_{\sigma, \psi_0}) = 0] = \frac{\prod_{a \in A} (nW(a))!^2}{n! \prod_{a,b \in A} (nW(a, b; i))!}.
\]

The proposition now follows from this equality and equation (2). \(\square\)

Let \( W \) be the set of all weights on \( \mathcal{G} \). It is a compact convex subset of \( \mathbb{R}^d \) for some \( d > 0 \). Define \( F : W \rightarrow \mathbb{R} \) by

\[
F(W) := -(\sum_{i=1}^{r} \sum_{a,b \in A} W(a, b; i) \log(W(a, b; i))) + (2r - 1) \sum_{a \in A} W(a) \log(W(a)).
\]

We follow the usual convention that \( 0 \log(0) = 0 \). Observe that \( F(T, \phi) = F(W_T) \).

Given a weight \( W \), let \( q_W \) denote the smallest positive integer such that \( W(a, b; i) \) is an integer for all \( a, b \in A \) and for all \( i \in \{1, \ldots, r\} \). If no such integer exists then set \( q_W := +\infty \). If \( p \) and \( q \) are integers, \( p \neq 0 \) and \( \frac{q}{p} \in \mathbb{Z} \) then we write \( p \mid q \). Otherwise we write \( p \nmid q \).

**Lemma 2.2.** \( F : W \rightarrow \mathbb{R} \) is continuous. Also, there exist constants \( 0 < c_1 < c_2 \) such that for every weight \( W \) with \( q_W < \infty \) and every \( n \geq 1 \) such that \( q_W \mid n \), if \( \sigma : G \rightarrow \text{Sym}(n) \) is chosen uniformly at random then

\[
c_1 n^{p_1} e^{F(W)n} \leq \mathbb{E}[(\{\psi : \{1, \ldots, n\} \rightarrow A \mid d_\sigma(W, W_{\sigma, \psi}) = 0\})] \leq c_2 n^{p_2} e^{F(W)n}.
\]

**Proof.** It is obvious that \( F \) is continuous. The second statement follows from the previous proposition and Stirling’s approximation. The constants depend only on \(|A|\) and the rank \( r \) of \( G \). \(\square\)

**Lemma 2.3.** There exists a constant \( k > 0 \) such that the following holds. Let \( W \) be a weight and let \( n > 0 \) be a positive integer. Then there exists a weight \( \tilde{W} \) such that \( q_{\tilde{W}} < \infty \), \( q_{\tilde{W}} \mid n \) and \( d_\sigma(W, \tilde{W}) < k/n \).

**Proof.** Choose \( a_0 \in A \). For \( b, c \in A - \{a_0\} \) and \( i \in \{1, \ldots, r\} \) define

\[
\tilde{W}(b) := \frac{|W(b)n|}{n},
\]

\[
\tilde{W}(a_0) := 1 - \sum_{b \in A - \{a_0\}} \tilde{W}(b).
\]
This proves that $j$ too, random. Given a weight $W$, let $d$ since $j$.

**Proof of Theorem 1.4**. Recall that $\phi: X \to A$ is an observable and $A$ is a finite set. Let $n \geq 0$ and let $\sigma_n: G \to \text{Sym}(n)$ be a homomorphism chosen uniformly at random. Given a weight $W$, let

$$Z_n(W) := |\{\psi: \{1, \ldots, n\} \to A \mid d_*(W_{\sigma_n, \psi}, W) = 0\}|.$$

For any $\varepsilon > 0$,

$$\mathbb{E}[|\{\psi: \{1, \ldots, n\} \to A \mid d_*(W_{\sigma_n, \psi}, W) \leq \varepsilon\}|] = \sum_{W: d_*(W, W_\mu) \leq \varepsilon} \mathbb{E}[Z_n(W)]. \quad (3)$$
Let $\delta > 0$. Since $F: \mathcal{W} \to \mathbb{R}$ is continuous, there exists $\varepsilon_0 > 0$ such that if $d_\ast(W, W_\mu) \leq \varepsilon_0$ then $|F(W) - F(W_\mu)| < \delta$. So let us fix $\varepsilon$ with $0 < \varepsilon < \varepsilon_0$.

By the previous lemma, if $n$ is sufficiently large then there exists a weight $W$ such that $d_\ast(W, W_\mu) \leq \varepsilon$ and $q_W \mid n$. Lemma 2.2 implies

$$
\mathbb{E}[\{\psi: \{1, \ldots, n\} \to A \mid d_{\sigma_n}^\ast(\phi, \psi) \leq \varepsilon\}] \geq \mathbb{E}[Z_n(W)] \geq c_1 n^{p_1} e^{F(W_\mu)n - \delta n},
$$

where $c_1 > 0$ and $p_1$ are constants.

If $W$ is a weight such that $q_W \uparrow n$ then $Z_n(W) = 0$. If $q_W \mid n$ then $W(a, b; i) \in \mathbb{Z}[1/n]$ for all $a, b \in A$ and $i \in \{1, \ldots, r\}$. The space of all weights lies inside the cube $[0, 1]^d \subset \mathbb{R}^d$ for some $d$. So the number of weights $W$ such that $Z_n(W) \neq 0$ is at most $n^d$. Lemma 2.2 and equation (3) now imply that

$$
\mathbb{E}[\{\psi: \{1, \ldots, n\} \to A \mid d_{\sigma_n}^\ast(\phi, \psi) \leq \varepsilon\}] \leq c_2 n^{p_2 + d} e^{F(W_\mu)n + \delta n}.
$$

(5)

Here $c_2 > 0$ and $p_2$ are constants. Equations (4) and (5) imply

$$
\limsup_{n \to \infty} \frac{1}{n} \log(\mathbb{E}[\{\psi: \{1, \ldots, n\} \to A \mid d_{\sigma_n}^\ast(\phi, \psi) \leq \varepsilon\}]) - F(W_\mu) \leq \delta.
$$

Since $\delta$ is arbitrary, it follows that

$$
\inf_{\varepsilon > 0} \lim_{n \to \infty} \frac{1}{n} \log(\mathbb{E}[\{\psi: \{1, \ldots, n\} \to A \mid d_{\sigma_n}^\ast(\phi, \psi) \leq \varepsilon\}]) = F(W_\mu) = F(T, \phi).
$$

$\square$

3. Proof of Theorem 1.3

As in the statement of Theorem 1.3, let $G = \langle s_1, \ldots, s_r \rangle$ be a free group or free semigroup of rank $r \geq 1$. Let $G \curvearrowright^T (X, \mu)$. Let $\phi: X \to A$ be a finite observable. Let $\Sigma = \{\mu_i\}_{i=1}^\infty$ where each $\mu_i$ is the uniform probability measure on the set of homomorphisms from $G$ to $\text{Sym}(i)$. Let $\sigma_i: G \to \text{Sym}(i)$ be a homomorphism chosen uniformly at random among all homomorphisms of $G$ into $\text{Sym}(i)$. Theorem 1.3 is an immediate consequence of the next two propositions.

**Proposition 3.1.** $h(\Sigma, T, \phi) \leq f(T, \phi)$.

**Proof.** Let $S = \{e, s_1, \ldots, s_r\}$. Observe that for any $n$, if $\psi: \{1, \ldots, n\} \to A$ is any function then $d_{\sigma_n}^S(\phi, \psi)r \geq d_{\sigma_n}^\ast(\phi, \psi)$. So if $\varepsilon > 0$ then

$$
\mathbb{E}[\{\psi: \{1, \ldots, n\} \to A \mid d_{\sigma_n}^S(\phi, \psi) \leq \varepsilon\}] \leq \mathbb{E}[\{\psi: \{1, \ldots, n\} \to A \mid d_{\sigma_n}^\ast(\phi, \psi) \leq r\varepsilon\}].
$$

This implies $h(\Sigma, T, \phi) \leq F(T, \phi)$. 
Recall that $B(e,n)$ denotes the ball of radius $n$ in $G$. Furthermore we have $f(T,\phi) = \inf_n F(T,\phi^{B(e,n)})$, and thus $\inf_n h(\Sigma, T, \phi^{B(e,n)}) \leq f(T,\phi)$. Since $\phi$ and $\phi^{B(e,n)}$ generate the same $\sigma$-algebra, Theorem 1.1 implies that $h(\Sigma, T, \phi) = h(\Sigma, T, \phi^{B(e,n)})$ for all $n$. This implies the proposition. \qed

**Proposition 3.2.** $h(\Sigma, T, \phi) \geq f(T, \phi)$.

**Proof.** Given a finite set $K \subseteq G$, define $h(\Sigma, T, \phi; K) := \inf_{\varepsilon > 0} \limsup_{n \to \infty} \frac{1}{n} \log(\mathbb{E}[\|\{\psi : \{1, \ldots, n\} \to A \mid d_{\sigma_n}^{K}(\phi, \psi) \leq \varepsilon\|])]$.

Claim 1. $h(\Sigma, T, \phi; B(e,m)) \geq F(T, \phi^{B(e,m)})$ for all $m \geq 0$.

Note that if $K \subseteq L$ then $h(\Sigma, T, \phi; K) \geq h(\Sigma, T, \phi; L)$ holds. It follows that $h(\Sigma, T, \phi) = \inf_m h(\Sigma, T, \phi; B(e,m))$. Thus claim 1 implies the proposition.

To simplify notation, let $B$ denote $B(e,m)$. To prove claim 1, for $m, n \geq 0$, let $P(m,n,\varepsilon)$ be the set of all pairs $(\sigma, \omega)$ with $\sigma : G \to \text{Sym}(n)$ a homomorphism and $\omega : \{1, \ldots, n\} \to A$ a map such that $d_{\sigma_n}^{B}(\phi, \omega) \leq \varepsilon$. Since there are $n!r$ homomorphisms from $G$ into $\text{Sym}(n)$,

\[
h(\Sigma, T, \phi; B) = \inf_{\varepsilon > 0} \limsup_{n \to \infty} \frac{1}{n} \log \left( \frac{|P(m,n,\varepsilon)|}{n!r} \right). \quad (6)
\]

Let $Q(m,n,\varepsilon)$ be the set of all pairs $(\sigma, \psi)$ with $\sigma : G \to \text{Sym}(n)$ a homomorphism and $\psi : \{1, \ldots, n\} \to A^B$ a map such that $d_{\sigma_n}^{B}(\phi^B, \psi) \leq \varepsilon$. By Theorem 1.4,

\[
F(T, \phi^B) = \inf_{\varepsilon > 0} \limsup_{n \to \infty} \frac{1}{n} \log \left( \frac{|Q(m,n,\varepsilon)|}{n!r} \right). \quad (7)
\]

For $g \in B$ let $\pi_g : A^B \to A$ denote the projection map $\pi_g((a_h)_{h \in B}) = a_g$. For $(\sigma, \psi) \in Q(m,n,\varepsilon)$, define $R(\sigma, \psi) = (\sigma, \pi_{\varepsilon} \circ \psi)$. Define $H(x) := -x \log(x) - (1-x) \log(1-x)$.

Claim 2. If $c = 1 + |B|$ then the image of $R$ is contained in $P(m,n,\varepsilon c)$.

Claim 3. There are constants $C, k > 0$ depending only on $m$ such that if $\varepsilon < \frac{1}{4|B|}$ then $R$ is at most $C \exp(n k \varepsilon + n H(2|B|\varepsilon))$ to 1, i.e., for any $(\sigma, \omega)$ in the image of $R$, $|R^{-1}(\sigma, \omega)| \leq C \exp(n k \varepsilon + n H(2|B|\varepsilon))$.

Claims 2 and 3 imply

\[
C \exp(k n \varepsilon + n H(2|B|\varepsilon)|P(m,n,\varepsilon c))) \geq |Q(m,n,\varepsilon)|.
\]

Together with equations (6) and (7), this implies claim 1 and hence the proposition.

Next we prove claim 2. For this purpose, fix a homomorphism $\sigma : G \to \text{Sym}(n)$. Observe that for any $x \in X$ and any $t \in \{s_1, \ldots, s_r\}$,

\[
\pi_g \phi^B(x) = \phi(T_g x) = \pi_{g^{t-1}} \phi^B(T_t x) \quad \text{for all } g \in B \cap B_t.
\]
Therefore if \( i \in \{1, \ldots, n\} \) and, for some \( g \in B \cap B_t \), \( \psi : \{1, \ldots, n\} \to A^B \) satisfies 

\[
\pi_g \psi(i) \neq \pi_{gt^{-1}} \psi(\sigma(t)i),
\]

then \( \psi \lor \psi^t(i) \neq \phi^B \lor \phi^{Bt}(x) \) for any \( x \in X \).

So let \( \mathcal{F} \) be the set of all \( i \in \{1, \ldots, n\} \) such that for all \( t \in \{s_1, \ldots, s_r\} \),

\[
\pi_g \psi(i) = \pi_{gt^{-1}} \psi(\sigma(t)i) \quad \text{for all } g \in B \cap B_t.
\]

Thus

\[
d^*_\sigma(\phi^B, \psi) \geq \frac{|\mathcal{F}^c|}{n} = \zeta(\mathcal{F}^c),
\]

where \( \mathcal{F}^c \) denotes the complement of \( \mathcal{F} \) and \( \zeta \) denotes the uniform probability measure on \( \{1, \ldots, n\} \).

Let \( \mathcal{F}_m \) be the set of all \( i \in \{1, \ldots, n\} \) such that \( \sigma(g)i \in \mathcal{F} \) for all \( g \in B \). Note that

\[
\zeta(\mathcal{F}_m^c) \leq |B| \zeta(\mathcal{F}^c) \leq |B| d^*_\sigma(\phi^B, \psi). \tag{8}
\]

If \( i \in \mathcal{F}_m \) then \( \psi(i) = (\pi_\varepsilon \circ \psi)^B(i) \). Therefore

\[
\sum_{a \in A^B} |\psi_\varepsilon^t(\zeta(a)) - (\pi_\varepsilon \circ \psi)^B_\varepsilon \zeta(a)| \leq \frac{1}{n} |\{i \mid \psi(i) \neq (\pi_\varepsilon \circ \psi)^B(i)\}| \leq \zeta(\mathcal{F}_m^c) \leq |B| d^*_\sigma(\phi^B, \psi).
\]

Suppose that \( d^*_\sigma(\phi^B, \psi) \leq \varepsilon \). Then

\[
d^B(\phi, \pi_\varepsilon \circ \psi) = \sum_{a \in A^B} |\phi^B_\varepsilon \mu(a) - (\pi_\varepsilon \circ \psi)^B_\varepsilon \zeta(a)|
\]

\[
\leq \sum_{a \in A^B} |\phi^B_\varepsilon \mu(a) - \psi_\varepsilon \zeta(a)| + |\psi_\varepsilon \zeta(a) - (\pi_\varepsilon \circ \psi)^B_\varepsilon \zeta(a)|
\]

\[
\leq d^*_\sigma(\phi^B, \psi)(1 + |B|) \leq \varepsilon(1 + |B|).
\]

This proves claim 2.

Let \((\sigma, \omega)\) be in the image of \( R \).

\textit{Claim 4.} For every \( \psi \) with \( R(\sigma, \psi) = (\sigma, \omega) \), there exists a set \( L(\psi) \subseteq \{1, \ldots, n\} \) of cardinality \( \lceil n(1 - |B|\varepsilon) \rceil \) such that \( \psi(i) = \omega^B(i) \) for all \( i \in L(\psi) \).

To prove claim 4, observe that if \( \mathcal{F}_m \) is defined as above, then \( \psi(i) = \omega^B(i) \) for all \( i \in \mathcal{F}_m \). By equation (8),

\[
|\mathcal{F}_m| = n(1 - \zeta(\mathcal{F}_m^c)) \geq n(1 - |B| d^*_\sigma(\phi^B, \psi)) \geq n(1 - |B|\varepsilon).
\]

So let \( L(\psi) \) be any subset of \( \mathcal{F}_m \) with cardinality \( \lceil n(1 - |B|\varepsilon) \rceil \). This proves claim 4.

Next we prove claim 3. Claim 4 implies

\[
|R^{-1}(\sigma, \omega)| \leq |A| |B|^{n - |n(1 - |B|\varepsilon)|} \sum_{i=1}^{n} \left( \begin{array}{c} n \\ |n(1 - |B|\varepsilon)| \end{array} \right) . \tag{9}
\]
This is because there are \( \binom{n}{[n(1-|B|\varepsilon)]} \) sets in \( \{1, \ldots , n\} \) with cardinality equal to \( [n(1-|B|\varepsilon)] \) and for each \( i \in \{1, \ldots , n\} - L(\psi) \), there are at most \( |A|^{|B|} \) possible values for \( \psi(i) \).

Because \( H \) is monotone increasing for \( 0 < x < 1/2 \) it follows from Stirling’s approximation that if \( \varepsilon < \frac{1}{4|B|} \) then
\[
\binom{n}{[n(1-|B|\varepsilon)]} \leq C \exp(nH(2|B|\varepsilon)),
\]
where \( C > 0 \) is a constant. This and equation (9) now imply claim 3 and hence the proposition. \( \square \)

References

[Bo08a] L. Bowen, A measure-conjugacy invariant for actions of free groups. Ann. of Math., to appear.
[Bo08b] L. Bowen, Measure conjugacy invariants for actions of countable sofic groups. J. Amer. Math. Soc. 23 (2010), 217–245. Zbl MR 2552252
[Bo08c] L. Bowen, Nonabelian free group actions: Markov processes, the Abramov-Rohlin formula and Yuzvinskii’s formula. Preprint 2008. arXiv:0806.4420
[FO70] N. A. Friedman and D. S. Ornstein, On isomorphism of weak Bernoulli transformations. Adv. Math. 5 (1970), 365–394. Zbl 0203.05801 MR 0274718
[Or70] D. Ornstein, Factors of Bernoulli shifts are Bernoulli shifts. Adv. Math. 5 (1970), 349–364. Zbl 0227.28015 MR 0274717
[Po08] S. Popa, Private communication.
[PS07] S. Popa and R. Sasyk, On the cohomology of Bernoulli actions. Ergodic Theory Dynam. Systems 27 (2007), 241–251. Zbl 05144564 MR 2297095

Received March 11, 2009; revised June 24, 2009

L. Bowen, University of Hawai‘i, Mānoa, 2565 McCarthy Mall, Keller 401A, Honolulu, Hawaii 96822, U.S.A., and Mathematics Department, MS-3368, Texas A&M University, College Station, TX 77843-3368, U.S.A.
E-mail: lpbowen@math.tamu.edu