Groups Geom. Dyn. 4 (2010), 419–432
DOI 10.4171/GGD/89

Groups, Geometry, and Dynamics

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 Σ-entropy (for actions of sofic groups). The purpose of this paper is to show that the f-invariant is essentially a special case of Σ-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 Σ-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 Σ-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 Σ-entropy and the f-invariant.

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

$$\lim_{i\to\infty} \frac{|\{1 \le j \le m_i \mid \sigma_i(g_1)j = \sigma_i(g_2)j\}|}{m_i} = 0.$$

for every pair $g_1, g_2 \in G$ with $g_1 \ne g_2$. The treatment of Σ-entropy given next differs from [Bo08b] in two respects: for simplicity, we assume that each σ_i is a homomorphism and we use observables rather than partitions to define it.

We will write $G \curvearrowright^T (X,\mu)$ to mean (X,μ) is a standard probability measure space and $T = (T_g)_{g\in G}$ is an action of G on (X,μ) by measure-preserving transformations. This means that for each $g \in G$, $T_g : X \to X$ is a measure-preserving transformation and $T_{g_1}T_{g_2} = T_{g_1g_2}$. An *observable* of (X,μ) is a measurable map $\phi: X \to A$ where A is a finite or countably infinite set. We will say that ϕ is *finite* if A is finite. Roughly speaking, the Σ-entropy rate of ϕ is the exponential rate of growth of the number of observables $\psi: \{1,\ldots,m_i\} \to A$ that approximate ϕ. 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 ϕ and ψ is the observable $\phi \vee \psi: X \to A \times B$ defined by $\phi \vee \psi(x) = (\phi(x), \psi(x))$. If $g \in G$ then $T_g\phi: X \to A$ is defined by $T_g\phi(x) = \phi(T_g x)$. If $H \subset G$ is finite, then let $\phi^H := \bigvee_{h\in H} T_h\phi$. ϕ^H maps X into A^H, the direct product of $|H|$ copies of A. Let $\phi^H_*\mu$ denote the pushforward of μ on A^H. In other words, $\phi^H_*(\mu)(S) = \mu((\phi^H)^{-1}(S))$ for $S \subset A^H$.

For each i, let ζ_i denote the uniform probability measure on $\{1,\ldots,m_i\}$. If $\psi: \{1,\ldots,m_i\} \to A$ 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, ψ^H depends on σ_i but, to keep the notation simple, we will leave this dependence implicit. Let $\psi^H_*\zeta_i$ be the pushforward of ζ_i on A^H. Finally, let $d^H_{\sigma_i}(\phi,\psi)$ be the l^1-distance between $\phi^H_*\mu$ and $\psi^H_*\zeta_i$. In other words,

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

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

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

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

Definition 2. Define the *entropy* of ϕ 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) π_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 ϕ is *generating* if the smallest G-invariant σ-algebra on X that contains $\{\phi^{-1}(a)\}_{a\in A}$ is equal to the σ-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 $\sigma_i : G \to \mathrm{Sym}(m_i)$ for a group G. Let $G \curvearrowright^T (X,\mu)$. If ϕ_1 and ϕ_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 Σ and T are as above then the Σ-entropy of the action T is defined by $h(\Sigma, T) := h(\Sigma, \phi)$, where ϕ is any finite-entropy generating observable (if one exists).

Next let us discuss a slight variation on Σ-entropy. Let $\{m_i\}_{i=1}^{\infty}$ be a sequence of natural numbers. For each $i \in \mathbb{N}$, let μ_i be a probability measure on the set of homomorphisms from G to $\mathrm{Sym}(m_i)$. Let $\sigma_i : G \to \mathrm{Sym}(m_i)$ be chosen at random according to μ_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 $g_1 \neq g_2$,

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

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

$$h(\Sigma, T, \phi) := \inf_{H\subset G} \inf_{\varepsilon>0} \limsup_{i\to\infty} \tfrac{1}{m_i} \log(\mathbb{E}[|\{\psi : \{1,\dots,m_i\} \to A \mid d^H_{\sigma_i}(\phi,\psi) \le \varepsilon\}|]).$$

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

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 $G = \langle s_1, \dots, s_r\rangle$ be either a free group or free semigroup of rank r. Let $G \curvearrowright^T (X,\mu)$. Let α be a partition of X into at most countably many measurable sets. The *entropy* of α is defined by

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

where, by convention, $0\log(0) = 0$. If α and β are partitions of X then the *join* is the partition $\alpha \vee \beta := \{A \cap B \mid A \in \alpha, B \in \beta\}$. 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 $\{s_1, \dots, s_r\}$ if G is a semigroup or $\{s_1^{\pm 1}, \dots, s_r^{\pm 1}\}$ is G is a group). Define

$$F(T,\alpha) := (1-2r)H(\alpha) + \sum_{i=1}^{r} H(\alpha \vee T_{s_i}^{-1}\alpha),$$
$$\alpha^n := \bigvee_{g \in B(e,n)} T_g^{-1}\alpha,$$
$$f(T,\alpha) := \inf_n F(T,\alpha^n).$$

The partition α is *generating* if the smallest G-invariant σ-algebra containing α equals the σ-algebra of all measurable sets up to sets of measure zero.

Theorem 1.2. *Let $G = \langle s_1, \dots, s_r \rangle$ be a free group or free semigroup. Let $G \curvearrowright^T (X,\mu)$. If α_1 and α_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 α_1, α_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 α is any finite-entropy generating partition of X (if one exists).

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

Theorem 1.3. *Let $G = \langle s_1, \dots, s_r \rangle$ be a free group or free semigroup of rank $r \geq 1$. Let $G \curvearrowright^T (X,\mu)$. Let ϕ be a finite observable. For $i \geq 1$, let μ_i be the uniform probability measure on the set of all homomorphisms from G to* 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 $d_{\sigma_i}^{H}(\phi,\psi)$ given above. Define

$$d_{\sigma_i}^{*}(\phi,\psi) := \sum_{i=1}^{r} d_{\sigma_i}^{\{e,s_i\}}(\phi,\psi).$$

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

$$F(T,\phi) = \inf_{\varepsilon > 0} \limsup_{i \to \infty} \tfrac{1}{i} \log(\mathbb{E}[|\{\psi : \{1,\dots,i\} \to A \mid d_{\sigma_i}^{*}(\phi,\psi) \leq \varepsilon\}|]).$$

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 $G \curvearrowright^T (X, \mu)$. Let $\omega\colon G \to G$ 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 $\phi\colon X \to X$ such that $\phi(T_g x) = T^\omega_g \phi(x)$ for a.e. $x \in X$ and all $g \in G$.

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

If $\sigma_i : G \to \mathrm{Sym}(i)$ is chosen uniformly at random, it follows that the law of $\sigma_i \circ \omega$ is the same as the law of σ_i. Therefore, if μ_i is the uniform probability measure on the set of homomorphisms from G to $\mathrm{Sym}(i)$ 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 $\omega\colon G \to G$ be an automorphism. Then $f(T, \phi) = f(T^\omega, \phi)$ for any finite observable ϕ.*

This implies that $f(T, \phi)$ does not depend on the choice of free generator set $\{s_1, \dots, s_r\}$ 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 $G \curvearrowright^T (X, \mu)$ and $G \curvearrowright^S (Y, \nu)$. Then S is a *factor* of T if there exists a measurable map $\phi\colon X \to Y$ 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, μ is the uniform measure on X and $T = (T_g)_{g\in G}$ is a measure-preserving action of $G = \langle s_1, \dots, s_r\rangle$ 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 $\sigma_i : G \to \mathrm{Sym}(m_i)$. 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 ϕ 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 $G = \langle s_1, \dots, s_r \rangle$ 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 α 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 κ a probability measure on K. Let (K^G, κ^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, κ^G). It is the *Bernoulli shift* over G with base measure κ. In [Bo08a] it was shown that $f(T) = H(\kappa)$ where

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

Let α 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 $\phi\colon X \to A$ be a finite observable.

We will need to consider certain perturbations of the measure μ with respect to the given observable $\phi\colon X \to A$. 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 $W\colon V \sqcup E \to [0, 1]$ satisfying

$$W(a) = \sum_{b \in A} W(a, b; i) = \sum_{b \in A} W(b, a; i) \quad \text{for all } i = 1 \ldots r,\ a \in A,$$
$$1 = \sum_{a \in A} W(a).$$

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 μ. For a homomorphism $\sigma\colon G \to \mathrm{Sym}(n)$ and a function $\psi\colon \{1, \ldots, n\} \to A$ 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

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

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 $\sigma\colon G \to \mathrm{Sym}(n)$ is chosen uniformly at random then*

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

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

$$|\psi^{-1}(a)| = nW(a) \quad \text{for all } a \in A. \tag{1}$$

The number of functions $\psi\colon \{1,\dots,n\} \to A$ that satisfy this requirement is

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

If ψ_1, ψ_2 are two different functions that satisfy equation (1) then there is a permutation $\tau \in \mathrm{Sym}(n)$ such that $\psi_1 = \psi_2 \circ \tau$. If $\sigma^\tau\colon G \to \mathrm{Sym}(n)$ is the homomorphism defined by $\sigma^\tau(g) = \tau\sigma(g)\tau^{-1}$ then $W_{\sigma,\psi_1} = W_{\sigma^\tau,\psi_2}$. Since $\sigma\colon G \to \mathrm{Sym}(n)$ 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 ψ_0 satisfying equation (1). Then

$$\mathbb{E}[|\{\psi\colon \{1,\dots,n\} \to A \mid d_*(W, W_{\sigma,\psi}) = 0\}|] = \frac{n!\,\mathrm{Prob}[d_*(W, W_{\sigma,\psi_0}) = 0]}{\prod_{a\in A}(nW(a))!}.$$

For any two weights W_1, W_2 and $1 \le i \le r$, 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 $\sigma\colon G \to \mathrm{Sym}(n)$ is determined by its values $\sigma(s_1),\dots,\sigma(s_r)$. The event $d_i(W, W_{\sigma,\psi_0}) = 0$ is determined by $\sigma(s_i)$. So if $i \ne j$ then the events $d_i(W, W_{\sigma,\psi_0}) = 0$ and $d_j(W, W_{\sigma,\psi_0}) = 0$ are independent. Therefore,

$$\begin{aligned}&\mathbb{E}[|\{\psi\colon \{1,\dots,n\} \to A \mid d_*(W, W_{\sigma,\psi}) = 0\}|]\\ &\quad = \frac{n!\prod_{i=1}^r \mathrm{Prob}[d_i(W, W_{\sigma,\psi_0}) = 0]}{\prod_{a\in A}(nW(a))!}.\end{aligned} \tag{2}$$

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

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

Also there is a bijection from $M_{a,b}\colon P_{a,b} \to Q_{a,b}$ 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 α, β 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

$$M_{a,b} \colon P_{a,b} \to Q_{a,b}$$

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

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

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

Let $\mathcal{W}$ be the set of all weights on $\mathcal{G}$. It is a compact convex subset of $\mathbb{R}^d$ for some $d > 0$. Define $F \colon \mathcal{W} \to \mathbb{R}$ 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 $F(T,\phi) = F(W_\mu)$.

Given a weight W, let q_W denote the smallest positive integer such that $W(a,b;i)\mathrm{q}_W \in \mathbb{Z}$ for all $a, b \in A$ and for all $i \in \{1,\dots,r\}$. If no such integer exists then set $\mathrm{q}_W := +\infty$. 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. *$F \colon \mathcal{W} \to \mathbb{R}$ is continuous. Also, there exist constants $0 < c_1 < c_2$ and $p_1 < p_2$ such that for every weight W with $\mathrm{q}_W < \infty$ and every $n \geq 1$ such that $\mathrm{q}_W \mid n$, if $\sigma \colon G \to \mathrm{Sym}(n)$ is chosen uniformly at random then*

$$c_1 n^{p_1} e^{F(W)n} \leq \mathbb{E}[|\{\psi \colon \{1,\dots,n\} \to A \mid d_*(W, W_{\sigma,\psi}) = 0\}|] \leq c_2 n^{p_2} e^{F(W)n}.$$

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. □

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 $\widetilde{W}$ such that $\mathrm{q}_{\widetilde{W}} < \infty$, $\mathrm{q}_{\widetilde{W}} \mid n$ and $d_*(W, \widetilde{W}) < k/n$.*

Proof. Choose $a_0 \in A$. For $b, c \in A - \{a_0\}$ and $i \in \{1,\dots,r\}$ define

$$\widetilde{W}(b) := \frac{\lfloor W(b)n \rfloor}{n},$$
$$\widetilde{W}(a_0) := 1 - \sum_{b\in A-\{a_0\}} \widetilde{W}(b),$$

$$\begin{aligned}
\widetilde{W}(b,c;i) &:= \frac{\lfloor W(b,c;i)n \rfloor}{n},\\
\widetilde{W}(a_0,b;i) &:= \widetilde{W}(b) - \sum_{a\in A-\{a_0\}} \widetilde{W}(a,b;i),\\
\widetilde{W}(b,a_0;i) &:= \widetilde{W}(b) - \sum_{a\in A-\{a_0\}} \widetilde{W}(b,a;i),\\
\widetilde{W}(a_0,a_0;i) &:= \widetilde{W}(a_0) - \sum_{b\in A-\{a_0\}} \widetilde{W}(a_0,b;i).
\end{aligned}$$

Let us check that $\widetilde{W}$ is a weight. It is clear that $\sum_{a\in A} \widetilde{W}(a) = 1$. If $b \in A - \{a_0\}$ then $\widetilde{W}(b) = \sum_{a\in A} \widetilde{W}(a,b;i) = \sum_{a\in A} \widetilde{W}(b,a;i)$. It is immediate that $\widetilde{W}(a_0) = \sum_{b\in A} \widetilde{W}(a_0,b;i)$. Also

$$\begin{aligned}
\sum_{b\in A} \widetilde{W}(b,a_0;i) &= \widetilde{W}(a_0,a_0;i) + \sum_{b\in A-\{a_0\}} \widetilde{W}(b,a_0;i)\\
&= \widetilde{W}(a_0) - \sum_{b\in A-\{a_0\}} \widetilde{W}(a_0,b;i) + \sum_{b\in A-\{a_0\}} \widetilde{W}(b,a_0;i)\\
&= \widetilde{W}(a_0) + \sum_{b\in A-\{a_0\}} \widetilde{W}(b,a_0;i) - \widetilde{W}(a_0,b;i)\\
&= \widetilde{W}(a_0) + \sum_{b\in A-\{a_0\}} (\widetilde{W}(b)\\
&\qquad - \sum_{a\in A-\{a_0\}} \widetilde{W}(b,a;i)) - (\widetilde{W}(b) - \sum_{a\in A-\{a_0\}} \widetilde{W}(a,b;i))\\
&= \widetilde{W}(a_0).
\end{aligned}$$

This proves that $\widetilde{W}$ is a weight. It is clear that $\mathrm{q}_{\widetilde{W}} < \infty$ and $\mathrm{q}_{\widetilde{W}}\,|n$. Lastly observe that if $a, b \in A-\{a_0\}$ then $|W(a,b;i)-\widetilde{W}(a,b;i)| \le 1/n$. Since $|W(b)-\widetilde{W}(b)| \le 1/n$ too, $|W(a_0,b;i) - \widetilde{W}(a_0,b;i)| \le |A|/n$ and $|W(b,a_0;i) - \widetilde{W}(b,a_0;i)| \le |A|/n$. Since $|W(a_0) - \widetilde{W}(a_0)| \le |A|/n$, $|W(a_0,a_0;i) - \widetilde{W}(a_0,a_0;i)| \le |A|^2/n$. Thus $d_*(W,\widetilde{W}) \le r|A|^2/n$. □

We are now ready to prove Theorem 1.4.

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

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

For any $\varepsilon > 0$,

$$\mathbb{E}[|\{\psi\colon \{1,\dots,n\} \to A \mid d^*_{\sigma_n}(\phi,\psi) \le \varepsilon\}|] = \sum_{W\colon d_*(W,W_\mu)\le\varepsilon} \mathbb{E}[Z_n(W)]. \tag{3}$$

Let $\delta > 0$. Since $F\colon \mathcal{W} \to \mathbb{R}$ is continuous, there exists $\varepsilon_0 > 0$ such that if $d_*(W, W_\mu) \le \varepsilon_0$ then $|F(W) - F(W_\mu)| < \delta$. So let us fix ε with $0 < \varepsilon < \varepsilon_0$.

By the previous lemma, if n is sufficiently large then there exists a weight W such that $d_*(W, W_\mu) \le \varepsilon$ and $\mathrm{q}_W \,|\, n$. Lemma 2.2 implies

$$\mathbb{E}[|\{\psi\colon \{1,\dots,n\} \to A \mid d^*_{\sigma_n}(\phi,\psi) \le \varepsilon\}|] \ge \mathbb{E}[Z_n(W)] \ge c_1 n^{p_1} e^{F(W_\mu)n - \delta n}, \tag{4}$$

where $c_1 > 0$ and p_1 are constants.

If W is a weight such that $\mathrm{q}_W \nmid n$ then $Z_n(W) = 0$. If $\mathrm{q}_W \mid n$ then $W(a,b;i) \in \mathbb{Z}[1/n]$ for all $a, b \in A$ and $i \in \{1,\dots,r\}$. 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 $Z_n(W) \ne 0$ is at most n^d. Lemma 2.2 and equation (3) now imply that

$$\mathbb{E}[|\{\psi\colon \{1,\dots,n\} \to A \mid d^*_{\sigma_n}(\phi,\psi) \le \varepsilon\}|] \le c_2 n^{p_2+d} e^{F(W_\mu)n + \delta n}. \tag{5}$$

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

$$\limsup_{n\to\infty} |\tfrac{1}{n} \log(\mathbb{E}[|\{\psi\colon \{1,\dots,n\} \to A \mid d^*_{\sigma_n}(\phi,\psi) \le \varepsilon\}|]) - F(W_\mu)| \le \delta.$$

Since δ is arbitrary, it follows that

$$\inf_{\varepsilon>0} \lim_{n\to\infty} \tfrac{1}{n} \log(\mathbb{E}[|\{\psi\colon \{1,\dots,n\} \to A \mid d^*_{\sigma_n}(\phi,\psi) \le \varepsilon\}|]) = F(W_\mu) = F(T,\phi).$$

□

3. Proof of Theorem 1.3

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

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

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

$$\begin{aligned}&\mathbb{E}[|\{\psi\colon \{1,\dots,n\} \to A \mid d^S_{\sigma_n}(\phi,\psi) \le \varepsilon\}|]\\ &\qquad \le \mathbb{E}[|\{\psi\colon \{1,\dots,n\} \to A \mid d^*_{\sigma_n}(\phi,\psi) \le r\varepsilon\}|].\end{aligned}$$

This implies $h(\Sigma, T, \phi) \le F(T,\phi)$.

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 $\inf_n h(\Sigma,T,\phi^{B(e,n)}) \le f(T,\phi)$. Since ϕ and $\phi^{B(e,n)}$ generate the same σ-algebra, Theorem 1.1 implies that $h(\Sigma,T,\phi) = h(\Sigma,T,\phi^{B(e,n)})$ for all n. This implies the proposition. □

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

Proof. Given a finite set $K \subset G$, define

$$h(\Sigma,T,\phi;K) := \inf_{\varepsilon>0}\limsup_{n\to\infty} \tfrac{1}{n}\log(\mathbb{E}[|\{\psi\colon \{1,\dots,n\}\to A \mid d^K_{\sigma_n}(\phi,\psi)\le\varepsilon\}|]).$$

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

Note that if $K \subset L$ then $h(\Sigma,T,\phi;K) \ge h(\Sigma,T,\phi;L)$ 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 $m,n,\varepsilon \ge 0$, let $P(m,n,\varepsilon)$ be the set of all pairs (σ,ω) with $\sigma\colon G\to \mathrm{Sym}(n)$ a homomorphism and $\omega\colon \{1,\dots,n\}\to A$ a map such that $d^{\mathrm{B}}_\sigma(\phi,\omega)\le\varepsilon$. Since there are $n!^r$ homomorphisms from G into $\mathrm{Sym}(n)$,

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

Let $Q(m,n,\varepsilon)$ be the set of all pairs (σ,ψ) with $\sigma\colon G\to\mathrm{Sym}(n)$ a homomorphism and $\psi\colon\{1,\dots,n\}\to A^{\mathrm{B}}$ a map such that $d^*_\sigma(\phi^{\mathrm{B}},\psi)\le\varepsilon$. By Theorem 1.4,

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

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

Claim 2. If $c = 1+|\mathrm{B}|$ 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 $C\exp(nk\varepsilon + nH(2|\mathrm{B}|\varepsilon))$ to 1, i.e., for any (σ,ω) in the image of R, $|R^{-1}(\sigma,\omega)| \le C\exp(nk\varepsilon + nH(2|\mathrm{B}|\varepsilon))$.

Claims 2 and 3 imply

$$C\exp(kn\varepsilon + nH(2|\mathrm{B}|\varepsilon))|P(m,n,\varepsilon c))| \ge |Q(m,n,\varepsilon)|.$$

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 $\sigma\colon G\to\mathrm{Sym}(n)$. Observe that for any $x\in X$ and any $t\in\{s_1,\dots,s_r\}$,

$$\pi_g\phi^{\mathrm{B}}(x) = \phi(T_gx) = \pi_{gt^{-1}}\phi^{\mathrm{B}}(T_tx) \quad \text{for all } g\in\mathrm{B}\cap\mathrm{B}t.$$

Therefore if $i \in \{1,\dots,n\}$ and, for some $g \in \mathrm{B} \cap \mathrm{B}t$, $\psi : \{1,\dots,n\} \to A^{\mathrm{B}}$ satisfies

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

then $\psi \vee \psi^t(i) \neq \phi^{\mathrm{B}} \vee \phi^{\mathrm{B}t}(x)$ for any $x \in X$.

So let $\mathcal{G}$ be the set of all $i \in \{1,\dots,n\}$ such that for all $t \in \{s_1,\dots,s_r\}$,

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

Thus

$$d^*_\sigma(\phi^{\mathrm{B}}, \psi) \geq \frac{|\mathcal{G}^c|}{n} = \zeta(\mathcal{G}^c),$$

where $\mathcal{G}^c$ denotes the complement of $\mathcal{G}$ and ζ denotes the uniform probability measure on $\{1,\dots,n\}$.

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

$$\zeta(\mathcal{G}_m^c) \leq |\mathrm{B}|\zeta(\mathcal{G}^c) \leq |\mathrm{B}| d^*_\sigma(\phi^{\mathrm{B}}, \psi). \tag{8}$$

If $i \in \mathcal{G}_m$ then $\psi(i) = (\pi_e \circ \psi)^{\mathrm{B}}(i)$. Therefore

$$\begin{aligned}\sum_{a \in A^{\mathrm{B}}} |\psi_* \zeta(a) - (\pi_e \circ \psi)^{\mathrm{B}}_* \zeta(a)| &\leq \tfrac{1}{n} |\{i \mid \psi(i) \neq (\pi_e \circ \psi)^{\mathrm{B}}(i)\}| \\ &\leq \zeta(\mathcal{G}_m^c) \leq |\mathrm{B}| d^*_\sigma(\phi^{\mathrm{B}}, \psi).\end{aligned}$$

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

$$\begin{aligned} d^{\mathrm{B}}_\sigma(\phi, \pi_e \circ \psi) &= \sum_{a \in A^{\mathrm{B}}} |\phi^{\mathrm{B}}_* \mu(a) - (\pi_e \circ \psi)^{\mathrm{B}}_* \zeta(a)| \\ &\leq \sum_{a \in A^{\mathrm{B}}} |\phi^{\mathrm{B}}_* \mu(a) - \psi_* \zeta(a)| + |\psi_* \zeta(a) - (\pi_e \circ \psi)^{\mathrm{B}}_* \zeta(a)| \\ &\leq d^*_\sigma(\phi^{\mathrm{B}}, \psi)(1 + |\mathrm{B}|) \leq \varepsilon(1 + |\mathrm{B}|). \end{aligned}$$

This proves claim 2.

Let (σ, ω) be in the image of R.

Claim 4. For every ψ with $R(\sigma, \psi) = (\sigma, \omega)$, there exists a set $L(\psi) \subset \{1,\dots,n\}$ of cardinality $\lfloor n(1 - |\mathrm{B}|\varepsilon) \rfloor$ such that $\psi(i) = \omega^B(i)$ for all $i \in L(\psi)$.

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

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

So let $L(\psi)$ be any subset of $\mathcal{G}_m$ with cardinality $\lfloor n(1 - |\mathrm{B}|\varepsilon) \rfloor$. This proves claim 4.

Next we prove claim 3. Claim 4 implies

$$|R^{-1}(\sigma, \omega)| \leq |A|^{|\mathrm{B}|(n - \lfloor n(1 - |\mathrm{B}|\varepsilon) \rfloor)} \binom{n}{\lfloor n(1 - |\mathrm{B}|\varepsilon) \rfloor}. \tag{9}$$

This is because there are $\binom{n}{\lfloor n(1-|\mathrm{B}|\varepsilon)\rfloor}$ sets in $\{1,\dots,n\}$ with cardinality equal to $\lfloor n(1-|\mathrm{B}|\varepsilon)\rfloor$ and for each $i \in \{1,\dots,n\} - L(\psi)$, there are at most $|A|^{|\mathrm{B}|}$ possible values for $\psi(i)$.

Because H is monotone increasing for $0 < x < 1/2$ it follows from Stirling's approximation that if $\varepsilon < \frac{1}{4|\mathrm{B}|}$ then

$$\binom{n}{\lfloor n(1-|\mathrm{B}|\varepsilon)\rfloor} \le C\exp(nH(2|\mathrm{B}|\varepsilon)),$$

where $C > 0$ is a constant. This and equation (9) now imply claim 3 and hence the proposition. □

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