ANNALS OF
MATHEMATICS

A measure-conjugacy invariant for free group actions

By Lewis Phylip Bowen

SECOND SERIES, VOL. 171, NO. 2

March, 2010

ANMAAH

Annals of Mathematics, **171** (2010), 1387–1400

A measure-conjugacy invariant for free group actions

By Lewis Phylip Bowen

Abstract

This paper introduces a new measure-conjugacy invariant for actions of free groups. Using this invariant, it is shown that two Bernoulli shifts over a finitely generated free group are measurably conjugate if and only if their base measures have the same entropy. This answers a question of Ornstein and Weiss.

1. Introduction

This paper is motivated by an old and central problem in measurable dynamics: given two dynamical systems, determine whether they are measurably-conjugate, i.e., isomorphic. Let us set some notation.

A *dynamical system* (or system for short) is a triple (G, X, μ) where (X, μ) is a probability space and G is a group acting by measure-preserving transformations on (X, μ). We will also call this a *dynamical system over* G, a G*-system* or an *action of* G. In this paper, G will always be a discrete countable group. Two systems (G, X, μ) and (G, Y, ν) are *isomorphic* (i.e., *measurably conjugate*) if and only if there exist conull sets $X' \subset X$, $Y' \subset Y$ and a bijective measurable map $\phi : X' \to Y'$ such that ϕ^{-1} is measurable, $\phi_* \mu = \nu$ and $\phi(gx) = g\phi(x) \forall g \in G$, $x \in X'$.

A special class of dynamical systems called Bernoulli systems or Bernoulli shifts has played a significant role in the development of the theory as a whole; it was the problem of trying to classify them that motivated Kolmogorov to introduce the mean entropy of a dynamical system over $\mathbb{Z}$ [Kol58], [Kol59]. That is, Kolmogorov defined for every system $(\mathbb{Z}, X, \mu)$ a number $h(\mathbb{Z}, X, \mu)$ called the *mean entropy* of $(\mathbb{Z}, X, \mu)$ that quantifies, in some sense, how "random" the system is. His definition was modified by Sinai [Sin59]; the latter has become standard. The lectures notes [Roh67] are a classical reference. Modern references include [Pet83], [Rud90] and [Gla03].

Bernoulli shifts also play an important role in this paper, so let us define them. Let (K, κ) be a standard Borel probability space. For a discrete countable group G, let $K^G = \prod_{g\in G} K$ be the set of all functions $x : G \to K$ with the product Borel structure and let κ^G be the product measure on K^G. G acts on K^G by $(gx)(f) = x(g^{-1}f)$ for $x \in K^G$ and $g, f \in G$. This action is measure-preserving. The system (G, K^G, κ^G) is the *Bernoulli shift over G with base* (K, κ). It is nontrivial if κ is not supported on a single point.

Before Kolmogorov's seminal work [Kol58], [Kol59], it was unknown whether all nontrivial Bernoulli shifts over $\mathbb{Z}$ were measurably conjugate to each other. He proved that $h(\mathbb{Z}, K^{\mathbb{Z}}, \kappa^{\mathbb{Z}}) = H(\kappa)$ where $H(\kappa)$, the *entropy of* κ is defined as follows. If there exists a finite or countably infinite set $K' \subset K$ such that $\kappa(K') = 1$ then

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

where we follow the convention $0\log(0) = 0$. Otherwise, $H(\kappa) = +\infty$. Thus two Bernoulli shifts over $\mathbb{Z}$ with different base measure entropies cannot be measurably conjugate.

The converse was proven by D. Ornstein in the groundbreaking papers [Orn70a], [Orn70b]. That is, he proved that if two Bernoulli shifts $(\mathbb{Z}, K^{\mathbb{Z}}, \kappa^{\mathbb{Z}})$, $(\mathbb{Z}, L^{\mathbb{Z}}, \lambda^{\mathbb{Z}})$ are such that $H(\kappa) = H(\lambda)$ then they are isomorphic.

In [Kie75], Kieffer proved a generalization of the Shannon-McMillan theorem to actions of a countable amenable group G. In particular, he extended the definition of mean entropy from $\mathbb{Z}$-systems to G-systems. This leads to the generalization of Kolmogorov's theorem to amenable groups.

In the landmark paper [OW87], Ornstein and Weiss extended most of the classical entropy theory from $\mathbb{Z}$-systems to G-systems where G is any countable amenable group. In particular, they proved that if two Bernoulli shifts (G, K^G, κ^G), (G, L^G, λ^G) over a countably infinite amenable group G are such that $H(\kappa) = H(\lambda)$ then they are isomorphic. Thus Bernoulli shifts over G are completely classified by base measure entropy.

Now let us say that a group G is *Ornstein* if $H(\kappa) = H(\lambda)$ implies (G, K^G, κ^G) is isomorphic to (G, L^G, λ^G) whenever (K, κ) and (L, λ) are standard Borel probability spaces. By the above, all countably infinite amenable groups are Ornstein. Stepin proved that any countable group that contains an Ornstein subgroup is itself Ornstein [Ste75]. It is unknown whether every countably infinite group is Ornstein.

In [OW87], Ornstein and Weiss asked whether all Bernoulli shifts over a non-amenable group are isomorphic. The next result shows that the answer is 'no':

THEOREM 1.1. *Let* $G = \langle s_1, \dots, s_r\rangle$ *be the free group of rank* r. *If* (K_1, κ_1), (K_2, κ_2) *are standard probability spaces with* $H(\kappa_1)+H(\kappa_2)<\infty$ *then* (G, K_1^G, κ_1^G) *is measurably conjugate to* (G, K_2^G, κ_2^G) *if and only if* $H(\kappa_1) = H(\kappa_2)$.

The reason Ornstein and Weiss thought the answer might be 'yes' is due to a curious example presented in [OW87]. It pertains to a well-known fundamental property of entropy: it is nonincreasing under factor maps. Let (G, X, μ) and (G, Y, ν) be two systems. A map $\phi : X \to Y$ is a factor if $\phi_*\mu = \nu$ and $\phi(gx) = g\phi(x)$ for almost every $x \in X$ and every $g \in G$. If G is amenable then the mean entropy of a factor is less than or equal to the mean entropy of the source. This is essentially due to Sinai [Si59]. So if $K_n = \{1, \dots, n\}$ and κ_n is the uniform probability measure on K_n then (G, K_2^G, κ_2^G), which has entropy log(2), cannot factor onto (G, K_4^G, κ_4^G), which has entropy log(4).

The argument above fails if G is nonamenable. Indeed, let $G = \langle a, b\rangle$ be a rank 2 free group. Identify K_2 with the group $\mathbb{Z}/2\mathbb{Z}$ and K_4 with $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$. Then

$$\phi(x)(g) := \big(x(g) + x(ga), x(g) + x(gb)\big) \ \forall x \in K_2^G, g \in G$$

is a factor map from (G, K_2^G, κ_2^G) onto (G, K_4^G, κ_4^G). This is Ornstein-Weiss' example. It is the main ingredient in the proof of the next theorem, which will appear in a separate paper.

THEOREM 1.2. *Let G be any countable group that contains a nonabelian free subgroup. Then every nontrivial Bernoulli shift over G factors onto every Bernoulli shift over G.*

To prove Theorem 1.1, the following invariant is introduced. Let (X, μ) be any probability space on which $G = \langle s_1, \dots, s_r\rangle$, the rank r free group, acts by measure-preserving transformations. Let $\alpha = \{A_1, \dots, A_n\}$ be a partition of X into finitely many measurable sets. Let $B(e, n) \subset G$ denote the ball of radius n centered at the identity element with respect to the word metric induced by $S = \{s_1^{\pm 1}, \dots, s_r^{\pm 1}\}$. The join of two partitions α, β of X is defined by $\alpha \vee \beta = \{A \cap B \mid A \in \alpha, B \in \beta\}$. Let

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

$$F(\alpha) := (1 - 2r)H(\alpha) + \sum_{i=1}^{r} H(\alpha \vee s_i\alpha),$$

$$\alpha^n := \bigvee_{g\in B(e,n)} g\alpha,$$

$$f(\alpha) := \inf_n F(\alpha^n).$$

A partition α is *generating* if the smallest G-invariant σ-algebra containing α is the σ-algebra of all measurable sets (up to sets of measure zero). The main theorem of this paper is:

THEOREM 1.3. *Let $G = \langle s_1, \dots, s_r \rangle$. Let (G, X, μ) be a system. If α and β are finite measurable generating partitions of X then $f(\alpha) = f(\beta)$. Hence this number, denoted $f(G, X, \mu)$, is a measure-conjugacy invariant.*

Theorem 5.2 below implies that if $|K| < \infty$ then $f(G, K^G, \kappa^G) = H(\kappa)$. This and Stepin's theorem proves Theorem 1.1. A simple exercise reveals that if $r = 1$, then $f(G, X, \mu) = h(G, X, \mu)$ is Kolmogorov's entropy.

Here is a brief outline of the paper. In the next section, standard entropy-theory definitions are presented. In Section 3, an equivalence relation, called combinatorial equivalence, is introduced on the space of finite partitions of X, where (X, μ) is a standard probability space on which a countable group G acts. We prove that the combinatorial equivalence class of a finite generating partition is dense in the space of all generating partitions. In Section 4, we introduce an operation on partitions called splitting and show that any two combinatorially equivalent partitions have a common splitting. This culminates in a condition sufficient for a function from the space of partitions to $\mathbb{R}$ to induce a measure-conjugacy invariant. In Section 5, this condition is shown to hold for the function F defined above. This proves Theorem 1.3. Then we prove Theorem 5.2 (that $f(G, K^G, \kappa^G) = H(\kappa)$ if $|K| < \infty$) and conclude Theorem 1.1.

2. Some standard definitions

For the rest of this section, fix a standard probability space (X, μ).

Definition 1. A *partition* $\alpha = \{A_1, \dots, A_n\}$ is a pairwise disjoint collection of measurable subsets A_i of X such that $\cup_{i=1}^n A_i = X$. The sets A_i are called the *partition elements* of α. Alternatively, they are called the *atoms* of α. Unless stated otherwise, all partitions in this paper are finite (i.e., $n < \infty$).

If α and β are partitions of X then we write $\alpha = \beta$ a.e. if for all $A \in \alpha$ there exists $B \in \beta$ with $\mu(A \Delta B) = 0$. Let $\mathcal{P}$ denote the set of all a.e.-equivalence classes of finite partitions of X. By a standard abuse of notation, we will refer to elements of $\mathcal{P}$ as partitions.

Definition 2. If $\alpha, \beta \in \mathcal{P}$ then the *join* of α and β is the partition $\alpha \vee \beta = \{A \cap B \mid A \in \alpha,\ B \in \beta\}$.

Definition 3. Let $\mathcal{F}$ be a σ-algebra contained in the σ-algebra of all measurable subsets of X. Given a partition α, define the *conditional information function* $I(\alpha|\mathcal{F}) : X \to \mathbb{R}$ by

$$I(\alpha|\mathcal{F})(x) = -\log\big(\mu(A_x|\mathcal{F})(x)\big)$$

where A_x is the atom of α containing x. Here $\mu(A_x|\mathcal{F}) : X \to \mathbb{R}$ is the conditional expectation of χ_{A_x}, the characteristic function of A_x, with respect to the σ-algebra $\mathcal{F}$.

The *conditional entropy of* α *with respect to* $\mathcal{F}$ is defined by

$$H(\alpha|\mathcal{F}) = \int_X I(\alpha|\mathcal{F})(x)\, d\mu(x).$$

If β is a partition then, by abuse of notation, we can identify β with the σ-algebra equal to the set of all unions of partition elements of β. Through this identification, $I(\alpha|\beta)$ and $H(\alpha|\beta)$ are well-defined. Let $I(\alpha) = I(\alpha|\{\varnothing, X\})$ and $H(\alpha) = H(\alpha|\{\varnothing, X\})$.

LEMMA 2.1. *For any two partitions* α, β *and for any two* σ*-algebras* $\mathcal{F}_1, \mathcal{F}_2$ *with* $\mathcal{F}_1 \subset \mathcal{F}_2$,

$$H(\alpha \vee \beta) = H(\alpha) + H(\beta|\alpha),$$
$$H(\alpha|\mathcal{F}_2) \le H(\alpha|\mathcal{F}_1).$$

Proof. This is well-known. For example, see [Gla03, Prop. 14.16, p. 255]. □

Definition 4 (Rokhlin distance). Define $d : \mathcal{P} \times \mathcal{P} \to \mathbb{R}$ by

$$d(\alpha, \beta) = H(\alpha|\beta) + H(\beta|\alpha) = 2H(\alpha \vee \beta) - H(\alpha) - H(\beta).$$

By [Par69, Th. 5.22, p. 62] this defines a distance function on $\mathcal{P}$. If G is a group acting by measure-preserving transformations on (X, μ) then the action of G on $\mathcal{P}$ is isometric. Thus, if $g \in G$, $\alpha, \beta \in \mathcal{P}$ then $d(g\alpha, g\beta) = d(\alpha, \beta)$. From now on, we consider $\mathcal{P}$ with the topology induced by $d(\cdot, \cdot)$.

Definition 5. Let G be a group acting on (X, μ). Let α be a partition of X. Let Σ_α be the smallest G-invariant σ-algebra containing α. Then α is *generating* if for any measurable set $A \subset X$ there exists a set $A' \in \Sigma_\alpha$ such that $\mu(A \Delta A') = 0$. Let $\mathcal{P}_{\text{gen}} \subset \mathcal{P}$ denote the set of all generating partitions.

3. **Combinatorially equivalent partitions**

For this section, fix a countable group G and an action of G on a standard probability space (X, μ) by measure-preserving transformations.

Definition 6. Given $\alpha \in \mathcal{P}$ and $F \subset G$ finite, let $\alpha^F = \bigvee_{f \in F} f\alpha$.

Definition 7. If $\alpha, \beta \in \mathcal{P}$ are such that for all $A \in \alpha$ there exists $B \in \beta$ with $\mu(A - B) = 0$ then we say that α *refines* β and denote it by $\alpha \ge \beta$. Equivalently, β is a *coarsening* of α.

Definition 8. Let $\alpha, \beta \in \mathcal{P}$. We say that α is *combinatorially equivalent* to β if there exist finite sets $L, M \subset G$ such that $\alpha \le \beta^L$ and $\beta \le \alpha^M$. Let $\mathcal{P}_{\text{eq}}(\alpha) \subset \mathcal{P}$ denote the set of partitions combinatorially equivalent to α

The goal of this section is to prove Theorem 3.3 below: If α is a generating partition then $\mathcal{P}_{\text{eq}}(\alpha)$ is dense in the subspace of all generating partitions.

LEMMA 3.1. *Let* α *be a generating partition and* $\beta = \{B_1, \dots, B_m\} \in \mathcal{P}$. *Let* $\varepsilon > 0$. *Then there exists a partition* $\beta' = \{B'_1, \dots, B'_m\}$ *and a finite set* $L \subset G$ *such that* $\alpha^L \geq \beta'$ *and for all* $i = 1 \dots m$, $\mu(B_i \Delta B'_i) \leq \varepsilon$.

Proof. Since α is generating, there exists a finite set $L \subset G$ such that for every $i \in \{1, \dots, m\}$, there is a set B''_i, equal to a finite union of atoms of α^L, such that $\mu(B_i \Delta B''_i) < \frac{\varepsilon}{m}$. For $i = 1 \dots m-1$, let

$$B'_i := B''_i - \bigcup_{j \neq i} B''_j.$$

$$B'_m := X - \bigcup_{i<m} B'_i = B''_m \cup \bigcup_{i \neq j} B''_i \cap B''_j.$$

Observe that for all $i = 1 \dots m$,

$$B_i - \bigcup_j B''_j \Delta B_j \subset B'_i \subset B_i \cup \bigcup_j B''_j \Delta B_j.$$

Thus

$$\mu(B'_i \Delta B_i) \leq m\Big(\frac{\varepsilon}{m}\Big) = \varepsilon.$$

By construction, $\beta' = \{B'_1, \dots, B'_m\} \leq \alpha^L$. □

LEMMA 3.2. *Let* $\alpha = \{A_1, \dots, A_n\} \in \mathcal{P}$ *and* $\beta \in \mathcal{P}_{\text{gen}}$. *Let* $\varepsilon > 0$. *Then there exists a finite set* $M \subset G$ *such that for all finite* $L \subset G$ *with* $M \subset L$, *the partition elements* $\{B^L_1, \dots, B^L_{m_L}\}$ *of* β^L *can be ordered so that there exists an* $r \in \{1, \dots, m_L\}$ *and a function* $f : \{1, 2, \dots, r\} \to \{1, 2, \dots, n\}$ *so that for all* $i \in \{1, \dots, r\}$,

$$\frac{\mu(B^L_i \cap A_{f(i)})}{\mu(B^L_i)} \geq 1 - \varepsilon$$

and

$$\mu\Big(\bigcup_{i>r} B^L_i\Big) < \varepsilon.$$

Proof. Let $\delta > 0$ be such that $\delta < \left(\frac{\varepsilon}{n}\right)^2$. By the previous lemma, there exists a partition $\alpha' = \{A'_1, \dots, A'_n\} \in \mathcal{P}$ and a finite set $M \subset G$ such that $\alpha' \leq \beta^M$ and $\mu(A'_i \Delta A_i) < \delta$ for all i. Let L be any finite subset of G with $M \subset L$.

Let $\beta^L = \{B^L_1, \dots, B^L_{m_L}\}$ and let $f : \{1, \dots, m_L\} \to \{1, \dots, n\}$ be the function $f(i) = j$ if $\mu(B^L_i - A'_j) = 0$. This is well-defined since β^L refines α'.

After reordering the partition elements of $\beta^L = \{B^L_1, \dots, B^L_{m_L}\}$ if necessary, we may assume that there is an $r \in \{0, \dots, m_L\}$ such that, if $r > 0$ then for all $i \leq r$,

$$\frac{\mu(B^L_i \cap A_{f(i)})}{\mu(B^L_i)} \geq 1 - \sqrt{\delta},$$

and if $i > r$ then

$$\frac{\mu(B_i^L \cap A_{f(i)})}{\mu(B_i^L)} < 1 - \sqrt{\delta}.$$

So if $i > r$ then

$$\mu(B_i^L \cap A_{f(i)}) < (1 - \sqrt{\delta})\mu(B_i^L).$$

Thus

$$\begin{aligned}\mu(B_i^L) &= \mu(B_i^L - A_{f(i)}) + \mu(B_i^L \cap A_{f(i)}) \\ &< \mu(B_i^L - A_{f(i)}) + (1 - \sqrt{\delta})\mu(B_i^L).\end{aligned}$$

Solve for $\mu(B_i^L)$ to obtain

$$\mu(B_i^L) < \frac{1}{\sqrt{\delta}}\mu(B_i^L - A_{f(i)}).$$

Since the atoms B_i^L are pairwise disjoint, it follows that

$$\mu\Big(\bigcup_{i>r} B_i^L\Big) < \frac{1}{\sqrt{\delta}}\mu\Big(\bigcup_{i>r} B_i^L - A_{f(i)}\Big).$$

Since $\mu(B_i^L - A'_{f(i)}) = 0$, it must be that $B_i^L - A_{f(i)} \subset A'_{f(i)} - A_{f(i)}$, up to a set of measure zero. So,

$$\begin{aligned}\mu\Big(\bigcup_{i>r} B_i^L\Big) &\le \frac{1}{\sqrt{\delta}}\mu\Big(\bigcup_{i>r} A'_{f(i)} - A_{f(i)}\Big) \\ &\le n\sqrt{\delta} < \varepsilon.\end{aligned}$$

□

THEOREM 3.3. *If α is a generating partition then*

$$\mathcal{P}_{\text{gen}} \subset \overline{\mathcal{P}_{\text{eq}}(\alpha)}.$$

In other words, the subspace of partitions combinatorially equivalent to α is dense in the space of all generating partitions.

Proof. Let $\alpha = \{A_1, \dots, A_n\}$ and $\beta = \{B_1, \dots, B_m\} \in \mathcal{P}_{\text{gen}}$. Without loss of generality, we assume that $\mu(A_i) > 0$ for all $i = 1 \dots n$. Let $\varepsilon > 0$. By the previous lemma, there exists a finite set $L \subset G$ such that the atoms of $\beta^L = \{B_1^L, \dots, B_{m_L}^L\}$ can be ordered so that there exists an $r \in \{1, \dots, m_L\}$ and a function $f : \{1, 2, \dots r\} \to \{1, 2, \dots, n\}$ so that for all $i \in \{1, \dots, r\}$,

$$\frac{\mu(B_i^L \cap A_{f(i)})}{\mu(B_i^L)} \ge 1 - \varepsilon$$

and

$$\mu\Big(\bigcup_{i>r} B_i^L\Big) < \varepsilon. \tag{1}$$

By choosing ε small enough (if necessary) we may assume that f maps onto $\{1, 2, \ldots, n\}$ (for example, by choosing ε to be smaller than $\frac{1}{2}\mu(A_j)$ over all $j = 1 \ldots n$). By definition of β^L, $m_L \leq m^{|L|}$. If necessary, we may assume that $m_L = m^{|L|}$ after modifying β^L by adding to it several copies of the empty set. That is, for some i, it may occur that $B_i^L = \varnothing$.

Let $\delta > 0$ be such that $\delta < \varepsilon$. By Lemma 3.1 there exists a partition $\gamma = \{C_1, \ldots, C_m\}$ and a finite set $M \subset G$ such that $\gamma \leq \alpha^M$ and $\mu(C_i \Delta B_i) < \delta$ for all i. By choosing δ small enough we may assume the following. Let $\gamma^L = \{C_1^L, \ldots, C_{m_L}^L\}$. Then, after reordering the atoms of γ^L if necessary,

$$\mu\Big(\bigcup_{j=1}^{m_L} C_j^L - B_j^L\Big) \leq \varepsilon. \tag{2}$$

Let

$$\begin{aligned} C_i' &= \{x \in C_i \mid \text{ if } x \in C_j^L \text{ for some } j \text{ then } x \in A_{f(j)}\} \\ &= \bigcup_{j=1}^{m_L} C_i \cap C_j^L \cap A_{f(j)}. \end{aligned}$$

Let $C_{i,j} = C_i \cap A_j - C_i'$. Let

$$\gamma_1 = \{C_i' \mid i = 1 \ldots m\} \cup \{C_{i,j} \mid 1 \leq i, j \leq m\}.$$

Note that $\gamma_1 \leq (\alpha^M)^L = \alpha^{LM}$ where $LM = \{lm \mid l \in L, m \in M\}$. We claim that γ_1 is combinatorially equivalent to α. Let Σ_1 be the smallest G-invariant collection of subsets of X that is closed under finite intersections and unions and contains the atoms of γ_1. It suffices to show that every atom of α is in Σ_1. Observe that, for each i, $C_i = C_i' \cup \bigcup_{j=1}^m C_{i,j}$. Hence, $C_i \in \Sigma_1$. Therefore the atoms of γ^L are also in Σ_1. Since f maps onto $\{1, 2, \ldots, n\}$, the definition of C_i' implies

$$C_i' \cap A_p = \cup\{C_i' \cap C_j^L \mid f(j) = p\}.$$

So $C_i' \cap A_p$ is in Σ_1 for all i, p. Now $C_i \cap A_p = C_{i,p} \cup (C_i' \cap A_p)$. So $C_i \cap A_p \in \Sigma_1$ for all i, p. Since

$$A_p = \bigcup_{i=1}^m C_i \cap A_p,$$

$A_p \in \Sigma_1$. Since p is arbitrary, this proves the claim. Thus $\gamma_1 \in \mathcal{P}_{\text{eq}}(\alpha)$.

We claim that $\mu(C_i' \Delta C_i) \leq 3\varepsilon$ for all i. By definition,

$$C_i' \Delta C_i = C_i - C_i' \subset \bigcup_{j=1}^{m_L} C_j^L - A_{f(j)}.$$

For each j,

$$C_j^L - A_{f(j)} \subset (C_j^L - B_j^L) \cup (B_j^L - A_{f(j)}).$$

Thus,

$$C_i' \Delta C_i \subset \bigcup_{j=1}^{m_L} (C_j^L - B_j^L) \cup \bigcup_{j=1}^{r} (B_j^L - A_{f(j)}) \cup \bigcup_{j>r} (B_j^L - A_{f(j)}). \tag{3}$$

If $j \le r$, then by definition of r,

$$\frac{\mu(B_j^L \cap A_{f(j)})}{\mu(B_j^L)} \ge 1 - \varepsilon.$$

This implies

$$\mu(B_j^L - A_{f(j)}^L) \le \varepsilon \mu(B_j^L).$$

Thus

$$\mu\Big(\bigcup_{j=1}^{r} B_j^L - A_{f(j)}^L\Big) \le \sum_j \varepsilon \mu(B_j^L) \le \varepsilon. \tag{4}$$

Equations (3), (2), (4) and (1) imply the claim.

Since $\delta < \varepsilon$ and $\mu(C_i \Delta B_i) < \delta$ for all i, the above claim implies that $\mu(C_i' \Delta B_i) \le 4\varepsilon$ for all i. Thus we have shown that for every $\varepsilon > 0$, there exists a partition $\gamma_1 = \{C_1', \dots, C_m', \dots\}$, combinatorially equivalent to α, containing at most $m + m^2$ partition elements and such that $\mu(C_i' \Delta B_i) < 4\varepsilon$ for $i = 1 \dots m$. This implies that β is in the closure of $\mathcal{P}_{\text{eq}}(\alpha)$. Since β is arbitrary this implies the theorem. □

4. Splittings

In this section, G can be any finitely generated group with finite symmetric generating set S. Let (X, μ) be a standard probability space on which G acts by measure-preserving transformations.

Definition 9. Let α be a partition. A *simple splitting* of α is a partition σ of the form $\sigma = \alpha \vee s\beta$ where $s \in S$ and β is a coarsening of α.

A *splitting* of α is any partition σ that can be obtained from α by a sequence of simple splittings. In other words, there exist partitions $\alpha_0, \alpha_1, \dots, \alpha_m$ such that $\alpha_0 = \alpha$, $\alpha_m = \sigma$ and α_{i+1} is a simple splitting of α_i for all $1 \le i < m$.

If σ is a splitting of α then α and σ are combinatorially equivalent. The splitting concept originated from Williams' work [Wil73] in symbolic dynamics.

Definition 10. The *Cayley graph* Γ of (G, S) is defined as follows. The vertex set of Γ is G. For every $s \in S$ and every $g \in G$ there is a directed edge from g to gs labeled s. There are no other edges.

The *induced subgraph* of a subset $F \subset G$ is the largest subgraph of Γ with vertex set F. A subset $F \subset G$ is *connected* if its induced subgraph in Γ is connected.

LEMMA 4.1. *If $\alpha, \beta \in \mathcal{P}$, α refines β and $F \subset G$ is finite, connected and contains the identity element e then*

$$\alpha \vee \bigvee_{f \in F^{-1}} f\beta$$

is a splitting of α.

Proof. We prove this by induction on $|F|$. If $|F| = 1$ then $F = \{e\}$ and the statement is trivial. Let $f_0 \in F - \{e\}$ be such that $F_1 = F - \{f_0\}$ is connected. To see that such an f_0 exists, choose a spanning tree for the induced subgraph of F. Let f_0 be any leaf of this tree that is not equal to e.

By induction, $\alpha_1 := \alpha \vee \bigvee_{f \in F_1^{-1}} f\beta$ is a splitting of α. Since F is connected, there exists an element $f_1 \in F_1$ and an element $s_1 \in S$ such that $f_1 s_1 = f_0$. Since $f_1 \in F_1$, α_1 refines $(f_1^{-1}\beta)$. Thus

$$\alpha \vee \bigvee_{f \in F^{-1}} f\beta = \alpha_1 \vee f_0^{-1}\beta = \alpha_1 \vee s_1^{-1}(f_1^{-1}\beta)$$

is a splitting of α. ◻

PROPOSITION 4.2. *Let α, β be two combinatorially equivalent generating partitions. Then there is an $n \geq 0$ such that*

$$\alpha^n = \bigvee_{g \in B(e,n)} g\alpha$$

is a splitting of β. Here $B(e, n)$ is the ball of radius n centered at the identity element in G with respect to the word metric induced by S. Of course, α^n is also a splitting of α.

This proposition is a variation of a result that is well-known in the case $G = \mathbb{Z}$ in the context of subshifts of finite-type. For example, see [LM95, Th. 7.1.2, p. 218]. It was first proven in [Wil73].

Proof. Let $L, M \subset G$ be finite sets such that $\alpha \leq \beta^L$ and $\beta \leq \alpha^M$. Let $l, m \in \mathbb{N}$ be such that $L \subset B(e, l)$ and $M \subset B(e, m)$. So $\alpha \leq \beta^l$ and $\beta \leq \alpha^m$. Since balls are connected and $\alpha \leq \beta^l$, the previous lemma implies $\beta^l \vee \alpha^{m+l}$ is a splitting of β^l, and therefore, is a splitting of β. But $\beta^l \vee \alpha^{m+l} = (\beta \vee \alpha^m)^l = \alpha^{m+l}$. ◻

THEOREM 4.3. *Let $\Phi : \mathcal{P} \to \mathbb{R}$ be any continuous function. Suppose that Φ is monotone decreasing under splittings*; *i.e., if σ is a splitting of α then $\Phi(\sigma) \leq \Phi(\alpha)$. Define $\phi : \mathcal{P} \to \mathbb{R}$ by*

$$\phi(\alpha) = \lim_{n\to\infty} \Phi(\alpha^n) = \inf_n \Phi(\alpha^n).$$

Then, for any two finite generating partitions α_1 and α_2, $\phi(\alpha_1) = \phi(\alpha_2)$. So we may define $\phi(G, X, \mu) = \phi(\alpha)$ for any finite generating partition α. The number $\phi(G, X, \mu)$ is a measure-conjugacy invariant.

Proof. Let α and β be two combinatorially equivalent finite partitions. We claim that $\phi(\alpha) = \phi(\beta)$. To see this, suppose, for a contradiction, that $\phi(\alpha) < \phi(\beta)$. Then there exists an $n \geq 0$ such that $\Phi(\alpha^n) < \phi(\beta)$. But by the previous proposition, there is an $m \geq 0$ such that β^m is a splitting of α^n which implies $\Phi(\alpha^n) \geq \Phi(\beta^m) \geq \phi(\beta)$, a contradiction. So $\phi(\alpha) = \phi(\beta)$.

For $n \geq 0$ and $\alpha \in \mathcal{P}$, let $\Phi_n(\alpha) = \Phi(\alpha^n)$. Since Φ is continuous and the map $\alpha \mapsto \alpha^n$ is also continuous, it follows that Φ_n is continuous. Since $\phi(\alpha) = \inf_n \Phi_n(\alpha)$, it follows that ϕ is upper semi-continuous, i.e., if $\{\beta_n\}$ is a sequence of partitions converging to α then $\limsup_n \phi(\beta_n) \leq \phi(\alpha)$.

Now let α, β be two finite generating partitions. By Theorem 3.3, there exist finite partitions $\{\beta_n\}_{n=1}^{\infty}$ combinatorially equivalent to β such that $\{\beta_n\}_{n=1}^{\infty}$ converges to α. So $\phi(\beta) = \limsup_n \phi(\beta_n) \leq \phi(\alpha)$. Similarly, $\phi(\alpha) \leq \phi(\beta)$. So $\phi(\alpha) = \phi(\beta)$. □

5. The f-invariant

In this section, $G = \langle s_1, \ldots, s_r \rangle$. Let (X, μ) be a standard probability space on which G acts by measure-preserving transformations and let $S = \{s_1^{\pm 1}, \ldots, s_r^{\pm 1}\}$. Note $|S| = 2r$. Let $F : \mathcal{P} \to \mathbb{R}$ be defined as in the introduction.

PROPOSITION 5.1. *Let $\alpha \in \mathcal{P}$. If σ is a splitting of α then $F(\sigma) \leq F(\alpha)$.*

Proof. By induction, it suffices to prove the proposition in the special case in which σ is a simple splitting of α. So let $\sigma = \alpha \vee t\beta$ for some $t \in S$ and coarsening β of α. For any $s \in S$,

$$\begin{aligned} H(\sigma \vee s\sigma) &= H(\alpha \vee s\alpha) + H(\sigma \vee s\sigma | \alpha \vee s\alpha) \\ &= H(\alpha \vee s\alpha) + H(s\sigma | \alpha \vee s\alpha) + H(\sigma | \alpha \vee s\alpha \vee s\sigma) \\ &\leq H(\alpha \vee s\alpha) + H(\sigma | \alpha \vee s^{-1}\alpha) + H(\sigma | \alpha \vee s\alpha). \end{aligned}$$

The last inequality occurs because μ is G-invariant, so

$$H(s\sigma | \alpha \vee s\alpha) = H(\sigma | \alpha \vee s^{-1}\alpha).$$

Since $H(\sigma) = H(\alpha) + H(\sigma|\alpha)$, the above implies

$$\begin{aligned} F(\sigma) &\le (1-2r)\big(H(\alpha) + H(\sigma|\alpha)\big) \\ &\quad + \sum_{i=1}^{r} H(\alpha \vee s\alpha) + H(\sigma|\alpha \vee s^{-1}\alpha) + H(\sigma|\alpha \vee s\alpha) \\ &= F(\alpha) + (1-2r)H(\sigma|\alpha) + \sum_{s\in S} H(\sigma|\alpha \vee s\alpha). \end{aligned}$$

Since $\sigma \le \alpha \vee t\alpha$, $H(\sigma|\alpha \vee t\alpha) = 0$. Hence

$$\begin{aligned} F(\sigma) - F(\alpha) &\le (1-2r)H(\sigma|\alpha) + \sum_{s\in S-\{t\}} H(\sigma|\alpha \vee s\alpha) \\ &= \sum_{s\in S-\{t\}} \Big(H(\sigma|\alpha \vee s\alpha) - H(\sigma|\alpha)\Big) \le 0. \end{aligned}$$ □

Theorem 1.3 now follows from the proposition above and Theorem 4.3.

Definition 11. Let K be a finite set and κ a probability measure on K. Let K^G be the product space with the product measure κ^G. The system (G, K^G, κ^G) is called the *Bernoulli shift* over G with base measure κ.

Let $A_k = \{x \in K^G \mid x(e) = k\}$ where e denotes the identity element in G. Then $\alpha = \{A_k \mid k \in K\}$ is the *Bernoulli partition* associated to K. It is generating and $H(\kappa) = H(\alpha)$, by definition.

THEOREM 5.2. *Let $G = \langle s_1, \dots, s_r\rangle$ be the free group of rank r. Let K be a finite set and κ a probability measure on K. Then*

$$f(G, K^G, \kappa^G) = H(\kappa).$$

Proof. Let α be the Bernoulli partition associated to K. Let $g_1, \dots, g_n$ be n distinct elements of G. It follows from the Bernoulli condition that the collection $\{g_i\alpha\}_{i=1}^n$ of partitions is independent. This means that if $j : \{1, \dots, n\} \to K$ is any function then

$$\kappa^G\Big(\bigcap_{i=1}^{n} g_i A_{j(i)}\Big) = \prod_{i=1}^{n} \kappa^G(A_{j(i)}).$$

It is well-known that this implies

$$H\Big(\bigvee_{i=1}^{n} g_i\alpha\Big) = \sum_{i=1}^{n} H(g_i\alpha) = nH(\alpha).$$

See, for example, [Gla03, Prop. 14.19, p. 257]. So for any $k \geq 1$,

$$F(\alpha^k) = \Big(\frac{1}{2}\sum_{s\in S} H(\alpha^k \vee s\alpha^k)\Big) - (|S|-1)H(\alpha^k)$$
$$= \Big(\frac{1}{2}\sum_{s\in S} |B(e,k)\cup B(s,k)|H(\alpha)\Big) - (|S|-1)|B(e,k)|H(\alpha).$$

Suppose $r > 1$. Then, since $G = \langle s_1, \dots, s_r\rangle$ is free, it is a short exercise to compute:

$$|B(e,k)| = 1 + |S|\frac{(|S|-1)^k - 1}{|S|-2},$$
$$|B(e,k)\cup B(s,k)| = 2\frac{(|S|-1)^{k+1} - 1}{|S|-2}$$

for all $s \in S$. Thus,

$$F(\alpha^k) = H(\alpha)\Big(|S|\frac{(|S|-1)^{k+1}-1}{|S|-2} - (|S|-1) - (|S|-1)|S|\frac{(|S|-1)^k-1}{|S|-2}\Big)$$
$$= H(\alpha).$$

If $r = 1$, then $|B(e,k)| = 2k+1$ and $|B(e,k)\cup B(s,k)| = 2k+2$. So it follows in a similar way that $F(\alpha^k) = H(\alpha)$. Thus $f(G,X,\mu) = \lim_{k\to\infty} F(\alpha^k) = H(\alpha) = H(\kappa)$. □

Proof of Theorem 1.1. By Stepin's theorem [Ste75], if (K_1,κ_1), (K_2,κ_2) are standard Borel probability spaces with $H(\kappa_1) = H(\kappa_2)$ then (G, K_1^G, κ_1^G) is measurably conjugate to (G, K_2^G, κ_2^G).

Suppose (K_1,κ_1), (K_2,κ_2) are Borel probability spaces such that (G, K_1^G, κ_1^G) is measurably conjugate to (G, K_2^G, κ_2^G). Let $(L_1,\lambda_1), (L_2,\lambda_2)$ be probability spaces with $|L_1| + |L_2| < \infty$ and $H(\lambda_i) = H(\kappa_i)$ for $i = 1,2$. By Stepin's theorem, (G, L_i^G, λ_i^G) is measurably conjugate to (G, K_i^G, κ_i^G). By the above theorem, $f(G, L_i^G, \lambda_i^G) = H(\lambda_i)$. Since f is a measure-conjugacy invariant, $H(\kappa_1) = H(\kappa_2)$. □

Acknowledgments. I am grateful to Russell Lyons for suggesting this problem, for encouragements and for many helpful conversations. I would like to thank Dan Rudolph for pointing out errors in a preliminary version. I would like to thank Benjy Weiss for many comments that helped me to greatly improve the exposition of this paper and simplify several arguments.

References

[Gla03] E. GLASNER, *Ergodic Theory via Joinings*, *Mathematical Surveys and Monographs* **101**, Amer. Math. Soc., Providence, RI, 2003. MR 2004c:37011 Zbl 1038.37002

[Kie75] J. C. KIEFFER, A generalized Shannon-McMillan theorem for the action of an amenable group on a probability space, *Ann. Probability* **3** (1975), 1031–1037. MR 52 #14232

[Kol58] A. N. KOLMOGOROV, A new metric invariant of transient dynamical systems and automorphisms in Lebesgue spaces, *Dokl. Akad. Nauk SSSR* (N.S.) **119** (1958), 861–864. MR 21 #2035a Zbl 0083.10602

[Kol59] ———, Entropy per unit time as a metric invariant of automorphisms, *Dokl. Akad. Nauk SSSR* **124** (1959), 754–755. MR 21 #2035b Zbl 0086.10101

[LM95] D. LIND and B. MARCUS, *An Introduction to Symbolic Dynamics and Coding*, Cambridge Univ. Press, Cambridge, 1995. MR 97a:58050

[Orn70a] D. ORNSTEIN, Bernoulli shifts with the same entropy are isomorphic, *Advances in Math.* **4** (1970), 337–352 (1970). MR 41 #1973

[Orn70b] ———, Two Bernoulli shifts with infinite entropy are isomorphic, *Advances in Math.* **5** (1970), 339–348 (1970). MR 43 #478a

[OW87] D. S. ORNSTEIN and B. WEISS, Entropy and isomorphism theorems for actions of amenable groups, *J. Analyse Math.* **48** (1987), 1–141. MR 88j:28014

[Par69] W. PARRY, *Entropy and Generators in Ergodic Theory*, W. A. Benjamin, Inc., New York-Amsterdam, 1969. MR 41 #7071

[Pet83] K. PETERSEN, *Ergodic Theory*, *Cambridge Studies in Advanced Mathematics* **2**, Cambridge University Press, Cambridge, 1983. MR 87i:28002

[Roh67] V. A. ROHLIN, Lectures on the entropy theory of transformations with invariant measure, *Uspehi Mat. Nauk* **22** (1967), 3–56. MR 36 #349

[Rud90] D. J. RUDOLPH, *Fundamentals of Measurable Dynamics*, *Oxford Science Publications*, The Clarendon Press Oxford Univ. Press, New York, 1990. MR 92e:28006

[Sin59] J. SINAĬ, On the concept of entropy for a dynamic system, *Dokl. Akad. Nauk SSSR* **124** (1959), 768–771. MR 21 #2036a

[Ste75] A. M. STEPIN, Bernoulli shifts on groups, *Dokl. Akad. Nauk SSSR* **223** (1975), 300–302. MR 53 #13521

[Wil73] R. F. WILLIAMS, Classification of subshifts of finite type, *Ann. of Math.* **98** (1973), 120–153; errata, ibid. **99** (1974), 380–381. MR 48 #9769

(Received February 14, 2009)

E-mail address: lpbowen@math.tamu.edu
MATHEMATICS DEPARTMENT, UNIVERSITY OF HAWAII AT MANOA, 2565 MCCARTHY MALL, HONOLULU HI 96822, UNITED STATES

and

MATHEMATICS DEPARTMENT, MAILSTOP 3368, TEXAS A&M UNIVERSITY, COLLEGE STATION, TX 77843-3368, UNITED STATES

http://www.math.tamu.edu/~lpbowen/