MIXING FOR DYADIC EQUIVALENCE

J. R. HASFURA-BUENAGA

Abstract. The notion of dyadic orbit equivalence for measure-preserving actions of \( \Gamma = \bigoplus_{\infty} \mathbb{Z}_2 \) on non-atomic probability spaces is introduced and it is shown that every dyadic equivalence class contains a mixing action. Also, a direct proof of a theorem of Stépin’s characterizing the values of entropy across an equivalence class is given.

1. Introduction

The orbit equivalence of groups acting on a Lebesgue space by ergodic, measure-preserving automorphisms has attracted the attention of numerous mathematicians in recent years. The earliest result in the subject is Dye’s theorem [D] which guarantees that any two ergodic actions of the integers are orbit equivalent. More recently, Connes, Feldman and Weiss [CFW] showed that in fact the result above still holds if we replace the group of integers by any two countable amenable groups. Furthermore, several authors have studied the equivalence relations that arise by considering only those orbit equivalences that satisfy some given restriction. For more precision, assume that \( G \) is a group, that \( T \) and \( S \) are \( G \)-actions and that a notion of restricted orbit equivalence is given. For many such notions the following relation is an equivalence relation in the class of \( G \)-actions: \( T \) and \( S \) are related if they are orbit equivalent by an orbit equivalence that satisfies the given restriction. Examples of equivalence relations defined by restricted orbit equivalences are isomorphism (an orbit equivalence that conjugates the actions), Kakutani equivalence and \( \alpha \)-equivalence. (For definitions and properties of these last two equivalences see [F], [Ka], [ORW], [dJR] and [FdJR].)

In [R], Rudolph has developed a very deep, general theory of restricted orbit equivalence of \( \mathbb{Z} \)-actions which contains as particular cases the result of Dye’s mentioned above, Ornstein’s isomorphism theorem and the Kakutani equivalence theorem. But, while an isomorphism theory that parallels the classical Bernoulli theory has been developed for a large class of amenable groups by Ornstein and Weiss [OW], difficulties arise when trying to lift Rudolph’s results beyond the \( \mathbb{Z}^d \) case (even in the case of discrete abelian groups, [KR], [K]).

Received April 28, 1994; revised November 7, 1994.

This paper gives a result concerning a special type of orbit equivalence for actions of the Abelian group \( \Gamma = \bigoplus_1^\infty \mathbb{Z}_2 = \{(g(n))_{n \in \mathbb{N}} \mid g(n) \in \mathbb{Z}_2 \text{ for all } n, \text{ and } g(n) = 1 \text{ for only finitely many } n\} \) (under componentwise addition mod 2). Recall that an action of \( \Gamma \) (or a \( \Gamma \)-action) on the non-atomic Lebesgue probability space \((X, \mathcal{B}, \mu)\) is a group homomorphism \( T \) from \( \Gamma \) to the group \( \text{Aut}(X) \) of measure-preserving automorphisms of \( X \). For \( g \in \Gamma \), the automorphism \( T(g) \) of \( X \) will be denoted by \( T^g \). Also, we will write \( \gamma_i, i \geq 1 \), for that element of \( \Gamma \) having every coordinate except the \( i \)-th one equal to 0 and \( \Gamma_i, i \geq 0 \), for the subgroup of \( \Gamma \) consisting of those \( g \in \Gamma \) with \( g(n) = 0 \) for \( n > i \). We say that the \( \Gamma \)-actions \( T \) and \( S \) on, respectively, \((X, \mathcal{B}, \mu)\) and \((Y, \mathcal{C}, \nu)\) are orbit equivalent if a measure-preserving isomorphism \( \phi \) exists with \( \phi(\text{Orb}_T(x)) = \text{Orb}_S(\phi(x)) \) for a.e. \( x \in X \). Finally, \( T \) and \( S \) are dyadically equivalent if they are orbit equivalent by an orbit equivalence \( \phi \) which for every \( x \in X \) and every \( i > 0 \) has \( \phi(\Gamma_i(x)) = \Gamma_i(\phi(x)) \). When actions \( T \) and \( S \) are dyadically equivalent we will write \( T \sim S \). It is immediately verified that dyadic equivalence is an equivalence relation in the class of \( \Gamma \)-actions and a result of Štepin’s [S, Theorem 2 below] guarantees that there are uncountably many dyadic equivalence classes. (The work of Vershik’s [V] shows that if we replace in the definition of dyadic equivalence the condition ‘\( i > 0 \)’ by ‘for infinitely many \( i \)’ we get an equivalence relation with only one equivalence class, and Belinskaya [B] used this result to prove an orbit equivalence theorem for \( \mathbb{Z} \)-actions.) This paper is partially motivated by a comment in [S] in which the author asserts that the dyadic equivalence class of a particular \( \Gamma \)-action given there contains actions with continuous spectrum. We will extend that assertion as follows.

**Theorem 1.** Given any ergodic \( \Gamma \)-action \( T_0 \) there is a mixing \( \Gamma \)-action \( T \) with \( T_0 \sim T \).

Further motivation for the study of dyadic equivalence comes, as suggested above, from the interest in extending the restricted orbit equivalence theory to large classes of amenable groups. The definition of dyadic equivalence is reminiscent of that of Kakutani equivalence of \( \mathbb{Z}^d \)-actions (see [dJR]). It is the belief of the author that dyadic equivalence is the analogue of Kakutani equivalence in the \( \Gamma \)-context and Theorems 1 and 2 (which says that entropy is invariant under dyadic equivalence) provide some evidence supporting this thesis. But, can the analogy be established more firmly by, for example, establishing an equivalence theorem like those in [ORW], [Ha] and [KR]? Attempting to do it seems worthwhile given the very simple structure of the group \( \Gamma \) (which includes the presence of nice Følner sequences) and the possibility that this might shed some light on the difficulties associated in lifting the restricted orbit equivalence theorem to other (than \( \mathbb{Z}^d \)) groups. We will proceed to the details now. In §2 we give definitions.
and preliminary results, in §3 we prove Theorem 1 and in §4 we state formally and prove the result of Stëpin’s on entropy mentioned above.

2. Basic Facts

A $\Gamma$-action $T$ on $(X, B, \mu)$ is said to be free if $\mu(\{x \in X \mid T^g x = x\}) = 0$ for all $g \in \Gamma$ and ergodic if the only sets in $B$ invariant under each $\gamma_i$ are trivial, i.e. have $\mu$-measure 0 or 1. It is easily seen that ergodicity is an invariant of dyadic equivalence and that it does not imply freeness. We also notice that the sequence of subgroups $\Gamma_i$, $i > 0$ is, in the terminology of Ornstein and Weiss [OW], a special averaging sequence and, according to the results in that paper, the pointwise ergodic theorem holds when we average along it.

The following Rokhlin-type result will be used repeatedly in the sequel. Its proof is standard, so we give only a sketch of it.

**Lemma 1.** Let $T$ be a free $\Gamma$-action and $n \geq 1$. There exists a set $B_n \in B$ such that $T^g B_n \cap T^h B_n = \emptyset$ for $g, h \in \Gamma, g \neq h$, and $\mu(\cup_g T^g B_n) = 1$.

*Sketch of proof.* The existence of $B_1$ satisfying $\cup_g T^g B_1 = X$ is given in [H, p. 70]. To construct $B_n$ proceed as follows. Assume that a set $B_{n-1}$ with the properties of the statement has been constructed. Use the result for $n = 1$ to find a set $B$ with $B \cap T^g B = \emptyset$ and $\mu(B \cup T^g B) = 1$. Without losing generality we can assume that $\mu(A = B \cap B_{n-1}) > 0$. Enumerate the elements of $\Gamma_{n-1}$ as $g_1, g_2, \ldots, g_t$. Now find a subset $A_i$ of $A$, $\mu(A_i) > 0$, such that $T^{g_i} A_i \cap A_i = \emptyset$. This can be achieved because $T$ is free. Proceed inductively to find $A_i \subset A_{i-1}$, with $\mu(A_i) > 0$ and $T^{g_i} A_i \cap A_i = \emptyset$, $i = 2, \ldots, t$. Again, freeness allows this. Now one checks that the set $A_t$ has $T^g A_t \cap T^h A_t = \emptyset$ if $g, h \in \Gamma, g \neq h$. If $\mu(\cup_g T^g A_t) < 1$ then carry out the above procedure in the $\Gamma_n$-invariant set $X \setminus \cup_g T^g A_t$ — transfinite if necessary, but at most countably many times. \[\square\]

In this paper we are interested in a property of actions which is stronger than ergodicity. We say that the $\Gamma$-action $T$ is mixing if given any two sets $A, B \in B$ and $\epsilon > 0$ one can find $n_0 \in \mathbb{N}$ such that if $n > n_0$ then $|\mu(A \cap T^g B) - \mu(A) \mu(B)| < \epsilon$, for all $g \in \Gamma \setminus \Gamma_{n_0}$. (Throughout this paper both the back and front slash symbols will be used. The back slash will always denote set difference, while the front slash will be used when writing factor groups.) Also, an eigenfunction of $T$ is a function $f : X \to C$ for which a group homomorphism $\lambda : \Gamma \to \{-1, 1\}$ exists satisfying $f(T^g(x)) = \lambda(g) f(x)$ for all $g \in \Gamma$, a.e. $x \in X$, and $T$ is said to have continuous spectrum if its only eigenfunctions are the constants. It is very easy to show that mixing implies ergodicity and continuous spectrum and that non-mixing, ergodic actions exist. For a sample, assume that $f$ is a non-constant eigenfunction of the ergodic action $T$. We can assume without losing generality that $\mu(\{x \in X : \Re f(x) > 0\}) > 0$. By the assumed ergodicity there exists $i$
such that $\lambda(\gamma_i) = -1$. Therefore, for infinitely many $n$ we can find $g \in \Gamma \setminus \Gamma_n$ with $\lambda(g) = -1$. But for such $g$, $\mu(P \cap T^g(P)) = 0 \neq \mu(P)\mu(P)$ and $T$ is not mixing.

Now assume that $T$ is a $\Gamma$-action and that $\mathcal{P}$ is a partition of the space $X$. If $|\text{dist}(T^g\mathcal{P} \vee \mathcal{P}) - \text{dist}(\mathcal{P}) \times \text{dist}(\mathcal{P})| < \epsilon$ for all $g \in \Gamma \setminus \Gamma_L$ (here $L < M$ are positive integers) we say that $T$ is $\epsilon$-mixing with respect to $\mathcal{P}$ on $[L, M]$. The following simple result will be used in the sequel.

**Lemma 2.** If $\mathcal{P}_i \uparrow \mathcal{B}$ is an increasing sequence of partitions of $X$, $\epsilon_i > 0$ decreases to 0, $L_i \in \mathbb{N}$, $L_i \leq L_{i+1}$, $i > 0$ and $T$ is $\epsilon_i$-mixing with respect to $\mathcal{P}_i$ on $[L_i, L_{i+1}]$ then $T$ is mixing.

For the proof of Theorem 1 we will adopt and adapt the strategy devised by Fieldsteel and Friedman in the proof of [FF, Thm. 1]. Thus, the action $T$ of Theorem 1 will be obtained as the limit of a sequence $T_i$ of actions, where $T_i$ is constructed from $T_{i-1}$ by an orbit change. To facilitate the description of these orbit changes and to explain the sense in which the limit above is taken, the following notion will be useful. (Here $T$ will be a $\Gamma$-action.) A **cocycle** $\alpha$ is a measurable function $\alpha: X \times \Gamma \to \Gamma$ which, for a.e. $x \in X$, satisfies

(i) $\alpha(x, \gamma) \gamma \to \Gamma$ is one-to-one and onto, and
(ii) $\alpha(x, gh) = \alpha(x, g) \alpha(T^g(x), h)$ (this is called the cocycle condition).

Two actions $T$ and $S$ orbit-equivalent by $\phi$ determine a cocycle $\alpha$ by setting, for a.e. $x$ and all $g \in \Gamma$, $\alpha(x, g) = \overline{g}$ if $\phi(T^g(x)) = S\overline{g}(\phi(x))$. Conversely, given the action $T$ and a cocycle $\alpha$ one can define another $\Gamma$-action $S$, orbit equivalent to $T$ by an orbit equivalence that preserves orbits, by setting $S\overline{g}(x) = T^g(x)$ if $\overline{g} = \alpha(x, g)$. We say that $\alpha$ is the cocycle for the pair $(T, S)$.

**Lemma 3.** If for a.e. $x \in X$ the cocycle $\alpha$ for the pair $(T, S)$ satisfies $\alpha(x, g) \in \Gamma_n$ for $g \in \Gamma_n$, all $n$, then $T \sim S$.

**Proof.** With $\phi: X \to X$ the identity map we get

$$\phi(T^g(x)) = T^g(x) = S\overline{g}(x) = S\overline{g}(\phi(x))$$

and by assumption $\overline{g} \in \Gamma_n$ if $g \in \Gamma_n$. \hfill \square

Cocycles satisfying the condition of the lemma will henceforth be referred to as **dyadic cocycles**.

If $\alpha$ and $\beta$ are cocycles their composition $\alpha \circ \beta$ is the cocycle defined by $\alpha \circ \beta(x, g) = \alpha(x, \beta(x, g))$. Now assume that $\alpha_i$ is a cocycle for each $i > 0$ and set $\beta_j = \alpha_j \circ \alpha_{j-1} \circ \cdots \circ \alpha_1$. If $\beta(x, g) = \lim_{j \to \infty} \beta_j(x, g)$ exists for a.e. $x \in X$ and $g \in \Gamma$ then $\beta$ satisfies the cocycle condition and, for a.e. $x$, $\beta(x, \gamma): \Gamma \to \Gamma$ is injective but need not be surjective. However, the following simple result holds.

**Lemma 4.** If $\alpha_i$ is a dyadic cocycle, $i > 0$, and $\lim_{j \to \infty} \beta_j = \beta$ (notation is as in the preceding paragraph) then $\beta$ is a dyadic cocycle.
Proof. Plainly, that $\alpha_t$ is dyadic for $i > 0$ implies that $\beta_j, j > 0$, is too. Given $g \in \Gamma$ let $n(g)$ be the smallest positive integer with $g \in \Gamma_{n(g)}$. Then, for a.e. $x$ and for all $j > 0$ there exists $g_j \in \Gamma_{n(g)}$ with $\beta_j(x, g_j) = g$. Since $\Gamma_{n(g)}$ is finite there must be a $\overline{g} \in \Gamma_{n(g)}$ such that $g_j = \overline{g}$ for infinitely many $j$. But $\beta_j \rightarrow \beta$ so that, in fact, $g_j = \overline{g}$ for all but finitely many $j$. Thus, $\beta(x, \overline{g}) = g$ and $\beta$ is a dyadic cocycle. \hfill \Box

Now we can be more precise regarding the limit above. Starting with $T_0$ we will construct a sequence of dyadic cocycles $\alpha_i$, and $\Gamma$-actions $T_i, i > 0$, with $\alpha_i$ the cocycle for the pair $(T_{i-1}, T_i)$ (alternatively, $\beta_j = \alpha_j \circ \cdots \circ \alpha_1$ will be the cocycle for the pair $(T_0, T_j)$). The sequence $\beta_j$ will converge to a dyadic cocycle $\beta$ and, as described above, this will give rise to an action $T$ dyadically equivalent to $T_0$. We will write that $T = \lim_{i \rightarrow \infty} T_i$. Thus, $T$ is the pointwise limit of the sequence $T_i$: for a.e. $x$ and all $g \in \Gamma$, $T^g_i(x)$ is eventually fixed. $T^g(x)$ is defined to be that point.

3. Proof of Theorem 1

The iterative step of the proof of Theorem 1 is contained in the following lemma. Before stating it, we make some comments on notation. If $\mathcal{P} = \{p_1, p_2, \ldots, p_t\}$ is a partition of $X$ and $A \subset X$ is measurable, by $\text{dist}_A(\mathcal{P})$ we will mean the distribution of the partition $\mathcal{P}_A = \{p_1 \cap A, p_2 \cap A, \ldots, p_t \cap A\}$ on the set $A$ with conditional measure $\mu_A(\cdot) = \mu(\cdot)/\mu(A)$. In symbols, $\text{dist}_A(\mathcal{P}) = (\mu_A(p_1 \cap A), \mu_A(p_2 \cap A), \ldots, \mu_A(p_t \cap A))$. The set $\{1, 2, \ldots, t\}$ is the index set of $\mathcal{P}$. If $\mathcal{P}'$ is another partition of $X$ with same index set, we identify the index set of $\mathcal{P} \vee \mathcal{P}'$ with the set of pairs $[i, j], 1 \leq i, j \leq t$, in the obvious way. Finally, we will write $\mathcal{P}(x) = i$ if $x \in p_i$, so we can think of a partition as a random variable with values in its index set. In particular, by the above identification, the random variable $(\mathcal{P} \vee \mathcal{P}')(x)$ can be thought of as a pair of random variables $[\mathcal{P}(x), \mathcal{P}'(x)]$, each with values in $\{1, 2, \ldots, t\}$.

Lemma 5. Given a free, ergodic $\Gamma$-action $T$, a finite partition $\mathcal{P} = \{p_1, p_2, \ldots, p_t\}$ of $X$ and $\epsilon > 0$ there is $L \in \mathbb{N}$, such that for all sufficiently large $M$ a $\Gamma$-action $S$ can be constructed which is dyadically equivalent to $T$ and $\epsilon$-mixing with respect to $\mathcal{P}$ on $[L, M]$.

Proof. Let $\eta > 0$, a small number, be given. We will see later that by making it sufficiently small, the conclusion of the lemma is obtained.

(1) Let $L$ be so large that for all $x$ in a measurable set $G$, $\mu(G) > 1 - \eta^2$, the distribution of $\mathcal{P}$ in the $(\mathcal{P}, T, L)$-name of $x$ is within $\eta$ of $\text{dist}(\mathcal{P})$. The ergodic theorem guarantees the existence of such $L$. (The $(\mathcal{P}, T, L)$-name of $x$ is the function $\mathcal{N} : \Gamma_L \rightarrow \mathcal{P}$ given by $\mathcal{N}(g) = p_i$ if $T^g(x) \in p_i$.)
(2) Let $M$ be so large that $M\eta > L$. 

(3) Use Lemma 1 to select a measurable set $B_M$ with $T^g B_M \cap T^h B_M = \emptyset$, $g, h \in \Gamma_M$, $g \neq h$ and $\mu(\cup_{g \in \Gamma_M} T^g B_M) = 1$. 

(4) Let $B$ be the set of $x \in B_M$ for which at least a fraction $> (1 - \eta)$ of the $g$ in $\Gamma_M$ has $T^g(x) \in G$. By (1), $\mu(B) > (1 - \eta) \mu(B_M)$. 

(5) Consider the factor group $\Gamma_M / \Gamma_L$. The symbol $[g]$ will denote that element of $\Gamma_M / \Gamma_L$ containing $g \in \Gamma_M$. Let $c : \Gamma_M / \Gamma_L \to \Gamma_L$ be any function satisfying $c([e]) = e$, where $e$ denotes the identity element in the appropriate group. The (finite) set of all such functions will be denoted by $\Xi$. Given one such $c$, the mapping $\pi : \Gamma_M \to \Gamma_M$ given by $\pi(g) = gc([g])$ is a permutation, as the following computation shows: $\pi(\pi(g)) = \pi(gc([g])) = gc([g]) = gc(gc([g])) = gc([g])c([g]) = g$. 

(6) Let $Q = \{q_1, q_2, \ldots, q_s\}$ be the partition of $B$ (see (4)) determined by $\mathcal{V}_{g \in \Gamma_M} T^g \mathcal{P}$. For $i = 1, 2, \ldots, s$ find a measurable, uniformly distributed map $f_i : q_i \to \Xi$. If $x \in q_i \subset B$, we denote $f_i(x)$ by $c_x$. Now fix $i$ and $[g] \in \Gamma_M / \Gamma_L$. The expression $c_x([g])$, regarded as a function of $x \in q_i$, is a random variable with values in $\Gamma_L$. Furthermore, by construction, if $[g], [gh] \in \Gamma_M / \Gamma_L$, $[g] \neq [gh]$ and $[g], [gh] \neq [e]$ then $c_x([g])$ and $c_x([gh])$ are independent random variables. We denote by $\pi_x$ the permutation of $\Gamma_M$ determined by $c_x$ as in (5). 

(7) For $x \in q_i$ and $g \in \Gamma_M$ set $\alpha_M(x, g) = \pi_x(g)$. If on the other hand, $g \notin \Gamma_M$ but $T^g(x) \in B_M$ we set $\alpha_M(x, g) = g$ and if $x \in B_M \setminus B$ and $g \in \Gamma_M$ we will put $\alpha_M(x, g) = g$. This partial function $\alpha_M$ extends to a measurable bijective cocycle $\alpha$ in the following way: if $x \in X$ and $g \in \Gamma$ are arbitrary, we set

$$
\alpha(x, g) = \pi_{x_0}(g_0)v(x_0, x_1)\pi_{x_1}(g_1)
$$

where $x_0, x_1 \in B_M$, $g_0, g_1 \in \Gamma_M$, $T^{g_0}(x) = x_0$, $T^{g_1}(T^{g_0}(x)) = x_1$ and $v = v(x_0, x_1)$ is that element of $\Gamma$ with $T^v(x_0) = x_1$. As explained above, $\alpha$ determines a $\Gamma$-action $S$ which is orbit equivalent to $T$.

(8) We now check that the cocycle $\alpha$ is dyadic. We let $x \in X$ be arbitrary and consider three possibilities for $g \in \Gamma$: (i) if $i > M$ and $g \in \Gamma_i \setminus \Gamma_M$ then (with notation as in (6)) $\alpha(x, g) = \pi_{x_0}(g_0)v(x_0, x_1)\pi_{x_1}(g_1) \in \Gamma_M \Gamma_i \Gamma_M \subset \Gamma_i$, (ii) if $L < i \leq M$ and $g \in \Gamma_i \setminus \Gamma_L$ then $\alpha(x, g) = \pi_{x_0}(g_0)^v x_0 \pi_{x_0}(g_0) g_0 g \Gamma_L \in \Gamma_i$ and (iii) if $g \in \Gamma_i$, $i \leq L$, then $\alpha(x, g) = \pi_{x_0}(g_0)^v x_0 \pi_{x_0}(g_0) g_0 g \Gamma_L = \pi_{x_0}(g_0)^v x_0 \pi_{x_0}(g_0) g_0 g \Gamma_L = \Gamma_i$, and in all three cases, $\alpha(x, g) \in \Gamma_i$, if $g \in \Gamma_i$ so that $\alpha$ is indeed dyadic. Lemma 3 implies that $T \sim S$. 

(9) Now fix $q_i \in Q$ and let $g$ be an element of $\Gamma_M \setminus \Gamma_L$. We say that $h \in \Gamma_M \setminus \Gamma_L$ is “good” for $g$ if

$$
\begin{align*}
(i) & \quad gh \not\in \Gamma_L \\
(ii) & \quad |\text{dist}_{T^h \Gamma_L q_i} (\mathcal{P}) - \text{dist} \mathcal{P}| < \eta \\
(iii) & \quad |\text{dist}_{T^{gh} \Gamma_L q_i} (\mathcal{P}) - \text{dist} \mathcal{P}| < \eta.
\end{align*}
$$
It follows from (2) and (4) that a fraction larger than $1 - 4\eta$ of $\Gamma_M$ consists of elements which are good for $g$.

(10) Again fix $q_i \in Q$ and $g \in \Gamma_M \setminus \Gamma_L$, and let $h$ be good for $g$ and $x$ be an element of $q_i$. Then $\pi_x(h) = hk$, with $k = c_x([h]) \in \Gamma_L$ uniformly distributed on $q_i$. Since $S^h(x) = T^{hk}(x)$ we get, using (9ii),

\[(i) \quad |\text{dist}_{S^h q_i}(\mathcal{P}) - \text{dist}(\mathcal{P})| < \eta.\]

Also, by (9i), $\pi_x(gh) = ghk$, with $k = c_x([gh]) \in \Gamma_L$ uniformly distributed on $q_i$. Thus, the element of $\mathcal{P}$ containing $S^{gh}(x)$ is the same one that contains $T^{ghk}(x)$ and, therefore, (9iii) gives

\[(ii) \quad |\text{dist}_{S^h q_i}(S^g \mathcal{P}) - \text{dist}(\mathcal{P})| < \eta.\]

(11) We now want to approximate $\text{dist}_{S^h q_i}(\mathcal{P} \vee S^g \mathcal{P})$. With the hypothesis of (10) still standing, we have that (recall the comments regarding notation that preceed the statement of the lemma)

\[(\mathcal{P} \vee S^g \mathcal{P})(S^h(x)) = [\mathcal{P}(S^h(x)), S^g \mathcal{P}(S^h(x))] = [\mathcal{P}(S^h(x)), \mathcal{P}(S^{gh}(x))]
= [\mathcal{P}(T^{hk}(x)), \mathcal{P}(T^{ghk}(x))].\]

Recalling (see (9)) that $[h] \neq [gh]$ and $[h], [gh] \neq [e]$ we use (6) to conclude that $c_x([h])$ and $c_x([gh])$ are independent. Thus,

\[(i) \quad \text{dist}_{S^h q_i}(\mathcal{P} \vee S^g \mathcal{P}) = \text{dist}_{T^{hk} \mathcal{L} q_i}(\mathcal{P}) \times \text{dist}_{T^{ghk} \mathcal{L} q_i}(\mathcal{P}).\]

(12) Combining (11i) with (10i) and (10ii) we obtain

\[|\text{dist}_{S^h q_i}(\mathcal{P} \vee S^g \mathcal{P}) - \text{dist}(\mathcal{P}) \times \text{dist}(\mathcal{P})| < 6\eta.\]

(13) Adding over $q_i$, $i = 1, 2, \ldots, s$ and over the appropriate $h$’s and using (3) and (9) we conclude that

\[|\text{dist}(\mathcal{P} \vee S^g \mathcal{P}) - \text{dist}(\mathcal{P}) \times \text{dist}(\mathcal{P})| < 10\eta.\]

Thus, if we take $\eta$ smaller than $\epsilon/10$ the conclusion of the lemma will hold. \(\square\)

We want to single out two features of the construction above that will be used in the conclusion of the proof of Theorem 1. Firstly, it is an immediate consequence of (8iii) that for a.e. $x$, $T^g(x) = S^g(x)$ for all $g \in \Gamma_L$. Secondly, since $T^{\Gamma^M}(x) = S^{\Gamma^M}(x)$ for a.e. $x \in X$ (see (8)), given any partition $\mathcal{R}$ of $X$, the frequencies of the atoms of $\mathcal{R}$ in the $(\mathcal{R}, T, M)$-name of $x$ are equal to their frequencies in its $(\mathcal{R}, S, M)$-name.
Proof of Theorem 1. We will use Lemma 5 repeatedly. Let $P_i \uparrow B$ and $\epsilon_i \downarrow 0$. Apply Lemma 5 to $T_0$, $P_1$ and $\epsilon_1$ to get a $\Gamma$-action $T_1$, and natural numbers $L_1 < L_2$ with $T_1 \epsilon_1$-mixing with respect to $P_1$ on $[L_1, L_2]$. Here, we make sure $L_2$ is so large that a fraction $(1-\eta_2^2)$ of $X$ is covered with points whose $(P_2, T_0, L_2)$-name has distribution of $P_2$ within $\eta_2$ of dist $(P_2)$, where $\eta_2$ is the value of $\eta$ in Lemma 5 that corresponds to $\epsilon_2$. By the comments in the preceding paragraph, the same is true of the $(P_2, T_1, L_2)$-names of those points and for all $g \in \Gamma_{L_2}$, $T_0^g = T_1^g$ a.e.

Now proceed inductively. Assume that natural numbers $L_1 < L_2 < \ldots < L_{i+1}$ and $\Gamma$-actions $T_j$, $1 \leq j \leq i$, have been found such that $T_j$ is $\epsilon_j$-mixing with respect to $P_j$ on $[L_j, L_{j+1}]$, $T_j^g = T_j^{g-1}$ a.e. for $g \in \Gamma_{L_j}$ and $L_{j+1}$ is so large that at least $1 - \eta_{j+1}^2$ of $X$ is covered by points whose $(P_{j+1}, T_j, L_{j+1})$-name has distribution of $P_{j+1}$ within $\eta_{j+1}$ of dist $(P_{j+1})$, where $\eta_{j+1}$ is the value of $\eta$ in Lemma 5 that corresponds to $\epsilon_{j+1}$. Then, further application of that lemma with $T_i, \epsilon_i+1$ and $P_{i+1}$ produces a $\Gamma$-action $T_{i+1}$ and a natural number $L_{i+2}, L_{i+1} < L_{i+2}$, with $T_{i+1} \epsilon_{i+1}$-mixing with respect to $P_{i+1}$ on $[L_{i+1}, L_{i+2}]$ and $T_i^g = T_{i+1}^g$ for all $g \in \Gamma_{L_{i+1}}$. Of course $L_{i+2}$ is chosen large enough to enable us to continue the induction.

What we have found is an increasing sequence of natural numbers $L_i$ and a convergent sequence $T_i$ of $\Gamma$-actions (whose limit $T$ coincides with $T_i$ on $\Gamma_{L_i}$) such that, for $1 \leq j \leq i \leq k$, $T_k$ — and consequently $T$ — is $\epsilon_j$-mixing with respect to $P_j$ on $[L_j, L_{j+1}]$. By choice of $P_i$ and $\epsilon_i$ it follows (Lemma 2) that $T$ is mixing while Lemma 4 guarantees that $T_0 \sim T$. \hfill $\square$

4. Entropy

In this section we prove the following theorem due to Stépin [S2] which says that, as in the case of even Kakutani equivalence, entropy is stable with respect to dyadic equivalence. Formally:

Theorem 2. If $T$ and $S$ are $\Gamma$-actions and $T \sim S$ then $h(T) = h(S)$.

The proof of Theorem 2 given in [S2] is somewhat indirect, relying on a Sinai-type result (an action of positive entropy has Bernoulli factors of full entropy) for periodic subgroups of $\mathbb{R}/\mathbb{Q}$. Here we give a direct proof which depends only on the usual name-counting characterization of entropy and on constructions like those of the previous section. Recall that for a $\Gamma$-action $T$ one defines $h(T) = \sup h(T, \mathcal{P})$, where $h(T, \mathcal{P}) = \lim_{n \to \infty} \frac{1}{2^n} H(\bigvee_{g \in \Gamma_n} T^g \mathcal{P})$ is the entropy of $T$ with respect to $\mathcal{P}$, and the supremum is taken over all finite partitions $\mathcal{P}$ of $X$. The Shannon-McMillan-Breiman theorem for $\Gamma$-actions [OW] implies that $h(T, \mathcal{P})$ is given by the infimum of all $h' \geq 0$ having the property that for every $\epsilon > 0$, if $n \in \mathbb{N}$ is sufficiently large, one can cover at least $1 - \epsilon$ of $X$ in measure with fewer than $2^{n(h' + \epsilon)} (T, \mathcal{P}, n)$-atoms. Now assume that $T$ and $S$ are $\Gamma$-actions on the space $(X, \mathcal{B}, \mu)$, $T \sim S$, that $\mathcal{P}$ is a finite partition of $X$, and that $0 < l \in \mathbb{N}$. Let
\(B_t \in \mathcal{B}\) be such that \(S^g B_t \cap S^h B_t = \emptyset\) for \(g, h \in \Gamma_t, g \neq h\) and \(\mu(\cup_{g \in \Gamma_t} S^g B_t) = 1\) (Lemma 1). The following two assignments define a new \(\Gamma\)-action \(S_t\) on \(X\):  
(a) if \(x \in B_t\) and \(g \in \Gamma_t\) then \((S_t)^g(x) = T^g(x)\),  
(b) if \(x, x' \in B_t\) and \(S^g(x) = x'\) then we set \((S_t)^g(x) = x'\).

That \(S\) and \(T\) are dyadically equivalent implies that \(S\) and \(S_t\) are too. Let \(\alpha\) be the cocycle for the pair \((S, S_t)\). Then, for a.e. \(x \in B_t\), the map \(\alpha(x, \cdot)\) restricted to \(\Gamma_t\) is a permutation. Call \(\mathcal{Q} = \{q_1, q_2, \ldots, q_l\}\) that partition of \(B_t\) obtained by refining \(\forall_{g \in \Gamma_t} S^g P | B_t\) (this is the trace of \(\forall_{g \in \Gamma_t} S^g P\) on \(B_t\)) according to the permutation of \(\Gamma_t\) used to get \(S_t\) from \(S\). That is, two elements \(x, x'\) of \(B_t\) will be in the same atom of \(\mathcal{Q}\) exactly when they have the same \((\mathcal{P}, S, l)\)-name and \(\alpha(x, \cdot)\) of \(\Gamma_t\) to be the partition of \(X\) whose atoms are the sets \(S^g q_i\), \(g \in \Gamma_t\), \(q_i \in \mathcal{Q}\). The defining properties of \(B_t\) guarantee that \(\mathcal{P}_t\) is indeed a partition, and it is easily seen that \(\mathcal{P}_t\) refines \(\mathcal{P}\). The following equalities also follow from the construction of \(\mathcal{P}_t\): \(\mathcal{P}_t = \forall_{g \in \Gamma_t} S^g \mathcal{P}_t = \forall_{g \in \Gamma_t} (S_t)^g \mathcal{P}_t\). Indeed, if \(g, h \in \Gamma_t\) and \(p, q \in \mathcal{P}_t\) (so that \(p = S^k q_i, q = S^n q_j\)) then \(S^g p \cap S^h q = S^{gk} q_i \cap S^{hm} q_j\) is empty if \(i \neq j\) or if \(i = j\) and \(gk \neq hm\). Of course, if \(i = j\) and \(gk = hm\) then \(S^g p \cap S^h q = S^{gk} q_i \in \mathcal{P}_t\) because \(gk \in \Gamma_t\). This proves the first equality. To get the second one, it suffices to recall that, by construction, if \(x, x' \in q_i\) then \(\alpha(x, \cdot)\) of \(\Gamma_t\), and to notice that this condition implies that, for \(g \in \Gamma_t, S^q q_i = (S_t)^h q_i\) for some \(h \in \Gamma_t\). It follows from these observations that for \(n > l\)

\[
(i) \quad \text{dist}\ (\forall_{g \in \Gamma_n} S^g \mathcal{P}_t) = \text{dist}\ (\mathcal{P}_t \cup \forall_{i=1}^{l+1} S^\gamma \mathcal{P}_t)
\]

and

\[
(ii) \quad \text{dist}\ (\forall_{g \in \Gamma_n} (S_t)^g \mathcal{P}_t) = \text{dist}\ (\mathcal{P}_t \cup \forall_{i=1}^{l+1} (S_t)^\gamma \mathcal{P}_t).
\]

(Recall from the Introduction that \(\gamma_i\) denotes that element of \(\Gamma\) having every coordinate except the \(i^{th}\) one equal to 0.)

We wish to show that \(h(S, \mathcal{P}_t) = h(S_t, \mathcal{P}_t)\). To that end, assume that \(n > l\) and that \(x, x' \in X\) have the same \((\mathcal{P}_t, S, n)\)-name. We prove that for \(l + 1 \leq i \leq n\), \((S_t)^\gamma_i(x)\) and \((S_t)^\gamma_i(x')\) are in the same atom of \(\mathcal{P}_t\). Fix such \(i\), and consider the following facts:

(a) Because \(x, x'\) are in the same atom of \(\mathcal{P}_t\), there exist \(x_0, x'_0 \in q_r \in \mathcal{Q}\) and a unique \(g \in \Gamma_t\) such that \(x_0 = (S_t)^g(x)\) and \(x'_0 = (S_t)^g(x')\).

(b) Since \(S^\gamma_i(x)\) and \(S^\gamma_i(x')\) are in the same atom of \(\mathcal{P}_t\) (and given that \(S \sim S_t\), there exist \(x_1, x'_1 \in q_r \in \mathcal{Q}\) and \(h, h' \in \Gamma_t\) such that \((S_t)^h(x_1) = (S_t)^\gamma_i(x)\) and \((S_t)^{h'}(x'_1) = (S_t)^\gamma_i(x')\).
(c) Since \( x \) and \( x' \) have the same \((P_l, S, n)\)-name there is a unique \( k \in \Gamma_l \) such that \( x_1 = S^k(x_0) \) and \( x'_1 = S^k(x'_0) \); but then, by construction, \( x_1 = (S_l)^k(x_0) \) and \( x'_1 = (S_l)^k(x'_0) \).

Now (a), (b) and (c) together give that \( (S_l)^{\gamma_l}(x) = (S_l)^{ghk}(x) \) and \( (S_l)^{\gamma_l}(x') = (S_l)^{ghk}(x') \). Thus, \( \gamma_l = ghk = gh'k \) so that \( h = h' \) and \( (S_l)^{\gamma_l}(x), (S_l)^{\gamma_l}(x') \in (S_l)^h q_5 \), an atom of \( P_l \). Similarly, if \( x, x' \) are in the same atom of \( P_l \) but have different \((P_l, S, n)\)-names then they have different \((P_l, S_l, n)\)-names. Combining these facts with (ii) we get that

\[
dist (\forall g \in \Gamma, S^g P_l) = dist (\forall g \in \Gamma, (S_l)^g P_l)
\]

and this certainly implies, because \( n > l \) is arbitrary, that \( h(S, P_l) = h(S_l, P_l) \).

We now use this construction to give the

\[\text{Proof of Theorem 2.} \text{ Let } \epsilon > 0, \text{ } P \text{ a finite partition of } X, \text{ and a positive integer } l \text{ be given. Construct } S_l \text{ and } P_l \text{ as in the preceding paragraph. Let } n > l \text{ be so large that at least } 1 - \epsilon \text{ in measure of } X \text{ can be covered with a collection of } (T, P_l, n)\text{-atoms whose cardinality is less than } 2^{2^n(h(T, P_l) + \epsilon)}. \text{ Let } A \text{ be one of the atoms in that collection. We claim that the number } a_n \text{ of } (S_l, P_l, n)\text{-names in } A \text{ is bounded by the solution to the recurrence relation (i) } a_n = a_{n-1}^2 2^{n-1}, n > l. \text{ We see this as follows: the value of } (S_l)^{\gamma_n}(x) \text{ and the } S_l \Gamma_{n-1}\text{-orbits of } x \text{ and } T^{\gamma_n}(x) \text{ unambiguously specify the } (S_l, P_l, n)\text{-name of } x. \text{ There are } 2^{n-1} \text{ choices for the first and } \leq a_{n-1} \text{ choices for each of the second and third, so the claim holds. Because } S_l \text{ and } T \text{ coincide on } \Gamma_l \text{ we have } a_l = 1. \text{ The solution to (i) subject to this initial condition is given by } a_n = 2^{b_n}, \text{ where }

\[
b_n = \sum_{k=1}^{n-l-1} k 2^{n-l-k-1} + (2^{n-l} - 1) l.
\]

It follows that at least \( 1 - \epsilon \) in measure of \( X \) can be covered with fewer than

\[
2^{2^n(h(T, P_l) + \epsilon)2b_n} = 2^{2^n(h(T, P_l) + \epsilon + \frac{1}{2l}(\sum_{k=1}^{n-l-1} k 2^{n-l-k-1} + (2^{n-l} - 1) l))}
\]

\((S_l, P_l, n)\)-atoms. But the last quantity is no larger than \( 2^{2^n(h(T, P_l) + \epsilon + f(l))} \) where

\[
f(l) = \frac{1}{2l} \left( \sum_{k=1}^{\infty} \frac{k}{2k+1} + l \right) = \frac{l + 1}{2l}.
\]

Thus, by the name-counting characterization of entropy,

\[
h(S, P_l) = h(S_l, P_l) \leq h(T, P_l) + f(l)
\]
where \( \lim_{l \to \infty} f(l) = 0 \). To finish the proof, let \( P_i \uparrow B \), an increasing generating sequence of finite partitions, and \( l_i \uparrow \infty \), an increasing sequence of positive integers, be given. Then

\[
h(S) = \lim_{i \to \infty} h(S, (P_i)_{l_i}) = \lim_{i \to \infty} h(S_{l_i}, (P_i)_{l_i}) \leq \lim_{i \to \infty} (h(T, (P_i)_{l_i}) + f(l_i)) = h(T)
\]

By symmetry \( h(T) \leq h(S) \) and the theorem follows. \( \square \)

We close by raising a couple of questions related to the ideas in this paper. First is the question of whether one can topologize the class of \( \Gamma \)-actions with a ‘size’ (see [R]) on orbit changes so that the dyadic equivalence class of an action \( T \) is the closure of the set of its coboundary changes. Secondly, and more along the lines of this paper, is the question of whether stronger statistical behavior exists in all appropriate equivalence classes, e.g. are there actions of completely positive entropy in all positive entropy equivalence classes?

Acknowledgment. The author thanks the referee for the very thorough examination of this paper. The detailed comments have greatly improved the exposition.

References


J. R. Hasfura-Buenaga, Department of Mathematics, Trinity University, San Antonio, TX 78212, U.S.A., e-mail: jhasfura@trinity.edu