

BERNOULLI SHIFTS WITH BASES OF EQUAL ENTROPY ARE ISOMORPHIC

BRANDON SEWARD

(Communicated by Lewis Bowen)

ABSTRACT. We prove that if G is a countably infinite group and (L, λ) and (K, κ) are probability spaces having equal Shannon entropy, then the Bernoulli shifts $G \curvearrowright (L^G, \lambda^G)$ and $G \curvearrowright (K^G, \kappa^G)$ are isomorphic. This extends Ornstein's famous isomorphism theorem to all countably infinite groups. Our proof builds on a slightly weaker theorem by Lewis Bowen in 2011 that required both λ and κ have at least 3 points in their support. We furthermore produce finitary isomorphisms in the case where both L and K are finite.

1. INTRODUCTION

Let G be a countably infinite group and let (L, λ) be a standard probability space. The *Bernoulli shift* over G with base (L, λ) is the probability space (L^G, λ^G) together with the left-shift action of G : for $g \in G$ and $x \in L^G$, $g \cdot x$ is defined by the rule $(g \cdot x)(t) = x(g^{-1}t)$ for all $t \in G$. Bernoulli shifts have been an important object of study since the very inception of ergodic theory by Birkhoff and von Neumann in the 1930s. In fact, in those early days von Neumann posed the question as to whether the Bernoulli shifts $\mathbb{Z} \curvearrowright (2^{\mathbb{Z}}, u_2^{\mathbb{Z}})$ and $\mathbb{Z} \curvearrowright (3^{\mathbb{Z}}, u_3^{\mathbb{Z}})$ are isomorphic (here n denotes $\{0, 1, \dots, n-1\}$ and u_n denotes the uniform probability measure on $\{0, 1, \dots, n-1\}$). This question turned out to be surprisingly difficult.

Today, a major open problem is to classify all Bernoulli shifts up to isomorphism for every countably infinite group. Progress on this problem has historically possessed two halves—on the one hand, using a notion of entropy to obtain the non-isomorphism of Bernoulli shifts of distinct entropies, and, on the other, constructing isomorphisms between Bernoulli shifts of equal entropy. Entropy was first developed by Kolmogorov [13, 14] (and corrected and improved by Sinai [26]) in 1958 for probability-measure-preserving (p.m.p.) actions of \mathbb{Z} . The Kolmogorov–Sinai entropy was extended to p.m.p. actions of countable amenable groups by Kieffer in 1975 [12]. In a major breakthrough in 2008, Bowen [2], together with improvements by Kerr and Li [9], developed the notion of sofic entropy for p.m.p. actions of sofic groups. Every countable

Received February 9, 2021; revised December 8, 2021.

2020 *Mathematics Subject Classification*: Primary: 37A35; Secondary: 28D20.

Key words and phrases: Bernoulli shift, isomorphism, finitary isomorphism, Ornstein, entropy, non-amenable.

The author was partially supported by ERC grant 306494 and Simons Foundation grant 328027 (PI. Tim Austin).

amenable group is sofic, and sofic entropy and Kolmogorov–Sinai entropy coincide for actions of countable amenable groups [3, 10].

For both Kolmogorov–Sinai entropy and sofic entropy, the entropy of Bernoulli shifts $G \curvearrowright (L^G, \lambda^G)$ for sofic G has been computed to be equal to the Shannon entropy $H(L, \lambda)$ of the base space (L, λ) [2, 11], where

$$H(L, \lambda) = \sum_{\ell \in L} -\lambda(\ell) \log \lambda(\ell)$$

if λ has countable support and $H(L, \lambda) = \infty$ otherwise. As entropy is an isomorphism invariant, it follows that $G \curvearrowright (L^G, \lambda^G)$ and $G \curvearrowright (K^G, \kappa^G)$ are non-isomorphic whenever G is sofic (so the sofic entropy is defined) and $H(L, \lambda) \neq H(K, \kappa)$. In particular, when G is a sofic group, such as \mathbb{Z} , an amenable group, a residually finite group, or a free group, $G \curvearrowright (2^G, u_2^G)$ and $G \curvearrowright (3^G, u_3^G)$ are non-isomorphic (indeed, it was the work of Kolmogorov that finally settled von Neumann’s question).

It’s not yet known if all countable groups are sofic. However, if non-sofic countable groups G exist, then it is still unknown whether $(2^G, u_2^G)$ and $(3^G, u_3^G)$ are isomorphic. One route to solving this problem may be to use Rokhlin entropy, which was introduced by the author in 2014 and is defined for p.m.p. actions of general countable groups [24]. Like sofic entropy, Rokhlin entropy also coincides with Kolmogorov–Sinai entropy for free actions of countable amenable groups [1]. For free actions of sofic groups, Rokhlin entropy is an upper-bound to sofic entropy and it is open whether the two coincide (excluding cases where the sofic entropy is minus infinity). The Rokhlin entropy of the Bernoulli shift $G \curvearrowright (L^G, \lambda^G)$ is $H(L, \lambda)$ when G is sofic but when G is not sofic its value is not yet known. The problem of computing the Rokhlin entropy of Bernoulli shifts over non-sofic groups has strong connections to Gottschalk’s surjunctivity conjecture¹ and Kaplansky’s direct finiteness conjecture [25].

Returning to the second half of the classification problem for Bernoulli shifts, the development of entropy raised the question whether Bernoulli shifts $G \curvearrowright (L^G, \lambda^G)$ and $G \curvearrowright (K^G, \kappa^G)$ are isomorphic whenever $H(L, \lambda) = H(K, \kappa)$. Initial positive results in special cases were obtained by Mešalkin, and further motivation was generated by Sinai’s factor theorem [27], which stated that Bernoulli shifts over \mathbb{Z} factor onto all Bernoulli shifts of lesser or equal entropy (so that in particular, Bernoulli shifts over \mathbb{Z} of equal entropy factor onto each other). The question was finally settled in the positive for \mathbb{Z} in 1970 by Ornstein, whose famous theorem has come to be known as “Ornstein’s isomorphism theorem” [19, 20]. In 1975, Stepin defined a group G to be *Ornstein* if $G \curvearrowright (L^G, \lambda^G)$ and $G \curvearrowright (K^G, \kappa^G)$ are isomorphic whenever $H(L, \lambda) = H(K, \kappa)$, and Stepin observed that any group containing an Ornstein subgroup is Ornstein [28]. Then in 1987,

¹Specifically, it was shown in [25] that for every finite set K , every countably infinite group G , and every G -invariant probability measure μ on K^G , if μ does not have full support and $G \curvearrowright (K^G, \mu)$ is essentially free, then the Rokhlin entropy of $G \curvearrowright (K^G, \mu)$ is strictly less than $\log|K|$. Thus Gottschalk’s surjunctivity conjecture will hold provided $G \curvearrowright (K^G, u_K^G)$ has Rokhlin entropy $\log|K|$, where u_K is the normalized counting measure on K .

Ornstein and Weiss proved that all countably infinite amenable groups are Ornstein [21]. Combined with Stepin's observation, it follows that any countable group containing \mathbb{Z} as a subgroup or containing an infinite amenable subgroup must be Ornstein. This covers the vast majority of groups that one tends to encounter, but excludes torsion groups not containing any infinite amenable subgroup (certain Tarski monsters [18] are examples of such groups; see the comment in [4] after Theorem 1.3).

In 2011, Lewis Bowen made a significant advancement by proving that if G is any countably infinite group and (L, λ) and (K, κ) are standard probability spaces with $H(L, \lambda) = H(K, \kappa)$, and one further assumes that the support of λ and the support of κ each have at least 3 points, then the Bernoulli shifts $G \curvearrowright (L^G, \lambda^G)$ and $G \curvearrowright (K^G, \kappa^G)$ are isomorphic [4]. In this paper, our main theorem removes Bowen's extra assumption and completes one-half of the classification problem for Bernoulli shifts. To use the terminology of Stepin, we prove that all countably infinite groups are Ornstein.

THEOREM 1.1. *Let G be a countably infinite group and let (L, λ) and (K, κ) be standard probability spaces. If $H(L, \lambda) = H(K, \kappa)$, then the Bernoulli shifts $G \curvearrowright (L^G, \lambda^G)$ and $G \curvearrowright (K^G, \kappa^G)$ are isomorphic.*

When combined with computations of sofic entropy [2, 11], this completes the classification of Bernoulli shifts up to isomorphism over sofic groups.

COROLLARY 1.2. *Let G be a countably infinite sofic group, and let (L, λ) and (K, κ) be standard probability spaces. Then the Bernoulli shifts $G \curvearrowright (L^G, \lambda^G)$ and $G \curvearrowright (K^G, \kappa^G)$ are isomorphic if and only if $H(L, \lambda) = H(K, \kappa)$.*

Recall that a measure-preserving map $\phi : (X, \mu) \rightarrow (Y, \nu)$ between two topological measure spaces (X, μ) and (Y, ν) is *finitary* if there is a conull set $X_0 \subseteq X$ such that the restriction of ϕ to X_0 is continuous with respect to the subspace topology. We say that (X, μ) and (Y, ν) are *finitarily isomorphic* if there is a measure-space isomorphism $\phi : (X, \mu) \rightarrow (Y, \nu)$ such that both ϕ and ϕ^{-1} are finitary. In case we are considering p.m.p. actions $G \curvearrowright (X, \mu)$ and $G \curvearrowright (Y, \nu)$, we further require that the map ϕ be G -equivariant.

Finitary isomorphisms between Bernoulli shifts $(L^{\mathbb{Z}}, \lambda^{\mathbb{Z}})$ and $(K^{\mathbb{Z}}, \kappa^{\mathbb{Z}})$ were constructed by Keane and Smorodinsky in 1979 when K and L are finite and $H(L, \lambda) = H(K, \kappa)$ [7], and were constructed by Petit in 1982 when L and K are countable and $H(L, \lambda) = \infty = H(K, \kappa)$ [22]. By modifying the proof of our main theorem, we generalize Keane–Smorodinsky and Petit theorems to all countably infinite groups.

THEOREM 1.3. *Let G be a countably infinite group and let (L, λ) and (K, κ) be standard probability spaces. Assume that either L and K are finite and $H(L, \lambda) = H(K, \kappa)$, or that L and K are countable and $H(L, \lambda) = \infty = H(K, \kappa)$. Then the Bernoulli shifts $G \curvearrowright (L^G, \lambda^G)$ and $G \curvearrowright (K^G, \kappa^G)$ are finitarily isomorphic.*

Outline. In Section 2 we show that it is sufficient to establish isomorphisms for pairs of Bernoulli shifts coming from certain specialized pairs of equal-entropy

base spaces. In Section 3 we prove the main theorem, and in Section 4 we construct isomorphisms that are finitary. Finally, in Section 5 we discuss the historical origins of the methods of our proof and in particular we articulate how our proof is different from and similar to the prior proof of Bowen that excluded two-atom base spaces.

2. REDUCTION TO A SPECIAL CASE

In this section we show that it is sufficient to produce isomorphisms between (L^G, λ^G) and (K^G, κ^G) for specific pairs of standard probability spaces (L, λ) and (K, κ) with $H(L, \lambda) = H(K, \kappa)$. Specifically, we will consider pairs that are related via the relation defined below.

DEFINITION 2.1. Fix a non-trivial finite group Γ . We define a symmetric relation R_Γ on the set of standard probability spaces as follows. For two standard probability spaces (L, λ) and (K, κ) we declare $(L, \lambda) R_\Gamma (K, \kappa)$ if

- (i) $H(L, \lambda) = H(K, \kappa)$;
- (ii) the sets L and K are not disjoint;
- (iii) there is a Borel Γ -invariant set $P \subseteq (L \cap K)^\Gamma$ satisfying the following:
 - (a) $\lambda^\Gamma \upharpoonright P = \kappa^\Gamma \upharpoonright P$ and $\lambda^\Gamma(P) > 0$;
 - (b) every element of P has trivial Γ -stabilizer.

We denote by $[R_\Gamma]^{\text{trans}}$ the equivalence relation obtained from the transitive closure of R_Γ . In other words, (L, λ) and (K, κ) are $[R_\Gamma]^{\text{trans}}$ -equivalent if and only if there is $n \in \mathbb{N}$ and probability spaces (M_i, μ_i) for $0 \leq i \leq n$ such that $(M_0, \mu_0) = (L, \lambda)$, $(M_n, \mu_n) = (K, \kappa)$, and $(M_i, \mu_i) R_\Gamma (M_{i+1}, \mu_{i+1})$ for every $0 \leq i < n$.

We will show in the next section that if $\Gamma \leq G$, then any pair of R_Γ -related probability spaces will produce isomorphic Bernoulli shifts over G . Since isomorphism is transitive, it will follow that any pair of $[R_\Gamma]^{\text{trans}}$ -equivalent probability spaces will produce isomorphic Bernoulli shifts. The purpose of this section is to prove the following proposition.

PROPOSITION 2.2. Fix a finite group Γ with $|\Gamma| \geq 5$. Then for any two standard probability space (L, λ) and (K, κ) we have $H(L, \lambda) = H(K, \kappa)$ if and only if $(L, \lambda) [R_\Gamma]^{\text{trans}} (K, \kappa)$.

We prove this in a few steps.

LEMMA 2.3. Let Γ be a finite group and let (L, λ) and (K, κ) be probability spaces with $H(L, \lambda) = H(K, \kappa)$. Suppose there are disjoint sets $A, B \subseteq L \cap K$ such that $\lambda \upharpoonright A \cup B = \kappa \upharpoonright A \cup B$ and $\lambda(A), \lambda(B) > 0$. Then $(L, \lambda) R_\Gamma (K, \kappa)$.

Proof. Set $P = \Gamma \cdot P_0$, where P_0 is the set of $x \in (L \cap K)^\Gamma$ such that $x(1_\Gamma) \in A$ and $x(\gamma) \in B$ for all $1_\Gamma \neq \gamma \in \Gamma$. Since A and B are disjoint, every point in P has trivial Γ -stabilizer. Also, $\lambda^\Gamma \upharpoonright P = \kappa^\Gamma \upharpoonright P$ since λ and κ agree on $A \cup B$. Thus all the conditions of Definition 2.1 are satisfied. \square

Below we write $\text{supp}(\lambda)$ for the support of λ . Also, for a countable Borel partition \mathcal{P} of (L, λ) we write $H_\lambda(\mathcal{P})$ for the Shannon entropy of \mathcal{P} with respect

to λ , i.e.,

$$H_\lambda(\mathcal{P}) = \sum_{P \in \mathcal{P}} -\lambda(P) \log(\lambda(P)).$$

Also, if $\bar{p} = (p_i)$ is a probability vector (i.e., a finite or countable tuple of non-negative real numbers that sum to 1), then we write $H(\bar{p}) = \sum_i -p_i \log(p_i)$.

LEMMA 2.4. *Let (L, λ) be a probability space with $|\text{supp}(\lambda)| \geq 4$. Then there is $\epsilon > 0$ so that for every $0 < \alpha, \beta < \epsilon$ there is a probability space (K, κ) that is R_Γ -related to (L, λ) and there are $a, b \in K$ with $\kappa(a) = \alpha$ and $\kappa(b) = \beta$.*

Proof. Since $|\text{supp}(\lambda)| \geq 4$ there exists a Borel partition $\{X, Y, M_1, M_2\}$ of L into four sets of positive measure. Set $M = M_1 \cup M_2$. Let λ_X, λ_Y , and λ_M denote the normalized restrictions of λ to X, Y , and M , respectively. Note that by construction $H(M, \lambda_M) > 0$. Fix $0 < \epsilon \leq \lambda(M)/2$ sufficiently small so that

$$(2.1) \quad H\left(\frac{\alpha}{\lambda(M)}, \frac{\beta}{\lambda(M)}, 1 - \frac{\alpha + \beta}{\lambda(M)}\right) \leq H(M, \lambda_M),$$

whenever $0 \leq \alpha, \beta < \epsilon$.

Now fix $0 < \alpha, \beta < \epsilon$. Then $\alpha + \beta < \lambda(M)$. Inequality (2.1) implies that we can construct a probability vector $\bar{p} = (p_0, p_1, \dots)$ with $p_0 = \alpha/\lambda(M)$, $p_1 = \beta/\lambda(M)$, and satisfying $H(\bar{p}) = H(M, \lambda_M)$. Set $K = X \cup Y \cup \mathbb{N}$ and define the probability measure κ by

$$\kappa = \lambda \upharpoonright (X \cup Y) + \lambda(M) \cdot \sum_{i \in \mathbb{N}} p_i \cdot \delta_i,$$

where δ_i is the single-point mass on i . Setting $a = 0$ and $b = 1$, we clearly have $\kappa(a) = \alpha$ and $\kappa(b) = \beta$. Also, by standard properties of Shannon entropy we have

$$\begin{aligned} H(K, \kappa) &= H_\kappa(\{X, Y, \mathbb{N}\}) + \lambda(X)H(X, \lambda_X) + \lambda(Y)H(Y, \lambda_Y) + \lambda(M)H(\bar{p}) \\ &= H_\lambda(\{X, Y, M\}) + \lambda(X)H(X, \lambda_X) + \lambda(Y)H(Y, \lambda_Y) + \lambda(M)H(M, \lambda_M) \\ &= H(L, \lambda). \end{aligned}$$

Since $\lambda \upharpoonright X \cup Y = \kappa \upharpoonright X \cup Y$ and $\lambda(X), \lambda(Y) > 0$, the previous lemma implies that (K, κ) is R_Γ -related to (L, λ) . □

The following is a technical lemma for handling the case of probability measures supported on only two or three points. We remind the reader that this is the key case to handle due to the prior work of Bowen [4].

LEMMA 2.5. *Let $\bar{p} = (p_0, p_1, \dots)$ be a probability vector with $p_0 \leq 1/2$ and $p_0 p_1 > 0$ and let $k \geq 4$ be any integer. Then there is a probability vector $\bar{r} = (r_0, r_1, \dots)$ having at least 4 strictly positive terms and satisfying $H(\bar{r}) = H(\bar{p})$ and $r_0^k \cdot r_1 = p_0^k \cdot p_1$.*

Proof. Set $r_0 = 1 - 2p_0^k$ and set

$$r_1 = \left(\frac{p_0}{r_0}\right)^k \cdot p_1 = \left(\frac{p_0}{1 - 2p_0^k}\right)^k \cdot p_1 \leq \left(\frac{p_0}{1 - 2p_0^k}\right)^k \cdot (1 - p_0).$$

Then $r_0^k \cdot r_1 = p_0^k \cdot p_1$ as required.

We claim that $r_0 + r_1 < 1$. As

$$r_0 + r_1 \leq 1 - 2p_0^k + \frac{p_0^k}{(1 - 2p_0^k)^k},$$

it suffices to show

$$1 - 2p_0^k + \frac{p_0^k}{(1 - 2p_0^k)^k} < 1,$$

or equivalently, $(1 - 2p_0^k)^k > \frac{1}{2}$. Since $p_0 \leq \frac{1}{2}$ it's enough to check $(1 - 2^{-k+1})^k > \frac{1}{2}$, or equivalently

$$2^{-1/k} + 2^{-k+1} < 1.$$

The derivative with respect to k of the expression on the left is

$$\ln(2)k^{-2}2^{-1/k} - \ln(2)2^{-k+1},$$

which is 0 when $k^2 2^{1/k} = 2^{k-1}$ or $k \approx 6.59$. So $2^{-1/k} + 2^{-k+1}$ is decreasing for $0 < k < 6.59$, and direct computation verifies that its value is below 1 when $k = 4$, and for $k > 6.59$ the function $2^{-1/k} + 2^{-k+1}$ is increasing to 1 as $k \rightarrow \infty$. Thus $2^{-1/k} + 2^{-k+1} < 1$ when $k \geq 4$, as required. We conclude that $r_0 + r_1 < 1$.

Next we claim that the probability vector $(r_0, r_1, 1 - r_0 - r_1)$ satisfies

$$H(r_0, r_1, 1 - r_0 - r_1) < H(p_0, 1 - p_0).$$

We have

$$\begin{aligned} H(r_0, r_1, 1 - r_0 - r_1) &= H(r_0, 1 - r_0) + (1 - r_0) \cdot H\left(\frac{r_1}{1 - r_0}, \frac{1 - r_0 - r_1}{1 - r_0}\right) \\ &\leq H(1 - 2p_0^k, 2p_0^k) + 2p_0^k \cdot \log(2). \end{aligned}$$

So it suffices to show that

$$H(1 - 2p_0^k, 2p_0^k) + 2p_0^k \cdot \log(2) < H(p_0, 1 - p_0).$$

Define $f : \{z \in \mathbb{C} : 0 < \operatorname{Re}(z) < 1\} \rightarrow \mathbb{C}$ by

$$f(z) = -z \log(z) - (1 - z) \log(1 - z),$$

where \log denotes the standard branch of the logarithm. Direct computation gives that the n^{th} derivative at $z = 1/2$ is

$$f^{(n)}(1/2) = \begin{cases} \log(2) & \text{if } n = 0 \\ 0 & \text{if } n \text{ is odd} \\ -2^n \cdot (n - 2)! & \text{if } n > 0 \text{ is even.} \end{cases}$$

The function f is a holomorphic function and thus for every z satisfying $|z - 1/2| < 1/2$ we have

$$\begin{aligned} f(z) &= \sum_{n=0}^{\infty} \frac{f^{(n)}(1/2)}{n!} \cdot (z - 1/2)^n \\ &= \log(2) - 2(z - 1/2)^2 + \sum_{m=2}^{\infty} \frac{-2^{2m}}{2m \cdot (2m - 1)} \cdot (z - 1/2)^{2m}. \end{aligned}$$

Plugging in $z = p_0$ and using the inequalities $k \geq 4$, $2p_0 \leq 1$, and $|p_0 - \frac{1}{2}| < |2p_0^k - \frac{1}{2}|$ we obtain

$$\begin{aligned} & H(p_0, 1 - p_0) \\ &= \log(2) - 2(p_0 - 1/2)^2 + \sum_{m=2}^{\infty} \frac{-2^{2m}}{2m \cdot (2m - 1)} \cdot (p_0 - 1/2)^{2m} \\ &= \log(2) - \frac{1}{2} + 2p_0 - 2p_0^2 + \sum_{m=2}^{\infty} \frac{-2^{2m}}{2m \cdot (2m - 1)} \cdot (p_0 - 1/2)^{2m} \\ &\geq \log(2) - \frac{1}{2} + p_0 + \sum_{m=2}^{\infty} \frac{-2^{2m}}{2m \cdot (2m - 1)} \cdot (p_0 - 1/2)^{2m} \\ &> \log(2) - \frac{1}{2} + 8p_0^k + \sum_{m=2}^{\infty} \frac{-2^{2m}}{2m \cdot (2m - 1)} \cdot (2p_0^k - 1/2)^{2m} \\ &> \log(2) - \frac{1}{2} + 4p_0^k - 8p_0^{2k} + 2p_0^k \cdot \log(2) + \sum_{m=2}^{\infty} \frac{-2^{2m}}{2m \cdot (2m - 1)} \cdot (2p_0^k - 1/2)^{2m} \\ &= \log(2) - 2(2p_0^k - 1/2)^2 + 2p_0^k \cdot \log(2) + \sum_{m=2}^{\infty} \frac{-2^{2m}}{2m \cdot (2m - 1)} \cdot (2p_0^k - 1/2)^{2m} \\ &= H(1 - 2p_0^k, 2p_0^k) + 2p_0^k \cdot \log(2), \end{aligned}$$

justifying our claim.

Finally, since $H(r_0, r_1, 1 - r_0 - r_1) < H(\bar{p})$, one can divide the remaining mass $1 - r_0 - r_1 > 0$ and extend (r_0, r_1) to a probability vector $\bar{r} = (r_i)$ having at least 4 strictly positive terms and satisfying $H(\bar{r}) = H(\bar{p})$. \square

Proof of Prop. 2.2. Fix a finite group Γ with $|\Gamma| \geq 5$. Let (L, λ) and (K, κ) be probability spaces with $H(L, \lambda) = H(K, \kappa)$. If $|\text{supp}(\lambda)|, |\text{supp}(\kappa)| \geq 4$, then pick $0 < \alpha, \beta < \min(\epsilon_L, \epsilon_K)$, where ϵ_L, ϵ_K come from applying Lemma 2.4 to (L, λ) and (K, κ) , respectively. Let (L', λ') be the resulting space R_Γ -related to (L, λ) , and let (K', κ') be the resulting space R_Γ -related to (K, κ) . Then $a, b \in L' \cap K'$, $\lambda'(a) = \kappa'(a) = \alpha > 0$, and $\lambda'(b) = \kappa'(b) = \beta > 0$, so (L', λ') and (K', κ') are R_Γ -related by Lemma 2.3. We conclude that $(L, \lambda) [R_\Gamma]^{\text{trans}} (K, \kappa)$ in this case. So it suffices to show that every non-trivial probability space (L, λ) is R_Γ -related to a probability space (M, μ) with $|\text{supp}(\mu)| \geq 4$.

So let's assume $|\text{supp}(\lambda)| < 4$, as otherwise we are done. Then λ is purely atomic and supported on either two or three points. So there are $\ell_0, \ell_1, \ell_2 \in L$ with $\lambda(\{\ell_0, \ell_1, \ell_2\}) = 1$. Set $\bar{p} = (p_0, p_1, p_2)$, where $p_i = \lambda(\ell_i)$. By rearranging indices we may assume that $p_0 \leq 1/2$ and $p_0 p_1 > 0$. Set $k = |\Gamma| - 1 \geq 4$. By Lemma 2.5, there is a probability vector $\bar{r} = (r_0, r_1, \dots)$ having at least 4 strictly positive terms and satisfying $r_0^k r_1 = p_0^k p_1$ and $H(\bar{r}) = H(\bar{p})$.

Define the probability space (M, μ) by setting $M = \{\ell_0, \ell_1, m_2, m_3, \dots\}$ and defining $\mu(\ell_0) = r_0$, $\mu(\ell_1) = r_1$, $\mu(m_i) = r_i$ ($i \geq 2$). Then $|\text{supp}(\mu)| \geq 4$ and

$$H(M, \mu) = H(\bar{r}) = H(\bar{p}) = H(L, \lambda).$$

Define

$$P = \{x \in \{\ell_0, \ell_1\}^\Gamma : \ell_1 \text{ occurs precisely once in } x\}.$$

Clearly each $x \in P$ has trivial Γ stabilizer and

$$\lambda^\Gamma(x) = \lambda(\ell_0)^{|\Gamma|-1} \lambda(\ell_1) = p_0^k p_1 = r_0^k r_1 = \mu(\ell_0)^{|\Gamma|-1} \mu(\ell_1) = \mu^\Gamma(x).$$

Thus $(M, \mu) R_\Gamma (L, \lambda)$. □

3. THE ISOMORPHISM THEOREM

In this section we prove the main theorem. We first note a group-theoretic fact that allows us, by the previous section, to significantly reduce the difficulty of the main theorem. We remark that, oddly, the lower-bound 5 in the following lemma happens to be important because we don't know how to prove Lemma 2.5, or a functionally equivalent lemma, without this lower-bound (recall that lemma was applied with $k = |\Gamma| - 1 \geq 4$).

LEMMA 3.1. *Let G be a countably infinite group. Then either G contains \mathbb{Z} as a subgroup or else there is a finite subgroup $\Gamma \leq G$ with $|\Gamma| \geq 5$.*

Proof. If any $g \in G$ has order at least 5 (or infinite) then we are done, as either $\langle g \rangle \cong \mathbb{Z}$ or else $\Gamma = \langle g \rangle$ is finite and has at least 5 elements. So assume that $|\langle g \rangle| \leq 4$ for all $g \in G$. Since every group element has order at most 4, it follows from work inspired by the Burnside problem² that every finitely generated subgroup of G must be finite [15, 16]. In particular, if $W \subseteq G$ is any set of 5 elements then $\Gamma = \langle W \rangle$ is a finite group of order at least 5. □

We are now ready to prove the main theorem.

THEOREM 3.2. *Let G be a countably infinite group. If (L, λ) and (K, κ) are standard probability spaces with $H(L, \lambda) = H(K, \kappa)$ then the corresponding Bernoulli shifts $G \curvearrowright (L^G, \lambda^G)$ and $G \curvearrowright (K^G, \kappa^G)$ are isomorphic.*

Proof. Stepin observed that the Ornstein isomorphism theorem for Bernoulli shifts over \mathbb{Z} extends immediately to Bernoulli shifts over any group containing \mathbb{Z} as a subgroup [28]. Thus we are done if $\mathbb{Z} \leq G$. So by Lemma 3.1 we can assume that G possesses a subgroup Γ with $5 \leq |\Gamma| < \infty$. Furthermore, since the relation of isomorphism is transitive, by Proposition 2.2 it will suffice to consider the case where $(L, \lambda) R_\Gamma (K, \kappa)$.

Let $P \subseteq (L \cap K)^\Gamma$ be as in Definition 2.1. In particular, $\lambda^\Gamma \upharpoonright P = \kappa^\Gamma \upharpoonright P$, $\lambda^\Gamma(P) > 0$, P is Γ -invariant, and every point in P has trivial Γ -stabilizer. If $\lambda^\Gamma(P) = 1$ then λ must have no atoms since every point in P has trivial Γ -stabilizer and Γ is finite. Thus when $\lambda^\Gamma(P) = 1$ we must have that (L, λ) and (K, κ) are already isomorphic and we are done. So we can assume $0 < \lambda^\Gamma(P) = \kappa^\Gamma(P) < 1$.

²Note that this does not follow from progress on the Burnside problem itself because we allow different group elements to have different orders, we just require that all orders be bounded by 4.

Fix a new symbol $*$ not in $L \cup K$, and set $M = (L \cap K) \cup \{*\}$. Define the G -equivariant map $\theta_L : L^G \rightarrow M^G$ by

$$\theta_L(x)(g) = \begin{cases} x(g) & \text{if } (g^{-1} \cdot x) \upharpoonright \Gamma \in P \\ * & \text{otherwise.} \end{cases}$$

Also define $\theta_K : K^G \rightarrow M^G$ by the exact same rule. One can easily check that θ_L and θ_K are G -equivariant.

We claim that these two maps push-forward λ^G and κ^G to the same measure, $(\theta_L)_*(\lambda^G) = (\theta_K)_*(\kappa^G)$. Indeed, let us explicitly describe this common measure. First define a probability measure μ_0 on M^Γ by

$$\mu_0 = \lambda^\Gamma \upharpoonright P + (1 - \lambda^\Gamma(P)) \cdot \delta_{*\Gamma} = \kappa^\Gamma \upharpoonright P + (1 - \kappa^\Gamma(P)) \cdot \delta_{*\Gamma},$$

where $*^\Gamma$ is the element of M^Γ that has constant value $*$, and $\delta_{*\Gamma}$ is the point-mass. Notice that μ_0 is Γ -invariant. Now for $g\Gamma \in G/\Gamma$ let $\mu_0^{g\Gamma}$ be the probability measure on $M^{g\Gamma} = \{z : g\Gamma \rightarrow M\}$ obtained by translating μ_0 , i.e., $\mu_0^{g\Gamma}(gA) = \mu_0(A)$. Since the maps θ_L and θ_K are determined by considering the labeling of each Γ -coset independently, and since λ^G and κ^G are i.i.d, it's straight-forward to see that

$$(\theta_L)_*(\lambda^G) = \prod_{g\Gamma \in G/\Gamma} \mu_0^{g\Gamma} = (\theta_K)_*(\kappa^G),$$

justifying our claim. Let's denote the common measure $\mu = \prod_{g\Gamma \in G/\Gamma} \mu_0^{g\Gamma}$.

We will define an isomorphism $\pi : (L^G, \lambda^G) \rightarrow (K^G, \kappa^G)$ that preserves the common factor (M^G, μ) , i.e., π will satisfy $\theta_K \circ \pi = \theta_L$. In other words, any occurrence of a Γ -coset whose labeling is in P will remain unchanged by π , while the labels of other Γ -cosets (the ones labeled with $*$'s in the factor (M^G, μ)) will be overwritten by π .

We claim that the action $G \curvearrowright (M^G, \mu)$ is essentially free, meaning that for every $1_G \neq s \in G$ the set of $z \in M^G$ satisfying $s \cdot z = z$ has μ -measure 0. Fix a group element $s \neq 1_G$ and fix $g \in G$. Suppose that $z \in M^G$ satisfies $(gs \cdot z) \upharpoonright \Gamma = (g \cdot z) \upharpoonright \Gamma$. If $gs = \gamma g$ with $\gamma \in \Gamma$, then it is necessary that $(g \cdot z) \upharpoonright \Gamma$ be fixed by $\gamma \neq 1_\Gamma$. Since each labeling in P has trivial stabilizer, it follows that $z \upharpoonright g^{-1}\Gamma$ must be the constant $*$ function. Alternatively, if $gs \notin \Gamma g$ then $s^{-1}g^{-1}\Gamma \neq g^{-1}\Gamma$ and the labeling of z on $g^{-1}\Gamma$ determines the labeling of z on $s^{-1}g^{-1}\Gamma$. Therefore the μ -probability that $z \in M^G$ satisfies $(gs \cdot z) \upharpoonright \Gamma = (g \cdot z) \upharpoonright \Gamma$ is bounded above by the maximum of $1 - \lambda^\Gamma(P)$ and $\mu_0 \times \mu_0(\{(r, r) : r \in M^\Gamma\})$, which is strictly less than 1 and independent of g . Notice that the functions $(gs \cdot z) \upharpoonright \Gamma$ and $(g \cdot z) \upharpoonright \Gamma$ are determined by $z \upharpoonright (s^{-1}g^{-1}\Gamma \cup g^{-1}\Gamma)$. Since it's easy to find an infinite collection g_i with the sets $s^{-1}g_i^{-1}\Gamma \cup g_i^{-1}\Gamma$ pairwise disjoint, and since the labelings on these sets are i.i.d. with respect to μ , it follows that for μ -almost-every $z \in M^G$ there is $g \in G$ with $(gs \cdot z) \upharpoonright \Gamma \neq (g \cdot z) \upharpoonright \Gamma$. This proves our claim.

Consider the set $\{z \in M^G : z(1_G) = *\}$. This set is Γ -invariant and has positive measure since $\lambda^\Gamma(P) < 1$. As Γ is finite, we can pick a Borel set

$$V \subseteq \{z \in M^G : z(1_G) = *\}$$

that meets every Γ -orbit in $\{z \in M^G : z(1_G) = *\}$ precisely once [8, Thm. 12.16]. Notice that the sets $\gamma \cdot V$, $\gamma \in \Gamma$, are pairwise-disjoint since G acts freely on (M^G, μ) . So M^G is partitioned by the sets $M^G \setminus \Gamma \cdot V$ and $\gamma \cdot V$, $\gamma \in \Gamma$, and we have

$$M^G \setminus \Gamma \cdot V = \{z \in M^G : z \upharpoonright \Gamma \in P\}.$$

Let $E_G^{M^G} = \{(z, g \cdot z) : z \in M^G, g \in G\}$ be the G -orbit-equivalence relation on M^G . Since μ is G -ergodic and $\mu(V) > 0$, the restricted equivalence relation $E_G^{M^G} \cap V \times V$ on V has infinite classes almost-everywhere. So we can fix an aperiodic element T in the full group $[E_G^{M^G} \cap V \times V]$ [6, Prop. 1.2 and Lem. 3.25] (i.e., $T : V \rightarrow V$ is a Borel bijection whose orbits are infinite almost-everywhere and for every $z \in V$ we have $(z, T(z)) \in E_G^{M^G}$). For each $z \in V$ define $t(z) \in G$ by the rule $t(z) \cdot z = T(z)$ (this is well-defined almost-everywhere).

Set $V_L = \theta_L^{-1}(V)$ and $V_K = \theta_K^{-1}(V)$. Similarly lift T to aperiodic transformations $T_L : V_L \rightarrow V_L$ and $T_K : V_K \rightarrow V_K$ by setting

$$T_L(x) = t \circ \theta_L(x) \cdot x \quad \text{and} \quad T_K(y) = t \circ \theta_K(y) \cdot y.$$

Notice that L^G is partitioned by the sets $L^G \setminus \Gamma \cdot V_L$ and $\gamma \cdot V_L$, $\gamma \in \Gamma$, and that $L^G \setminus \Gamma \cdot V_L = \{x \in L^G : x \upharpoonright \Gamma \in P\}$ (and similarly for K^G and V_K).

Define probability spaces (A, α) and (B, β) by setting $A = L^\Gamma \setminus P$, $B = K^\Gamma \setminus P$, and letting α and β be the normalized restrictions of λ^Γ to A and of κ^Γ to B , respectively. If we write $(\lambda^\Gamma)_P$ and $(\kappa^\Gamma)_P$ for the normalized restrictions of λ^Γ and κ^Γ to P , respectively, then we have

$$\begin{aligned} H(L^\Gamma, \lambda^\Gamma) &= H_{\lambda^\Gamma}(\{P, L^\Gamma \setminus P\}) + \lambda^\Gamma(P)H(P, (\lambda^\Gamma)_P) + (1 - \lambda^\Gamma(P))H(A, \alpha), \\ H(K^\Gamma, \kappa^\Gamma) &= H_{\kappa^\Gamma}(\{P, K^\Gamma \setminus P\}) + \kappa^\Gamma(P)H(P, (\kappa^\Gamma)_P) + (1 - \kappa^\Gamma(P))H(B, \beta). \end{aligned}$$

Our assumptions on P imply that $\lambda^\Gamma(P) = \kappa^\Gamma(P)$ and $(\lambda^\Gamma)_P = (\kappa^\Gamma)_P$. Additionally,

$$H(L^\Gamma, \lambda^\Gamma) = |\Gamma| \cdot H(L, \lambda) = |\Gamma| \cdot H(K, \kappa) = H(K^\Gamma, \kappa^\Gamma).$$

So it follows from these equations that $H(A, \alpha) = H(B, \beta)$. Therefore by Ornstein's original isomorphism theorem there is a \mathbb{Z} -equivariant isomorphism $\zeta : (A^\mathbb{Z}, \alpha^\mathbb{Z}) \rightarrow (B^\mathbb{Z}, \beta^\mathbb{Z})$ [19, 20].

We can now intuitively describe the isomorphism π . For each point $x \in L^G$, we leave the Γ -cosets that are labeled with a pattern from P unchanged. For each remaining coset $g\Gamma$ there corresponds a unique point in $(\Gamma g^{-1} \cdot x) \cap V_L$, and the transformation T_L arranges these distinguished points into various \mathbb{Z} -orbits. Each distinguished point's labeling of Γ is an element of A . So along the \mathbb{Z} -orbits created by T_L , we have bi-infinite sequences of elements from A which we can apply ζ to in order to obtain bi-infinite sequences of elements of B . Each element of $B = K^\Gamma \setminus P$ is a Γ -labeling, and the newly acquired B -labels of the distinguished points tell each distinguished point what its new labeling of Γ should be after applying π . We now proceed to define π explicitly and verify that it is an isomorphism.

First define $f_L : V_L \rightarrow A^\mathbb{Z}$ by

$$f_L(x)(n)(\gamma) = T_L^{-n}(x)(\gamma) \quad (n \in \mathbb{Z}, \gamma \in \Gamma),$$

and similarly define $f_K : V_K \rightarrow B^{\mathbb{Z}}$. Notice that f_L is equivariant for T_L and S :

$$(f_L \circ T_L(x))(n)(\gamma) = T_L^{1-n}(x)(\gamma) = f_L(x)(n-1)(\gamma) = (S \circ f_L(x))(n)(\gamma).$$

Similarly f_K is equivariant for T_K and S . Now define $\pi = \pi_L : (L^G, \lambda^G) \rightarrow (K^G, \kappa^G)$ by

$$\pi_L(x)(g) = \begin{cases} x(g) & \text{if } g^{-1} \cdot x \notin \Gamma \cdot V_L \\ \zeta \circ f_L(\gamma^{-1} g^{-1} \cdot x)(0)(\gamma^{-1}) & \text{if } g^{-1} \cdot x \in \gamma \cdot V_L, \gamma \in \Gamma. \end{cases}$$

We similarly define $\pi_K : (K^G, \kappa^G) \rightarrow (L^G, \lambda^G)$, replacing ζ with ζ^{-1} . It is easily checked that $\pi_L(x)(g) = \pi_L(g^{-1} \cdot x)(1_G)$ and therefore π_L (and π_K) is G -equivariant. Note that if $x \notin \Gamma \cdot V_L$ then $\pi_L(x) \upharpoonright \Gamma = x \upharpoonright \Gamma \in P$, while if $x \in \Gamma \cdot V_L$ then $x \upharpoonright \Gamma \in A = L^\Gamma \setminus P$ and $\pi_L(x) \upharpoonright \Gamma \in B = K^\Gamma \setminus P$ (and similarly if we swap L 's and K 's). In particular, $\theta_K \circ \pi_L = \theta_L$ and $\theta_L \circ \pi_K = \theta_K$ as originally desired.

It only remains to check that $\pi = \pi_L$ is an isomorphism. In fact we'll see that $\pi_L^{-1} = \pi_K$. The simplest way to do this is to check that π_L induces measure-space isomorphisms between the fibers over θ_L and the corresponding fibers over θ_K , and that fiber-wise $\pi_K = \pi_L^{-1}$. So let

$$\lambda^G = \int_{M^G} (\lambda^G)_z d\mu(z) \quad \text{and} \quad \kappa^G = \int_{M^G} (\kappa^G)_z d\mu(z)$$

be the disintegrations of λ^G and κ^G over μ , respectively, and fix $z \in M^G$. We will check that π_L induces an isomorphism between $(\lambda^G)_z$ and $(\kappa^G)_z$ with inverse π_K .

Due to the symmetry of information involved, we point out that anything we assert is true will remain true when L and K (and all of their associated objects and roles) are swapped. To enhance readability, in the remainder of the proof we will often avoid explicitly noting this.

Label the elements of the set $\{g \in G : g \cdot z \in V\}$ as w_n^i ($n \in \mathbb{Z}$, $i \in I$), where I is finite or countable and $T(w_n^i \cdot z) = w_{n+1}^i \cdot z$. Note that $(w_n^i \cdot z) \upharpoonright \Gamma$ is the constant $*$ function, while if $g \notin \Gamma \{w_n^i : n \in \mathbb{Z}, i \in I\}$ then $(g \cdot z) \upharpoonright \Gamma \in P$. When $g \notin \Gamma \{w_n^i : n \in \mathbb{Z}, i \in I\}$ we have $(g \cdot x) \upharpoonright \Gamma = (g \cdot z) \upharpoonright \Gamma$ for $(\lambda^G)_z$ -almost-every $x \in L^G$. Moreover, this agreement is preserved by π_L . So we only need to pay attention to the joint distributions under $(\lambda^G)_z$ of the labelings $(w_n^i \cdot x) \upharpoonright \Gamma$, $(n, i) \in \mathbb{Z} \times I$, and similarly for $(\kappa^G)_z$.

By construction, for $(\lambda^G)_z$ -almost-every $x \in L^G$, the w_n^i ($n \in \mathbb{Z}$, $i \in I$) are precisely those group elements g for which $g \cdot x \in V_L$, and we have $T_L(w_n^i \cdot x) = w_{n+1}^i \cdot x$. Since under $(\lambda^G)_z$ the labelings $(w_n^i \cdot x) \upharpoonright \Gamma$, $(n, i) \in \mathbb{Z} \times I$, are i.i.d. with distribution α , and since

$$f_L(w_0^i \cdot x)(m) = T_L^{-m}(w_0^i \cdot x) \upharpoonright \Gamma = (w_{-m}^i \cdot x) \upharpoonright \Gamma,$$

we have that, if $x \in L^G$ is a $(\lambda^G)_z$ -random point, then the random variables $f_L(w_0^i \cdot x)$, $i \in I$, and $\zeta \circ f_L(w_0^i \cdot x)$, $i \in I$, are i.i.d. with distribution $\alpha^{\mathbb{Z}}$ and i.i.d. with distribution $\beta^{\mathbb{Z}}$, respectively.

By inspecting the values $\pi_L(x)((w_n^i)^{-1}\gamma)$ in the definition of π_L , with $\gamma \in \Gamma$, we see that for $(\lambda^G)_z$ -almost-every $x \in L^G$

$$(w_n^i \cdot \pi_L(x)) \upharpoonright \Gamma = \zeta \circ f_L(w_n^i \cdot x)(0).$$

However, f_L is equivariant for (T_L, S) and ζ is S -equivariant, so we have

$$(3.1) \quad \begin{aligned} (w_n^i \cdot \pi_L(x)) \upharpoonright \Gamma &= \zeta \circ f_L \circ T_L^n(w_0^i \cdot x)(0) \\ &= S^n \circ \zeta \circ f_L(w_0^i \cdot x)(0) = \zeta \circ f_L(w_0^i \cdot x)(-n). \end{aligned}$$

So it follows from the previous paragraph that under $(\lambda^G)_z$ the labelings

$$(w_n^i \cdot \pi_L(x)) \upharpoonright \Gamma \in B$$

are i.i.d. with distribution β . This proves that π_L pushes $(\lambda^G)_z$ forward to $(\kappa^G)_z$.

Lastly we check that $\pi_K = \pi_L^{-1}$. From (3.1) we have

$$f_K(w_0^i \cdot \pi_L(x))(m) = T_K^{-m}(w_0^i \cdot \pi_L(x)) \upharpoonright \Gamma = (w_{-m}^i \cdot \pi_L(x)) \upharpoonright \Gamma = \zeta \circ f_L(w_0^i \cdot x)(m).$$

Therefore $f_K(w_0^i \cdot \pi_L(x)) = \zeta \circ f_L(w_0^i \cdot x)$. Now, by applying (3.1) with the roles of L and K swapped, with ζ^{-1} in place of ζ , and with $\pi_L(x)$ in place of x , we obtain

$$\begin{aligned} (w_n^i \cdot \pi_K \circ \pi_L(x)) \upharpoonright \Gamma &= \zeta^{-1} \circ f_K(w_0^i \cdot \pi_L(x))(-n) \\ &= f_L(w_0^i \cdot x)(-n) = T_L^n(w_0^i \cdot x) \upharpoonright \Gamma = (w_n^i \cdot x) \upharpoonright \Gamma \end{aligned}$$

We conclude that $\pi_K = \pi_L^{-1}$. □

4. FINITARY ISOMORPHISMS

We will modify the proof of our main theorem by applying the following two lemmas.

LEMMA 4.1. *The set V referred to in the proof of Theorem 3.2 can be chosen such that there is a G -invariant conull set $Z \subseteq M^G$ with $V \subseteq Z$ clopen in the subspace topology on Z .*

LEMMA 4.2. *Let V be as given by the previous lemma. The transformation $T : V \rightarrow V$ referred to in the proof of Theorem 3.2 can be chosen such that there is a G -invariant conull set $Z' \subseteq M^G$ and a continuous function $t : \mathbb{Z} \times (V \cap Z') \rightarrow G$ such that $t(n, v) \cdot v = T^n(v)$ for all $n \in \mathbb{Z}$ and all $v \in V \cap Z'$.*

Before we prove these lemmas, let's first use them to prove the finitary isomorphism theorem while the proof of Theorem 3.2 is still fresh in our minds. Recall that a measure-preserving map between topological measure spaces is finitary if there is a conull set on which the restricted function is continuous.

THEOREM 4.3. *Let G be a countably infinite group and let (L, λ) and (K, κ) be standard probability spaces. Assume that either L and K are finite and $H(L, \lambda) = H(K, \kappa)$, or that L and K are countable and $H(L, \lambda) = \infty = H(K, \kappa)$. Then the Bernoulli shifts $G \curvearrowright (L^G, \lambda^G)$ and $G \curvearrowright (K^G, \kappa^G)$ are finitarily isomorphic.*

Proof. The proof mostly follows the proof of Theorem 3.2 but with a few modifications. First suppose that $\mathbb{Z} \leq G$. We will detail, and modify, the observation of Stepin [28] we used before. Say $u \in G$ is of infinite order. Define $q : L^G \rightarrow L^{\mathbb{Z}}$ by $q(x)(n) = x(u^n)$. By applying the Keane–Smorodinsky theorem when L and K are finite [7], or applying Petit’s theorem when L and K are countable [22], we obtain a finitary isomorphism $\psi : (L^{\mathbb{Z}}, \lambda^{\mathbb{Z}}) \rightarrow (K^{\mathbb{Z}}, \kappa^{\mathbb{Z}})$. Now define $\pi : (L^G, \lambda^G) \rightarrow (K^G, \kappa^G)$ by

$$\pi(x)(g) = \psi \circ q(g^{-1} \cdot x)(0).$$

It is easily checked that π is G -equivariant, and since for $w \in G$ and $m \in \mathbb{Z}$ we have

$$\begin{aligned} \pi(x)(wu^m) &= \psi \circ q(u^{-m}w^{-1} \cdot x)(0) \\ &= S^{-m} \circ \psi \circ q(w^{-1} \cdot x)(0) = \psi \circ q(w^{-1} \cdot x)(m), \end{aligned}$$

it’s straight-forward to see that π pushes λ^G forward to κ^G . Using ψ^{-1} one can make an analogous definition to obtain the inverse to π . So π is an isomorphism. Finally, since q is continuous and ψ is finitary, it follows that π is finitary as well.

In the remaining case we assume that G has a subgroup Γ satisfying $5 \leq |\Gamma| < \infty$. Looking back at the proof of Proposition 2.2, our argument actually shows that when K and L are finite (respectively countable), (L, λ) and (K, κ) can be connected by a finite sequence of *finite* (respectively *countable*) probability spaces that are sequentially R_Γ -related. Since finitary isomorphism is transitive, it follows that we can once again assume that $(L, \lambda) R_\Gamma (K, \kappa)$. The remainder of the proof is the same as before except that we use Lemmas 4.1 and 4.2 to obtain V and T , and instead of obtaining $\zeta : (A^{\mathbb{Z}}, \alpha^{\mathbb{Z}}) \rightarrow (B^{\mathbb{Z}}, \beta^{\mathbb{Z}})$ from Ornstein’s isomorphism theorem, we instead obtain it from either the Keane–Smorodinsky theorem or Petit’s theorem so that it is finitary. We now must only check that the map π constructed as before is finitary.

Recall that the maps $\theta_L : (L^G, \lambda^G) \rightarrow (M^G, \mu)$, $f_L : V_L \rightarrow A^{\mathbb{Z}}$, and $\pi : (L^G, \lambda^G) \rightarrow (K^G, \kappa^G)$ were defined as:

$$\begin{aligned} \theta_L(x)(g) &= \begin{cases} x(g) & \text{if } (g^{-1} \cdot x) \upharpoonright \Gamma \in P \\ * & \text{otherwise} \end{cases} \\ f_L(x)(n) &= T_L^{-n}(x) \upharpoonright \Gamma \\ \pi(x)(g) &= \begin{cases} x(g) & \text{if } g^{-1} \cdot x \notin \Gamma \cdot V_L \\ \zeta \circ f_L(\gamma^{-1}g^{-1} \cdot x)(0)(\gamma^{-1}) & \text{if } g^{-1} \cdot x \in \gamma \cdot V_L, \gamma \in \Gamma. \end{cases} \end{aligned}$$

It is evident that θ_L is continuous. Therefore, since V comes from Lemma 4.1, there is a G -invariant conull set $X \subseteq L^G$ so that $V_L = \theta_L^{-1}(V) \subseteq X$ is relatively clopen in X . Similarly, as T comes from Lemma 4.2, there is a G -invariant conull set $X' \subseteq X$ such that the map $x \in V_L \cap X' \mapsto t(n, \theta_L(x)) \in G$ is continuous for every $n \in \mathbb{Z}$. Since for $x \in V_L \cap X'$ we have

$$f_L(x)(n) = T_L^{-n}(x) \upharpoonright \Gamma = (t(-n, \theta_L(x)) \cdot x) \upharpoonright \Gamma,$$

it follows that f_L is continuous on $V_L \cap X'$.

We saw in the proof of Theorem 3.2 that the normalized restriction of λ^G to V_L gets pushed forward by f_L to $\alpha^{\mathbb{Z}}$. Since ζ is finitary there is a conull \mathbb{Z} -invariant set $\Omega \subseteq A^{\mathbb{Z}}$ on which ζ is continuous. Now let X'' be the set of $x \in X' \subseteq L^G$ such that $f_L(g \cdot x) \in \Omega$ whenever $g \cdot x \in V_L$. Then X'' is G -invariant and conull, and $\zeta \circ f_L$ is continuous on $X'' \cap V_L$. It follows that π is continuous on X'' and thus π is finitary. \square

We now turn to the proofs of Lemma 4.1 and 4.2. We will rely upon a well-known marker lemma that has appeared in various works. The version we state below comes from [23, Lem. 3.1] which itself is a mild generalization of [5, Lem. 2.1].

LEMMA 4.4. *Let $G \curvearrowright X$ be a free continuous action on a zero-dimensional space X , let $Y \subseteq X$ be clopen, and let $F \subseteq G$ be finite with $1_G \in F$. Then there is a clopen set $D \subseteq Y$ such that*

- (i) $Fd \cap D = \{d\}$ for all $d \in D$; and
- (ii) $Y \subseteq F \cdot D$.

We can now prove the two lemmas from the start of the section.

Proof of Lemma 4.1. Let $G \curvearrowright (M^G, \mu)$ be the factor action constructed in the proof of Theorem 3.2. We argued in that proof that this action is essentially free. So let $Z \subseteq M^G$ be a G -invariant conull set on which the action is free everywhere, not just almost-everywhere. Since M^G is zero-dimensional and the action of G on M^G is continuous, we get that Z is zero-dimensional and $G \curvearrowright Z$ is continuous. Set $Y = \{z \in Z : z(1_G) = *\}$ and set $F = \Gamma$ and apply Lemma 4.4 to get a relatively clopen set $V = D \subseteq Y$. Clause (i) of that lemma implies that V meets every Γ -orbit in Y at most once, and clause (ii) implies that V meets every Γ -orbit in Y at least once. Thus V meets every Γ -orbit in $\{z \in Z : z(1_G) = *\}$ precisely once. So V has the properties required of it in the proof of Theorem 3.2 and it is relatively clopen as a subset of the G -invariant conull set Z . \square

Proof of Lemma 4.2. Consider for the moment a free continuous action $G \curvearrowright X$. Let's call a function ϕ *cocycle-continuous* if $\text{dom}(\phi) \subseteq X$ and there is a continuous function $\bar{\phi} : \text{dom}(\phi) \rightarrow G$ satisfying $\phi(z) = \bar{\phi}(z) \cdot z$ for all $z \in \text{dom}(\phi)$. Notice that, since G acts continuously, cocycle-continuous functions are continuous. Additionally, if $\text{dom}(\phi)$ is clopen then ϕ is both an open and a closed map (in particular, the image under ϕ of a clopen set is clopen). Also notice that if ϕ and ψ are both cocycle-continuous maps and $\text{rng}(\phi) \subseteq \text{dom}(\psi)$ then $\psi \circ \phi$ is cocycle-continuous as well.

As in the previous proof, let $Z \subseteq M^G$ be a G -invariant conull set on which the action is free. Let $V \subseteq Z$ be the set constructed in Lemma 4.1. Using the language of the previous paragraph, our goal is to find a G -invariant conull set $Z' \subseteq Z$ and build a bijection $T : V \cap Z' \rightarrow V \cap Z'$ having infinite orbits such that every power of T and every power of T^{-1} is cocycle-continuous. By the previous paragraph, it suffices to check that T and T^{-1} are cocycle-continuous.

Set $D_0 = V$ and $F_0 = \{1_G\}$. Fix an increasing sequence $(F_n)_{n \geq 1}$ of finite subsets of G containing the identity and such that, for every n , F_{n+1} contains two disjoint translates of $F_n^{-1}F_n$. For each $n \geq 1$, apply Lemma 4.4 to obtain a Z -relatively clopen set $D_n \subseteq V$ satisfying $F_n^{-1}F_n d \cap D_n = \{d\}$ for all $d \in D_n$ and $V \subseteq F_n^{-1}F_n D_n$. Notice that $F_n d \cap F_n d' = \emptyset$ for $d \neq d' \in D_n$. Also notice that since $F_n^{-1}F_n \cdot v \cap D_n \neq \emptyset$ for all $v \in V$ and since F_{n+1} contains two disjoint translates of $F_n^{-1}F_n$, it follows that $|D_n \cap F_{n+1} d| \geq 2$ for every $d \in D_{n+1}$ and every n .

Fix an enumeration $1_G = g_0, g_1, \dots$ of G . For each $n \in \mathbb{N}$ we define a cocycle-continuous map $\ell_n^{n+1} : D_n \rightarrow D_{n+1}$. Fix $d \in D_n$. First, if there is $f \in F_{n+1}$ with $f^{-1} \cdot d \in D_{n+1}$ (by construction there can be at most one such f) then set $\ell_n^{n+1}(d) = f^{-1} \cdot d$. Otherwise, set $\ell_n^{n+1}(d) = g_i \cdot d$ where i is least with $g_i \in F_{n+1}^{-1}F_{n+1}$ and $g_i \cdot d \in D_{n+1}$. Since D_n is contained in the G -invariant conull set Z , since each D_{n+1} is relatively clopen in Z , and since G acts continuously, it's immediate that ℓ_n^{n+1} is cocycle-continuous. Also, notice that by the last statement of the previous paragraph, ℓ_n^{n+1} is everywhere at least two-to-one.

Through the function ℓ_n^{n+1} we view each point in D_n as belonging to some point in D_{n+1} . Set $\ell_0^n = \ell_{n-1}^n \circ \ell_{n-2}^{n-1} \circ \dots \circ \ell_0^1$. Define equivalence relations E_n on $V = D_0$ by setting $(v, v') \in E_n$ if and only if $\ell_0^n(v) = \ell_0^n(v')$. Notice that every class of E_n possesses precisely one point from D_n . The equivalence relations E_n are Borel, have finite classes, and are increasing, so the union $E = \bigcup_n E_n$ is hyperfinite. Moreover, since ℓ_n^{n+1} is everywhere at least two-to-one, each class of E_n has cardinality at least 2^n , and each class of E is infinite.

Let us intuitively describe how we will construct T . On each class of E_1 we will choose an ordering and let T move each point to the next one in the ordering. This will leave each class with a “first” element where T^{-1} is not defined, and a “last” element where T is not defined. Inductively, each class of E_n will have a first and a last point, and repeated application of T will move the first point through all of its E_n -class until it reaches the last point. Each E_n class has a unique point in D_n , and the function ℓ_n^{n+1} indicates how the E_n -classes will be combined into E_{n+1} -classes, indexed by elements of D_{n+1} . For each $d \in D_{n+1}$ we will order the D_n points belonging to d , in effect ordering the E_n -classes in the E_{n+1} -class of d . T will then take the last point of an E_n -class to the first point of the next E_n -class. This will only leave T undefined at the last point of the last E_n -class (i.e., the last point of the E_{n+1} -class of d), and T^{-1} undefined at the first point of the first E_n -class (i.e., the first point of the E_{n+1} -class of d). We then go up one scale and repeat this construction by induction.

For $n \geq 1$ and $d \in D_n$ set $W_n(d) = \{g \in F_n^{-1}F_n : g \cdot d \in D_{n-1} \text{ and } \ell_{n-1}^n(g \cdot d) = d\}$. Note that $W_n(d)$ is a continuous function of d . Using our fixed enumeration of G we obtain an ordering of the set $W_n(d)$. Set $a_n^{n-1}(d) = g \cdot d \in D_{n-1}$ where g is the least element of $W_n(d)$, and set $b_n^{n-1}(d) = g \cdot d \in D_{n-1}$ where g is the largest element of $W_n(d)$. If $g, h \in W_n(d)$ and h is the next element in $W_n(d)$ after g , then set $s_n(g \cdot d) = h \cdot d$. So s_n is a function from $D_{n-1} \setminus a_n^{n-1}(D_n)$ to $D_{n-1} \setminus b_n^{n-1}(D_n)$. Note that a_n^{n-1} , b_n^{n-1} , and s_n are cocycle-continuous functions

and that, by remarks in the first paragraph, $a_n^{n-1}(D_n)$ and $b_n^{n-1}(D_n)$ are relatively clopen in Z .

Set $a_n^0 = a_1^0 \circ a_2^1 \circ \dots \circ a_n^{n-1}$ and $b_n^0 = b_1^0 \circ \dots \circ b_n^{n-1}$ and define $A_n = a_n^0(D_n)$ and $B_n = b_n^0(D_n)$ for $n \geq 1$. Set $B_0 = A_0 = V$. Notice that the A_n 's are decreasing, the B_n 's are decreasing, and that all of these sets are relatively clopen in Z . Now for $v \in B_{n-1} \setminus B_n$ set

$$T(v) = a_{n-1}^0(s_n(\ell_0^{n-1}(v))) \in A_{n-1} \setminus A_n.$$

Equivalently, $T(v) = a_{n-1}^0 \circ s_n \circ (b_{n-1}^0)^{-1}(v)$. In words, if v is the last point of its E_{n-1} -class but not the last point of its E_n -class, then we lift via ℓ_0^{n-1} to the representative $d' \in D_{n-1}$ of its E_{n-1} -class, we then apply s_n to shift to the representative $d'' \in D_{n-1}$ of the next E_{n-1} -class in its E_n -class, and then we apply a_{n-1}^0 to travel down from d'' to the first point in its E_{n-1} -class. The formula for the inverse to T is similarly described: when $v \in A_{n-1} \setminus A_n$ we have

$$T^{-1}(v) = b_{n-1}^0(s_n^{-1}(\ell_0^{n-1}(v))) = b_{n-1}^0 \circ s_n^{-1} \circ (a_{n-1}^0)^{-1}(v).$$

Since T is a composition of cocycle-continuous functions it is cocycle-continuous on its domain, and we see directly from the formula for T^{-1} that it is cocycle continuous on its domain as well. Lastly, we inspect the domain of T . The set B_n meets every E_n -class precisely once, and since each E_n class has cardinality at least 2^n , it follows that $\mu(B_n) \leq 2^{-n}$. The sets B_n are decreasing so their intersection has measure 0. Similarly the intersection of the A_n 's is null. Setting $Z' = Z \setminus G \cdot ((\bigcap_n B_n) \cup (\bigcap_n A_n))$, we have that Z' is G -invariant, conull in M^G , and that T is a bijection on $V \cap Z'$. \square

5. A COMPARATIVE DISCUSSION OF THE PROOF

Given that a weaker, but quite similar, version of our main theorem was proven by Bowen a few years ago, we feel it's proper to openly identify how our proof both builds upon and differs from Bowen's proof. We'll also take this opportunity to more generally acknowledge the historical roots of some of the methods involved. Specifically, beyond Ornstein's isomorphism theorem, which is a highly celebrated, deep, and very technical achievement, the proof of our main theorem involves two additional main ingredients.

One is to abstractly apply Ornstein's theorem (or a variation of it). This trick was first used by Stepin in 1975 in the case of groups G that contain \mathbb{Z} as a subgroup. One of Bowen's main contributions was to do this in a more abstract way by using a copy of \mathbb{Z} in the full-group of the orbit-equivalence relation rather than a copy of \mathbb{Z} in G . We too used this technique (though our copy of \mathbb{Z} was in the pseudo-group, not the full-group; it was generated by the transformation $T : V \rightarrow V$). We remark that in order for this technique to work, it is necessary that (L^G, λ^G) and (K^G, κ^G) admit a non-trivial common factor and that they each have strong independence properties over this common factor.

The second main ingredient is the reduction to considering only special pairs of equal-entropy probability spaces (like we did with R_T). This clever reduction was first used in Keane and Smorodinsky's proof of the finitary isomorphism

theorem for Bernoulli shifts over \mathbb{Z} in 1979, and it was used by Bowen and again by us. We remark that for Bowen and ourselves, the choice of special pairs had to be carefully made in conjunction with the construction of common factors with strong independence properties, as required for the method described in the previous paragraph.

The biggest difference between our proof and Bowen's, and the source of all the smaller differences, is in this last ingredient. In contrast with our relation R_Γ , Bowen considered pairs (L, λ) and (K, κ) such that $H(L, \lambda) = H(K, \kappa)$ and such that there exists a third space (M, μ) that is non-trivial and admits measure-preserving maps $(L, \lambda) \rightarrow (M, \mu)$ and $(K, \kappa) \rightarrow (M, \mu)$. By applying these maps coordinate-wise, one then sees that (M^G, μ^G) is a common factor of (L^G, λ^G) and (K^G, κ^G) , just as we similarly constructed a common factor, and just as is required as we mentioned above. However, it is evident that this requirement that the base spaces have a non-trivial factor immediately rules out the possibility of considering two-atom base spaces. Our main contribution here, therefore (ignoring the result on finitary isomorphisms), is in the creation of the relations R_Γ , the use of lesser-known group theoretic facts (Lemma 3.1), and in particular the technical proof of Lemma 2.5.

Lastly, differences in the construction of the common factor M^G led Bowen to use a transformation $T : M^G \rightarrow M^G$ on all of M^G , while we used a transformation on a subset $T : V \rightarrow V$. Additionally, due to differences in the independence properties present, Bowen had to invoke Thouvenot's relative isomorphism theorem and apply it to the transformations lifted from T , while our independence properties were cleaner and we were able to simply apply Ornstein's isomorphism theorem to the lifted transformations. As an unintended bonus, our cleaner independence properties furthermore allowed us to invoke the Keane–Smorodinsky and Petit finitary isomorphism results and obtain finitary isomorphisms in general.

REFERENCES

- [1] A. Alpeev and B. Seward, [Krieger's finite generator theorem for ergodic actions of countable groups III](#), *Ergodic Theory Dynam. Systems*, **41** (2021), 2881–2917.
- [2] L. Bowen, [Measure conjugacy invariants for actions of countable sofic groups](#), *J. Amer. Math. Soc.*, **23** (2010), 217–245.
- [3] L. Bowen, [Sofic entropy and amenable groups](#), *Ergodic Theory Dynam. Systems*, **32** (2012), 427–466.
- [4] L. Bowen, [Every countably infinite group is almost Ornstein](#), *Dynamical Systems and Group Actions*, 67–78, Contemp. Math., 567, Amer. Math. Soc., Providence, RI, 2012.
- [5] S. Gao and S. Jackson, [Countable abelian group actions and hyperfinite equivalence relations](#), *Invent. Math.*, **201** (2015), 309–383.
- [6] S. Jackson, A. S. Kechris and A. Louveau, [Countable Borel equivalence relations](#), *J. Math. Logic*, **2** (2002), 1–80.
- [7] M. Keane and M. Smorodinsky, [Bernoulli schemes of the same entropy are finitarily isomorphic](#), *Ann. of Math. (2)*, **109** (1979), 397–406.
- [8] A. Kechris, *Classical Descriptive Set Theory*, Springer-Verlag, New York, 1995.
- [9] D. Kerr and H. Li, [Entropy and the variational principle for actions of sofic groups](#), *Invent. Math.*, **186** (2011), 501–558.

- [10] D. Kerr and H. Li, Soficity, amenability, and dynamical entropy, *Amer. J. Math.*, **135** (2013), 721–761.
- [11] D. Kerr and H. Li, Bernoulli actions and infinite entropy, *Groups Geom. Dyn.*, **5** (2011), 663–672.
- [12] J. C. Kieffer, A generalized Shannon–McMillan theorem for the action of an amenable group on a probability space, *Ann. Prob.*, **3** (1975), 1031–1037.
- [13] A. N. Kolmogorov, New metric invariant of transitive dynamical systems and endomorphisms of Lebesgue spaces, (Russian) *Dokl. Akad. Nauk SSSR*, **119** (1958), 861–864.
- [14] A. N. Kolmogorov, Entropy per unit time as a metric invariant for automorphisms, (Russian) *Dokl. Akad. Nauk SSSR*, **124** (1959), 754–755.
- [15] D. V. Lytkina, Structure of a group with elements of order at most 4, *Siberian Math. J.*, **48** (2007), 283–287.
- [16] A. S. Mamontov, Groups of exponent 12 without elements of order 12, *Siberian Math. J.*, **54** (2013), 114–118.
- [17] L. D. Mešalkin, A case of isomorphism of Bernoulli schemes, *Dokl. Akad. Nauk SSSR*, **128** (1959), 41–44.
- [18] A. Yu. Olshanskii, An infinite group with subgroups of prime orders, *Izv. Akad. Nauk SSSR Ser. Mat.*, **44** (1980), 309–321.
- [19] D. Ornstein, Bernoulli shifts with the same entropy are isomorphic, *Advances in Math.*, **4** (1970), 337–348.
- [20] D. Ornstein, Two Bernoulli shifts with infinite entropy are isomorphic, *Advances in Math.*, **5** (1970), 339–348.
- [21] D. Ornstein and B. Weiss, Entropy and isomorphism theorems for actions of amenable groups, *J. Analyse Math.*, **48** (1987), 1–141.
- [22] B. Petit, Deux schémas de Bernoulli d’alphabet dénombrable et d’entropie infinie sont finitairement isomorphes, *Z. Wahrsch. Verw. Gebiete*, **59** (1982), 161–168.
- [23] S. Schneider and B. Seward, Locally nilpotent groups and hyperfinite equivalence relations, preprint, [arXiv:1308.5853](https://arxiv.org/abs/1308.5853), 2013.
- [24] B. Seward, Krieger’s finite generator theorem for ergodic actions of countable groups I, *Invent. Math.*, **215** (2019), 265–310.
- [25] B. Seward, Krieger’s finite generator theorem for ergodic actions of countable groups II, *J. Mod. Dyn.*, **15** (2019), 1–39.
- [26] Ya. G. Sinai, On the concept of entropy for a dynamical system, (Russian) *Dokl. Akad. Nauk SSSR*, **124** (1959), 768–771.
- [27] Ya. G. Sinai, A weak isomorphism of transformations with invariant measure, (Russian) *Dokl. Akad. Nauk SSSR*, **147** (1962), 797–800.
- [28] A. M. Stepin, Bernoulli shifts on groups, (Russian) *Dokl. Akad. Nauk SSSR*, **223** (1975), 300–302.

BRANDON SEWARD <b.m.seward@gmail.com>: Department of Mathematics, University of California San Diego, 9500 Gilman Dr #0112, La Jolla, CA, 92093, USA