CHOQUET-DENY GROUPS AND THE INFINITE CONJUGACY CLASS PROPERTY

JOSHUA FRISCH, YAIR HARTMAN, OMER TAMUZ, AND POOYA VAHIDI FERDOWSI

Abstract. A countable discrete group G is called Choquet-Deny if for every non-degenerate probability measure μ on G it holds that all bounded μ-harmonic functions are constant. We show that a finitely generated group G is Choquet-Deny if and only if it is virtually nilpotent. For general countable discrete groups, we show that G is Choquet-Deny if and only if none of its quotients has the infinite conjugacy class property. Moreover, when G is not Choquet-Deny, then this is witnessed by a symmetric, finite entropy, non-degenerate measure.

1. Introduction

Let G be a countable discrete group. A probability measure μ on G is *non-degenerate* if its support generates G as a semigroup.[1] A function $f\colon G \to \mathbb{R}$ is *μ-harmonic* if $f(k) = \sum_{g\in G} \mu(g)f(kg)$ for all $k \in G$. We say that the *measured group* (G,μ) is *Liouville* if all the bounded μ-harmonic functions are constant; this is equivalent to the triviality of the Poisson boundary $\Pi(G,\mu)$ [12–14] (also called the Furstenberg-Poisson boundary; for formal definitions see also, e.g., Furstenberg and Glasner [11], Bader and Shalom [1], or a survey by Furman [10]).

When G is non-amenable, (G,μ) is not Liouville for every non-degenerate μ [14]. Conversely, when G is amenable, then there exists some non-degenerate μ such that (G,μ) is Liouville, as shown by Kaimanovich and Vershik [21] and Rosenblatt [24]. It is natural to ask for which groups G it holds that (G,μ) is Liouville for *every* non-degenerate μ. We call such groups *Choquet-Deny* groups; as we discuss in §1.1, there are a few variants of this definition (see, e.g., [15–17], or [19]), which, however, we show to be equivalent.

The classical Choquet-Deny Theorem (which was first proved for $\mathbb{Z}^d$ by Blackwell [3]) states that abelian groups are Choquet-Deny [4]; the same holds for virtually nilpotent groups [6]. There are many examples

Date: February 21, 2019.

[1] In the context of Markov chains such measures are called *irreducible*.

of amenable groups that are not Choquet-Deny: first examples of such groups[2] are due to Kaimanovich [20] and Kaimanovich and Vershik [21], they include locally finite groups; Erschler shows that finitely generated solvable groups that are not virtually nilpotent are not Choquet-Deny [8], and that even some groups of intermediate growth are not Choquet-Deny [7]. Kaimanovich and Vershik [21, p. 466] conjecture that "Given an exponential group G, there exists a symmetric (non-finitary, in general) measure with non-trivial boundary." See Bartholdi and Erschler [2] for additional related results and further references and discussion.

Our main result is a characterization of Choquet-Deny groups. We say that G has the *infinite conjugacy class* property (ICC) if it is non-trivial, and if each of its non-trivial elements has an infinite conjugacy class. We say that μ is *fully supported* if $\operatorname{supp}\mu = G$; obviously this implies that μ is non-degenerate.

Theorem 1. *A countable discrete group G is Choquet-Deny if and only if it has no ICC quotients. Moreover, when G does have an ICC quotient, then there exists a fully supported, symmetric, finite entropy probability measure μ on G such that (G, μ) is not Liouville. In particular, if G is finitely generated, then it is Choquet-Deny if and only if it is virtually nilpotent.*

That a group with no ICC quotients is Choquet-Deny was shown by Jaworski [18, Theorem 4.8].[3] Our contribution is therefore in the proof of the converse, which appears in §2.

Groups with no ICC quotients are known as FC-hypercentral (see, e.g., [5, 22], or [23, §4.3]). This class is closed under forming subgroups, quotients, direct products and finite index extensions, and includes all virtually nilpotent groups. Among finitely generated groups, virtually nilpotent groups are precisely those with no ICC quotients (see [22, Theorem 2] and [5, Theorem 2]); this implies the result in Theorem 1 for finitely generated groups. Since finitely generated groups of exponential growth are not virtually nilpotent, Theorem 1 implies that the above mentioned conjecture of Kaimanovich and Vershik [21] is correct.

A very recent result by three of the authors of this paper shows that a countable discrete group is strongly amenable if and only if it has no ICC quotients [9]. This implies that G is strongly amenable if and only if (G, μ) is Liouville for every non-degenerate μ, paralleling the

[2]In the Lie group setting, an example of an amenable group that is not Choquet-Deny was already known to Furstenberg [12].

[3]In fact, Jaworski proves there a stronger statement; see the discussion in §1.1.

above mentioned characterization of amenability as equivalent to the existence of a non-degenerate μ such that (G, μ) is Liouville. While the proofs of these two similar results are different, it is natural to ask whether there is some deeper connection between strong amenability and the Choquet-Deny property.

1.1. **Different possible definitions of Choquet-Deny groups.** Our definition of Choquet-Deny groups is not the usual one, which states that a group is Choquet-Deny if (G, μ) is Liouville for every *adapted* measure μ, where μ is called adapted if its support generates G as a *group* (rather than as a semigroup, as in the non-degenerate case) [15–17]. Yet another definition used in the literature requires that for *every* μ, every bounded μ-harmonic function is constant on the left cosets of G_μ, where G_μ is the subgroup of G generated by the support of μ [19].

While a priori these are different definitions, they are equivalent, as demonstrated by our result and by Jaworski's Theorem 4.8 in [18]. Jaworski's result shows that groups with no ICC quotients are Choquet-Deny according to any of these definitions. Since our construction of μ with a non-trivial boundary yields measures that are supported on all of G (hence non-degenerate, hence adapted), it shows that groups with ICC quotients are not Choquet-Deny according to any of these definitions. Moreover, our result shows that the class of Choquet-Deny groups (whether defined with adapted or with non-degenerate measures) is closed under taking subgroups, which, to the best of our knowledge, was also not previously known.

Acknowledgments. We would like to thank Anna Erschler and Vadim Kaimanovich for many useful comments on the first draft of this paper. We thank Wojciech Jaworski for bringing a number of errors to our attention and suggesting many improvements. We likewise thank an anonymous referee for many helpful suggestions.

2. Proofs

In this section we prove the main result of our paper, Theorem 1. Unless stated otherwise, we will assume that all groups are countable and discrete.

Recall that a probability measure μ on G is symmetric if $\mu(g) = \mu(g^{-1})$ for all $g \in G$. Its Shannon entropy (or just entropy) is $H(\mu) = -\sum_{g\in G} \mu(g) \log \mu(g)$.

Our Theorem 1 is a direct consequence of [18, Theorem 4.8], which proves it for the case of groups with no ICC quotients, and of the following proposition, which handles the case of groups with ICC quotients.

Proposition 2.1. *Let G be a group with an ICC quotient. Then there exists a fully-supported, symmetric, finite entropy probability measure μ on G such that $\Pi(G,\mu)$ is non-trivial.*

The main technical effort in the proof of Proposition 2.1 is in the proof of the following proposition.

Proposition 2.2. *Let G be an amenable ICC group. For every $h \in G \setminus \{e\}$ there exists a fully supported, symmetric, finite entropy probability measure μ such that*

$$\lim_{m\to\infty} \|h\mu^{*m} - \mu^{*m}\| > 0. \tag{2.1}$$

Here μ^{*m} is the m-fold convolution $\mu * \cdots * \mu$. We will prove this Proposition later, and now turn to the proof of Proposition 2.1.

Proof of Proposition 2.1. The case of non-amenable G is known, so assume that G is amenable and has an ICC quotient Q. Let h be a non-identity element of Q. Applying Proposition 2.2 to Q and h yields a finite entropy, symmetric measure $\bar{\mu}$ on Q that is fully supported, and satisfies (2.1).

Since $\bar{\mu}$ has full support and satisfies (2.1), it follows from [15, Theorem 2] that $(Q,\bar{\mu})$ has a non-trivial Poisson boundary. Let μ be any symmetric, finite entropy non-degenerate probability measure on G that is projected to $\bar{\mu}$; the existence of such a μ is straightforward. Then (G,μ) has a non-trivial Poisson boundary. □

2.1. **Switching Elements.**

Here we introduce two notions: switching elements and super-switching elements. We will use these notions in the proof of Proposition 2.2.

Definition 2.3. Let X be a finite symmetric subset of a group G.

- We call $g \in G$ a *switching element for X* if
$$X \cap gXg^{-1} \subseteq \{e\}.$$
- We call $g \in G$ a *super-switching element for X* if
$$X \cap \left(gXg \cup gXg^{-1} \cup g^{-1}Xg \cup g^{-1}Xg^{-1}\right) \subseteq \{e\}.$$

Note that since X is symmetric, $g \in G$ is a switching element for X if and only if g^{-1} is a switching element for X.

Claim 2.4. *Let X be a finite symmetric subset of a group G and let $g \in G$ be a super-switching element for X. If $g^{w_1} x g^{w_2} = y$ for $x, y \in X$ and $w_1, w_2 \in \{-1, +1\}$, then $x = y = e$.*

Proof. Let $g^{w_1} x g^{w_2} = y$ for $x, y \in X$ and $w_1, w_2 \in \{-1, +1\}$. Since

$$y = g^{w_1} x g^{w_2} \in \left(gXg \cup gXg^{-1} \cup g^{-1}Xg \cup g^{-1}Xg^{-1} \right)$$

and $y \in X$, it follows from the definition of a super-switching element for X that $y = e$.

From $g^{w_1} x g^{w_2} = y$, we get $g^{-w_1} y g^{-w_2} = x$. So, by symmetry, the same argument shows $x = e$. □

Proposition 2.5. *Let G be a discrete (not necessarily countable) amenable ICC group, and let X be a finite symmetric subset of G. The set of super-switching elements for X is infinite.*

Proof of Proposition 2.5. Fix an invariant finitely additive probability measure d on G. For $A \subseteq G$, we call $d(A)$ the density of A. We will need the fact that infinite index subgroups have zero density, and that $d(A) = 0$ for every finite subset $A \subset G$.

Let $C_G(x)$ be the centralizer of a non-identity $x \in X$. Then, since X is finite, there is a finite set of cosets of $C_G(x)$ that includes all $g \in G$ such that $g^{-1} x g \in X$. So, non-switching elements for X are in the union of finitely many cosets of subgroups with infinite index, since G is ICC. This means that the set of non-switching elements for X has zero density, and so the set S of switching elements for X has density one.

Let T be the set of all super-switching elements for X. Let $A \subseteq G$ be the set of involutions $\{g \in G \mid g^2 = e\}$.

If $d(A) > 0$, then $d(A \cap S) > 0$. On the other hand, for any $g \in A \cap S$, since g is switching for X and $g^{-1} = g$, g is super-switching for X. Hence $A \cap S \subseteq T$. This shows that if $d(A) > 0$, then $d(T) \geq d(A \cap S) > 0$, and so we are done.

So, we can assume that $d(A) = 0$. For any $x, y \in X$, let $S_{x,y} = \{g \in S \mid gxg = y\}$. Note that

$$T = S \setminus \bigcup_{\substack{x,y \in X \\ (x,y) \neq (e,e)}} S_{x,y}.$$

It is thus enough to be shown that each $S_{x,y}$ has zero density when $(x, y) \neq (e, e)$. So assume for the sake of contradiction that $d(S_{x,y}) > 0$.

Fix $g \in S_{x,y}$. We have the following for all $h \in g^{-1}S_{x,y}$.

$$\begin{aligned} gxg = y = ghxgh &\implies (xg) = h(xg)h \\ &\implies (xg)^{-1}h^{-1}(xg) = h \\ &\implies h = (xg)^{-1}h^{-1}(xg) \\ &\qquad = (xg)^{-1}[(xg)^{-1}h^{-1}(xg)]^{-1}(xg) \\ &\qquad = (xg)^{-2}h(xg)^2 \\ &\implies h \text{ is in the centralizer of } (xg)^2. \end{aligned}$$

So, the centralizer of $(xg)^2$ includes $g^{-1}S_{x,y}$, which has a positive density. So, the centralizer of $(xg)^2$ has finite index. This implies that $(xg)^2 = e$, because in an ICC group only the identity can have a finite index centralizer. Hence $xg \in A$ for all $g \in S_{x,y}$. So $xS_{x,y} \subseteq A$. Hence $S_{x,y}$ also has zero density, which is a contradiction. □

2.2. **A Heavy-Tailed Probability Distribution on $\mathbb{N}$.** Here we state and prove a lemma about the existence of a probability distribution on $\mathbb{N} = \{1, 2, \ldots\}$ such that infinite i.i.d. samples from this measure have certain properties. We will use this distribution in the proof of Proposition 2.2.

Lemma 2.6. *Let p be the following probability measure on $\mathbb{N}$: $p(n) = cn^{-5/4}$, where $1/c = \sum_{n=1}^{\infty} n^{-5/4}$. Then p has finite entropy and the following property: for any $\varepsilon > 0$ there exist constants $K_\varepsilon, N_\varepsilon \in \mathbb{N}$ such that for any natural number $m \geq K_\varepsilon$ there exists an $E_{\varepsilon,m} \subseteq \mathbb{N}^m$ such that:*

(1) $p^{\times m}(E_{\varepsilon,m}) \geq 1 - \varepsilon$, where $p^{\times m}$ is the m-fold product measure $p \times \cdots \times p$.
(2) For any $s = (s_1, \ldots, s_m) \in E_{\varepsilon,m}$, the maximum of $\{s_1, \ldots, s_{K_\varepsilon}\}$ is at most N_ε.
(3) For any $s = (s_1, \ldots, s_m) \in E_{\varepsilon,m}$ and for any $K_\varepsilon \leq k \leq m$, the maximum of $\{s_1, \ldots, s_k\}$ is at least k^2.
(4) For any $s = (s_1, \ldots, s_m) \in E_{\varepsilon,m}$ and for any $K_\varepsilon \leq k \leq m$, the maximum of $\{s_1, \ldots, s_k\}$ appears in $(s_1, \ldots, s_k)$ only once.

Proof. It is straightforward to see that p has finite entropy.

Let $s = (s_1, s_2, \ldots) \in \mathbb{N}^\infty$ have distribution $p^{\times\infty}$; i.e., s is a sequence of i.d.d. random variables with distribution p. Since each s_i has distribution p, for each $n \in \mathbb{N}$ we have:

(2.2)
$$\mathbb{P}\left[s_i \geq n\right] = \sum_{m=n}^{\infty} p(m) = c \sum_{m=n}^{\infty} m^{-5/4} \geq c \int_n^\infty x^{-5/4} \mathrm{d}x = 4cn^{-1/4}.$$

For $k \geq 1$, let

$$M_k \coloneqq \max\{s_1, \ldots, s_k\},$$

and let

$$\mathrm{next}(k) \coloneqq \min\{i > k \mid s_i \geq M_k\}.$$

In words, $\mathrm{next}(k)$ is the first index $i > k$ for which s_i matches or exceeds M_k.

We first show that with probability one, $M_k \geq k^2$ for all k large enough. To this end, let A_k be the event that $M_k < k^2$. We have:

$$\begin{aligned}
\mathbb{P}\left[A_k\right] &= \mathbb{P}\left[s_i < k^2 \ \forall i \in \{1, \ldots, k\}\right] \\
&= (1 - \mathbb{P}\left[s_1 < k^2\right])^k \\
&\leq (1 - 4c(k^2)^{-1/4})^k \\
&\leq e^{-4ck^{1/2}}.
\end{aligned}$$

Since the sum of these probabilities is finite, by Borel-Cantelli we get that

$$\mathbb{P}\left[A_k \text{ infinitely often}\right] = 0.$$

Hence $M_k \geq k^2$ for all k large enough, almost surely. Furthermore, the expectation of $1/M_k$ is small:

(2.3)
$$\mathbb{E}\left[\frac{1}{M_k}\right] = \mathbb{E}\left[\frac{1}{M_k}\middle| A_k\right]\mathbb{P}\left[A_k\right] + \mathbb{E}\left[\frac{1}{M_k}\middle| \neg A_k\right]\mathbb{P}\left[\neg A_k\right] \leq e^{-4ck^{1/2}} + \frac{1}{k^2}.$$

Next, we show that, with probability one, $s_{\mathrm{next}(k)} > M_k$ for all k large enough. That is, for large enough k, the first time that M_k is matched or exceeded after index k, it is in fact exceeded.

Let B_k be the event that $s_{\mathrm{next}(k)} = M_k$. We would like to show that this occurs only finitely often. Note that

$$\begin{aligned}
\mathbb{P}\left[B_k | M_k\right] &= \mathbb{P}\left[s_{\mathrm{next}(k)} = M_k \middle| M_k\right] \\
&= \sum_{i=k+1}^{\infty} \mathbb{P}\left[s_i = M_k, \mathrm{next}(k) = i | M_k\right].
\end{aligned}$$

Applying the definition of $\mathrm{next}(k)$ yields

$$\mathbb{P}\left[B_k|M_k\right] = \sum_{i=k+1}^{\infty} \mathbb{P}\left[s_i = M_k, s_{k+1}, \ldots, s_{i-1} < M_k|M_k\right].$$

By the independence of the s_i's we can write this as

$$\begin{aligned}\mathbb{P}\left[B_k|M_k\right] &= \sum_{i=k+1}^{\infty} \mathbb{P}\left[s_i = M_k|M_k\right] \prod_{n=1}^{i-(k+1)} \mathbb{P}\left[s_{k+n} < M_k|M_k\right] \\ &= \sum_{i=k+1}^{\infty} \frac{c}{M_k^{5/4}} \mathbb{P}\left[s_{k+1} < M_k|M_k\right]^{i-(k+1)}.\end{aligned}$$

By (2.2), $\mathbb{P}\left[s_{k+1} < M_k|M_k\right] \leq 1 - 4cM_k^{-1/4}$. Hence

$$\mathbb{P}\left[B_k|M_k\right] \leq \frac{c}{M_k^{5/4}} \cdot \frac{1}{4cM_k^{-1/4}} = \frac{1}{4M_k}.$$

Using (2.3) it follows that

$$\mathbb{P}\left[B_k\right] = \mathbb{E}\left[\mathbb{P}\left[B_k|M_k\right]\right] \leq \mathbb{E}\left[\frac{1}{4M_k}\right] \leq \frac{1}{4}e^{-4ck^{1/2}} + \frac{1}{4k^2}.$$

Hence $\sum_k \mathbb{P}\left[B_k\right] < \infty$, and so by Borel-Cantelli B_k occurs only finitely often.

Since A_k and B_k both occur for only finitely many k, the (random) index ind' at which they stop occurring is almost surely finite, and is given by

$$\mathrm{ind}' = \min\{\ell \in \mathbb{N} \,:\, s \notin A_k \cup B_k \text{ for all } k \geq \ell\}.$$

Let

$$\mathrm{ind} = \mathrm{next}(\mathrm{ind}').$$

Hence for $k \geq \mathrm{ind}$, $M_k \geq k^2$ and M_k appears in $(s_1, \ldots, s_k)$ only once.

Fix $\varepsilon > 0$. Since ind is almost surely finite, then for large enough constants $K_\varepsilon \in \mathbb{N}$ and $N_\varepsilon \in \mathbb{N}$ the event

$$E_\varepsilon = \{\mathrm{ind} \leq K_\varepsilon \text{ and } M_{K_\varepsilon} \leq N_\varepsilon\}$$

has probability at least $1 - \varepsilon$, and additionally, conditioned on E_ε it holds that $k \geq \mathrm{ind}$ for all $k \geq K_\varepsilon$, and hence $M_k \geq k^2$ and M_k appears in $(s_1, \ldots, s_k)$ only once. Therefore, if for $m \geq K_\varepsilon$ we let $E_{\varepsilon,m}$ be the projection of E_ε to the first m coordinates, then $E_{\varepsilon,m}$ satisfies the desired properties. $\square$

2.3. **Proof of Proposition 2.2.** Let $\frac{1}{8} > \varepsilon > 0$. Let p, $K_\varepsilon \in \mathbb{N}, N_\varepsilon \in \mathbb{N}$, and $E_{\varepsilon,m} \subseteq \mathbb{N}^m$ be the probability measure, the constants, and the events from Lemma 2.6. To simplify notation let $N = N_\varepsilon$ and $K = K_\varepsilon$.

Let $G = \{a_1, a_2, \ldots\}$, where $a_1 = a_2 = \cdots = a_N = e$. We define $(g_n)_n$, $(A_n)_n$, $(B_n)_n$ and $(C_n)_n$ recursively. Given $g_1, \ldots, g_n$, let $A_n = \{g_n, g_n^{-1}, a_n, a_n^{-1}\}$ and $B_n = \cup_{i\leq n} A_i$. Denote $C_n = B_n \cup \{h^{-1}, h\}$. Note that A_n, B_n, and C_n are finite and symmetric for any $n \in \mathbb{N}$. Let $g_1 = g_2 = \ldots = g_N = e$. For $n+1 > N$, given C_n, let $g_{n+1} \in G$ be a super-switching element for $(C_n)^{2n+1}$ which is not in $(C_n)^{8n+1}$. The existence of such a super-switching element is guaranteed by Proposition 2.5 and the facts that $(C_n)^{2n+1}$ is a finite symmetric subset of G and that $(C_n)^{8n+1}$ is finite.

For $n \in \mathbb{N}$, define a symmetric probability measure μ_n on A_n by

$$\mu_n = \varepsilon 2^{-n}(\frac{1}{2}\delta_{a_n} + \frac{1}{2}\delta_{a_n^{-1}}) + (1-\varepsilon 2^{-n})(\frac{1}{2}\delta_{g_n} + \frac{1}{2}\delta_{g_n^{-1}}).$$

Here δ_g is the point mass on $g \in G$. Finally, let

$$\mu = \sum_{n=1}^{\infty} p(n)\mu_n.$$

Obviously μ is symmetric and $\operatorname{supp}\mu = G$. Since p has finite entropy and each μ_n has support of size at most 4, it follows easily that μ has finite entropy.

We want to show that

$$\lim_{m\to\infty} \|h\mu^{*m} - \mu^{*m}\| > 0.$$

Fix $m \in \mathbb{N}$ larger than K and N. For each $n \in \mathbb{N}$ define $f_n : \{1,2,3,4\} \to A_n$ by

$$f_n(1) = a_n,\ f_n(2) = a_n^{-1},\ f_n(3) = g_n,\ f_n(4) = g_n^{-1},$$

and define $\nu_n : \{1,2,3,4\} \to [0,1]$ by

$$\nu_n(1) = \nu_n(2) = \frac{1}{2}\varepsilon 2^{-n},\ \nu_n(3) = \nu_n(4) = \frac{1}{2}(1-\varepsilon 2^{-n}).$$

Let

$$\Omega = \{(s,w) \mid s \in \mathbb{N}^m,\ w \in \{1,2,3,4\}^m\}.$$

We define the measure η on the countable set Ω by specifying its values on the singletons:

$$\eta(\{(s,w)\}) = p^{\times m}(s)\ \nu_{s_1}(w_1)\ \nu_{s_2}(w_2)\ \ldots\ \nu_{s_m}(w_m).$$

It follows immediately from this definition that η is a probability measure.

Define $r : \Omega \to G$ by

$$r(s,w) = f_{s_1}(w_1) f_{s_2}(w_2) \dots f_{s_m}(w_m).$$

It is not difficult to see that $r_*\eta = \mu^{*m}$, and so we need to show that $\|hr_*\eta - r_*\eta\|$ is uniformly bounded away from zero for m larger than K and N.

Recall that $E_{\varepsilon,m} \subseteq \mathbb{N}^m$ is the event given by Lemma 2.6. Fix $s \in E_{\varepsilon,m}$. Define

$$\begin{aligned} i_{s,1} &= \min\{j \in \{1\dots,m\} \mid s_j > N\}, \\ i_{s,2} &= \min\{j > i_{s,1} \mid s_j \geq s_{i_{s,1}}\}, \\ &\vdots \\ i_{s,l(s)} &= \min\{j > i_{s,l(s)-1} \mid s_j \geq s_{i_{s,l(s)-1}}\}. \end{aligned}$$

Note that by the second property of $E_{\varepsilon,m}$ in Lemma 2.6, we know that

$$K < i_{s,1} < i_{s,2} < \cdots < i_{s,l(s)},$$

and by the fourth property,

$$N < s_{i_{s,1}} < s_{i_{s,2}} < \cdots < s_{i_{s,l(s)}} = \max\{s_1,\dots,s_m\}.$$

Let

$$W_\varepsilon^s = \{w \in \{1,2,3,4\}^m \mid \forall k \leq l(s)\ w_{i_{s,k}} = 3,4\}.$$

For $s \in \mathbb{N}^m$ let η_s be the measure η, conditioned on the first coordinate equalling s. I.e., let

$$\eta_s(A) = \frac{\eta(A \cap \Omega^s)}{\eta(\Omega^s)},$$

where $\Omega^s = \{s\} \times \{1,2,3,4\}^m \subseteq \Omega$.

Then

$$\begin{aligned} \eta_s(\{s\} \times W_\varepsilon^s) &= 1 - \eta_s(\{w_{i_{s,1}} = 1,2;\ \text{or } w_{i_{s,2}} = 1,2;\ \dots;\ \text{or } w_{i_{s,l(s)}} = 1,2\ \}) \\ &\geq 1 - \sum_{k=1}^{l(s)} \eta_s(\{w_{i_{s,k}} = 1,2\}) \\ &= 1 - \sum_{k=1}^{l(s)} \varepsilon 2^{-s_{i_{s,k}}} \\ &\geq 1 - \sum_{j=1}^{\infty} \varepsilon 2^{-j} \\ &= 1 - \varepsilon, \end{aligned}$$

where the first inequality follows from the union bound, and the last inequality holds since $s_{i_{s,1}} < s_{i_{s,2}} < \cdots < s_{i_{s,l(s)}}$.

Finally, let

$$\Omega_\varepsilon = \{(s, w) \in \Omega \mid s \in E_{\varepsilon,m},\ w \in W^s_\varepsilon\}.$$

By the above, and since $\eta(E_{\varepsilon,m} \times \{1,2,3,4\}^m) \geq 1 - \varepsilon$ by Lemma 2.6, we have shown that

$$\eta(\Omega_\varepsilon) \geq\ (1-\varepsilon)(1-\varepsilon) >\ 1 - 2\varepsilon.$$

Claim 2.7. *For any $\alpha, \beta \in \Omega_\varepsilon$, we have $hr(\alpha) \neq r(\beta)$.*

We prove this claim after we finish the proof of the Proposition.

Let η_1 be equal to η conditioned on Ω_ε, and η_2 be equal to η conditioned on the complement of Ω_ε. We have $\eta = \eta(\Omega_\varepsilon)\eta_1 + (1-\eta(\Omega_\varepsilon))\eta_2$, and by the above claim we know $\|hr_*\eta_1 - r_*\eta_1\| = 2$. So for m larger than K and N

$$\begin{aligned}
\|h\mu^{*m} - \mu^{*m}\| &=\ \|hr_*\eta - r_*\eta\| \\
&=\ \|\eta(\Omega_\varepsilon)(hr_*\eta_1 - r_*\eta_1) + (1-\eta(\Omega_\varepsilon))(hr_*\eta_2 - r_*\eta_2)\| \\
&\geq \eta(\Omega_\varepsilon)\,\|hr_*\eta_1 - r_*\eta_1\| - 2(1-\eta(\Omega_\varepsilon)) \\
&\geq 2(1-2\varepsilon) - 2(2\varepsilon) = 2 - 8\varepsilon,
\end{aligned}$$

which is uniformly bounded away from zero since $\varepsilon < \frac{1}{8}$. Since $\|h\mu^{*m} - \mu^{*m}\|$ is a decreasing sequence, this completes the proof of Proposition 2.2.

Proof of Claim 2.7. Let $\alpha = (s, w)$, $\beta = (t, v) \in \Omega_\varepsilon$. Hence $\max\{K, N\} < m$, $s \in E_{\varepsilon,m}$, $t \in E_{\varepsilon,m}$, $w \in W^s_\varepsilon$, and $v \in W^t_\varepsilon$. Assume that $hr(\alpha) = r(\beta)$. So, we have

$$hf_{s_1}(w_1)\cdots f_{s_m}(w_m) = f_{t_1}(v_1)\cdots f_{t_m}(v_m).$$

Let $K < i_1 < i_2 < \cdots < i_{l(s)}$ and $K < j_1 < j_2 < \cdots < j_{l(t)}$ be the indices we defined for s and t in the proof of Proposition 2.2. We remind the reader that the unique maximum of $(s_1, \ldots, s_m)$ is attained at $i_{l(s)}$, with a corresponding statement for $(t_1, \ldots, t_m)$ and $j_{l(t)}$. So we have

$$\begin{aligned}
& h\,\overbrace{f_{s_1}(w_1)\cdots f_{s_{i_{l(s)}-1}}(w_{i_{l(s)}-1})}^{b_1}\, f_{s_{i_{l(s)}}}(w_{i_{l(s)}})\, \overbrace{f_{s_{i_{l(s)}+1}}(w_{i_{l(s)}+1})\cdots f_{s_m}(w_m)}^{b_2} \\
=\ & \underbrace{f_{t_1}(v_1)\cdots f_{t_{j_{l(t)}-1}}(v_{j_{l(t)}-1})}_{c_1}\, f_{t_{j_{l(t)}}}(v_{j_{l(t)}})\, \underbrace{f_{t_{j_{l(t)}+1}}(v_{j_{l(t)}+1})\cdots f_{t_m}(v_m)}_{c_2}.
\end{aligned}$$

Let $p = s_{i_{l(s)}} = \max\{s_1, \ldots, s_m\}$ and $q = t_{j_{l(t)}} = \max\{t_1, \ldots, t_m\}$. Since $w \in W^s_\varepsilon$ and $v \in W^t_\varepsilon$, we know $f_{s_{i_{l(s)}}}(w_{i_{l(s)}}) = g_p^{\pm 1}$ and $f_{t_{j_{l(t)}}}(v_{j_{l(t)}}) =$

$g_q^{\pm 1}$, so

$$hb_1 g_p^{\pm 1} b_2 = c_1 g_q^{\pm 1} c_2. \tag{2.4}$$

Since $p = \max\{s_1, \ldots, s_m\}$, and since $m \geq K$, we know that $m \leq m^2 \leq p$. So $b_1, b_2 \in (B_{p-1})^{p-1} \subseteq (C_{p-1})^{p-1}$. Similarly $c_1, c_2 \in (C_{q-1})^{q-1}$.

Consider the case that $p > q$. Then $c_1, c_2, g_q^{\pm 1} \in (C_q)^q \subseteq (C_{p-1})^{p-1}$. Hence $g_p^{\pm 1} = [b_1^{-1}]h^{-1}[c_1 g_q^{\pm 1} c_2 b_2^{-1}]$ by (2.4), and so

$$g_p \in (C_{p-1})^{4(p-1)}\{h, h^{-1}\}(C_{p-1})^{4(p-1)} \subseteq (C_{p-1})^{8(p-1)+1},$$

which is a contradiction with our choice of g_p, since $p > N$. Similarly, if $p < q$, we get a contradiction. So we can assume that $p = q$.

If $p = q$, then by (2.4) we have

$$hb_1 g_p^{\pm 1} b_2 = c_1 g_p^{\pm 1} c_2,$$

and $c_1, c_2, b_1, b_2 \in (C_{p-1})^{p-1}$. So, for $x = c_1^{-1} h b_1 \in (C_{p-1})^{2(p-1)+1}$ we have $g_p^{\pm 1} x g_p^{\pm 1} = c_2 b_2^{-1} \in (C_{p-1})^{2(p-1)} \subseteq (C_{p-1})^{2(p-1)+1}$. By the fact that g_p is a super-switching element for $(C_{p-1})^{2(p-1)+1}$ and from Claim 2.4, we get that x is the identity.

So $hb_1 = c_1$, i.e.

$$hf_{s_1}(w_1) \cdots f_{s_{i_{l(s)}-1}}(w_{i_{l(s)}-1}) = f_{t_1}(v_1) \cdots f_{t_{j_{l(t)}-1}}(v_{j_{l(t)}-1}).$$

By the exact same argument, we can see this leads to a contradiction unless

$$hf_{s_1}(w_1) \cdots f_{s_{i_{l(s)-1}-1}}(w_{i_{l(s)-1}-1}) = f_{t_1}(v_1) \cdots f_{t_{j_{l(t)-1}-1}}(v_{j_{l(t)-1}-1}).$$

And again, this leads to a contradiction unless

$$hf_{s_1}(w_1) \cdots f_{s_{i_{l(s)-2}-1}}(w_{i_{l(s)-2}-1}) = f_{t_1}(v_1) \cdots f_{t_{j_{l(t)-2}-1}}(v_{j_{l(t)-2}-1}).$$

Note that if $l(s) \neq l(t)$, at some point in this process we get that either all the s_i's or all the t_i's are at most N while the other string has characters strictly greater than N. This leads to a contradiction similar to the case $p \neq q$, which we explained before. So, by continuing this process, we get a contradiction unless

$$hf_{s_1}(w_1) \cdots f_{s_{i_1-1}}(w_{i_1-1}) = f_{t_1}(v_1) \cdots f_{t_{j_1-1}}(v_{j_1-1}). \tag{2.5}$$

Note that $s_1, \ldots, s_{i_1-1} \leq N$, which implies

$$f_{s_1}(w_1) = \cdots = f_{s_{i_1-1}}(w_{i_1-1}) = e.$$

Similarly, $t_1, \ldots, t_{j_1-1} \leq N$ implies that

$$f_{t_1}(v_1) = \cdots = f_{t_{j_1-1}}(v_{j_1-1}) = e.$$

So, from (2.5) we get $h = e$, which is a contradiction.

□

References

[1] Uri Bader and Yehuda Shalom, *Factor and normal subgroup theorems for lattices in products of groups*, Invent. Math. **163** (2006), no. 2, 415–454.

[2] Laurent Bartholdi and Anna Erschler, *Poisson–Furstenberg boundary and growth of groups*, Probability Theory and Related Fields **168** (2017), no. 1-2, 347–372.

[3] David Blackwell, *On transient markov processes with a countable number of states and stationary transition probabilities*, The Annals of Mathematical Statistics (1955), 654–658.

[4] Gustave Choquet and Jacques Deny, *Sur l'équation de convolution* $\mu = \mu * \sigma$, C. R. Acad. Sci. Paris **250** (1960), 799–801.

[5] AM Duguid and DH McLain, *FC-nilpotent and FC-soluble groups*, Mathematical proceedings of the cambridge philosophical society, 1956, pp. 391–398.

[6] Evgeniı B Dynkin and MB Maljutov, *Random walk on groups with a finite number of generators*, Dokl. akad. nauk sssr, 1961, pp. 1042–1045.

[7] Anna Erschler, *Boundary behavior for groups of subexponential growth*, Annals of Mathematics (2004), 1183–1210.

[8] ———, *Liouville property for groups and manifolds*, Inventiones mathematicae **155** (2004), no. 1, 55–80.

[9] Joshua Frisch, Omer Tamuz, and Pooya Vahidi Ferdowsi, *Strong amenability and the infinite conjugacy class property*, arXiv preprint arXiv:1801.04024 (2018).

[10] Alex Furman, *Random walks on groups and random transformations*, Handbook of dynamical systems, 2002, pp. 931–1014.

[11] H. Furstenberg and E. Glasner, *Stationary dynamical systems*, Contemp. Math., vol. 532, Amer. Math. Soc., Providence, RI, 2010.

[12] Harry Furstenberg, *Noncommuting random products*, Transactions of the American Mathematical Society **108** (1963), no. 3, 377–428.

[13] ———, *Random walks and discrete subgroups of Lie groups*, Advances in Probability and Related Topics **1** (1971), 1–63.

[14] ———, *Boundary theory and stochastic processes on homogeneous spaces*, Harmonic analysis on homogeneous spaces **26** (1973), 193–229.

[15] Shmuel Glasner, *On Choquet-Deny measures*, Ann. Inst. H. Poincaré B **12** (1976), 1–10.

[16] ———, *Proximal flows*, Lecture Notes in Mathematics, Vol. 517, Springer-Verlag, Berlin-New York, 1976.

[17] Yves Guivarc'h, *Croissance polynomiale et périodes des fonctions harmoniques*, Bull. Soc. Math. France **101** (1973), no. 333, 379.

[18] Wojciech Jaworski, *Countable amenable identity excluding groups*, Canadian Mathematical Bulletin **47** (2004), no. 2, 215–228.

[19] Wojciech Jaworski and C Robinson Edward Raja, *The Choquet–Deny theorem and distal properties of totally disconnected locally compact groups of polynomial growth*, New York J. Math **13** (2007), 159–174.

[20] Vadim A Kaimanovich, *Examples of non-abelian discrete groups with non-trivial exit boundary*, Zapiski Nauchnykh Seminarov POMI **123** (1983), 167–184.

[21] Vadim A Kaimanovich and Anatoly M Vershik, *Random walks on discrete groups: boundary and entropy*, The annals of probability (1983), 457–490.

[22] DH McLain, *Remarks on the upper central series of a group*, Glasgow Mathematical Journal **3** (1956), no. 1, 38–44.
[23] Derek S Robinson, *Finiteness conditions and general soluble groups*, Springer, Berlin, 1972.
[24] Joseph Rosenblatt, *Ergodic and mixing random walks on locally compact groups*, Mathematische Annalen **257** (1981), no. 1, 31–42.

(J. Frisch, O. Tamuz, P. Vahidi Ferdowsi) California Institute of Technology

(Y. Hartman) Northwestern University