SOFICITY AND VARIATIONS ON HIGMAN'S GROUP

MARTIN KASSABOV, VIVIAN KUPERBERG, AND TIMOTHY RILEY

Abstract. A group is sofic when every finite subset can be well approximated in a finite symmetric group. No example of a non-sofic group is known. Higman's group, which is a circular amalgamation of four copies of the Baumslag–Solitar group, is a candidate. Here we contribute to the discussion of the problem of its soficity in two ways.

We construct variations on Higman's group replacing the Baumslag–Solitar group by other groups G. We give an elementary condition on G, enjoyed for example by $\mathbb{Z} \wr \mathbb{Z}$ and the integral Heisenberg group, under which the resulting group is sofic.

We then use soficity to deduce that there exist permutations of $\mathbb{Z}/n\mathbb{Z}$ that are seemingly pathological in that they have order dividing four and yet locally they behave like exponential functions over most of their domains. Our approach is based on that of Helfgott and Juschenko, who recently showed the soficity of Higman's group would imply some the existence of some similarly pathological functions. Our results call into question their suggestion that this might be a step towards proving the existence of a non-sofic group.
2010 Mathematics Subject Classification: 03C20, 20F69
Key words and phrases: sofic, Higman's group, residually finite

1. Our results

The word *sofic*, derived from the Hebrew for *finite*, was applied to a group by Weiss in [20] when every finite subset can be well approximated in a finite symmetric group or, equivalently, when the group is a subgroup of a metric ultraproduct of finite symmetric groups. The focus of this article is the outstanding open question about soficity, posed by Gromov in his 1999 paper [10]: is every group sofic? We will give more background on soficity in Section 2.

It is not known whether *Higman's group*

$$H_4 \;=\; \left\langle a,b,c,d \;\middle|\; b^a = b^2, c^b = c^2, d^c = d^2, a^d = a^2 \right\rangle$$

is sofic. This group can be constructed as follows. First amalgamate two copies of the Baumslag–Solitar group $\mathrm{BS}(1,2) = \langle a, b \mid b^a = b^2\rangle$ to give $\langle a,b,c \mid b^a = b^2, c^b = c^2\rangle$. By properties of the free products with amalgamation, its subgroup $\langle a, c\rangle$ is free of rank 2. Amalgamate with a second copy $\langle c,d,a \mid d^c = d^2, a^d = a^2\rangle$ along the common $\langle a,c\rangle$ subgroup to give H_4.

Again, properties of free products with amalgamation tell us that the subgroups $\langle a,b\rangle$, $\langle b,c\rangle$, $\langle c,d\rangle$, and $\langle a,d\rangle$ are copies of $\mathrm{BS}(1,2)$, and that $\langle a,c\rangle$ is free of rank 2. In particular, H_4 is not amenable, since it contains a non-abelian free subgroup. And H_4 is not residually finite, because it has no finite quotients [13]. These properties make H_4 a candidate for a non-sofic group. The case is made all the more compelling because H_4 fails to have a property slightly more restrictive than soficity: Thom proved in [19] that it does not embed

Date: May 24, 2018.

into a metric ultraproduct of finite groups with a commutator-contractive invariant length function.

The building blocks for our variations on Higman's group (explored in more detail in Section 3) are a group G, subgroups A and B, an isomorphism $\phi : B \to A$, and a $k \in \mathbb{N}$. For $1 \leq i \leq k$, let G_i be copies of G, let $A_i, B_i \leq G_i$ be copies of its subgroups A and B, and let $\phi_i : B_i \to A_{i+1}$ (indices mod k) be the map naturally induced by ϕ. We define

$$\overline{\mathrm{Hig}}_k(G,\phi) \ := \ \langle G_1, \ldots, G_k \mid b_i = \phi_i(b_i) \text{ for all } i \text{ and all } b_i \in B_i \rangle,$$

which is k copies of G assembled in a cyclic analog of a free product with amalgamation. If $G = \mathrm{BS}(1,2) = \langle a, b \mid b^a = b^2 \rangle$ and $\phi : \langle b \rangle \to \langle a \rangle$ maps $b \mapsto a$, then $\overline{\mathrm{Hig}}_4(G,\phi) = H_4$.

Next we define $\mathrm{Hig}_k(G,\phi)$ to be the semi-direct product of $\overline{\mathrm{Hig}}_k(G,\phi)$ with a cyclic group $\langle t \rangle$ of order k in which t conjugates G_i to G_{i+1} (indices mod k). Taking $G = G_1$ and eliminating all the other G_i from the natural presentation, we have that

$$\mathrm{Hig}_k(G,\phi) \ := \ \langle G, t \mid t^k = 1, \ b^t = \phi(b); \forall b \in B \rangle.$$

The index of $\overline{\mathrm{Hig}}_k(G,\phi)$ in $\mathrm{Hig}_k(G,\phi)$ is k, so one is sofic if and only if the other is; see [17].

(Monod has generalized Higman's construction in a different direction in [16].)

In contrast to H_4, we can often prove soficity for these groups. Indeed, in many cases they are residually solvable, and so sofic. We will prove in Section 4:

Theorem 1.1. *Suppose G is a residually solvable group and ϕ is an isomorphism $B \to A$ between subgroups $A, B \leq G$. Suppose there exists a group homomorphism $\pi : G \to A \times B$ such that $\pi(a) = (a, 1)$ for all $a \in A$ and $\pi(b) = (1, b)$ for all $b \in B$. Then* $\mathrm{Hig}_k(G,\phi)$ *and* $\overline{\mathrm{Hig}}_k(G,\phi)$ *are residually solvable for all $k \geq 4$.*

A number of examples are presented in Section 4.

With a view to showing that H_4 is not sofic, Helfgott and Juschenko proved:

Theorem 1.2 (Helfgott–Juschenko [12])**.** *If Higman's group H_4 is sofic, then for all $\epsilon > 0$ there exists $N \in \mathbb{N}$, such that for all odd $n > N$ there exists $f \in \mathrm{Sym}(\mathbb{Z}/n\mathbb{Z})$ of order dividing 4 with $f(x+1) = 2f(x)$ for at least $(1-\varepsilon)n$ elements $x \in \mathbb{Z}/n\mathbb{Z}$.*

The f of Helfgott and Juschenko's theorem behave locally like an exponential function over most of $\mathbb{Z}/n\mathbb{Z}$ but nevertheless are permutations of order dividing four. They gave a heuristic argument as to why such f are unlikely to exist, based on the assumption that these two properties are independent (an intuition that they backed up with comparisons to prominent conjectures in analytic number theory). In Section 7 we give further analysis as to why one might have expected such f not to exist.

Since Helfgott and Juschenko's paper first appeared (as a preprint on the arXiv in December 2015) doubt has been cast on this intuition by the following two very similar theorems.

Theorem 1.3. *For all $\varepsilon > 0$ and $k \geq 3$, there exists C such that if n is coprime to m and $|m| > C$ and $n > |m|^C$, then there exists $f \in \mathrm{Sym}(\mathbb{Z}/n\mathbb{Z})$ which has order dividing k and the property that $f(x+1) = mf(x)$ for at least $(1-\varepsilon)n$ values of $x \in \mathbb{Z}/n\mathbb{Z}$.*

Theorem 1.4 (Helfgott–Juschenko [12], also Glebsky [8])**.** *For all $m > 2$ and all $\varepsilon > 0$, there exists C such that for all $n > C$ coprime to m, there exists $f \in \mathrm{Sym}(\mathbb{Z}/n\mathbb{Z})$ of order dividing 4 with $f(x+1) = mf(x)$ for at least $(1-\varepsilon)n$ values of $x \in \mathbb{Z}/n\mathbb{Z}$.*

For instance, taking N sufficiently large that $\ln\ln N > C$ and $N > (\ln N)^C$, Theorem 1.3 tells us that there exists $N \in \mathbb{N}$ such that for all coprime integers m and n with $n > N$ and $\ln\ln n < m < \ln n$, there exists f with the given properties.

Theorems 1.3 and 1.4 both run counter to Helfgott and Juschenko's heuristics. (However, neither theorem addresses the case $m = 2$ directly, so the existence of the functions f of Helfgott and Juschenko's theorem remains open.)

We will prove Theorem 1.3 in Section 6.4.

Theorems 1.2–1.4 all arise from a relationship between soficity and the existence of particular permutations of $\mathbb{Z}/n\mathbb{Z}$ set out in Theorem 1.5 below, which is a generalization of a result of Helfgott and Juschenko [12]. In the case of Theorem 1.2, soficity of H_4 is a hypothesis. For Theorem 1.3, we use the soficity of $\mathrm{Hig}_4(\mathbb{Z} \wr \mathbb{Z})$ established as a consequence of Theorem 1.1. Theorem 1.4 uses a theorem of Glebsky [8] which says that for $m \geq 3$, $\overline{\mathrm{Hig}_4}(\mathrm{BS}(1,m))$ has sofic quotients into which $\mathrm{BS}(1,m)$ embeds.

Theorem 1.5. *Suppose G is a group, ϕ is an isomorphism $B \to A$ between subgroups $A, B \leq G$, and $k \geq 1$ is an integer. The following two conditions are equivalent.*

(1) $\mathrm{Hig}_k(G,\phi)$ *has a sofic quotient Q such that the composition $G \to \mathrm{Hig}_k(G,\phi) \to Q$ is injective.*
(2) Sofic approximations of G exist for which there are permutations of order dividing k that almost conjugate the action of A to the action of B.

If G is amenable, then these are also equivalent to:

(3) For all sofic approximations of G into sufficiently large symmetric groups, there are permutations of order dividing k which almost conjugate the action of A to the action of B.

We will present a precise version of this theorem in Section 5. Conditions *(1)* and *(2)* each imply that G is sofic, as does the hypothesis of amenability under which we consider Condition *(3)*. The map $G \to \mathrm{Hig}_k(G,\phi)$ in *(1)* can fail to be injective—see Section 3.

We will prove Theorem 1.5 in Section 5, building on the arguments in [12]. The equivalence between Conditions *(1)* and *(2)* is analogous to that between the two definitions of soficity outlined at the start of this article—see Proposition 2.2. The idea behind the implication *(2)* ⇒ *(1)* is that the sofic approximations together with the almost-conjugating functions can be assembled into a homomorphism from $\mathrm{Hig}_k(G,\phi)$ to an ultraproduct of finite symmetric groups with image Q. For the implication *(1)* ⇒ *(2)*, we obtain the requisite sofic approximation of $S \subseteq G$ and the almost-conjugating permutation from a sofic approximation for the image in Q of a suitably constructed finite subset $S' \subseteq \mathrm{Hig}_k(G,\phi)$ with $S \cup \{t\} \subseteq S'$.

The equivalence of *(3)* is significantly more complicated. The additional assumption that the group G is amenable gives better control of the sofic approximations. The key result is a theorem which is due to Helfgott and Juschenko [12] in the form we will use and has origins in Elek and Szabo [6] and Kerr and Li [15]. It spells out a manner in which any two sofic approximations of an amenable group are almost conjugate.

In Section 6 we give applications of Theorem 1.5. We look at $G = \mathbb{Z}^2 = \langle a, b \mid ab = ba \rangle$ and $\phi : b \mapsto a$, which we view as an introductory example—in this case, $\mathrm{Hig}_k(G, \phi)$ will be a right-angled Artin group. We review the case of $G = \mathrm{BS}(1, m) = \langle a, b \mid a^b = a^m \rangle$ addressed by Helfgott and Juschenko and by Glebsky, where soficity of $\mathrm{Hig}_k(G, \phi)$ remains unknown for $m \geq 2$. We present our most novel applications which are when G is the 3-dimensional integral Heisenberg group $\mathcal{H}$, or $\mathbb{Z} \wr \mathbb{Z}$, or the free metabelian group $\mathcal{M}$ on two generators. In these cases, $\mathrm{Hig}_k(G, \phi)$ will be sofic by Theorem 1.1. We explain how the $G = \mathbb{Z} \wr \mathbb{Z}$ case leads to Theorem 1.3.

We do not know how to construct functions f explicitly satisfying the conditions of Theorems 1.3 or 1.4. In principle one could follow the constructions in the proofs, however this would require constructing several Følner sets for G and switching between sofic approximations several times. (In the case of Theorem 1.3, where the quotients could be taken to be the metabelian groups of Proposition 4.6, sofic approximations could be constructed explicitly; for Theorem 1.4 the quotients are residually nilpotent and constructing explicit sofic approximations is again possible, but significantly more difficult.) It seems unlikely that this will lead to an enlightening description of f.

By the same token, we do not know how C and N depend on ε in Theorems 1.2–1.4. One could obtain explicit estimates from our proofs, but they will be very weak. We give some examples in Remarks 6.3, 6.6, and 6.14. Sufficiently strong estimates (which may well not exist) could have important applications, including a proof that Higman's group H_4 is sofic.

2. Soficity

The *normalized Hamming distance* d on the symmetric group $\mathrm{Sym}(n)$ is

$$d(\rho, \sigma) \; = \; \frac{1}{n} |\{1 \leq i \leq n \mid \rho(i) \neq \sigma(i)\}|.$$

This metric is invariant under both the left and right action of $\mathrm{Sym}(n)$—i.e.,

$$d(\rho, \sigma) \; = \; d(\tau \rho \tau', \tau \sigma \tau')$$

for all $\rho, \sigma, \tau, \tau' \in \mathrm{Sym}(n)$. It follows that:

Lemma 2.1. *For $\sigma, \tau, \mu, \sigma_1, \ldots, \sigma_m \in \mathrm{Sym}(n)$,*

(i) $d(\mathrm{id}, \sigma_1 \cdots \sigma_m) \leq \sum_{i=1}^m d(\mathrm{id}, \sigma_i)$,
(ii) $d(\tau^{-1} \sigma \tau, \mathrm{id}) = d(\sigma, \mathrm{id})$,
(iii) $d(\tau^{-1} \sigma \tau, \mu^{-1} \sigma \mu) \leq 2 d(\tau, \mu)$.

For $n \in \mathbb{N}$, $\delta > 0$, and S a finite subset of a group G, an *(S, δ, n)-approximation* is a map $\psi : G \to \mathrm{Sym}(n)$ such that

- $d(\psi(g)\psi(h), \psi(gh)) < \delta$ for all $g, h \in S$ such that $gh \in S$, and
- $d(\psi(g), \mathrm{id}) > 1 - \delta$ for all $g \in S \smallsetminus \{e\}$.

It is a technical convenience that ψ is defined on all of G instead of just on S. Its values on $G \smallsetminus S$ are irrelevant to the definition. Note that if $e \in S$, then $d(\psi(e), \mathrm{id}) = d(\psi(e)\psi(e), \psi(e)) < \delta$, and so e must be mapped to within δ of the identity permutation.

A *filter* $\mathcal{F}$ on a set I is a nonempty set of subsets of I such that $\emptyset \notin \mathcal{F}$; for all $U, V \in \mathcal{U}$, $U \cap V \in \mathcal{F}$; and if $U \in \mathcal{F}$ and $U \subseteq V$, then $V \in \mathcal{F}$. An *ultrafilter* $\mathcal{U}$ on I is a maximal filter; equivalently, for all $U \subseteq I$, either $U \in \mathcal{U}$ or $(I \smallsetminus U) \in \mathcal{U}$.

Suppose $\mathcal{U}$ is an ultrafilter on a set I. To each $i \in I$ associate some $n_i \in \mathbb{N}$. For $x = (x_i)_{i\in I}$ and $y = (y_i)_{i\in I}$ in the direct product $\prod_{i\in I} \mathrm{Sym}(n_i)$, we write $x \approx_{\mathcal{U}} y$ when $\{i \in I \mid d(x_i, y_i) < \delta\} \in \mathcal{U}$ for all $\delta > 0$. Let $\mathrm{id} = (\mathrm{id}_{n_i})_{i\in I}$. Define $\mathcal{N} := \{x \in \prod_{i\in I} \mathrm{Sym}(n_i) \mid x \approx_{\mathcal{U}} \mathrm{id}\}$, which is called the normal subgroup of *infinitesimals*. Define the *(metric) ultraproduct* $\prod_{\mathcal{U}} \mathrm{Sym}(n_i) := (\prod_{i\in I} \mathrm{Sym}(n_i))/\mathcal{N}$. See [17] for further background.

A group G is *sofic* when it satisfies either of the conditions of the following proposition.

Proposition 2.2. *For a group G, the following are equivalent.*

(1) The group G is isomorphic to a subgroup of some metric ultraproduct of finite symmetric groups—that is, there exist an ultrafilter $\mathcal{U}$ on a set I, natural numbers $\{n_i\}_{i\in I}$, and an injective homomorphism

$$G \hookrightarrow \prod_{\mathcal{U}} \mathrm{Sym}(n_i).$$

(2) For all finite subsets $S \subseteq G$ and all $\delta > 0$, there exists an (S, δ, n)-approximation for some n.

Proofs of this result can be found in [5, 17]. We include the following account in order to prepare the reader for our proof of Theorem 5.1, which is much more involved, but contains many of the same elements.

Sketch proof of Proposition 2.2. For *(1)* ⇒ *(2)*, a homomorphic embedding $G \hookrightarrow \prod_{\mathcal{U}} \mathrm{Sym}(n_i)$ can be lifted (non-uniquely) to a map $\psi = (\psi_i) : G \to \prod_{i\in I} \mathrm{Sym}(n_i)$, where $\psi_i : G \to \mathrm{Sym}(n_i)$. However, ψ may fail to be a group homomorphism. For all $a, b \in G$, $\psi(a)\psi(b)\psi(ab)^{-1}$ is an infinitesimal. This implies that for each finite set S and each $\delta > 0$, the set of i such that ψ_i is an (S, δ, n_i)-approximation is in the ultrafilter $\mathcal{U}$, and so is not empty. The second condition of the approximation is not immediately satisfied—one only gets that $d(\psi_i(g), \mathrm{id}) > \delta$ for $g \in S \smallsetminus \{e\}$. An 'amplification trick' improves this to $1 - \delta$.

For *(2)* ⇒ *(1)*, let $I = \{(S, \delta) \mid \text{ finite } S \subseteq G,\ \delta > 0\}$. For $(S, \delta) \in I$, define

$$\overline{(S, \delta)} := \{(S', \delta') \in I \mid \text{ finite } S' \supseteq S,\ \delta' \leq \delta\}.$$

The family $\mathcal{F}$ of all subsets $\overline{(S, \delta)}$ of I where $(S, \delta) \in I$ enjoys the finite intersection property since $\bigcap_{i=1}^{k} \overline{(S_i, \delta_i)} = \overline{(\bigcup_{i=1}^{k} S_i, \max_{i=1}^{k} \delta_i)}$. So there is an ultrafilter $\mathcal{U}$ on I with $\mathcal{F} \subseteq \mathcal{U}$. For all $i = (S, \delta) \in I$, let $\psi_i : G \to \mathrm{Sym}(n_i)$ be an (S, δ, n_i)-approximation. These maps combine in $g \mapsto (\psi_i(g))_{i\in I}$ to induce a monomorphism $G \hookrightarrow \prod_{\mathcal{U}} \mathrm{Sym}(n_i)$: it is a homomorphism because for all $g, h \in G$,

$$(\psi_i(g)\psi_i(h)\psi_i(gh)^{-1})_{i\in I} \in \mathcal{N}$$

since for all $\delta > 0$, $\{i \in I \mid d(\psi_i(g)\psi_i(h), \psi_i(gh)) < \delta\} \in \mathcal{U}$ as it is a superset of $\overline{(\{g, h, gh\}, \delta)} \in \mathcal{U}$; and it is injective because likewise for $\delta > 0$ and $g \in G \smallsetminus \{e\}$, the set

$$\{i \in I \mid d(\psi_i(g), \mathrm{id}_{n_i}) > 1 - \delta\} \in \mathcal{U}$$

and so $(\psi_i(g))_{i\in I} \notin \mathcal{N}$. □

Amenable groups are sofic, as are residually finite groups and, more generally, residually sofic groups. In particular, residually solvable groups are sofic—a fact we will use for Corollary 4.5. For further background, we refer to the surveys [2, 17, 18].

3. Variations on Higman's group

Our notation is $b^a = a^{-1}ba$ and $[a, b] = a^{-1}b^{-1}ab$.

As explained in Section 1, for a group G, subgroups A and B, an isomorphism $\phi : B \to A$, and a $k \in \mathbb{N}$, we define G_i, where $1 \le i \le k$, to be copies of G and $A_i, B_i \le G_i$ to be copies of its subgroups A and B. Then ϕ induces an isomorphism $\phi_i : B_i \to A_{i+1}$ and we define

$$\overline{\mathrm{Hig}}_k(G, \phi) \ := \ \langle G_1, \ldots, G_k \mid b_i = \phi_i(b_i) \text{ for all } i \text{ and all } b_i \in B_i\rangle.$$

Thus, $\overline{\mathrm{Hig}}_k(G, \phi)$ is the quotient of the free product of k copies of G in which B in the i-th is identified with A in the $(i + 1)$-st for $i = 0, \ldots, k - 1$ (indices modulo k).

By construction there are maps $\iota_1, \ldots, \iota_k$ from the group G to $\overline{\mathrm{Hig}}_k(G, \phi)$. We regard $\iota := \iota_1$ as the *natural map* $G \to \overline{\mathrm{Hig}}_k(G, \phi)$. We will often work in settings where these maps are injective, and then for simplicity we will suppress them and consider G as a subgroup of $\overline{\mathrm{Hig}}_k(G, \phi)$ via ι.

For example, if $G = \langle a_1, a_2 \mid R\rangle$ is a 2-generator group such that a_1 and a_2 have the same order, then $\overline{\mathrm{Hig}}_k(G, \phi)$, where $\phi : a_2 \mapsto a_1$, is the *cyclically presented group*

$$\langle a_1, \ldots, a_k \mid \sigma^i(r)\,;\ r \in R,\ i = 0, \ldots, k - 1\rangle,$$

where σ cycles the indices of the letters of r.

The semi-direct product of $\overline{\mathrm{Hig}}_k(G, \phi)$ with the cyclic group C_k of order k in which a generator t of C_k conjugates G_i to G_{i+1} (indices mod k) is

$$\mathrm{Hig}_k(G, \phi) \ = \ \langle G, t \mid t^k = 1,\ b^t = \phi(b); \forall b \in B\rangle.$$

Then $\overline{\mathrm{Hig}}_k(G, \phi)$ is the normal closure of $\iota(G)$ in $\mathrm{Hig}_k(G, \phi)$ and is the kernel of $\mathrm{Hig}_k(G, \phi) \to C_k$.

In the case when G is a group generated by two elements $a, b \in G$ of the same order, with $A = \langle a\rangle$, $B = \langle b\rangle$, and $\phi : B \to A$ given by $\phi(b) = a$, we will write $\mathrm{Hig}_k(G)$ and $\overline{\mathrm{Hig}}_k(G)$ in place of $\mathrm{Hig}_k(G, \phi)$ and $\overline{\mathrm{Hig}}_k(G, \phi)$.

The cases $k = 1, 2$ are degenerate:

Lemma 3.1. $\overline{\mathrm{Hig}}_1(G, \phi)$ *is a quotient of G. If G is generated by the subgroups A and B, then* $\overline{\mathrm{Hig}}_2(G, \phi)$ *is a quotient of G.*

For large k one expects G generally to embed in $\mathrm{Hig}_k(G, \phi)$, but this can fail:

Example 3.2. When $G = B = \mathbb{Z}$, $A = 2\mathbb{Z}$, and ϕ is multiplication by 2, $\mathrm{Hig}_k(G, \phi)$ is finite for all k, and so $\iota : G \not\hookrightarrow \mathrm{Hig}_k(G, \phi)$.

When $k \ge 4$, here is a sufficient condition:

Lemma 3.3. *If $A \cap B = \{1\}$ and $k \ge 4$, then G and $A * A$ both embed in* $\overline{\mathrm{Hig}}_k(G, \phi)$ *and so in* $\mathrm{Hig}_k(G, \phi)$*. In particular, if $G \ne \{1\}$, then* $\mathrm{Hig}_k(G, \phi)$ *is not amenable.*

The key idea behind this lemma is that $\overline{\mathrm{Hig}}_k(G,\phi)$ is the amalgamated free product of $G_1 *_{\phi_1} G_2 *_{\phi_2} \cdots *_{\phi_{k-3}} G_{k-2}$ and $G_{k-1} *_{\phi_{k-1}} G_k$ along the subgroup $\langle A_1, B_{k-2}\rangle = A_1 * B_{k-2}$, which is identified with $B_k * A_{k-1}$. A proof proceeds via the ideas we will use to establish Lemma 4.2.

The case $k = 3$ is trickier. Sometimes G does not embed in $\overline{\mathrm{Hig}}_3(G,\phi)$ because the latter group is very small—for example, $\overline{\mathrm{Hig}}_3(\mathrm{BS}(1,2)) = \{1\}$—but it is also possible that G embeds in $\overline{\mathrm{Hig}}_3(G,\phi)$, which is the case for most other examples considered in this paper.

4. Soficity via residual solvability

Here we prove:

Theorem 1.1. *Suppose G is a residually solvable group and ϕ is an isomorphism $B \to A$ between subgroups $A, B \leq G$. Suppose there exists a group homomorphism $\pi : G \to A \times B$ such that $\pi(a) = (a,1)$ for all $a \in A$ and $\pi(b) = (1,b)$ for all $b \in B$. Then* $\mathrm{Hig}_k(G,\phi)$ *and* $\overline{\mathrm{Hig}}_k(G,\phi)$ *are residually solvable for all $k \geq 4$.*

We use an approach which, similarly to our proof of Lemma 3.3, is based on viewing the amalgamated products as a combination of a free product and a semidirect product.

Define $G_A = \pi^{-1}(1,*)$ or, equivalently, $G_A = \ker(\phi_A \circ \pi)$, where ϕ_A is the projection $A \times B \to A$. So G_A is a normal subgroup of G and $G/G_A \simeq A$. The hypothesis that $\pi(a) = (a,1)$ for all $a \in A$ implies that A is a complement of G_A in G, and so G can be expressed as a semidirect product $G = A \ltimes G_A$. And $B \subseteq G_A$ because $\pi(b) = (1,b)$ for all $b \in B$. Likewise, $G = B \ltimes G_B$ with $A \subseteq G_B$.

As (3)–(5) of the following examples show, the hypotheses of Theorem 1.1 do not imply that A and B commute. Rather, they imply that $[A,B] \subseteq G_A \cap G_B = \ker \pi$.

Examples 4.1. In each case take $A = \langle a\rangle = \mathbb{Z}$ and $B = \langle b\rangle = \mathbb{Z}$:

(1) $G = \mathbb{Z}^2 = \langle a, b \mid ab = ba\rangle$. Take π to be the identity. The semi-direct products are direct products $\mathbb{Z} \times \mathbb{Z}$.
(2) $G = \mathrm{BS}(1,m) = \langle a, b \mid a^b = a^m\rangle$. In this case there is no map π for $m \neq 1$ because $[a,b] = a^{m-1}$, and so cannot be in $\ker \pi$.
(3) $G = \mathcal{H} = \langle a, b \mid [a,[a,b]] = [b,[a,b]] = 1\rangle$, the three-dimensional integral Heisenberg group. Take π to be the map onto $\mathbb{Z}^2 = \langle a, b \mid ab = ba\rangle$ quotienting by the center $\langle [a,b]\rangle$ of $\mathcal{H}$. Then $G_A = \langle b, [a,b]\rangle \simeq \mathbb{Z}^2$ and $G_B = \langle a, [a,b]\rangle \simeq \mathbb{Z}^2$.
(4) $G = \mathbb{Z} \wr \mathbb{Z} = \left\langle a, b \,\middle|\, \left[a^{b^i}, a^{b^j}\right] = 1 \text{ for all } i, j\right\rangle$, which is $\mathbb{Z} \ltimes \bigoplus_{i\in\mathbb{Z}} \mathbb{Z} = \langle b\rangle \ltimes \bigoplus_{i\in\mathbb{Z}} \langle a_i\rangle$ where $a_i = a^{b^i}$ and b acts so as to map $a_i \mapsto a_{i+1}$. Again, take π to be the abelianization map onto $\mathbb{Z}^2 = \langle a, b \mid ab = ba\rangle$. Then G_B is the kernel of the map $\mathbb{Z} \wr \mathbb{Z} \twoheadrightarrow \langle b\rangle$ given by quotienting by a, which is $\bigoplus_{i\in\mathbb{Z}} \mathbb{Z} = \bigoplus_{i\in\mathbb{Z}} \langle a_i\rangle$. And G_A is the kernel of the map $\mathbb{Z} \wr \mathbb{Z} \twoheadrightarrow \langle a\rangle$ given by quotienting by b, which is $\langle b\rangle \ltimes \bigoplus_{i\in\mathbb{Z}} \langle a_i^{-1} a_{i+1}\rangle$ and is isomorphic to G.
(5) $G = \mathcal{M} = \left\langle a, b \,\middle|\, \left[[a,b], [a,b]^{a^i b^j}\right] = 1 \ \forall i, j \in \mathbb{Z}\right\rangle$, the free metabelian group on two generators. Again, we take π to be the abelianization map onto $\mathbb{Z}^2 = \langle a, b \mid ab = ba\rangle$.

We will use the following description of amalgamated products over subgroups which have a normal complement.

Lemma 4.2. *Suppose G_1 and G_2 are groups having subgroups H_1 and H_2 respectively with normal complements—i.e., $G_1 = H_1 \ltimes N_1$ and $G_2 = H_2 \ltimes N_2$ for some N_1 and N_2. For any isomorphism $\phi : H_1 \to H_2$, the amalgamated product $G_1 *_\phi G_2$ can be expressed as a semidirect product $H_1 \ltimes (N_1 * N_2)$, where the action of H_1 on N_2 comes from that of H_2 via the isomorphism ϕ.*

Proof. An arbitrary element of $G_1 *_\phi G_2$ can be represented as a product

$$w = x_1 y_1 x_2 y_2 \dots x_r y_r,$$

where $x_1, \dots, x_r \in G_1$ and $y_1, \dots, y_r \in G_2$. Express x_1 as $m_1 h_1$ where $m_1 \in G_1$ and $h_1 \in H_1$. Since we are working in the amalgamated product, we can move h_1 to G_2 and write

$$w = m_1\, (\phi(h_1) y_1)\, x_2 y_2 \dots x_r y_r.$$

The element $\phi(h_1)y_1$ in G_2 can then be expressed as $n_1 g_1$ where $n_1 \in N_2$ and $g_1 \in H_2$. Continuing this process, moving elements from H_1 or H_2 to the right, expresses w as

$$w = m_1 n_1 m_2 n_2 \dots m_r n_r h, \tag{1}$$

where $m_1, \dots, m_r \in N_1 \subseteq G_1$, $n_1, \dots, n_r \in N_2 \subseteq G_2$, and $h \in H_1$. The product $m_1 n_1 \dots m_r n_r$ can be considered as an element in $N_1 * N_2$. Such elements form a normal subgroup in $G_1 *_\phi G_2$, with quotient H_1. All that remains to check is that the action of H_1 on the free product is the one described. □

Corollary 4.3. *Suppose G_1 and G_2 are residually solvable groups satisfying the conditions of Lemma 4.2. Then the amalgamated product $G_1 *_\phi G_2$ is residually solvable.*

Proof. Free products of residually solvable groups are residually solvable (see e.g. [11]), but semidirect products of residually solvable groups can fail to be residually solvable. Nevertheless we will see that the semidirect products of Lemma 4.2 are residually solvable.

Let $w = m_1 n_1 m_2 n_2 \dots m_r n_r h$ be a non-trivial element in $G_1 *_\phi G_2$ as per (1), where all $m_i \in G_1$ and $n_i \in G_2$ are non-identity, with the possible exceptions of m_1 and n_r, with $h \in H_1$. If $h \neq 1$, then there is a solvable quotient $H_1 \to \overline{H}$ of H_1 where h survives (since subgroups of residually solvable groups are residually solvable), which leads to a quotient $G_1 *_\phi G_2$, where w has a nontrivial image. Therefore it suffices to consider the case $h = 1$.

Take k such that for $i = 1, 2$, $G_i \to \overline{G}_i := G_i / G_i^{(k)}$ are quotients of G_i by some derived subgroup such that all the m_i and n_i have nontrivial images in $\overline{G}_1$ and $\overline{G}_2$. Let $\overline{H}_1$, $\overline{N}_1$ and $\overline{N}_2$ denote the (necessarily solvable) images of H_1, N_1 and N_2, respectively, in $\overline{G}_1$ and $\overline{G}_2$. We can view w as element in the free product $\overline{N}_1 * \overline{N}_2$. Therefore, by the argument that free products of solvable groups are residually solvable (see e.g. [11]), there exists a quotient $\overline{\overline{N}_1 * \overline{N}_2}$ of $\overline{N}_1 * \overline{N}_2$ by one of its derived subgroups where w is non-trivial. Since this quotient is characteristic, it has a natural action of $\overline{H}_1$ which extends the actions of $\overline{H}_1$ on $\overline{N}_1$ and on $\overline{N}_2$. This allows us to map

$$G_1 *_\phi G_2 = H_1 \ltimes (N_1 * N_2) \to \overline{H}_1 \ltimes (\overline{N}_1 * \overline{N}_2) \to \overline{H}_1 \ltimes \overline{\overline{N}_1 * \overline{N}_2},$$

where $\overline{H}_1 \ltimes \overline{\overline{N}_1 * \overline{N}_2}$ is a solvable quotient of $G *_\phi G$ in which w has a nontrivial image. □

Lemma 4.4. *Suppose G is a group satisfying the conditions in Theorem 1.1. Then the amalgamated product $G *_\phi G$ can be written as a semidirect product $(A * B) \ltimes H$ for some normal subgroup H, and therefore there is a projection $G *_\phi G \to A * B$.*

Proof. Express $G *_\phi G$ as $B \ltimes (G_A * G_B)$ as per Lemma 4.2. Then via annihilating the first factor in $B \ltimes (G_A * G_B)$ and using the maps $G_A \to B$ and $G_B \to A$ induced by π, $G *_\phi G$ maps to $A * B$. This map is clearly surjective with some kernel H and restricts to the identity on $A * B$ (viewed as a subgroup of $G_A * G_B$ via $B \le G_A$ and $A \le G_B$), so splits $G *_\phi G$ into a semidirect product. □

Proof of Theorem 1.1. Applying Corollary 4.3 and Lemma 4.4 repeatedly, we find that if $k \ge 4$, then the groups $J := G_1 *_{\phi_1} G_2 *_{\phi_2} \cdots *_{\phi_{k-3}} G_{k-2}$ and $K := G_{k-1} *_{\phi_{k-1}} G_k$ (in the notation of Section 3) are both residually solvable, and both contain $A * B$ in such a way that they both split over this group as semidirect products, and $\overline{\mathrm{Hig}}_k(G, \phi) = J *_{A*B} K$. So the hypotheses of Lemma 4.2 are met and $\overline{\mathrm{Hig}}_k(G, \phi)$ is residually finite by a final application of Corollary 4.3.

Finally, $\mathrm{Hig}_k(G, \phi) = \overline{\mathrm{Hig}}_k(G, \phi) \rtimes C_k$, so it is also residually solvable. (Semidirect products $H \rtimes A$ of residually solvable groups H and solvable groups A are residually solvable.) □

Theorem 1.1 may also hold when 'residually solvable' is replaced with 'residually nilpotent' or 'residually finite'; however, our proof would need further ideas and the given theorem suffices for our application:

Corollary 4.5. *When G is $\mathbb{Z}^2$, $\mathcal{H}$, $\mathbb{Z} \wr \mathbb{Z}$, or $\mathcal{M}$ as per Examples 4.1, $\overline{\mathrm{Hig}}_k(G)$ is residually solvable, and so sofic, for all $k \ge 4$.*

Finally, we remark on an alternative route. The following result is weaker than Theorem 1.1 in that it does not tell us that $\mathrm{Hig}_k(G, \phi)$ is sofic or residually solvable. But this proposition would suffice for our applications in Section 6 because it tells us that when G is sofic, $\mathrm{Hig}_k(G, \phi)$ has a sofic quotient into which G injects (condition (1) of Theorem 1.5). Moreover, it makes sofic approximations of that quotient easy to construct explicitly from sofic approximations of G.

Proposition 4.6. *Suppose there exists a homomorphism $\pi : G \to A \times B$ as per Theorem 1.1. Then for all $k \ge 3$, there are homomorphisms $\mu : \mathrm{Hig}_k(G, \phi) \to C_k \ltimes G^k$ and $\overline{\mu} : \overline{\mathrm{Hig}}_k(G, \phi) \to G^k$. Moreover, the restrictions of μ and $\overline{\mu}$ to any copy G_i of G inside $\mathrm{Hig}_k(G, \phi)$ are injective.*

Proof. Let $\pi_A : G \to A$ (respectively, $\pi_B : G \to B$) be the composition of π with projection onto A (respectively, B). Define the homomorphism $\overline{\mu} : \overline{\mathrm{Hig}}_k(G, \phi) \to G^k$, given by

$$\overline{\mu}(\iota_l(g)) = (1, \ldots, 1, \phi^{-1}(\pi_A(g)), g, \phi(\pi_B(g)), 1, \ldots, 1),$$

where g is an arbitrary element of G, and $\iota_l(g)$ is the element in $\overline{\mathrm{Hig}}_k(G)$ corresponding to g sitting in the l-th copy of G. The elements $\phi(\pi_B(g))$, g, and $\phi^{-1}(\pi_A(g))$ are sitting in coordinates $l-1$, l and $l+1$. Clearly $\overline{\mu}$ is well defined on each copy G_i appearing in the presentation of $\overline{\mathrm{Hig}}_k(G)$, so it suffices to verify that $\overline{\mu}$ identifies the l-th copy of B with the $l+1$-st copy of A. By definition we have

$$\overline{\mu}(\iota_l(b)) = (1, \ldots, 1, \phi^{-1}(\pi_A(b)), b, \phi(\pi_B(b)), 1, \ldots, 1) = (1, \ldots, 1, 1, b, \phi(b), 1, \ldots, 1)$$

$$\overline{\mu}(\iota_{l+1}(a)) = (1, \ldots, 1, \phi^{-1}(\pi_A(a)), a, \phi(\pi_B(a)), 1, \ldots, 1) = (1, \ldots, 1, \phi^{-1}(a), a, 1, 1, \ldots, 1),$$

and thus $\overline{\mu}(\iota_l(b)) = \overline{\mu}(\iota_{l+1}(\phi(b)))$, i.e., $\overline{\mu}$ extends to the group $\mathrm{Hig}_k(G)$. By construction, the restriction of $\overline{\mu}$ on each copy of G is injective. (Unless we are in a degenerate case, the maps μ and $\overline{\mu}$ are not injective.) □

Remark 4.7. If we remove the defining relator $t^k = 1$ from our presentation for $\mathrm{Hig}_k(G, \phi)$, then we get the HNN-extension $\langle G, t \mid b^t = \phi(b)\ \forall b \in B\rangle$, which by [4, Corollary 3.6] is sofic when G is solvable. When $G = \langle a, b \mid b^a = b^2\rangle$ and $\phi : b \mapsto a$, for example, this is Baumslag's one-relator group $\langle b, t \mid b^{b^t} = b^2\rangle$.

5. Sofic quotients and almost conjugation

This section is devoted to proving Theorem 1.5 relating soficity in the context of $\mathrm{Hig}_k(G, \phi)$ to seemingly pathological permutations f. These f come from permutations approximating $t \in \mathrm{Hig}_k(G, \phi)$. They will have order dividing k since $t^k = 1$ and, for all $b \in B$, will 'almost conjugate' permutations approximating b to permutations approximating $\phi(b)$ since $b^t = \phi(b)$ in $\mathrm{Hig}_k(G, \phi)$. When G is amenable and we have explicit sofic approximations for G, the permutations approximating b and $\phi(b)$ in $\mathrm{Hig}_k(G, \phi)$ essentially have to be those sofic approximations. In examples, the 'almost conjugate' conclusion will then amount to a local recurrence such as $f(x + 1) = mf(x)$ holding for most values of x.

We make Theorem 1.5 precise as:

Theorem 5.1. *Suppose G is a group, ϕ is an isomorphism $B \to A$ between subgroups $A, B \leq G$, and $k \geq 1$ is an integer. The following two conditions are equivalent.*

(1) $\mathrm{Hig}_k(G, \phi)$ *has a sofic quotient Q such that the composition $G \xrightarrow{\iota} \mathrm{Hig}_k(G, \phi) \xrightarrow{\pi} Q$ is injective.*
(2) *For all finite subsets $S \subseteq G$ and all $\delta, \varepsilon > 0$, there exists an (S, δ, n)-approximation ψ of G and a permutation $f \in \mathrm{Sym}(n)$ of order dividing k such that for all $b \in S \cap \phi^{-1}(A \cap S)$,*

$$d(\psi(b) \circ f, f \circ \psi(\phi(b))) < \varepsilon.$$

If G is amenable, then these are also equivalent to:

(3) *For all finite sets $S \subseteq G$ and all $\varepsilon > 0$, there exist a finite set $S' \subseteq G$ with $S \subseteq S'$ and $\delta > 0$ and an integer N such that if ψ is an (S', δ, n)-sofic approximation of G with $n > N$, then there exists a permutation $f \in \mathrm{Sym}(n)$ of order dividing k such that for all $b \in S \cap \phi^{-1}(A \cap S)$,*

$$d(\psi(b) \circ f, f \circ \psi(\phi(b))) < \varepsilon.$$

Proof of Theorem 5.1, (1) ⇒ (2). We have that there is a sofic quotient Q such that the composition of the natural map $G \xrightarrow{\iota} \mathrm{Hig}_k(G, \phi)$ with the quotient map $\pi : \mathrm{Hig}_k(G, \phi) \to Q$ is injective. In particular, the map ι is injective.

Suppose $S \subseteq G$ is a finite subset and $\varepsilon, \delta > 0$. We seek an n and an (S, δ, n)-approximation ψ of G together with a permutation $f \in \mathrm{Sym}(n)$ of order dividing k such that $d(\psi(\phi(b)) \circ f, f \circ \psi(b)) < \varepsilon$ for all $b \in S \cap \phi^{-1}(A \cap S)$.

Let

$$S' = \left\{\mathrm{id}, t, \ldots, t^{k-1}\right\} \cup \iota(S) \cup \iota\left(S \cap \phi^{-1}(A \cap S)\right) t \subseteq \mathrm{Hig}_k(G, \phi).$$

Let $\delta' = \min\{\delta, \varepsilon\}/6k$. Then $\pi(S')$ is a finite subset of the sofic group Q, so there exists an $n \in \mathbb{N}$ and an $(\pi(S'), \delta', n)$-approximation $Q \to \mathrm{Sym}(m)$. Via π this gives a map $\psi' : \mathrm{Hig}_k(G, \phi) \to \mathrm{Sym}(n)$ which enjoys the first defining property of an (S', δ', n)-approximation, but may fail the second as it could map some elements of S' to the identity. Since G naturally maps into $\mathrm{Hig}_k(G, \phi)$ and $\delta' < \delta$, the composition ψ of ι and ψ' is an (S, δ, n)-approximation of G, as required.

We will obtain the requisite permutation $f \in \mathrm{Sym}(n)$ from the action of t under ψ'. First set $\tilde{f} = \psi'(t)$. The order of this permutation may fail to divide k since ψ' is not necessarily a homomorphism. However,

$$d(\tilde{f}^k, \mathrm{id}) = d(\psi'(t)^k, \mathrm{id}) \leq d(\psi'(t)^k, \psi'(t^k)) + d(\psi'(t^k), \mathrm{id}) < (k-1)\delta' + \delta' = k\delta',$$

where the second inequality holds because $t^k = \mathrm{id}$ and $t^i \in S'$ for all i. Therefore the set of points which are not part of a cycle of length dividing k under the action of $\tilde{f}$ is correspondingly small and we can find a permutation f of order dividing k such that $d(f, \tilde{f}) < k\delta'$.

Suppose $b \in S \cap \phi^{-1}(A \cap S)$. It remains to show that

$$d\left(\psi(b) \circ f, f \circ \psi(\phi(b))\right) \leq \varepsilon.$$

As $d(f, \tilde{f}) \leq k\delta'$, Lemma 2.1 (iii) yields

$$d\left(f^{-1} \circ \psi(b) \circ f, \tilde{f}^{-1} \circ \psi(b) \circ \tilde{f}\right) < 2k\delta'. \tag{2}$$

By definition of $\tilde{f}$,

$$\tilde{f}^{-1} \circ \psi(b) \circ \tilde{f} = \psi'(t)^{-1} \circ \psi'(\iota(b)) \circ \psi'(t). \tag{3}$$

Now, as $t, t^{-1}, \mathrm{id} \in S'$ and ψ' is an (S', δ', n)-approximation,

$$d\left(\psi'(t)^{-1} \circ \psi'(\iota(b)) \circ \psi'(t), \psi'(t^{-1}) \circ \psi'(\iota(b)) \circ \psi'(t)\right) = d\left(\psi'(t)^{-1}, \psi'(t^{-1})\right) \leq 2\delta'. \tag{4}$$

And, likewise, as $t^{-1}, b, t, bt, t^{-1}bt \in S'$ and $\phi(\iota(b)) = t^{-1}\iota(b)t$ in $\mathrm{Hig}_k(G, \phi)$,

$$d(\psi'(t^{-1}) \circ \psi'(\iota(b)) \circ \psi'(t), \psi'(\iota(\phi(b)))) = d\left(\psi'(t^{-1}) \circ \psi'(\iota(b)) \circ \psi'(t), \psi'(t^{-1}\iota(b)t)\right) \leq 2\delta'. \tag{5}$$

In combination, (2)–(5) yield the first inequality of:

$$d\left(\psi(b) \circ f, f \circ \psi(\phi(b))\right) = d\left(f^{-1} \circ \psi(b) \circ f, \psi(\phi(b))\right) \leq (2k+4)\delta' \leq 6k\delta' \leq \varepsilon.$$

□

The following lemma will provide the heart of our proof that (2) ⇒ (1). Let $\{G, t\}^*$ denote the set of words on $G \cup \{t\}$. Since $\mathrm{Hig}_k(G, \phi)$ is generated by $\iota(G)$ and t, we can choose a section $\sigma : \mathrm{Hig}_k(G, \phi) \to \{G, t\}^*$ for the evaluation map $\{G, t\}^* \to \mathrm{Hig}_k(G, \phi)$—that is, for every $g \in \mathrm{Hig}_k(G, \phi)$ we choose a way of expressing g as a product $\sigma(g) = \iota(g_1)t^{j_1} \cdots \iota(g_r)t^{j_r}$ of elements of G and powers of t.

Given a map $\psi : G \to \mathrm{Sym}(n)$ (not necessarily a homomorphism) and a permutation $f \in \mathrm{Sym}(n)$, define a map $\psi^f : \mathrm{Hig}_k(G, \phi) \to \mathrm{Sym}(n)$ by

$$\psi^f(g) := \psi(g_1)f_i^{j_1} \cdots \psi(g_r)f_i^{j_r},$$

where $\sigma(g) = \iota(g_1)t^{j_1} \cdots \iota(g_r)t^{j_r}$. The lemma will tell us that if ψ and f are suitably compatible then ψ^f is close to a homomorphism.

Lemma 5.2. *For all finite subsets $S \subseteq \mathrm{Hig}_k(G, \phi)$ and $\overline{S} \subseteq G$ such that $\iota(\overline{S}) \subseteq S$ and all $\delta > 0$, there exists a finite set $S_0 \subseteq G$ with $\overline{S} \subseteq S_0$ and an $\varepsilon > 0$ satisfying the following. Suppose $\psi : G \to \mathrm{Sym}(n)$ is an (S_0, ε, n)-approximation and $f \in \mathrm{Sym}(n)$ is a permutation of order dividing k such that for all $b \in S_0 \cap B$*

$$d(\psi(\phi(b)) \circ f, f \circ \psi(b)) \ < \ \varepsilon.$$

Then for all $s_1, s_2 \in S$ for which $s_1 s_2 \in S$,

$$d\left(\psi^f(s_1)\psi^f(s_2), \psi^f(s_1 s_2)\right) \ < \ \delta \tag{6}$$

and for all $g \in \overline{S}$

$$d\left(\psi^f(\iota(g)), \psi(g)\right) \ < \ \delta. \tag{7}$$

Proof. Since S is finite there exists an integer m and a finite subset $S' \subseteq G$ containing $\overline{S}$ such that $\sigma(S) \subseteq \{\iota(S'), t\}^m$. Then S sits inside the subgroup $\Gamma = \langle \iota(S'), t \rangle$ of $\mathrm{Hig}_k(G, \phi)$. As Γ is finitely generated, there exists a finitely presented group $\Gamma' = \langle S', t \mid R' \rangle$ which projects onto Γ—that is, the composition $S' \hookrightarrow \Gamma' \twoheadrightarrow \Gamma$ is the identity.

By construction, every relation in R' is also satisfied in $\mathrm{Hig}_k(G, \phi)$, and so can be deduced from the defining relations in the presentation of $\mathrm{Hig}_k(G, \phi)$. These defining relations come in three types: relations in G, the relation $t^k = 1$, and relations of the form $\iota(b)^t = \iota(\phi(b))$ for some $b \in B$. We can enlarge the set S' to another finite subset $S'' \subseteq G$ by gathering all elements in G needed to deduce all the relations $r \in R'$, so as to view S as a subset of a finitely presented group

$$\Gamma'' \ = \ \langle S'', t \mid t^k,\ R'',\ b^t \phi(b)^{-1} \text{ for } b \in B'' \rangle$$

where R'' is a finite set of relations satisfied in the subgroup $\langle S'' \rangle$ of G, and B'' is a finite subset of B. Let $N \geq k$ be a number such that every defining relation in R'' has length at most N in the generating set S'' and all elements in B'' and $\phi(B'')$ can be expressed as words in S'' of length at most $N - 1$. By construction, there exists a constant M such that each relator in Γ'' of the form $s_3^{-1} s_1 s_2$ for $s_1, s_2, s_3 \in S$ or of the form or $g^{-1}\sigma(\iota(g))$ for $g \in \overline{S}$ can be written as product of at most M conjugates of the defining relators in the above presentation.

Define $S_0 = (S'')^N$ and $\varepsilon = \delta/8MN$. Suppose that ψ is an (S_0, ε, n)-approximation of G and $f \in \mathrm{Sym}(n)$ is a permutation of order dividing k such that $d(\psi(\phi(b)) \circ f, f \circ \psi(b)) < \varepsilon$ for all $b \in S_0 \cap \phi^{-1}(A \cap S)$. Extend $\psi|_{S''}$ to a homomorphism $\tilde{\psi}$ from the free group generated by S'' and t to $\mathrm{Sym}(n)$ by mapping t to the permutation f. Defining relations $r = t^k$ or $r \in R''$ or $r = b^t \phi(b)^{-1}$ in our presentation of Γ'' have lengths at most k, N and $2N$, respectively, and so $d(\tilde{\psi}(r), \mathrm{id}) \leq 2N\varepsilon$ by Lemma 2.1 (i). It then follows from Lemma 2.1 (i) and (ii) that for all relators r' of the form $s_3^{-1} s_1 s_2$ or $g^{-1}\sigma(\iota(g))$, we have

$$d(\tilde{\psi}(r'), \mathrm{id}) \ < \ 2MN\varepsilon \ < \ \delta/4. \tag{8}$$

For the relators of the first type this gives us that $d(\tilde{\psi}(s_3^{-1} s_1 s_2), \mathrm{id}) \ < \ \delta/4$. For those of second type we get both (7) for all $g \in \overline{S}$, and

$$d(\tilde{\psi}(s_i), \psi^f(s_i)) \ < \ \delta/4 \tag{9}$$

for $i = 1, 2, 3$. Then (8) applied to $r' = s_3^{-1} s_1 s_2$ and (9) give

$$\begin{aligned} d(\psi^f(s_3), \psi^f(s_1)\psi^f(s_2)) &= d(\psi^f(s_3)^{-1}\psi^f(s_1)\psi^f(s_2), \mathrm{id}) \\ &< d(\tilde{\psi}(s_3)^{-1}\tilde{\psi}(s_1)\tilde{\psi}(s_2), \mathrm{id}) + \frac{3\delta}{4} \\ &= d(\tilde{\psi}(s_3^{-1} s_1 s_2), \mathrm{id}) + \frac{3\delta}{4} \\ &< \delta, \end{aligned}$$

which yields inequality (6). □

Proof of Theorem 5.1, (2) ⇒ (1). This proof is similar to that of (2) ⇒ (1) of Proposition 2.2. Define

$$I := \left\{ \left(S, \overline{S}, \delta\right) \middle| \text{ finite } S \subseteq \mathrm{Hig}_k(G, \phi), \text{ finite } \overline{S} \subseteq G \text{ with } \iota(\overline{S}) \subseteq S,\ \delta > 0 \right\}.$$

For $\left(S, \overline{S}, \delta\right) \in I$, define

$$\overline{\left(S, \overline{S}, \delta\right)} := \left\{ \left(S', \overline{S}', \delta'\right) \in I \middle| S' \supseteq S,\ \overline{S}' \supseteq \overline{S},\ \delta' \leq \delta \right\}.$$

As in our proof of Proposition 2.2, the family $\mathcal{F}$ of all subsets $\overline{\left(S, \overline{S}, \delta\right)}$ enjoys the finite intersection property, and so there is an ultrafilter $\mathcal{U}$ on I with $\mathcal{F} \subseteq \mathcal{U}$.

Suppose $i = \left(S, \overline{S}, \delta\right) \in I$. Let $S_0 \subseteq G$ and $\varepsilon > 0$ be as per Lemma 5.2. Let ψ_i be an (S_0, ε, n_i)-approximation of G and $f_i \in \mathrm{Sym}(n_i)$ a permutation as per condition (2). Together ψ_i and f_i define maps $\psi_i^{f_i} : \mathrm{Hig}_k(G, \phi) \to \mathrm{Sym}(n_i)$; Lemma 5.2 tells us that these $\psi_i^{f_i}$ enjoy conditions (6) and (7).

If $g \in \overline{S} \subseteq S_0$, then (7) gives us that $d\left(\psi_i^{f_i}(\iota(g)), \psi(g)\right) < \delta$. If, additionally, $g \neq e$, then $d(\psi_i(g), \mathrm{id}) > 1 - \varepsilon$ because ψ_i is an (S_0, ε, n_i)-approximation. Together these give

$$d(\psi_i^{f_i}(\iota(g)), \mathrm{id}) > 1 - \delta - \varepsilon \tag{10}$$

for all $g \in \overline{S} \smallsetminus \{e\}$.

The $\left\{\psi_i^{f_i}\right\}_{i \in I}$ combine to induce a map

$$\Psi^{\boldsymbol{f}} : \mathrm{Hig}_k(G, \phi) \to \prod_{\mathcal{U}} \mathrm{Sym}(n_i).$$

This is a group homomorphism because of condition (6). Its image $Q = \Psi^{\boldsymbol{f}}(\mathrm{Hig}_k(G, \phi))$ is a sofic quotient of $\mathrm{Hig}_k(G, \phi)$. In general, $\Psi^{\boldsymbol{f}}$ might not be injective, but (10) tells us that the composition $G \xrightarrow{\iota} \mathrm{Hig}_k(G, \phi) \xrightarrow{\Psi^{\boldsymbol{f}}} Q$ is injective. In both cases, the details are similar to our derivations of corresponding statements in our proof of Proposition 2.2. □

Proof of Theorem 5.1, (3) ⇒ (2). This implication is immediate since G is sofic. □

The remaining implication (2) ⇒ (3) is significantly more complicated and uses that for an amenable group, any two approximations into the same $\mathrm{Sym}(n)$ are almost conjugate. This result is due to Helfgott and Juschenko in the form given but, as they explain, has origins in Elek and Szabo [6], builds on a lemma from Kerr and Li [15], and is also comparable to

Arzhantseva and Păunescu [1]. Helfgott and Juschenko's proof is a delicate analysis of the interplay between sofic approximations and the Følner characterization of amenability.

Theorem 5.3 (Helfgott–Juschenko [12])**.** *Suppose G is an amenable group, $\varepsilon > 0$, and S is a finite subset of G. Then there is a finite subset $S' \subseteq G$ with $S \subseteq S'$ and constants $N \in \mathbb{Z}^+, \delta > 0$ such that for any two (S', δ, n)-approximations ρ_1, ρ_2 of G with $n \geq N$, there exists $\tau \in \mathrm{Sym}(n)$ such that, for every $s \in S$,*

$$d(\tau^{-1} \circ \rho_1(s) \circ \tau, \rho_2(s)) < \varepsilon.$$

We will also use the following lemma which essentially says that the n in the definition of an (S, δ, n)-approximation is irrelevant provided it is sufficiently large. We omit its proof, which is routine.

Lemma 5.4. *Suppose $n = qm + r$ where m, n, q, r are non-negative integers with $m, n \geq 1$ and $q = \lfloor n/m \rfloor$. If $\alpha : S \to \mathrm{Sym}(m)$ is an (S, η, m)-approximation of a finite subset S of a group, then composing*

$$(\overbrace{\alpha, \dots, \alpha}^{q}, 1) : S \to \overbrace{\mathrm{Sym}(m) \times \cdots \times \mathrm{Sym}(m)}^{q} \times \mathrm{Sym}(r)$$

with the diagonal embedding into $\mathrm{Sym}(n)$ gives an $(S, \eta + \dfrac{1}{q+1}, n)$-approximation β.

Proof of Theorem 5.1 (2) ⇒ (3). We are given a finite set $S \subseteq G$ and some $\varepsilon > 0$.

We aim to show that for a suitable finite set $S' \subseteq G$ with $S \subseteq S'$ and suitable $\delta > 0$ and N, every (S', δ, n)-sofic approximation ψ of G with $n > N$ admits some $f \in \mathrm{Sym}(n)$ of order dividing k almost conjugating the action of A under ψ to the action of B under ψ. The idea will be to apply condition (2) and Lemma 5.4 to obtain some approximation $\overline{\psi}$ of G together with a permutation $\overline{f}$ of order k which will almost conjugate the action of A to the action of B. A priori ψ and $\overline{\psi}$ will be unrelated, but in fact by Theorem 5.3 will essentially be conjugate. We will apply this conjugation to $\overline{f}$ to obtain the requisite permutation f.

Here are the details. Let $\tilde{\varepsilon} = \varepsilon/3$. By Theorem 5.3 there exits a finite subset $S' \subseteq G$ with $S \cup \phi(S \cap B) \subseteq S'$ and $\tilde{\delta} > 0$ and $N_0 \in \mathbb{Z}^+$ such that any two $(S', \tilde{\delta}, n)$-approximations ρ_1 and ρ_2 of G with $n \geq N_0$ are *almost conjugate* in that there exists $\tau \in \mathrm{Sym}(n)$ such that for all $s \in S \cup \phi(S \cap B)$,

$$d(\tau^{-1} \circ \rho_1(s) \circ \tau,\ \rho_2(s)) < \tilde{\varepsilon}. \tag{11}$$

Let $\delta = \min\{\tilde{\delta}, \tilde{\varepsilon}\}$.

By Condition (2), there exists an $(S', \delta/2, m)$-approximation ψ' of G together with a permutation $f \in \mathrm{Sym}(m)$ of order dividing k such that for all $b \in S' \cap \phi^{-1}(A \cap S')$,

$$d(\psi(b) \circ f,\ f \circ \psi(\phi(b))) < \tilde{\varepsilon}. \tag{12}$$

Let $N = \max\{N_0, 2m/\delta\}$. With S' and δ as defined above, suppose ψ is an (S', δ, n)-approximation of G with $n > N$.

Via Lemma 5.4, we can use ψ' to construct another (S', δ, n)-approximation $\overline{\psi}$ of G and an associated permutation $\overline{f}$ which almost conjugates the action of A to B with the same error $\tilde{\varepsilon}$:

$$d(\overline{\psi}(b) \circ \overline{f},\ \overline{f} \circ \overline{\psi}(\phi(b))) < \tilde{\varepsilon} \tag{13}$$

for all $b \in S' \cap B$. This is possible because

$$\frac{\delta}{2} + \frac{1}{\lfloor n/m \rfloor + 1} \ \leq \ \frac{\delta}{2} + \frac{1}{\lfloor 2/\delta \rfloor + 1} \ < \ \frac{\delta}{2} + \frac{\delta}{2} \ \leq \ \delta$$

and given how $\overline{\psi}$ is assembled from copies of ψ and the identity (and correspondingly $\overline{f}$ from copies of f and the identity), the error $\tilde{\varepsilon}$ of (13) does not increase and $\overline{f}$, like f, has order dividing k.

By Theorem 5.3 there is a permutation $\tau \in \mathrm{Sym}(n)$ which almost conjugates ψ to $\overline{\psi}$—i.e.,

$$d\left(\tau^{-1} \circ \psi(s) \circ \tau,\ \overline{\psi}(s)\right) \ < \ \tilde{\varepsilon} \tag{14}$$

for all $s \in S'$.

Define $f = \tau \circ \overline{f} \circ \tau^{-1}$, which is a permutation of order dividing k since $\overline{f}$ has order dividing k. Suppose $b \in S \cap \phi^{-1}(A \cap S)$. We will complete our proof by showing that

$$d(\psi(b) \circ f,\ f \circ \psi(\phi(b)) \ < \ \varepsilon.$$

By definition of f,

$$f^{-1} \circ \psi(b) \circ f \ = \ \tau \circ \overline{f}^{-1} \circ \tau^{-1} \circ \psi(b) \circ \tau \circ \overline{f} \circ \tau^{-1}. \tag{15}$$

Since $b \in S \subseteq S'$, by (14),

$$d\left(\tau \circ \overline{f}^{-1} \circ \tau^{-1} \circ \psi(b) \circ \tau \circ \overline{f} \circ \tau^{-1}, \tau \circ \overline{f}^{-1} \circ \overline{\psi}(b) \circ \overline{f} \circ \tau^{-1}\right) \ < \ \tilde{\varepsilon}. \tag{16}$$

By (13), $d\left(\overline{f}^{-1} \circ \overline{\psi}(b) \circ \overline{f},\ \overline{\psi}(\phi(b))\right) < \tilde{\varepsilon}$, and therefore

$$d\left(\tau \circ \overline{f}^{-1} \circ \overline{\psi}(b) \circ \overline{f} \circ \tau^{-1}, \tau \circ \overline{\psi}(\phi(b)) \circ \tau^{-1}\right) \ < \ \tilde{\varepsilon}. \tag{17}$$

Since $\phi(b) \in S \cap B \subseteq S'$, by (14) again,

$$d\left(\tau \circ \overline{\psi}(\phi(b)) \circ \tau^{-1}, \psi(\phi(b))\right) \ < \ \tilde{\varepsilon}. \tag{18}$$

Together, (15)–(18) yield the first inequality of:

$$d(\psi(b) \circ f,\ f \circ \psi(\phi(b)) \ = \ d\left(f^{-1} \circ \psi(b) \circ f, \psi(\phi(b))\right) \ < \ 3\tilde{\varepsilon} \ = \ \varepsilon,$$

which completes the proof. □

6. Applications of Theorem 5.1

In this section we will examine the groups $\mathbb{Z}^2$, the 3-dimensional integral Heisenberg group $\mathcal{H}$, $\mathrm{BS}(1, m)$, $\mathbb{Z} \wr \mathbb{Z}$, and the 2-generator metabelian group in the context of Theorem 1.5 (or, in its precise form, Theorem 5.1). Each of these groups is amenable. We will exhibit families of maps witnessing to their soficity, and will then explain what Theorem 5.1 allows us to conclude about the existence of seemingly pathological permutations. In particular, we will explain how the case of $\mathrm{BS}(1, m)$ yields Theorems 1.2 and 1.4, and how $\mathbb{Z} \wr \mathbb{Z}$ yields Theorem 1.3.

We begin with $\mathbb{Z}^2$, which we view as an introductory example.

6.1. $\mathbb{Z}^2$**.** We present $\mathbb{Z}^2$ as $\langle a, b \mid ab = ba \rangle$, so $\phi : \mathbb{Z} \to \mathbb{Z}$, given by $b \mapsto a$, is the map defining $\mathrm{Hig}_k(\mathbb{Z}^2)$.

To obtain a family of functions witnessing to the soficity of $\mathbb{Z}^2$, we identify $\mathrm{Sym}(n)$ with $\mathrm{Sym}(\mathbb{Z}/n\mathbb{Z})$ and then for $p, q \in \mathbb{N}$, define $\psi_{n,p,q} : G \to \mathrm{Sym}(n)$ by

$$\psi_{n,p,q}(a) : x \mapsto x + p, \text{ and}$$
$$\psi_{n,p,q}(b) : x \mapsto x + q.$$

Lemma 6.1. *For any finite set $S \subseteq \mathbb{Z}^2$ and any $\delta > 0$, there exists a constant C such that $\psi_{n,p,q}$ is an (S, δ, n)-approximation of $\mathbb{Z}^2$ provided that $p > Cq$ and $n > Cp$.*

Proof. Take C sufficiently large that $S \subseteq \left\{ a^\lambda b^\mu \,\middle|\, |\lambda| < C/3,\ |\mu| < C/3 \right\}$. Since the map $\psi_{n,p,q}$ is a group homomorphism, we only need to show that $d(\psi_{n,p,q}(s), \mathrm{id}) > 1 - \delta$ for all $s \in S \setminus \{1\}$ provided that $p > Cq$ and $n > Cp$. Then for $s = a^\lambda b^\mu \in S$ we find $\psi_{n,p,q}(s)$ is translation by $\lambda p + \mu q$, which is not divisible by n (unless $\lambda = \mu = 0$), and therefore $d(\mathrm{id}, \psi_{n,p,q}(s)) = 1$. □

The equivalence *(1)* ⇔ *(3)* of Theorem 5.1 tells us that for $k \in \mathbb{N}$, the group $\mathrm{Hig}_k(\mathbb{Z}^2)$ has a sofic quotient Q such that the composition $\mathbb{Z}^2 \to \mathrm{Hig}_k(\mathbb{Z}^2) \to Q$ is injective if and only if for any n, p, q such that n/p and p/q are sufficiently large, there is a permutation $f : \mathbb{Z}/n\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$ of order k with $d(\psi_{n,p,q}(b) \circ f, f \circ \psi_{n,p,q}(a)) < \varepsilon$. As $(\psi_{n,p,q}(b) \circ f)(x) = f(x) + q$ and $(f \circ \psi_{n,p,q}(a))(x) = f(x + p)$, the latter condition amounts to $f(x + p) = f(x) + q$ for at least $(1 - \varepsilon)n$ elements $x \in \mathbb{Z}/n\mathbb{Z}$.

However, for $k \geq 1$, the group $\overline{\mathrm{Hig}}_k(\mathbb{Z}^2)$ is a right-angled Artin group, so it is linear and thus residually finite (see [14]). Thus $\mathrm{Hig}_k(\mathbb{Z}^2)$ and $\overline{\mathrm{Hig}}_k(\mathbb{Z}^2)$ are sofic. (For $k \geq 4$, we reached the same conclusion in Corollary 4.5 via the residual solvability established in Theorem 1.1. For $k \leq 3$ the group is abelian, and thus also sofic.) And for $k \geq 2$, $\mathbb{Z}^2 \hookrightarrow \overline{\mathrm{Hig}}_k(\mathbb{Z}^2)$. Thus:

Theorem 6.2. *Suppose $k \geq 2$ and $\varepsilon > 0$. Then there exists $C > 0$ such that for all n, p, q satisfying $n \geq Cp$ and $p \geq Cq$, there exists a permutation $f \in \mathrm{Sym}(\mathbb{Z}/n\mathbb{Z})$ of order dividing k such that*

$$f(x + p) = f(x) + q$$

for at least $(1 - \varepsilon)n$ elements $x \in \mathbb{Z}/n\mathbb{Z}$.

In some cases this is straight-forward. If n is a prime congruent to 1 modulo k, then there exists $l \in \mathbb{Z}/n\mathbb{Z}$ such that $l^k = 1$ and $q = lp$ in $\mathbb{Z}/n\mathbb{Z}$, and then $f : \mathbb{Z}/n\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$ mapping $x \mapsto lx$ satisfies the given conditions, because $f(x + p) = lx + lp = lx + q = f(x) + q$ for all $x \in \mathbb{Z}/n\mathbb{Z}$ and $f^k = \mathrm{id}$. Indeed, such f arise from a natural sofic quotient of $\mathrm{Hig}_k(\mathbb{Z}^2)$—take the semidirect product of the cyclic group of order k and the abelianization of $\overline{\mathrm{Hig}}_k(G)$. Then $\mathrm{Hig}_k(\mathbb{Z}^2)$ maps onto $C_k \ltimes \mathbb{Z}/n\mathbb{Z}$, where the action is by multiplication by l.

But in most cases errors are inevitable. Suppose $f : \mathbb{Z}/n\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$ satisfies $f(x + p) = f(x) + q$ for all $x \in \mathbb{Z}/n\mathbb{Z}$. Then $f^l(x + p^l) = f^l(x) + q^l$ for all $x \in \mathbb{Z}/n\mathbb{Z}$ and all $l \in \mathbb{N}$. So if $f^k = \mathrm{id}$, then n divides $q^k - p^k$.

Whether or not n divides $q^k - p^k$, by Theorem 6.2, there exist such functions f satisfying $f(x + p) = f(x) + q$ for most $x \in \mathbb{Z}/n\mathbb{Z}$. Such f could be constructed explicitly by carefully following the arguments in our proofs of Theorems 5.1 and 5.3 (using that there are Følner sets for $\mathbb{Z}^2$ of a very simple form), but doing this in general would be quite technical.

Remark 6.3. In such a simple example it is possible to determine the dependance of the constant C in Theorem 6.2: one can take $C = O(\varepsilon^{-k})$.

6.2. **The Heisenberg group.** The Heisenberg group $\mathcal{H}$ has presentation

$$\mathcal{H} = \langle a, b \mid [a,[a,b]] = [b,[a,b]] = 1\rangle.$$

It is nilpotent and so is amenable and residually finite. Identify $\mathrm{Sym}(n^2)$ with $\mathrm{Sym}\left((\mathbb{Z}/n\mathbb{Z})^2\right)$. Define $\psi_n : \mathcal{H} \to \mathrm{Sym}(n^2)$ for $n \in \mathbb{N}$ by

$$\psi_n(a) : (x,y) \mapsto (x, y+1), \text{ and}$$
$$\psi_n(b) : (x,y) \mapsto (x+y, y),$$

which extends to $\mathcal{H}$ since $\psi_n(a)$ and $\psi_n(b)$ satisfy the defining relations of $\mathcal{H}$. This action of $\mathcal{H}$ arises from the finite quotient $\mathcal{H}_n := \mathcal{H}/\langle a^n, b^n\rangle$ acting on cosets of the subgroup $\langle a\rangle$. The following lemma can be proved in a similar manner to Lemmas 6.1, 6.10, and 6.15.

Lemma 6.4. *For all finite sets $S \subseteq \mathcal{H}$ and all $\delta > 0$, there exists $C > 0$ such that ψ_n is (S, δ, n^2)-approximation of $\mathcal{H}$ for all $n > C$.*

So Theorem 5.1 tells us that for all $k \in \mathbb{N}$ the group $\mathrm{Hig}_k(\mathcal{H})$ has a sofic quotient Q such that $\mathcal{H} \to \mathrm{Hig}_k(\mathcal{H}) \to Q$ is injective if and only if there exist infinitely many n and functions $f : (\mathbb{Z}/n\mathbb{Z})^2 \to (\mathbb{Z}/n\mathbb{Z})^2$ of order dividing k which conjugate the action of b to the action of a under ψ_n up to error ε.

For $k \geq 4$, Corollary 4.5 tells us that $\mathrm{Hig}_k(\mathcal{H})$ is sofic. As for the case $k = 2$, we have that $\overline{\mathrm{Hig}_2}(\mathcal{H}) \simeq \mathcal{H}$ is also sofic. And for $k = 3$, it is not hard to construct a surjective map from $\overline{\mathrm{Hig}_3}(\mathcal{H})$ onto the sofic group $\mathrm{SL}_3(\mathbb{Z})$ such that the composition $\mathcal{H} \to \overline{\mathrm{Hig}_3}(\mathcal{H}) \to \mathrm{SL}_3(\mathbb{Z})$ is injective. A sofic quotient of $\mathrm{Hig}_3(\mathcal{H})$ into which $\mathcal{H}$ injects could also be obtained via Proposition 4.6. (We do not know whether the group $\mathrm{Hig}_3(\mathcal{H})$ itself is sofic, but see no reason it should not be.)

So applying the condition $d(\psi(b) \circ f, f \circ \psi(\phi(b))) < \varepsilon$ of Theorem 5.1 (3) to ψ_n and to b and $a = \phi(b)$ we get the following.

Theorem 6.5. *For all $\varepsilon > 0$ and all $k \geq 2$, there exists an integer C such that for all $n > C$ there exists a permutation $f \in \mathrm{Sym}(n^2)$ of order dividing k which, when expressed as $f(x,y) = (f_1(x,y), f_2(x,y))$ so that $f_1, f_2 : \mathbb{Z}/n\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$ are its coordinate functions, satisfies*

$$f_1(x, y+1) = f_1(x,y) + f_2(x,y) \quad \text{and} \quad f_2(x, y+1) = f_2(x,y)$$

for at least $(1-\varepsilon)n^2$ pairs (x,y).

Remark 6.6. There is no f such that the above equality holds for all pairs (x, y), since $\psi_n(a)$ and $\psi_n(b)$ are not conjugate inside $\mathrm{Sym}(n^2)$—one of them has (a few) fixed points and the other has none. The generators a and b play asymmetric roles in the definition of ψ_n: for example, $d(\psi_n(a), \mathrm{id}) = 1$, but $d(\psi_n(b), \mathrm{id}) = 1 - 1/n < 1$. One can instead take the action of $\mathcal{H}_n$ on itself which will lead to a permutation representation $\mathcal{H} \to \mathrm{Sym}(n^3)$ in which the roles of a and b are symmetric. The functions f of the resulting analogue of Theorem 6.5 can be constructed in such a way that the equations are satisfied for *all* points if and only if there is an nontrivial semisimple element of order dividing k in the group $\mathrm{SL}_2(\mathbb{Z}/n\mathbb{Z})$.

Remark 6.7. It is again possible to estimate the dependance of the constant C. One can show that we can take $C = O(\varepsilon^{-3k})$.

Remark 6.8. One can also consider a Higman-like construction from $\mathcal{H}$ in which ϕ maps one of the standard generators to a generator of the center: define

$$G = \mathcal{H} = \langle b, c \mid [b,[b,c]] = [c,[b,c]] = 1\rangle, \quad \phi : b \mapsto a,$$

where $a := [b,c]$ so that

$$\overline{\mathrm{Hig}}_k(G,\phi) = \langle\, b_1, c_1, \ldots, b_k, c_k \mid [b_i,[b_i,c_i]] = [c_i,[b_i,c_i]] = 1, b_i = [b_{i+1}, c_{i+1}]\ \forall i\ (\mathrm{mod}\ k)\rangle.$$

We do not know whether the group $\overline{\mathrm{Hig}}_k(G,\phi)$ is sofic, since Theorem 1.1 does not apply. However $b_i \mapsto \mathrm{id} + e_{i,k+1}$ and $c_i \mapsto \mathrm{id} - e_{i+1,i}$ (indices mod k) defines a homomorphism

$$\overline{\mathrm{Hig}}_k(G,\phi) \ \to\ \mathrm{SL}_k(\mathbb{Z}) \ltimes \mathbb{Z}^k \ \subseteq\ \mathrm{SL}_{k+1}(\mathbb{Z}),$$

and for $k \geq 2$ the group $\mathcal{H}$ injects into this quotient (i.e., image) of $\overline{\mathrm{Hig}}_k(G,\phi)$ which is linear and thus sofic. As before (but with a and b now changed to b and c, respectively) define $\psi_n : \mathcal{H} \to \mathrm{Sym}(n^2)$ for $n \in \mathbb{N}$ by

$$\psi_n(b) : (x,y) \mapsto (x, y+1), \text{ and}$$
$$\psi_n(c) : (x,y) \mapsto (x+y, y).$$

Then, as $\psi_n([b,c]) : (x,y) \mapsto (x-1,y)$, applying Theorem 5.1 leads to:

Theorem 6.9. *Functions f exist exactly as per Theorem 6.5, except with the displayed equations replaced by:*

$$f_1(x-1,y) = f_1(x,y) + f_2(x,y) \quad \textit{and} \quad f_2(x-1,y) = f_2(x,y).$$

Despite their similarity, we do not see a way to derive one of Theorems 6.5 and 6.9 immediately from the other. Defining $g(x,y) := f(y,x)$ transforms one set of recurrences to the other, but the condition that the function's order divides k is lost.

6.3. The Baumslag–Solitar group BS(1, m).

This is the case addressed by Helfgott and Juschenko in [12]. Here we explain how it fits into our framework and give our own account of how it relates to recent work of Glebsky.

The Baumslag–Solitar group $\mathrm{BS}(1,m)$ has presentation

$$\mathrm{BS}(1,m) \ = \ \langle a, b \mid a^b = a^m\rangle.$$

It is a residually finite solvable group, and so is amenable. If $m \neq \pm 1$, then the image of a in any proper quotient of $\mathrm{BS}(1,m)$ is finite. (Every element can be expressed as $b^\mu a^\nu b^{-\lambda}$ for some $\mu, \lambda \geq 0$ and $\nu \in \mathbb{Z}$. The result then follows from consequences of a non-trivial $b^\mu a^\nu b^{-\lambda}$ mapping to the identity and the relation $a^b = a^m$.)

Identify $\mathrm{Sym}(n)$ with $\mathrm{Sym}(\mathbb{Z}/n\mathbb{Z})$. For all $n \in \mathbb{N}$ relatively prime to m, define a map $\psi_n : \mathrm{BS}(1,m) \to \mathrm{Sym}(n)$ by

$$\psi_n(a) : \ x \mapsto x+1, \text{ and}$$
$$\psi_n(b) : \ x \mapsto m^{-1}x,$$

which extends to a homomorphism defined on the whole of $\mathrm{BS}(1,m)$ since $\psi_n(a)$ and $\psi_n(b)$ satisfy the defining relation of $\mathrm{BS}(1,m)$. This action of $\mathrm{BS}(1,m)$ arises from the quotient $\mathrm{BS}(1,m)_n := \mathrm{BS}(1,m)/\langle a^n\rangle$ acting on cosets of the subgroup $\langle b\rangle$.

Lemma 6.10. *For all finite sets $S \subseteq \mathrm{BS}(1,m)$ and all $\delta > 0$, there exists an integer C such that for all $n > C$, the map ψ_n is an (S, δ, n)-approximation of $\mathrm{BS}(1,m)$, provided that $|m| \geq 2$.*

Proof. The group $\mathrm{BS}(1,m)$ can be represented by 2×2 matrices via

$$a \mapsto \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \qquad b \mapsto \begin{bmatrix} 1 & 0 \\ 0 & m \end{bmatrix}.$$

The image of this embedding is

$$\left\{ \begin{bmatrix} 1 & \lambda \\ 0 & m^{\mu} \end{bmatrix} \,\middle|\, \mu \in \mathbb{Z},\ \lambda \in \mathbb{Z}\left[\frac{1}{m}\right] \right\}.$$

For every finite set $S \subseteq \mathrm{BS}(1,m)$ there exists a positive integer N such that every $s \in S$ is sent to

$$\begin{bmatrix} 1 & \lambda_s m^{-N} \\ 0 & m^{\mu_s} \end{bmatrix}$$

where λ_s and μ_s are integers such that $|\mu_s| \leq N$ and $|\lambda_s| \leq \left|m^{2N}\right|$. One computes that

$$\psi_n(s) : \left(x \mapsto m^{-\mu_s}(x + \lambda_s m^{-N})\right).$$

If this permutation is non-trivial (i.e., $\mu_s \neq 0$ or $n \nmid \lambda_s$), then it has at most m^N fixed points. Therefore if $n > C := \max\left\{\left|\frac{m^N}{\delta}\right|, \left|m^{2N}\right|\right\}$, then ψ_n is an (S, δ, n)-approximation. □

Theorem 5.1 now tells us that for all $k \in \mathbb{N}$ the group $\mathrm{Hig}_k(\mathrm{BS}(1,m))$ has a sofic quotient into which $\mathrm{BS}(1,m)$ naturally embeds if and only there exist infinitely many n and functions $f : \mathbb{Z}/n\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$ of order dividing k which conjugate addition to multiplication by m up to error ε—that is, $m^{-1}f(x) = f(x+1)$ for at least $(1-\varepsilon)n$ values of $x \in \mathbb{Z}/n\mathbb{Z}$. Define $\tilde{f} \in \mathrm{Sym}(\mathbb{Z}/n\mathbb{Z})$ by $\tilde{f}(x) = -f(-x)$, which has order dividing k if and only if f does. The equation $m^{-1}f(x) = f(x+1)$ can be re-expressed as $f(x) = mf(x+1)$, and then as

$$\tilde{f}(x+1) = -f(-x-1) = -mf(-x) = m\tilde{f}(x).$$

So $m^{-1}f(x) = f(x+1)$ is satisfied by at least $(1-\varepsilon)n$ values of $x \in \mathbb{Z}/n\mathbb{Z}$ if and only if the same is true of $\tilde{f}(x+1) = m\tilde{f}(x)$. Theorem 1.2 then follows, or, in more detail, we have:

Theorem 6.11 (Helfgott–Juschenko [12])**.** *The group $\mathrm{Hig}_k(\mathrm{BS}(1,m))$ has a sofic quotient Q such that the composition $\mathrm{BS}(1,m) \to \mathrm{Hig}_k(\mathrm{BS}(1,m)) \to Q$ is injective if and only if for all $\varepsilon > 0$ there exists an integer C such that for all $n > C$ coprime to m, there exist a permutation $f \in \mathrm{Sym}(\mathbb{Z}/n\mathbb{Z})$ of order dividing k such that $f(x+1) = mf(x)$ for at least $(1-\varepsilon)n$ values of x.*

We do not know whether $\mathrm{Hig}_k(\mathrm{BS}(1,m))$ is sofic for $k \geq 4$. Both $\mathrm{Hig}_k(\mathrm{BS}(1,m))$ and $\overline{\mathrm{Hig}}_k(\mathrm{BS}(1,m))$ are finite for $k \leq 3$ (assuming $m \neq \pm 1$), and so cannot have a sofic quotient into which $\mathrm{BS}(1,m)$ injects. A beautiful argument due to Higman [13] shows that $\overline{\mathrm{Hig}}_k(\mathrm{BS}(1,2))$ has no finite quotients. Glebsky [8,9] shows that, by contrast, if $k \geq 4$ and p is a prime dividing $m-1$, then the groups $\overline{\mathrm{Hig}}_k(\mathrm{BS}(1,m))$ have many quotients which are finite p-groups. His main theorem in [8] amounts to the following. His proof is more combinatorial than the one we sketch below via Golod–Shafarevich machinery.

Theorem 6.12 (Glebsky [8])**.** *Suppose $k \geq 4$ and p is a prime dividing $m-1$. Then the pro-p completion of $\overline{\mathrm{Hig}}_k(\mathrm{BS}(1,m))$ is infinite. If, moreover, k is even and $m \neq \pm 1$, then* $\mathrm{BS}(1,m)$ *embeds into this pro-p completion.*

Sketch of a proof. The defining relator of $\mathrm{BS}(1,m)$ can be written in the from $[a,b]^{-1}a^{m-1}$ and lies in the p-Frattini subgroup of the free group. This implies that the pro-p completion $\hat{G}$ of $\overline{\mathrm{Hig}}_k(\mathrm{BS}(1,m))$ has a minimal pro-p presentation with k generators and k relations. Such a presentation satisfies the Golod–Shafarevich condition (since $k \leq k^2/4$) and therefore it defines an infinite pro-p group (see, for example, [7]).

Since a has finite order in any proper quotient of $\mathrm{BS}(1,m)$, to prove that $\mathrm{BS}(1,m)$ embeds into this pro-p completion, it suffices to show that the images of the generators $a_1,\dots,a_k$ of $\overline{\mathrm{Hig}}_k(\mathrm{BS}(1,m))$ have infinite order—indeed, that one of them has infinite order. In the case $k=4$, the defining relations of $\hat{G}$ are similar to the relations of $F_2 \times F_2$, where the first copy of the free group F_2 is generated by a_1 and a_3 and the second copy is generated by a_2 and a_4. It can be shown that $\hat{G}$ contains the free pro-p groups Γ_1 and Γ_2 generated by $\Gamma_1 = \langle a_1, a_3\rangle$ and $\Gamma_2 = \langle a_2, a_4\rangle$, and moreover that any element in $\hat{G}$ can be written uniquely as a product of two elements, one from Γ_1 and one from Γ_2. This shows that the order of a_i in $\hat{G}$ is infinite and that $\mathrm{BS}(1,m)$ embeds in $\hat{G}$. (When $k>4$ the group $\hat{G}$ does not have such nice combinatorial description, but there is a quotient of $\hat{G}$, which has similar structure, provided that k is even.) ◻

Theorems 6.11 and 6.12 together imply Theorem 1.4: if $|m|>2$ and $\varepsilon>0$, then there exists C such that for all $n>C$ coprime to m, there are permutations $g \in \mathrm{Sym}(\mathbb{Z}/n\mathbb{Z})$ with $g^4 = \mathrm{id}$ and with $g(x+1) = mg(x)$ for at least $(1-\varepsilon)n$ values of $x \in \mathbb{Z}/n\mathbb{Z}$.

Remark 6.13. Theorem 1.4 applies to all integers $m \neq 0,2$. The above proof via Theorem 6.12 works for $m \neq -1,0,1,2$. For $m=1$, when $\mathrm{BS}(1,1) = \mathbb{Z}\times\mathbb{Z}$, the analogue is a special case of Theorem 6.2. The case $m=-1$ only requires a minor strengthening of Theorem 6.12. The case $m=0$ is degenerate since a and b have different orders. Also, the order of g can divide any given even integer $k \geq 4$, not just 4. We stress that the analogue of Theorem 1.4 is unknown when $m=2$.

Remark 6.14. Estimating the dependance of the constant C in Theorem 1.4 is quite hard because it involves explicitly constructing the sets S' in Theorem 5.3. We believe that by carefully tracking all bounds one gets that $C = O\left(2^{K\varepsilon^{-2}}\right)$.

6.4. $\mathbb{Z}\wr\mathbb{Z}$**.** The wreath product $\mathbb{Z}\wr\mathbb{Z}$ has presentation

$$\mathbb{Z}\wr\mathbb{Z} = \left\langle a,b \;\middle|\; \left[a, a^{b^i}\right] = 1 \;\forall i \in \mathbb{N}\right\rangle.$$

It is a residually finite solvable group, and so is amenable. For any $n \in \mathbb{N}$ and any m coprime to n, define a homomorphism $\psi_{n,m} : \mathbb{Z}\wr\mathbb{Z} \to \mathrm{Sym}(\mathbb{Z}/n\mathbb{Z})$ by

$$\psi_{n,m}(a) = (x \mapsto x+1)\,,\text{ and}$$
$$\psi_{n,m}(b) = \left(x \mapsto m^{-1}x\right),$$

which is well-defined since the permutations $\psi_{n,m}(a)$ and $\psi_{n,m}(b)$ satisfy the defining relations of $\mathbb{Z}\wr\mathbb{Z}$. This action of $\mathbb{Z}\wr\mathbb{Z}$ arises from the quotient $(\mathbb{Z}\wr\mathbb{Z})_{n,m} = (\mathbb{Z}\wr\mathbb{Z})/\langle a^n = 1, a^b = a^m\rangle$ acting on cosets of the subgroup $\langle b\rangle$.

Lemma 6.15. *For all finite sets $S \subseteq \mathbb{Z} \wr \mathbb{Z}$ and all $\delta > 0$, there exists $C > 0$ such that if $n \in \mathbb{N}$ and $m \in \mathbb{Z}$ are coprime and satisfy $|m| > 2C + 1$ and $n > |m|^{C+1}$, then $\psi_{n,m}$ is an (S, δ, n)-approximation of $\mathbb{Z} \wr \mathbb{Z}$.*

Proof. The group $\mathbb{Z} \wr \mathbb{Z}$ can be represented by the group of matrices

$$\left\{ \begin{bmatrix} 1 & \tilde{t}(x) \\ 0 & x^k \end{bmatrix} \;\middle|\; k \in \mathbb{Z},\ \ \tilde{t}(x) \in \mathbb{Z}[x, x^{-1}] \right\}$$

via

$$a \mapsto \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \qquad b \mapsto \begin{bmatrix} 1 & 0 \\ 0 & x^{-1} \end{bmatrix}.$$

Suppose S is a finite subset of $\mathbb{Z} \wr \mathbb{Z}$. Then there exists an integer N with the following property. Every $s \in S$ can be represented by

$$\begin{bmatrix} 1 & \tilde{t}_{(s)}(x) \\ 0 & x^{\mu_{(s)}} \end{bmatrix}$$

where

$$\tilde{t}_{(s)}(x) \;=\; \sum_{i=-N}^{N} \tilde{t}_i x^i \;\in\; \mathbb{Z}[x, x^{-1}]$$

and $|\mu_{(s)}| \leq N$ and $|\tilde{t}_i| \leq N$ for all i. One computes that for s as above,

$$\psi_{n,m}(s) : \left(x \mapsto m^{-\mu_{(s)}} \left(x + \tilde{t}_{(s)}(m) \right) \right).$$

Define $t_{(s)}(x) := x^N \tilde{t}_{(s)}(x) \in \mathbb{Z}[x]$.

Fix $\delta > 0$. Let $C = \max\left\{ 2N + 1, \left\lfloor \frac{1}{\delta} + 1 \right\rfloor \right\}$.

Assume $n \in \mathbb{N}$ and $m \in \mathbb{Z}$ are coprime and satisfy $|m| > C + 1$ and $n > |m|^{C+1}$. In particular, $|m| \neq 1$. Suppose $t(x) = \sum t_i x^i \in \mathbb{Z}[x]$ is a nonzero polynomial whose degree d is most C and whose coefficients all satisfy $|t_i| < C/2$. (Our polynomial $t_{(s)}(x)$ has these properties.) Then

$$|t(m)| \;\leq\; \sum_{i=0}^{d} |t_i|\, |m|^i \;\leq\; C \frac{|m|^{d+1} - 1}{|m| - 1} \;<\; |m|^{C+1} \;<\; n.$$

If $d = 0$, then $0 \neq |t(m)| = |t_0| < C/2 < n$, so $n \nmid t(m)$, as required. Assume $d \geq 1$. Then

$$|t(m)| \;=\; \left| t_d m^d + \sum_{i=0}^{d-1} t_i m^i \right| \;\geq\; \left| t_d m^d \right| - \left| \sum_{i=0}^{d-1} t_i m^i \right| \;\geq\; |m|^d - C \frac{|m|^d - 1}{|m| - 1} \;\geq\; |m|^d / 2,$$

where the final inequality holds because $|m| > C$. As $d \geq 1$, we now have that $0 < |m|/2 \leq |t(m)| < n$, which implies that $n \nmid t(m)$.

Conclude that $n \nmid t_{(s)}(m)$.

If $\mu_{(s)} = 0$, then $\psi_{n,m}(s)$ maps x to $x + \tilde{t}_{(s)}(m) = x + m^{-N} t_{(s)}(m)$ and so has no fixed points as m^{-1} is coprime to n and $n \nmid t_{(s)}(m)$.

Suppose $\mu_{(s)} \neq 0$. If $n | (m^{\mu_{(s)}} - 1)$, then $\psi_{n,m}(s) : x \mapsto x + \tilde{t}_{(s)}(m)$ and so is either the identity ($n \mid \tilde{t}_{(s)}(m)$ would contradict $n \nmid t_{(s)}(m)$, and so does not occur) or has no fixed points. If $n \nmid (m^{\mu_{(s)}} - 1)$, then the permutation $\psi_{n,m}(s)$ has at most $\gcd(m^{\mu_{(s)}} - 1, n)$ fixed points. But

$\gcd(m^{\mu_{(s)}} - 1, n) \le \delta n$, else the polynomial $M(x^{|\mu_{(s)}|} - 1)$ will have m as a root mod n for some $M < 1/\delta$. So $\psi_{n,m}(s)$ has at most δn fixed points.

In every case we have $d(\psi_{n,m}(s), \mathrm{id}) > 1 - \delta$. And the *almost homomorphism* condition is immediate since $\psi_{n,m}$ is, in fact, a homomorphism. So $\psi_{n,m}$ is an (S, δ, n)-approximation of $\mathbb{Z} \wr \mathbb{Z}$, as required. □

So Theorem 5.1 tells us that for all $k \in \mathbb{N}$ the group $\mathrm{Hig}_k(\mathbb{Z} \wr \mathbb{Z})$ has a sofic quotient into which $\mathbb{Z} \wr \mathbb{Z}$ naturally embeds if and only there exist infinitely many n and functions $f : \mathbb{Z}/n\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$ of order dividing k which conjugate addition of 1 to multiplication by m up to error ε. After conjugating by a minus sign (in the manner of replacing f by $\tilde{f}$ in Section 6.3) this latter condition becomes $f(x + 1) = mf(x)$ for at least $(1 - \varepsilon)n$ values of $x \in \mathbb{Z}/n\mathbb{Z}$. However, for $k \ge 4$ we know by Lemma 3.3 that $\mathbb{Z} \wr \mathbb{Z}$ naturally embeds in $\mathrm{Hig}_k(\mathbb{Z} \wr \mathbb{Z})$ and by Corollary 4.5 that $\mathrm{Hig}_k(\mathbb{Z} \wr \mathbb{Z})$ is sofic. We also know by Proposition 4.6 that $\mathrm{Hig}_3(\mathbb{Z} \wr \mathbb{Z})$ has a sofic quotient into which $\mathbb{Z} \wr \mathbb{Z}$ naturally embeds. So we have that such n and f do exist for $k \ge 3$. Adjusting the constant C of Lemma 6.15 suitably, we have:

Theorem 1.3. *For all $\varepsilon > 0$ and $k \ge 3$, there exists C such that if n is coprime to m and $|m| > C$ and $n > |m|^C$, then there exists $f \in \mathrm{Sym}(\mathbb{Z}/n\mathbb{Z})$ which has order dividing k and the property that $f(x + 1) = mf(x)$ for at least $(1 - \varepsilon)n$ values of $x \in \mathbb{Z}/n\mathbb{Z}$.*

As mentioned above, $\mathrm{Hig}_k(\mathbb{Z} \wr \mathbb{Z})$ is sofic for all $k \ge 4$ by Corollary 4.5. For $k = 1$ and $k = 2$ the groups $\mathrm{Hig}_k(\mathbb{Z} \wr \mathbb{Z})$ are sofic since

$$\overline{\mathrm{Hig}}_1(\mathbb{Z} \wr \mathbb{Z}) = \mathrm{Hig}_1(\mathbb{Z} \wr \mathbb{Z}) \cong \mathbb{Z} \quad \text{and} \quad \overline{\mathrm{Hig}}_2(\mathbb{Z} \wr \mathbb{Z}) \cong \mathcal{H}$$

(but $\mathbb{Z} \wr \mathbb{Z}$ does not embed in $\mathrm{Hig}_1(\mathbb{Z} \wr \mathbb{Z})$ or $\mathrm{Hig}_2(\mathbb{Z} \wr \mathbb{Z})$). As in the case of $\mathcal{H}$, we do not know whether $\mathrm{Hig}_3(\mathbb{Z} \wr \mathbb{Z})$ is sofic, but we see no reason it should not be.

Remark 6.16. Before proving Theorem 1.1, which implies that $\mathrm{Hig}_4(\mathbb{Z} \wr \mathbb{Z})$ is sofic, we constructed finite quotients which can be combined to give a residually finite quotient Q of $\mathrm{Hig}_4(\mathbb{Z} \wr \mathbb{Z})$ in which $\mathbb{Z} \wr \mathbb{Z}$ embeds. We find these quotients interesting on their own, and will briefly describe them. Pick a prime p and two functions $f, \lambda : \mathbb{F}_p \to \mathbb{F}_p^*$. Then there is an action $\psi_{p,f,\lambda}$ of $\mathrm{Hig}_4(\mathbb{Z} \wr \mathbb{Z})$ on $S = \mathbb{F}_p \times \mathbb{F}_p \times \mathbb{F}_p \times \mathbb{F}_p$ defined by

$$t : (x, y, z, w) \mapsto (y, z, w, x), \qquad a : (x, y, z, w) \mapsto (x\lambda(z), y, z, w + f(z)).$$

These actions $\psi_{p,f,\lambda}$ are quite different from those arising in our proof of Theorem 1.1. We expect that for generic functions f and λ, the image of $\psi_{p,f,\lambda}$ will either be the full symmetric group, or the alternating group, and so will be very far from (residually) solvable. It is intriguing question whether the actions $\psi_{p,f,\lambda}$ distinguish all elements in $\mathrm{Hig}_4(\mathbb{Z} \wr \mathbb{Z})$. We see no reason why this should not the case, but without having an easily understandable combinatorial model of $\mathrm{Hig}_4(\mathbb{Z} \wr \mathbb{Z})$, it is hard to prove such claim.

6.5. **The free metabelian group on two generators.** The free metabelian group on two generators $\mathcal{M}$ has a presentation

$$\mathcal{M} = \left\langle a, b \;\middle|\; \left[[a, b], [a, b]^{a^i b^j}\right] = 1 \ \forall i, j \in \mathbb{Z} \right\rangle.$$

It is a residually finite solvable group, so is amenable. For $n \in \mathbb{N}$ and p and q relatively prime to n, define a map $\psi_{n,p,q} : \mathbb{Z} \wr \mathbb{Z} \to \mathrm{Sym}(\mathbb{Z}/n\mathbb{Z})$ by

$$\psi_{n,p,q}(a) = \left(x \mapsto q^{-1}(x+1)\right), \text{ and}$$
$$\psi_{n,p,q}(b) = \left(x \mapsto p^{-1}x\right),$$

which extends to the whole of $\mathcal{M}$ since the permutations $\psi_{n,p,q}(a)$ and $\psi_{n,p,q}(b)$ satisfy the defining relations of $\mathcal{M}$. We get these analogues of Lemma 6.15 and Theorem 1.3:

Lemma 6.17. *For all finite sets $S \subseteq \mathcal{M}$ and all $\delta > 0$, there exists a constant C such that for all integers n, p, and q with n coprime to p and q and $|q| > 2C + 1$, $|p| > |q|^{C+1}$ and $n > |p|^{C+1}$, we have $\psi_{n,p,q}$ is an (S, δ, n)-approximation of $\mathcal{M}$.*

Theorem 6.18. *For all $\varepsilon > 0$ and $k \geq 2$, there exists C such that if n is coprime to p and q and $|q| > C$, $|p| > |q|^C$ and $n > |p|^C$, then there exists $f \in \mathrm{Sym}(\mathbb{Z}/n\mathbb{Z})$ such that $f^k = \mathrm{id}$ and $f(qx + 1) = pf(x)$ for at least $(1 - \varepsilon)n$ values of $x \in \mathbb{Z}/n\mathbb{Z}$.*

7. Heuristic

Here we will explain why the existence of the permutations $f \in \mathrm{Sym}(n)$ proved in Theorem 5.1 is surprising. We will focus on instances where $A = \langle a \rangle \cong B = \langle b \rangle \cong \mathbb{Z}$ and $k = 4$ which is the case in most of our examples. (We could generalize to $k \geq 4$ without significantly changing the following argument, but the assumption that $A \cong B \cong \mathbb{Z}$ is essential.)

Denote $\alpha = \psi(a)$ and $\beta = \psi(b)$. Each permutation f satisfies the global condition $f^4 = \mathrm{id}$; and many local conditions: the condition concerning $d(\psi(b) \circ f, f \circ \psi(\phi(b)))$ is equivalent to $f(\alpha(x)) = \beta(f(x))$ for at least $(1 - \varepsilon)n$ points x.

One can estimate the probability that a permutation f chosen uniformly at random from $\mathrm{Sym}(n)$ satisfies that global condition: by considering cycle structure, one counts the number of elements of order dividing 4 in the symmetric group $\mathrm{Sym}(n)$ (see [3]), which leads to

$$P = \mathrm{Prob}(f^4 = \mathrm{id}) \approx \frac{1}{\sqrt[4]{|\mathrm{Sym}(n)|}} \approx n^{-n/4}.$$

(Here we are only describing the leading term of the expansion of $\log P$.)

It is also quite easy to estimate the probability that a local condition is satisfied. A local condition asserts that $f(\alpha(x))$ is determined by $f(x)$—the probability that this happens at a given x is approximately $1/n$ (only approximately since x might be a fixed point of α or $f(x)$ might be a fixed point for β). However, this probability is irrelevant since we want $f(\alpha(x)) = \beta(f(x))$ for the majority of x (for at least $(1 - \varepsilon)n$ points x, to be precise). Informally, a small number of local conditions are almost independent from each other, but this is not true if we consider many local conditions.

The number of permutations in $\mathrm{Sym}(n)$ satisfying $f(\alpha(x)) = \beta(f(x))$ for at least $(1 - \varepsilon)n$ points x is at most $n^{2\varepsilon n+k}$ where k is the number cycles in the action of α—such permutations are determined by the following: the points x where the local condition is not satisfied; values of $f(\alpha(x))$ at each of these points; and the values of f at a single point on each cycle on the action of α. (This information specifies a function f but it may not be a permutation.) For all $\varepsilon' > 0$, we have $k < \varepsilon' n$ for large n, otherwise a small power of

α will be close to the identity permutation, which would contradict the fact that ψ detects the soficity of G. Thus, the number of permutations satisfying the majority of the local conditions is at most $K = n^{(2\varepsilon+\varepsilon')n}$.

If the global condition is almost independent from the local conditions, then the expected number of permutation satisfying both, should be around

$$P.K \;=\; n^{-n/4}.n^{(2\varepsilon+\varepsilon')n} \;=\; n^{(2\varepsilon+\varepsilon'-1/4)n} \ll 1,$$

if $\varepsilon, \varepsilon' < 1/20$. Thus, one should expect that there are no such permutations when n is sufficiently large.

The independence assumption is somewhat justified by the observation that the global condition is independent form each of the local conditions (it is also almost independent from any fixed number of local conditions). Notice that this heuristic does not really depend on the group G.

The main weakness of this heuristic is the assumption that the global condition is almost independent from the majority of the local conditions. One can interpret Theorem 5.1 as saying there is a connection between the soficity of the group $\mathrm{Hig}_k(G)$ and the independence of global versus local conditions.

Acknowledgements

We thank Harald Helfgott for discussions on this project and are grateful for partial support from NSF grants DMS 1303117 and 1601406 (Kassabov), the Rawlings College Presidential Research Scholars program at Cornell (Kuperberg), and Simons Collaboration Grant 318301 (Riley). We also thank an anonymous referee for a careful reading and some helpful comments.

References

1. Goulnara Arzhantseva and Liviu Păunescu, *Almost commuting permutations are near commuting permutations*, J. Funct. Anal. **269** (2015), no. 3, 745–757. MR 3350728
2. Valerio Capraro and Martino Lupini, *Introduction to sofic and hyperlinear groups and Connes' embedding conjecture*, Lecture Notes in Mathematics, vol. 2136, Springer International Publishing, 2015, arXiv:1309.2034.
3. S. Chowla, I. N. Herstein, and W. R. Scott, *The solutions of $x^d = 1$ in symmetric groups*, Norske Vid. Selsk. Forh., Trondheim **25** (1952), 29–31 (1953). MR 0054605
4. Benoit Collins and Kenneth J. Dykema, *Free products of sofic groups with amalgamation over monotileably amenable groups*, Münster J. Math **4** (2011), 101–118, arXiv:1003.1675.
5. Gábor Elek and Endre Szabó, *Sofic groups and direct finiteness*, J. Algebra **280** (2004), no. 2, 426–434. MR 2089244
6. ———, *Sofic representations of amenable groups*, Proc. Amer. Math. Soc. **139** (2011), no. 12, 4285–4291. MR 2823074
7. Mikhail Ershov, *Golod-Shafarevich groups: a survey*, Internat. J. Algebra Comput. **22** (2012), no. 5, 1230001, 68. MR 2949205
8. Lev Glebsky, *p-quotients of the G. Higman group*, arXiv:1604.06359 [math.GR].
9. ———, *Cycles in repeated exponentiation modulo p^n*, Integers **13** (2013), Paper No. A66, 7. MR 3118384
10. Mikhail Gromov, *Endomorphisms of symbolic algebraic varieties*, J. Eur. Math. Soc. (JEMS) **2** (1999), 109–197.
11. K. W. Gruenberg, *Residual properties of infinite soluble groups*, Proc. London Math. Soc. (3) **7** (1957), 29–62. MR 0087652
12. Harald Helfgott and Kate Juschenko, *Soficity, short cycles and the Higman group*, 2015, arXiv:1512:02135.

13. Graham Higman, *A finitely generated infinite simple group*, J. London Math. Soc. **26** (1951), 61–64.
14. Tim Hsu and Daniel T. Wise, *On linear and residual properties of graph products*, Michigan Math J. **46** (1999), 251–259.
15. David Kerr and Hanfeng Li, *Soficity, amenability, and dynamical entropy*, Amer. J. Math. **135** (2013), no. 3, 721–761. MR 3068400
16. Nicolas Monod, *Variations on a theme by Higman*, Expo. Math. **35** (2017), no. 2, 226–235. MR 3654077
17. Vladimir G. Pestov, *Hyperlinear and sofic groups: A brief guide*, The Bulletin of Symbolic Logic **14** (2008), 449–480, arXiv:0804.3968.
18. Andreas Thom, *Finitary approximations of groups and their applications*, arXiv:1712.01052.
19. ______, *About the metric approximation of Higman's group*, J. Group Theory **15** (2012), no. 2, 301–310. MR 2900231
20. Benjamin Weiss, *Sofic groups and dynamical systems*, Sankhyā **62** (2000), 350–359.

Martin Kassabov, kassabov@math.cornell.edu
Vivian Kuperberg, viviank@stanford.edu
Timothy Riley, tim.riley@math.cornell.edu

MK and TR: Dept. of Mathematics, 310 Malott Hall, Cornell, Ithaca, NY 14853, USA
VK: Dept. of Mathematics, 450 Serra Mall, Stanford University, Stanford, CA 94305, USA