Séminaire BOURBAKI Novembre 2013
66ème année, 2013-2014, n$^{\text{O}}$ 1077

APPROXIMATE GROUPS
[after Hrushovski, and Breuillard, Green, Tao]

by **Lou van den DRIES**

1. INTRODUCTION

Throughout G is an ambient group. Let $X, Y \subseteq G$, and set

$$XY := \{xy : x \in X,\ y \in Y\}, \qquad X^{-1} := \{x^{-1} : x \in X\},$$
$$X^0 := \{1\} \subseteq G, \quad X^1 := X, \quad X^2 := XX, \quad X^3 := XXX, \text{ and so on.}$$

Let $\langle X\rangle$ denote the subgroup of G generated by X. A *left coset* of X is a translate $gX \subseteq G$ (even if X is not a subgroup of G). We use the term *right coset* in the same way. Call X *symmetric* if $1 \in X$ and $X^{-1} = X$. Throughout, $m, n \in \mathbb{N} = \{0, 1, 2, \ldots\}$ and K, L are real numbers ≥ 1. Note that if X is symmetric, then $\langle X\rangle = \bigcup_n X^n$. When we say that Y is covered by K left (respectively, right) cosets of X we mean that there exists $E \subseteq G$ of cardinality $|E| \leq K$ such that $Y \subseteq EX$ (respectively, $Y \subseteq XE$).

Call X an *approximate group* (in G) if X is symmetric and X^2 can be covered by finitely many left cosets of X (equivalently, by finitely many right cosets of X). Of course, this notion is trivial for finite X. Any compact symmetric neighborhood of the identity in a locally compact group is clearly an approximate group. Call X a *K-approximate group* if X is symmetric and X^2 can be covered by K left cosets of X (equivalently, by K right cosets). This notion is of particular interest when X is finite. It is easy to check that 1-approximate groups in G are subgroups of G.

We think of K as small and fixed, and are interested in the structure of finite K-approximate groups X when its cardinality $|X|$ is large compared to K. On this we have the following result due to Breuillard, Green, Tao [2] and much of it conjectured by H. Helfgott and also by E. Lindenstrauss:

THEOREM 1.1. — *If $X \subseteq G$ is a finite K-approximate group, then there is a K^6-approximate*[1] *group $Y \subseteq X^4$, such that:*

(i) *X is covered by L left cosets of Y, where L depends only on K;*

(ii) *$\langle Y\rangle$ has a d-nilpotent subgroup of finite index, with $d \leq 3\log_2 K$.*

1. The K^6-bound is not in [2]. The bounds in (i) and (ii) are more important.

Here a group H is called *d-nilpotent* ($d \in \mathbb{N}$) if H is generated by elements $u_1, \dots, u_d$ such that $[u_i, u_j] \in \langle u_1, \dots, u_{i-1} \rangle$ whenever $1 \le i < j \le d$, in particular, $u_1 \in \mathrm{Z}(H)$; note that then H is nilpotent of class $\le d$. We also call $u_1, \dots, u_d$ a *nilpotent base* of H if the above holds.

The proof of Theorem 1.1 uses Hrushovski's modeling [13] of limits of finite K-approximate groups by compact neighborhoods of the identity in Lie groups. This may remind you of Gromov [8] on groups of polynomial growth, and among the applications of Theorem 1.1 are indeed strengthenings of Gromov's result. These are derived in Section 3, which also includes a generalized "Margulis Lemma" conjectured by Gromov; for more on this, see the paper by Courtois [3] in this volume.

Theorem 1.1 says that finite K-approximate groups are largely controlled by nilpotent groups. A more detailed version of this theorem in [2] gives even tighter control by so-called *coset nilprogressions*, which generalize symmetric arithmetic progressions in $\mathbb{Z}$. This amounts to a qualitative generalization of earlier "inverse" theorems by Freiman and Ruzsa in *additive combinatorics*, the study of set addition in abelian groups; see Tao and Van Vu [24]. *Multiplicative combinatorics* is its extension to arbitrary groups, and we start with some basic facts from this subject in Section 2 after sketching the proof of Theorem 1.1. That theorem as stated is trivial for finite G (take $Y = X$), but Breuillard showed me a remedy for this: using that $[G, G]$ is finite if $\mathrm{Z}(G)$ has finite index in G, one can replace (ii) in Theorem 1.1 by the following strengthening:

(ii)* there is an m depending only on K and a (finite) normal subgroup $N \subseteq Y^m$ of $\langle Y \rangle$ such that $\langle Y \rangle / N$ is d-nilpotent, with $d \le 3 \log_2 K$.

This is still weaker than the detailed main result in [2], but the proof is almost the same as in the present paper and avoids the more complicated *local* group setting of [2]. How to bound L and m in (i) and (ii)* explicitly in terms of K is not known. Such explicit bounds are known for various natural classes of finite groups; see Helfgott [9, 10, 11].

Sketch of proof for Theorem 1.1

For fixed K, finite K-approximate groups $X_i \subseteq G_i$ as $|X_i| \to \infty$ behave roughly like their (logical) limits $X \subseteq G$ where X is now a *pseudofinite* K-approximate group and the model-theoretic structure (G, X) is *rich* in a certain logical sense. (See Section 4 for the logical notions involved.) The properties of the (pseudo)counting measure on G, normalized so that X has measure 1, lead by a fundamental result in [13] to a group morphism $\pi : \langle X \rangle \to \mathcal{G}$ onto a locally compact group $\mathcal{G}$ with good properties such as $\ker(\pi) \subseteq X^4$.

Yamabe's theorem on approximating locally compact groups by Lie groups permits changing π to a group morphism $\rho : \langle Y \rangle \to \mathcal{H}$ onto a connected Lie group $\mathcal{H}$ for some definable symmetric $Y \subseteq X^4$ such that $\ker(\rho) \subseteq Y$ and finitely many left cosets of Y cover X. (See Section 4 on "definability".) Let H be the smallest definable subgroup

of G containing Y. We use induction on $d := \dim \mathcal{H}$ to construct definable normal subgroups H_i of H such that

$$\{1\} \;=\; H_0 \;\subseteq\; H_1 \subseteq\; H_2 \;\subseteq\; \cdots \;\subseteq\; H_{2d+1} \;=\; H$$

and the quotient H_{i+1}/H_i is pseudofinite for even i, and pseudocyclic and central in H/H_i for odd i. (On general logical grounds and by a group theoretic lemma this gives a weak version of Theorem 1.1, with bounds depending only on K instead of the specific bounds K^6 and $3\log_2 K$. The latter require additional steps.) To prepare for this induction we first use the "no small subgroups" property of Lie groups to shrink Y, without changing $\langle Y\rangle$ or H, so that the image of Y^2 in $\mathcal{H}$ contains no nontrivial subgroup of $\mathcal{H}$. Next, with $\mathbb{N}^* \supseteq \mathbb{N}$ and $\mathbb{R}^*$ in the role of $\mathbb{N}$ and $\mathbb{R}$, we define for $g \in G$ its exit norm (or escape norm) $|g| = |g|_Y \in \mathbb{R}^*$ by

$$|g| \;:=\; \begin{cases} 0 & \text{if } g^\nu \in Y \text{ for all } \nu \in \mathbb{N}^*, \\ 1/\nu & \text{if } \nu \in \mathbb{N}^* \text{ is minimal with } g^\nu \notin Y. \end{cases}$$

Thus $0 \le |g| \le 1$, and $|g| < 1 \;\Leftrightarrow\; g \in Y$. The Lie group $\mathcal{H}$ is controlled near the identity by its Lie algebra via the exponential map, and this allows [2] to adapt arguments stemming from Gleason [6] to show that for some $C \in \mathbb{N}$ and all $g, h \in Y$ we have

$$|gh| \;\le\; C\cdot(|g|+|h|), \qquad |ghg^{-1}| \;\le\; C|h|, \qquad |[g,h]| \;\le\; C\cdot|g|\cdot|h|.$$

This yields a definable normal subgroup of H, namely

$$H_1 \;:=\; \{h \in H :\; |h| = 0\} \;=\; \{h \in H :\; h^\nu \in Y \text{ for all } \nu \in \mathbb{N}^*\}$$

with $H_1 \subseteq \ker(\rho) \subseteq Y$. If $d = 0$, then $H = H_1 = Y = \langle Y\rangle$ and we are done, so assume $d > 0$. Replacing H by H/H_1 and Y by its image in H/H_1 without changing $\mathcal{H}$, we arrange that $|h| > 0$ for all $h \neq 1$ in H. Since Y is pseudofinite, we have $u \in Y$ with minimal $|u| > 0$. Then $|u|$ is infinitesimal, and the bound on the exit norm of commutators $[g,h]$ yields that u lies in the center of H. Let $H_2 := u^{\mathbb{Z}^*}$ be the smallest definable subgroup of H containing u. Replacing H by H/H_2 and Y by its image in H/H_2, we can replace $\mathcal{H}$ by the lower dimensional Lie group $\mathcal{H}/\mathcal{H}_2$, where $\mathcal{H}_2$ is the closure of the central subgroup $\rho(H_2 \cap \langle Y\rangle)$ in $\mathcal{H}$. This decrease in dimension gives by induction the desired result.

Comments

The proof uses the Gleason-Yamabe results [6, 27] around Hilbert's 5th Problem in more than one way. In fact, [2] uses Goldbring's extension [7] of these results to *local* groups where not all products xy may exist. The induction in [2] agrees with our *sketch* in having $H_1 \subseteq X^4$ as an actual pseudofinite group, but differs from it in having H_{i+1}/H_i for all $i > 0$ as a *pseudocyclic local* group quotient. The local group setting in [2] gives sharper results, but it complicates some statements and proofs. We shall avoid local groups and recover indirectly some of the lost information.

The main use of (pseudo)finiteness is in being able to pick as in the *sketch* an element $u \in Y$ with minimal $|u| > 0$. This is similar to a device in Bieberbach's proof [1] of Jordan's theorem on finite complex matrix groups.

After the logical preliminaries in Section 4 we present Hrushovski's Lie modeling in Section 5 ; consider the exit norm in Section 6, and establish Theorem 1.1 in Section 7. This includes the $3\log_2 K$ bound: following [13] and [2], the key point is that the dimension of the above connected nilpotent Lie group $\mathcal{H}$ modulo its largest compact central subgroup is $\leq 3\log_2 K$.

There are plenty of further interesting results in [13] and [2], where the reader can also find references to earlier work on special cases.

I thank Emmanuel Breuillard, Harald Helfgott, and Terence Tao for useful comments on a preliminary version of this paper.

2. MULTIPLICATIVE COMBINATORICS

For a more detailed account of this topic we refer to [23] and to Section 2.7 of [24]. One contrast with additive combinatorics is that the sizes of XY and YX for finite X, Y can be widely different, even when $Y = X^{-1}$. Nevertheless, some basic facts originating in the additive (abelian) setting do go through. Straightforward inductions on n give the following:

LEMMA 2.1. — *Suppose $X \subseteq G$ is symmetric and $X^2 \subseteq EX$, $E \subseteq G$. Then $X^{n+1} \subseteq E^n X$ and $X^{2n} \subseteq E^n X^n$. In particular, if X is a K-approximate group, then X^n is a K^n-approximate group.*

The condition that $X \subseteq G$ is a finite K-approximate group implies that $|X^2| \leq K|X|$. Theorem 2.8 below says that conversely, such a set of small doubling yields a related approximate group. Converting "small doubling" to efficient covering typically goes via the following very useful observation, often called Ruzsa's Covering Lemma.

LEMMA 2.2. — *Let $X, Y \subseteq G$ be finite and nonempty such that $|XY| \leq K|Y|$. Then $X \subseteq EYY^{-1}$ for some $E \subseteq X$ with $|E| \leq K$.*

Proof. — Let $E \subseteq X$ be such that $eY \cap e'Y = \emptyset$ for all distinct $e, e' \in E$. Then $|E| \leq K$, and so by taking E maximal, we get for each $x \in X$ an $e \in E$ with $xY \cap eY \neq \emptyset$, so $x \in eYY^{-1}$. □

COROLLARY 2.3. — *Suppose $X \subseteq G$ is a finite K-approximate group and $S \subseteq G$ is symmetric, $S^4 \subseteq X^4$, $|S| \geq c|X|$, $0 < c \leq 1$. Then X^4 is covered by K^7/c left cosets of S^2 and thus S^2 is a (K^7/c)-approximate group.*

Proof. — Take $E \subseteq G$ such that $X^2 \subseteq EX$ and $|E| \leq K$. Then $X^4 S \subseteq X^8 \subseteq E^7 X$, so $|X^4 S| \subseteq K^7|X| \leq (K^7/c)|S|$. Then Ruzsa's covering lemma provides $F \subseteq X^4$ with $|F| \leq K^7/c$ and $X^4 \subseteq FS^2$. □

COROLLARY 2.4. — *Suppose $X \subseteq G$ is symmetric, finite, and $|X^5| \leq K|X|$. Then X^2 is a K-approximate group.*

Proof. — From $|X^4X| \leq K|X|$ we obtain by Lemma 2.2 that X^4 is covered by K left cosets of X^2. □

For finite nonempty $X, Y \subseteq G$ we define their *Ruzsa distance*

$$d(X,Y) \;:=\; \log \frac{|XY^{-1}|}{|X|^{1/2}|Y|^{1/2}}.$$

It is easy to check that $d(X,X) = 0$ iff X is a right coset of a finite subgroup H of G, namely $H = XX^{-1}$. So $d(X,X) > 0$ is more typical, but in other respects the Ruzsa distance does behave like a metric:

LEMMA 2.5. — *Let $X, Y, Z \subseteq G$ be finite and nonempty. Then*

$$d(X,Y) \geq 0, \quad d(X,Y) = d(Y,X), \quad d(X,Z) \leq d(X,Y) + d(Y,Z).$$

For the triangle inequality, use $xz^{-1} = xy^{-1} \cdot yz^{-1}$. We can now derive an analogue of Corollary 2.4 for symmetric sets of "small tripling":

LEMMA 2.6. — *Let $X \subseteq G$ be symmetric and finite with $|X^3| \leq K|X|$. Then $|X^n| \leq K^{c_n}|X|$ for all $n \geq 1$, with $c_n \geq 0$ depending only on n. Moreover, X^2 is a K^5-approximate group.*

Proof. — Clearly the lemma holds for $n = 1, 2, 3$ with $c_1 = 0$ and $c_2 = c_3 = 1$. Note also that $d(X^2, X) \leq \log K$. Assume $n \geq 3$ and $|X^{n-1}| \leq K^{c_{n-1}}|X|$ and $|X^n| \leq K^{c_n}|X|$ with $c_{n-1},\ c_n \geq 0$. Then $d(X^{n-1}, X) \leq c_n \log K$, so

$$\begin{aligned} d(X^{n-1}, X^2) \;&\leq\; (c_n + 1)\log K, \text{ so} \\ |X^{n+1}| \;&\leq\; K^{c_n+1}|X^{n-1}|^{1/2}|X^2|^{1/2} \\ &\leq\; K^{c_n+1}K^{c_{n-1}/2}K^{1/2}|X|, \end{aligned}$$

so the lemma holds with $c_{n+1} = c_n + 1 + (c_{n-1} + 1)/2$. This gives $c_4 = 3$ and $c_5 = 5$, so $|X^5| \leq K^5|X|$, and thus X^2 is a K^5-approximate group. □

Corollary 2.4 and Lemma 2.6 might suggest that if $X \subseteq G$ is finite symmetric with $|X^2| \leq K|X|$, then X^2 is an L-approximate group where L depends only on K. This holds with $L = K^5$ if the ambient group G is commutative: [24], 6.29. But it fails in general, [24], p. 94: for a finite subgroup H of G and $X = H \cup \{a, a^{-1}\}$, $a \in G$, we have $|X^2|/|X| \leq 4$, but the size of $|X^3|/|X|$ can be arbitrarily large. Fortunately, Theorem 2.8 provides a good substitute. The proof of this theorem rests on the following result.

PROPOSITION 2.7. — *Let $X \subseteq G$ be finite and symmetric with $|X^2| \leq K|X|$. Then $S := \{s \in G :\ |X \cap Xs| > |X|/2K\}$ is symmetric, $S \subseteq X^2$ and*

$$|S| \geq |X|/2K, \qquad |XS^nX| \;\leq\; 2^n K^{2n+1}|X| \quad \textit{for all } n.$$

I omit the proof, which takes about a page in [23].

THEOREM 2.8. — *Suppose $X \subseteq G$ is symmetric, finite, and $|X^2| \leq K|X|$. Then there exists a $64K^{12}$-approximate group $Y \subseteq X^4$, such that X can be covered by $4K^4$ left cosets of Y, and $|Y| \leq 4K^5|X|$.*

Proof. — From Proposition 2.7 we get a symmetric $S \subseteq X^2$ such that

$$\begin{aligned} |S| \;&\geq\; |X|/2K, \qquad & |XSX| \;&\leq\; 2K^3|X|, \\ |XS^2X| \;&\leq\; 4K^5|X|, \qquad & |XS^5X| \;&\leq\; 32K^{11}|X|. \end{aligned}$$

In particular,

$$|S^5| \;\leq\; 32K^{11}|X| \;\leq\; 64K^{12}|S|,$$

so $Y := S^2$ is a $64K^{12}$-approximate group by Lemma 2.4. Therefore,

$$\begin{aligned} |Y| \;&= |S^2| \;\leq\; 4K^5|X|, \\ |XS| \;&\leq\; |XSX| \;\leq\; 2K^3|X| \;\leq\; 4K^4|S|. \end{aligned}$$

From $|XS| \leq 4K^4|S|$ we get by Ruzsa's covering lemma a set $D \subseteq X$ such that $|D| \leq 4K^4$ and $X \subseteq DS^2 = DY$. □

The setting in [23] for results like those above is more general in that G can be a unimodular locally compact group equipped with a Haar measure, with nonempty open precompact subsets of G and their measure instead of nonempty finite subsets of G and their cardinality. Some bounds in [23] are given as being polynomial in K, but not in explicit form like $64K^{12}$. New proofs of results like Theorem 2.8 were recently given by Ruzsa [20] and Petridis [17].

The following "slicing" lemma due to Helfgott [10] will also be very useful:

LEMMA 2.9. — *Let $X \subseteq G$ be a K-approximate group, and let H be a subgroup of G. Then $Y := X^2 \cap H$ is a K^3-approximate group in H, and $X^4 \cap H$ can be covered by K^3 left cosets of Y.*

Proof. — We have $Y^2 \subseteq X^4 \cap H$, so it is enough to show that $X^4 \cap H$ can be covered by K^3 left cosets of Y. Now X^4 can be covered by K^3 left cosets of X. Consider a left coset gX of X that has an element h in common with $X^4 \cap H$. Then $h^{-1}g \in X$, so $h^{-1}gX \cap H \subseteq Y$, and thus

$$gX \cap (X^4 \cap H) \;\subseteq\; gX \cap H \;\subseteq\; hY.$$

Thus $X^4 \cap H$ can indeed be covered by K^3 left cosets of Y. □

Sanders-Croot-Sisask

These names refer to [21] and [4], both of which prove (more than) Theorem 2.11 below. In Section 5 we use its consequence Corollary 2.13 to construct certain locally compact groups. (In this we follow [2] rather than [13].)

LEMMA 2.10. — *Let $f : (0,1] \to [1,K]$ be any function and $m \geq 2$. Then there is an ϵ with $0 < \epsilon < 1$ depending only on K, m such that*

$$f\big(\frac{t^2}{2K}\big) \ > \big(1 - \frac{1}{m}\big) f(t) \quad \textit{for some } t \textit{ with } \epsilon < t \leq 1.$$

Proof. — Take $n \geq 1$ such that $\big(1 - \frac{1}{m}\big)^n K < 1$. For example, this holds for

$$n \ := \ \text{integral part of} \quad \frac{\log K}{\log(m/(m-1))} + 1.$$

Then the nth iterate of the map $t \mapsto t^2/2K : (0,1] \to (0,1]$ evaluated at $t_0 = 1$ gives a number $\epsilon := t_n = (1/2K)^{2^n - 1}$ such that the lemma holds for this ϵ and $t = t_i = (1/2K)^{2^i - 1}$ for some $i < n$. □

We only need this lemma and the next theorem for $m = 96$.

THEOREM 2.11. — *Let $X \subseteq G$ be finite and symmetric, with $|X^2| \leq K|X|$, and let $m \geq 2$. Then there exists a symmetric $S \subseteq G$ such that $|S| \geq c|X|$ and $S^m \subseteq X^4$, where c with $0 < c < 1$ depends only on K, m.*

Proof. — Let $Y \subseteq G$ be nonempty and finite, and set

$$S \ = \ S(Y) \ := \ \{s \in G : \ |Y \setminus sY| < \frac{1}{m}|Y|\},$$

so S is symmetric. For $s_1, s_2 \in S$ we have $|Y \setminus s_1 Y| < \frac{1}{m}|Y|$, $|Y \setminus s_2 Y| < \frac{1}{m}|Y|$, so $|s_1 Y \setminus s_1 s_2 Y| < \frac{1}{m}|Y|$, and thus $|Y \setminus s_1 s_2 Y| < \frac{2}{m}|Y|$. Iterating this we get $|Y \setminus gY| < |Y|$ for all $g \in S^m$, so $Y \cap gY \neq \emptyset$ for those g, hence $S^m \subseteq YY^{-1}$. Thus it is enough to find a nonempty finite $Y \subseteq X^2$ such that $|S^m| \geq c|X|$ for $S = S(Y)$, with $c > 0$ depending only on K, m. For $0 < t \leq 1$, let the real number $f(t) \in [1, K]$ be given by

$$f(t) := \min\{\frac{|X'X|}{|X|} : \ X' \subseteq X, \ |X'| \geq t|X|\}.$$

By Lemma 2.10 we can take t such that

$$f(\frac{t^2}{2K}) \ > \ \big(1 - \frac{1}{m}\big) f(t), \qquad \epsilon < t \leq 1$$

with ϵ in the interval $0 < \epsilon < 1$ depending only on K, m. Take $X' \subseteq X$ such that $|X'| \geq t|X|$ and $|X'X| = f(t)|X|$. Then $S := S(Y)$ with $Y := X'X$ can be shown to have the desired properties with $c = \epsilon^2/2K$: this takes about a page of computations in [2]. □

We focused on the construction of S, in order to exhibit S as "definable" when this becomes relevant in Section 5. What we really need there is a "normal" variant, Corollary 2.13. Towards its proof we first establish an approximate version of the well-known fact that if $G_1, G_2, \dots, G_n$ are subgroups of finite index in G, $n \geq 1$, then so is $G_1 \cap \cdots \cap G_n$, with

$$[G : G_1 \cap \cdots \cap G_n] \ \leq \ [G : G_1] \cdots [G : G_n].$$

LEMMA 2.12. — *Let $X \subseteq G$ be finite and symmetric with $|X^2| \leq K|X|$, let $n \geq 1$, and let $X_1, \dots, X_n$ be subsets of X such that $|X_i| \geq \delta_i |X|$ and $\delta_i > 0$ for $i = 1, \dots, n$. Then there is a set $D \subseteq X$ such that*

$$DD^{-1} \ \subseteq \ X_1 X_1^{-1} \cap \cdots \cap X_n X_n^{-1}, \quad |D| \ \geq \ \frac{\delta_1 \cdots \delta_n}{K^{n-1}} |X|.$$

Proof. — For $n = 2$, use $|X_1^{-1} X_2| \leq K|X|$ to pick $g \in X_1^{-1} X_2$ such that

$$|D| \ \geq \ (\delta_1 \delta_2 / K)|X|, \quad D \ := \ \{x_2 \in X_2 : \ x_1^{-1} x_2 = g \text{ for sone } x_1 \in X_1\}.$$

The general case follows by induction on n. □

Notation: for $x, a \in G$ and $X, Y \subseteq G$ we set:

$$x^a \ := \ a^{-1} x a, \qquad X^Y \ := \ \{x^y : \ x \in X, \ y \in Y\}.$$

COROLLARY 2.13. — *Let $X \subseteq G$ be a finite K-approximate group and let $Y \subseteq X$ be symmetric, $|Y| \geq \delta |X|$, $\delta > 0$. Then for some symmetric $E \subseteq G$,*

$$|E| \ \geq \ \varepsilon |X|, \quad (E^{16})^X \ \subseteq \ Y^4,$$

where $\varepsilon > 0$ depends only on K and δ.

Proof. — We have $|Y^2| \leq K|X| \leq (K/\delta)|Y|$, so by Theorem 2.11 applied to Y in the role of X we get a symmetric $S \subseteq G$ with $S^{96} \subseteq Y^4$ and $|S| \geq c|X|$, where $c > 0$ depends only on K and δ. Then $S \subseteq X^4$, so

$$|XS| \ \leq \ |X^5| \ \leq \ K^4 |X| \ \leq \ (K^4/c)|S|,$$

so Ruzsa's covering lemma gives $X \subseteq \bigcup_{i=1}^n a_i S^2$ with $1 \leq n \leq (K^4/c) + 1$, $a_1, \dots, a_n \in X$, $a_1 = 1$. Now $a_i S a_i^{-1} \subseteq X^6$ for $i = 1, \dots, n$, so Lemma 2.12 gives $D \subseteq X^6$ with $DD^{-1} \subseteq a_i S^2 a_i^{-1}$ for $i = 1, \dots, n$, and $|D| \geq \varepsilon |X|$ where $\varepsilon > 0$ depends only on K, δ. Then $E := DD^{-1}$ is symmetric, $E^{a_i} \subseteq S^2$ for $i = 1, \dots, n$, and $|E| \geq \varepsilon |X|$, so $E^X \subseteq S^6$ in view of $X \subseteq \bigcup_{i=1}^n a_i S^2$. Thus

$$(E^{16})^X \ = \ (E^X)^{16} \ \subseteq \ S^{96} \ \subseteq \ Y^4.$$ □

3. APPLICATIONS

The applications of Theorem 1.1 we present here are from Section 11 in [2], which has further elaborations and also connections to geometry.

LEMMA 3.1. — *Let $G = \langle S \rangle$ with symmetric $S \subseteq G$, and let G_1 be a subgroup of G. If $[G : G_1] = \infty$, then S^n meets at least $n+1$ different left cosets of G_1, for each n. If $[G : G_1] = d < \infty$, then S^n meets at least $n+1$ different left cosets of G_1, for each $n < d$.*

Proof. — Note that $S^0 = \{1\}$ meets exactly one left coset of G_1. Consider first the case that $[G : G_1] = d < \infty$, and suppose towards a contradiction that S^m meets at most m left cosets of G_1, where $m < d$. This gives $i < m$ such that S^i and S^{i+1} meet the same left cosets of G_1, so $S^i G_1 = S^{i+1} G_1$. Then $S^i G_1 = S^j G_1$ for all $j \geq i$, so $S^i G_1 = \langle S \rangle G_1 = G$, so S^i must meet at least d left cosets of G_1, a contradiction. Likewise with $[G : G_1] = \infty$. □

Call a group *virtually d-nilpotent* (where $d \in \mathbb{N}$) if it has a d-nilpotent subgroup of finite index. Note: the next result does not assume $X \subseteq \langle S \rangle$.

COROLLARY 3.2. — *Let K be given, and let $L \in \mathbb{N}^{\geq 1}$ be as in Theorem 1.1. Suppose $X \subseteq G$ is a finite K-approximate group and $S \subseteq G$ is symmetric with $S^L \subseteq X$. Then $\langle S \rangle$ is virtually d-nilpotent with $d \leq 3 \log_2 K$.*

Proof. — Take Y as in Theorem 1.1, so X is covered by L left cosets of Y. Then S^L is covered by L left cosets of $\langle Y \rangle$, so S^L is covered by L left cosets of $G_1 := \langle Y \rangle \cap \langle S \rangle$, and hence $[\langle S \rangle : G_1] \leq L$ by Lemma 3.1. Now $\langle Y \rangle$ has a d-nilpotent subgroup G_0 of finite index with $d \leq 3 \log_2 K$. Then $G_0 \cap G_1$ has finite index in $\langle S \rangle$ and is d-nilpotent. □

The generalized Margulis Lemma conjectured by Gromov follows from:

THEOREM 3.3. — *Let $K \in \mathbb{N}^{\geq 1}$. Then there is an $\varepsilon = \varepsilon(K)$, $0 < \varepsilon \leq 2$, with the following property. Let (M, d) be a metric space and let $a \in M$ be such that the closed ball of radius 4 centered at a can be covered by K closed balls of radius 1. Let Γ be a subgroup of the isometry group of (M, d) such that there are only finitely many $\gamma \in \Gamma$ with $d(\gamma a, a) \leq 2$. Then the finite set*

$$S_\varepsilon(a) \ := \ \{\gamma \in \Gamma : \ d(\gamma a, a) \leq \varepsilon\}$$

generates a virtually d-nilpotent subgroup of Γ with $d \leq 3 \log_2 K$.

Proof. — For $x \in M$ and r a positive real number, set

$$\begin{aligned} B_r(x) \ &:= \ \{y \in M : \ d(x, y) \leq r\}, \\ S_r(x) \ &:= \ \{\gamma \in \Gamma : \ d(\gamma x, x) \leq r\} \qquad \text{(a symmetric subset of } \Gamma\text{)}. \end{aligned}$$

Take $a_1, \ldots, a_K \in M$ such that $B_4(a) \subseteq \bigcup_{i=1}^K B_1(a_i)$. We can assume that for $i = 1, \ldots, k$ with $k \leq K$ there exists $\gamma_i \in S_4(a)$ with $\gamma_i a \in B_1(a_i)$, and for $k < i \leq K$

there is no such γ_i. Let $\gamma \in S_4(a)$ be arbitrary. Then we have $i \in \{1, \dots, k\}$ with $d(\gamma a, \gamma_i a) \leq 2$, so $\gamma_i^{-1}\gamma \in X := S_2(a)$. This gives

$$X^2 \ \subseteq \ S_4(a) \ \subseteq \ \bigcup_{i=1}^{k} \gamma_i X,$$

so $X \subseteq \Gamma$ is a K-approximate group. With L as in Corollary 3.2, set $\varepsilon := 2/L$, and apply that corollary with $G := \Gamma$, $S := S_\varepsilon(a)$. □

See [3] for other versions of the Margulis Lemma and its role in geometry. Next an easy lemma for functions of polynomial growth:

LEMMA 3.4. — *If $c, d \in \mathbb{N}$ and $f : \mathbb{N} \to \mathbb{R}^{>0}$ is an increasing function with $f(n) \leq cn^d$ for all n, and $K > 5^d$, then $\{n : \ f(5n) \leq Kf(n)\}$ is infinite.*

COROLLARY 3.5 (Gromov). — *If G is a finitely generated group of polynomial growth, then G has a nilpotent subgroup of finite index.*

Proof. — Let $G = \langle S \rangle$ with finite symmetric $S \subseteq G$ such that $|S^n| \leq cn^d$ for all n, with constants $c, d \in \mathbb{N}$. Let $K > 5^d$. By Lemma 3.4 applied to

$$n \mapsto |S^n| \ : \ \mathbb{N} \to \mathbb{R}^{>0}$$

we have $|S^{5n}| \leq K|S^n|$ for infinitely many n. Take L as in Theorem 1.1. and take $n \geq L/2$ such that $|S^{5n}| \leq K|S^n|$. Then $X := S^{2n}$ is a K-approximate group by Corollary 2.4, and thus satisfies the hypothesis of Corollary 3.2. □

Next we derive a finitary refinement of Gromov's theorem using likewise a finitization of Lemma 3.4 on functions of polynomial growth:

LEMMA 3.6. — *Let $d \in \mathbb{N}$, set $K := 5^{d+1}$, and let $L \geq K$. Then for every increasing function $f : \mathbb{N} \to \mathbb{R}^{>0}$ and any $n \geq 5^{d+2}L^{d+1}$,*

$$f(n) \ \leq \ f(1)n^d \implies f(5m) \ \leq \ Kf(m) \text{ for some } m \text{ with } L \leq m \leq n/5.$$

The finitary refinement is of the kind that a degree d bound at just one large enough scale is enough for virtual nilpotence; cf. [15, 22].

COROLLARY 3.7. — *Let $d \in \mathbb{N}$. Then there is a positive integer $N(d)$ with the following property: if $G = \langle S \rangle$ with finite symmetric $S \subseteq G$ and $|S^n| \leq |S|n^d$ for some $n \geq N(d)$, then G is virtually $7(d+1)$-nilpotent.*

Proof. — Let $K := 5^{d+1}$, and take $L \geq K$, $L \in \mathbb{N}$ as in Theorem 1.1. Set

$$N(d) \ := \ 5^{d+2}L^{d+1}.$$

Suppose $G = \langle S \rangle$ with finite symmetric $S \subseteq G$ and $|S^n| \leq |S|n^d$ for some $n \geq N(d)$. Then Lemma 3.6 gives $m \geq L$ such that $|S^{5m}| \leq K|S^m|$. It remains to apply Corollary 3.2 to the K-approximate group $X := S^{2m}$. □

We finish this section by weakening slightly the hypothesis on X and the conclusion on $\langle S \rangle$ in Corollary 3.2. This requires the following lemma.

LEMMA 3.8. — *Let $X \subseteq G$ be symmetric and finite with $|X^2| \leq K|X|$. Then for some L depending only on K and some symmetric $Z \subseteq X^{16}$,*

(i) *X can be covered by L left cosets of Z;*

(ii) *$\langle Z \rangle$ is virtually d-nilpotent with $d \leq 18 + 36 \log_2 K$.*

Proof. — Theorem 2.8 gives a $64K^{12}$-approximate group $Y \subseteq X^4$ such that X can be covered by $4K^4$ left cosets of Y. Now apply Theorem 1.1 to Y and $64K^{12}$ in the role of X and K. □

COROLLARY 3.9. — *Let K be given, and let $L \in \mathbb{N}^{\geq 1}$ be as in Corollary 3.8. Suppose $S, X \subseteq G$ are finite symmetric such that $S^L \subseteq X$ and $|X^2| \leq K|X|$. Then $\langle S \rangle$ is virtually d-nilpotent, with $d \leq 18 + 36 \log_2 K$.*

Proof. — Take Z as in Lemma 3.8, and argue as in the proof of Corollary 3.2, with Z instead of Y. □

4. LOGICAL PRELIMINARIES

The *sketch* in the introduction refers to *definable*, *pseudofinite* and *rich*. In this section we define these notions and give enough background to work with them. Most of this section is written for those unfamiliar with model theory and consists of foundational generalities, but the "bounded quotients" at the end are of recent vintage; cf. Pillay [18], Section 2. The logical setting is essentially that of [13] and yields algorithms to compute some bounds that in the ultraproduct setting of [2] are purely existential.

Some notation: for a relation $R \subseteq P \times Q$, $p \in P$, and $X \subseteq P$ we set

$$R(p) \;:=\; \{q \in Q :\; (p,q) \in R\}, \qquad R(X) \;:=\; \bigcup_{x \in X} R(x) \;\subseteq\; Q.$$

In particular, if E is an equivalence relation on a set P and $p \in P$, then $E(p)$ is the equivalence class of p.

A (model-theoretic) *structure* $\mathcal{M}$ consists of a family $(M_i)_{i \in I}$ of nonempty sets M_i, and of a family $(R_j)_{j \in J}$ of relations $R_j \subseteq M_{i_1} \times \cdots \times M_{i_m}$ on these sets, with the finite sequence $i_1, \ldots, i_m$ in I depending on j; notation:

$$\mathcal{M} \;=\; \Big((M_i); (R_j)\Big).$$

Often these relations are (graphs of) functions $M_{i_1} \times \cdots \times M_{i_n} \to M_{i_{n+1}}$. Call the M_i the *underlying sets* of the structure, and the R_j its *primitives*. More precisely, $\mathcal{M}$ above is an I-sorted structure, and many texts only consider 1-sorted structures where $I = \{1\}$ and so doesn't need to be mentioned. But the extra generality adds useful flexibility and is natural. Virtually anything that mathematicians consider as a structure can be viewed as a structure in the above sense: for example groups as 1-sorted structures, with the product operation as primitive, a group G acting on a set S as a 2-sorted structure,

with G and S as underlying sets [2], and the product operation of G and the action map $G \times S \to S$ as primitives. Also a topological space is naturally a 2-sorted structure, with the underlying set of the space as first underlying set [3] and the collection of open subsets of the space as second underlying set, with the membership relation between points and open sets as primitive.

We construe a finite approximate group $X \subseteq G$ as a 2-sorted structure as follows. The first underlying set is G, the second underlying set is $\mathbb{R}$. The primitives are: the product operation of G, the subset X of G, the addition, multiplication, and ordering on $\mathbb{R}$, the subset $\mathbb{Z}$ of $\mathbb{R}$, and a bijection

$$X \to \{1, \ldots, n\} \quad \text{with } n = |X| \qquad \text{(a witness to } X \text{ being finite).}$$

The graph of this bijection is regarded as a subset of $G \times \mathbb{R}$.

Definable sets

Let $\mathcal{M} = \big((M_i); (R_j)\big)$ be an I-sorted structure. The main role of the primitives of $\mathcal{M}$ is to generate the definable sets of $\mathcal{M}$. Let $\boldsymbol{i} = i_1, \ldots, i_m$ and $\boldsymbol{j} = j_1, \ldots, j_n$ range over finite tuples from I, set $M_{\boldsymbol{i}} := M_{i_1} \times \cdots \times M_{i_m}$, and identify $M_{\boldsymbol{i}} \times M_{\boldsymbol{j}}$ with $M_{\boldsymbol{i},\boldsymbol{j}}$ in the obvious way. Then the 0-definable (or absolutely definable) sets of $\mathcal{M}$ are the relations $X \subseteq M_{\boldsymbol{i}}$, with $\boldsymbol{i}$ part of the specification, obtained recursively as follows:

(1) the primitives $R_j \subseteq M_{\boldsymbol{i}}$ of $\mathcal{M}$ are 0-definable;

(2) for $i \in I$ the diagonal $\{(x, y, z) \in M_{i,\boldsymbol{j},i} : \ x = z\}$ is 0-definable;

(3) if $X \subseteq M_{\boldsymbol{i}}$ is 0-definable, then so is its complement in $M_{\boldsymbol{i}}$;

(4) if $X, Y \subseteq M_{\boldsymbol{i}}$ are 0-definable, then so are $X \cup Y, X \cap Y \subseteq M_{\boldsymbol{i}}$;

(5) if $X \subseteq M_{\boldsymbol{i}}$ and $Y \subseteq M_{\boldsymbol{j}}$ are 0-definable, then so is $X \times Y \subseteq M_{\boldsymbol{i},\boldsymbol{j}}$;

(6) if $X \subseteq M_{\boldsymbol{i},\boldsymbol{j}}$ is 0-definable, then so is the projection $\pi(X) \subseteq M_{\boldsymbol{i}}$ where $\pi : M_{\boldsymbol{i},\boldsymbol{j}} \to M_{\boldsymbol{i}}$ is the obvious projection map.

This can be traced back to Weyl [26] (1910). [4] We extend this notion to *A-definability*. Here A is a so-called *parameter set*, that is, A is a family $(A_i)_{i \in I}$ with $A_i \subseteq M_i$ for all i; notation: $A \subseteq (M_i)$. Then the structure $\mathcal{M}_A$ is obtained from $\mathcal{M}$ by adding for each $i \in I$ and $a_i \in A_i$ the set $\{a_i\} \subseteq M_i$ as primitive; the A-definable sets of $\mathcal{M}$ are just the 0-definable sets of $\mathcal{M}_A$, equivalently: a set $X \subseteq M_{\boldsymbol{j}}$ with $\boldsymbol{j} = j_1, \ldots, j_n$ is A-definable (in $\mathcal{M}$) iff for some $\boldsymbol{i} = i_1, \ldots, i_m$ and 0-definable relation $R \subseteq M_{\boldsymbol{i}} \times M_{\boldsymbol{j}} = M_{\boldsymbol{i},\boldsymbol{j}}$ of $\mathcal{M}$ and $a \in A_{\boldsymbol{i}}$ we have $X = R(a)$. For $A = (M_i)$ we just write "definable" instead of "A-definable" and so all finite sets $X \subseteq M_{\boldsymbol{i}}$ are definable. For any set $S \subseteq M_{\boldsymbol{i}}$, not necessarily definable, $\mathrm{Def}(S)$ is the collection of sets $X \subseteq S$ that are definable in $\mathcal{M}$; thus $X, Y \in \mathrm{Def}(S) \Rightarrow X \cup Y, X \cap Y, X \setminus Y \in \mathrm{Def}(S)$. A parameter set $A \subseteq (M_i)$ is said to be *countable* if all A_i are countable, and $A_i \neq \emptyset$ for only countably many $i \in I$.

2. To fit our definition of *structure* we also require $S \neq \emptyset$.
3. This underlying set should be nonempty to conform with our notion of *structure*.
4. I don't know if Weyl's paper had any influence, for example on Tarski.

An example

In general the 0-definable relations of a structure cannot be described in a significantly more explicit way than by the above recursive definition. One case where a more explicit description does exist is the field $\mathbb{C}$ of complex numbers. Let us construe $\mathbb{C}$ as a 1-sorted structure with addition and multiplication as the primitives. (Including also as primitive, say, $x \mapsto x^{-1} : \mathbb{C}^\times \to \mathbb{C}$, wouldn't add to the 0-definable relations of $\mathbb{C}$.) By the Chevalley-Tarski constructibility theorem the 0-definable subsets of $\mathbb{C}^n$ are just the finite unions of sets of the form

$$\{a \in \mathbb{C}^n : \ f_1(a) = \cdots = f_m(a) = 0, \ g(a) \neq 0\}$$

with $f_1, \ldots, f_m, g \in \mathbb{Q}[x_1, \ldots, x_n]$. For a subfield A of $\mathbb{C}$ as parameter set we get the same description of the A-definable subsets of $\mathbb{C}^n$ but now the polynomials have their coefficients in A. The above goes through for any algebraically closed field instead of $\mathbb{C}$, with the prime field in place of $\mathbb{Q}$.

So in this case the notion of "A-definable" is akin to Weil's notion of an algebraic variety being defined over A. Model-theoretic notions are often similar to foundational items in Weil's algebraic geometry. For example, the rich structures considered below are like Weil's universal domains.

Definable quotients

Given $\mathcal{M}$ as above, a 0-definable set $X \subseteq M_{\boldsymbol{i}}$, and a 0-definable equivalence relation $E \subseteq X \times X \subseteq M_{\boldsymbol{i},\boldsymbol{i}}$ on X, let some map $\pi : X \to Q$ onto a set Q be given with kernel E, that is,

$$\pi(x) \ = \ \pi(y) \iff xEy, \qquad (x, y \in X).$$

The usual choice would be to take $Q = X/E$ with π the natural quotient map. Let $\mathcal{M}_\pi$ be the structure $\mathcal{M}$ with one more underlying set, namely $Q = \pi(X)$, and with the graph of π as extra primitive. Then a set $Y \subseteq M_{\boldsymbol{j}}$ is 0-definable in $\mathcal{M}$ iff it is 0-definable in $\mathcal{M}_\pi$. Also, a set $Y \subseteq M_{\boldsymbol{j}}$ is definable in $\mathcal{M}$ iff it is definable in $\mathcal{M}_\pi$. We consider $\mathcal{M}$ to be *expanded* (as the terminology goes) by such a quotient to $\mathcal{M}_\pi$, whenever convenient.

Formulas

An I-sorted language is a set $\mathcal{L}$ whose elements are so-called *relation symbols*, and each relation symbol $\underline{R} \in \mathcal{L}$ is equipped with a finite sequence $\boldsymbol{i} = i_1, \ldots, i_m$ in I, its *sort*. An $\mathcal{L}$-structure is an I-sorted structure $\mathcal{M}$ as above, together with a bijection $\mathcal{L} \to J$ onto the index set of the family of primitives $(R_j)_{j \in J}$; for $\underline{R} \in \mathcal{L}$ of sort $\boldsymbol{i}$ corresponding to j under this bijection we require that $R_j \subseteq M_{\boldsymbol{i}}$ and we say that $\underline{R}$ names R_j. In this way we can construe for example all groups as $\mathcal{L}$-structures for a single 1-sorted language $\mathcal{L}$ with just one ternary relation symbol (naming in each group the graph of its product operation).

Besides the symbols from $\mathcal{L}$ we also have the logical relation symbol $=$, and, for each $i \in I$, infinitely many *variables* of sort i (just symbols). With these we form *atomic* $\mathcal{L}$-formulas $x = y$ where x, y are variables of the same sort, and $\underline{R}x_1 \dots x_n$ where $\underline{R}$ is a relation symbol in $\mathcal{L}$ of sort $\boldsymbol{i} = i_1, \dots, i_n$ and $x_1, \dots, x_n$ are variables of sort $i_1, \dots, i_n$, respectively. Starting with these atomic formulas, we now use the logical symbols $\neg$, $\wedge$, $\vee$, $\exists$, $\forall$ in the familiar way to form arbitrary $\mathcal{L}$-formulas. (Strictly speaking, a formula is a finite sequence of symbols formed according to certain recursive rules, but going into more detail here would be distracting.)

Let $\phi(x_1, \dots, x_n)$ be an $\mathcal{L}$-formula; the notation indicates that $x_1, \dots, x_n$ are distinct variables, and that any variable occurring free (not bound by a quantifier $\exists$ or $\forall$) in the formula is among $x_1, \dots, x_n$. Let $x_1, \dots, x_n$ be of sort $i_1, \dots, i_n$, respectively, and $\boldsymbol{i} := i_1, \dots, i_n$. Then $\phi(x_1, \dots, x_n)$ defines in any I-sorted structure $\mathcal{M}$ a certain 0-definable set $\phi(M_{\boldsymbol{i}}) \subseteq M_{\boldsymbol{i}}$, consisting of the tuples $(a_1, \dots, a_n) \in M_{\boldsymbol{i}}$ for which the formula becomes true in $\mathcal{M}$ when $a_1, \dots, a_n$ are substituted for the free occurrences of $x_1, \dots, x_n$ in the formula. (The reader can supply a precise recursive definition of $\phi(M_{\boldsymbol{i}})$. It is easy to check that every 0-definable set $X \subseteq M_{\boldsymbol{i}}$ has the above form $\phi(M_{\boldsymbol{i}})$.)

Sentences

Let $\mathcal{M}$ be an $\mathcal{L}$-structure. When no variables occur free in an $\mathcal{L}$-formula, we call it a *sentence*, and such a sentence σ is either true in $\mathcal{M}$ or not (in the latter case its negation $\neg\sigma$ is true in $\mathcal{M}$). *Example*: consider groups as $\mathcal{L}$-structures where $\mathcal{L} = \{\underline{R}\}$ and the ternary relation symbol $\underline{R}$ names in each group the graph of its product operation. Let $xy = yx$ denote the formula $\exists z\big(\underline{R}xyz \wedge \underline{R}yxz\big)$. Let G be any group. Then the formula $xy = yx$ defines in G the set

$$\{(a, b) \in G \times G : \ ab = ba\} \ \subseteq \ G \times G,$$

the formula $\phi(x) := \forall y(xy = yx)$ defines in G its center, while the sentence $\forall x \forall y(xy = yx)$ is true in G iff G is commutative.

Logical compactness

Let Σ be a set of $\mathcal{L}$-sentences. Then a *model* of Σ is by definition an $\mathcal{L}$-structure in which all sentences of Σ are true. "Logical compactness" is the fact that Σ has a model iff each finite subset of Σ has a model. (It is a consequence of a more precise result, namely Gödel's completeness theorem.) Suppose all models of Σ are known to have a certain property that can be expressed as an infinite disjunction of $\mathcal{L}$-sentences, more precisely, we have a sequence $\sigma_0, \sigma_1, \sigma_2, \dots$ of $\mathcal{L}$-sentences such that in every model of Σ one of the σ_i is true. Then, by logical compactness, there is n such that in every model of Σ one of the σ_i with $i \leq n$ is true. If in addition we can effectively enumerate Σ as well as the sequence $\sigma_0, \sigma_1, \dots$, then we can find such n: by Gödel's completeness theorem, there will be for some n a formal proof of $\sigma_0 \vee \dots \vee \sigma_n$ from Σ, and by

systematically listing proofs from Σ, we eventually find such a proof.[5] Whenever we claim a (computable) bound "by logical compactness," this is what we have in mind.

Pseudofiniteness

Let $\mathcal{M}$ be a structure with a definable ordered field $\mathbb{R}^*$, that is, $\mathbb{R}^*$ is an ordered field whose underlying set is a definable set of $\mathcal{M}$ such that the ordering, addition, and multiplication of $\mathbb{R}^*$ are definable in $\mathcal{M}$ as well. We also assume given a set $\mathbb{Z}^* \subseteq \mathbb{R}^*$, definable in $\mathcal{M}$, such that the following conditions are satisfied:

(i) $\mathbb{R}^*$ is *definably complete*: every nonempty definable $S \subseteq \mathbb{R}^*$ with an upper bound in $\mathbb{R}^*$ has a least upper bound in $\mathbb{R}^*$,

(ii) $\mathbb{Z}^*$ is a subring of $\mathbb{R}^*$ and *discrete*: there is no $k \in \mathbb{Z}^*$ with $0 < k < 1$.[6]

Of course, "definability" in (i) refers to $\mathcal{M}$. It follows from (i) and (ii) that $\mathbb{Z}^*$ is cofinal in $\mathbb{R}^*$: for each $r \in \mathbb{R}^*$ there is $k \geq r$ in $\mathbb{Z}^*$, more precisely, for each $r \in \mathbb{R}^*$ there is $k \in \mathbb{Z}^*$ such that $k \leq r < k + 1$. It also follows that $\mathbb{N}^* := (\mathbb{Z}^*)^{\geq 0}$ is *definably wellordered*: each nonempty definable set $S \subseteq \mathbb{N}^*$ has a least element. The ordinary mathematics based on $\mathbb{N}$ (and induction) and $\mathbb{R}$ (and completeness), goes through with the definably wellordered $\mathbb{N}^*$ in the role of $\mathbb{N}$ and the definably complete $\mathbb{R}^*$ instead of $\mathbb{R}$, provided we stick to definable relations. We shall freely avail ourselves of this principle, and refer to [5] for a result that justifies it: There is a definable relation $E \subseteq \mathbb{R}^* \times \mathbb{R}^*$ such that $\{E(r) : r \in \mathbb{R}^*\} = \mathrm{Def}(\mathbb{N}^*)$. In effect, this allows us to use the membership relation and (universally and existentially) quantify over $\mathrm{Def}(\mathbb{N}^*)$ without getting out of the realm of definable relations. Put

$$[N] := \{\nu \in \mathbb{N}^* : 1 \leq \nu \leq N\} \qquad (N \in \mathbb{N}^*).$$

Let Y be a definable set of our structure $\mathcal{M}$. We declare Y to be *pseudofinite*[7] if there is a definable bijection $Y \to [N]$ for some $N \in \mathbb{N}^*$; such N is uniquely determined, and we call it the pseudocardinality of Y, and set $|Y| := N$. This behaves just like ordinary finite cardinality:

(1) if Y is pseudofinite, then so is every definable subset of Y;

(2) if $Y, Z \subseteq M_{\boldsymbol{i}}$ are pseudofinite, then so is $Y \cup Z \subseteq M_{\boldsymbol{i}}$, and

$$|Y \cup Z| + |Y \cap Z| = |Y| + |Z|;$$

(3) if $Y \subseteq M_{\boldsymbol{i}}$ and $Z \subseteq M_{\boldsymbol{j}}$ are pseudofinite, then so is $Y \times Z \subseteq M_{\boldsymbol{i},\boldsymbol{j}}$, with $|Y \times Z| = |Y| \cdot |Z|$.

We stress that "pseudofinite" includes being definable. Of course, if Y is finite, it is pseudofinite and $|Y|$ has the usual meaning. Note: "pseudofinite" is relative to an ambient $\mathcal{M}$ with distinguished definable ordered field $\mathbb{R}^*$ and definable $\mathbb{Z}^* \subseteq \mathbb{R}^*$; in later use these will be clear from the context.

5. This algorithm to find n is easy to program, but its actual use is of course not practical.
6. As an aside, given (i) there is at most one definable $\mathbb{Z}^* \subseteq \mathbb{R}^*$ satisfying (ii).
7. In other contexts "pseudofinite" can have a different meaning.

Pseudofinite approximate groups

In the beginning of this section we construed finite approximate groups as 2-sorted structures. A *pseudofinite* approximate group is likewise a 2-sorted structure consisting of:

(i) a group G with a distinguished approximate group $X \subseteq G$,

(ii) an ordered field $\mathbb{R}^*$ equipped with a discrete subring $\mathbb{Z}^*$,

and an additional primitive: a bijection $X \to [N]$ with $N \in \mathbb{N}^* := (\mathbb{Z}^*)^{\geq 0}$; in the resulting 2-sorted structure $\mathcal{M}$ we require $\mathbb{R}^*$ to be definably complete.

A finite approximate group is just a pseudofinite approximate group with $\mathbb{R}^* = \mathbb{R}$ and $\mathbb{Z}^* = \mathbb{Z}$. We fix a single (finite) language $\mathcal{L}$ such that all pseudofinite approximate groups are $\mathcal{L}$-structures. It is easy to specify a set Σ_K of $\mathcal{L}$-sentences whose models are exactly the pseudofinite K-approximate groups: expressing definable completeness needs infinitely many sentences, and the other conditions can be expressed by finitely many sentences.

Elementary equivalence

Two $\mathcal{L}$-structures $\mathcal{M}$ and $\mathcal{N}$ are said to be *elementarily equivalent* if for every $\mathcal{L}$-sentence σ we have:

$$\sigma \text{ is true in } \mathcal{M} \iff \sigma \text{ is true in } \mathcal{N}.$$

For example, two algebraically closed fields are elementarily equivalent iff they have the same characteristic. Much deeper is the result (Sela) that any two noncommutative free groups are elementarily equivalent. We are not going to use these facts, and just mention them by way of illustrating the notion of elementary equivalence. What we need, for countable $\mathcal{L}$, is the rather elementary fact that any $\mathcal{L}$-structure is elementarily equivalent to some *rich* $\mathcal{L}$-structure. We define "rich" in the next subsection.

Rich structures

Working in rich structures is a way to make efficient use of logical compactness. We call an $\mathcal{L}$-structure $\mathcal{M} = \big((M_i); (R_j)\big)$ *rich*[8] if $\mathcal{L}$ is countable and for each countable $A \subseteq (M_i)$ and each $i \in I$, every family of A-definable subsets of M_i with the finite intersection property has nonempty intersection in M_i. ("Finite intersection property": every finite subfamily has nonempty intersection.) This automatically extends to the cartesian products $M_{\boldsymbol{i}}$, as the reader may easily verify.

In topological terms: if $\mathcal{M}$ is rich, and $A \subseteq (M_i)$ is countable, then the (countable) collection of A-definable sets $X \subseteq M_{\boldsymbol{i}}$ is a basis for a quasi-compact topology on $M_{\boldsymbol{i}}$, the A-topology, and the A-definable sets $X \subseteq M_{\boldsymbol{i}}$ are exactly the open-and-closed sets in this topology. We frequently use this as follows: if $\mathcal{M}$ is rich and $X_m, Y_n \subseteq M_{\boldsymbol{i}}$

8. The more usual terminology is "$\aleph_1$-saturated" instead of "rich" except that the former does not require that $\mathcal{L}$ be countable. In our situation countability of $\mathcal{L}$ is convenient.

are definable for $m, n = 0, 1, 2, \ldots$ and $\bigcap_m X_m \subseteq \bigcup_n Y_n$, then there are m, n such that $\bigcap_{i=0}^{m} X_i \subseteq \bigcup_{j=0}^{n} Y_j$.

Assume $\mathcal{M}$ is rich. Let $X \subseteq M_{\boldsymbol{i}}$. Then X is Σ-*definable* if $X = \bigcup_n X_n$ for some definable sets $X_n \subseteq M_{\boldsymbol{i}}$, equivalently, there is a countable $A \subseteq (M_i)$ such that X is A-open in $M_{\boldsymbol{i}}$. We say that X is Π-*definable* if its complement in $M_{\boldsymbol{i}}$ is Σ-definable, that is, $X = \bigcap_n X_n$ for some definable $X_n \subseteq M_{\boldsymbol{i}}$. Thus:

$$X \text{ is definable } \iff X \text{ is } \Sigma\text{-definable and } \Pi\text{-definable.}$$

If $\mathcal{M}$ is rich, then so is $\mathcal{M}_A$ for any countable parameter set $A \subseteq (M_i)$ and any expansion $\mathcal{M}_\pi$ of $\mathcal{M}$ by a definable quotient.

Suppose the language $\mathcal{L}$ is countable. As mentioned before, every $\mathcal{L}$-structure is elementarily equivalent to a rich $\mathcal{L}$-structure. For model-theorists this is a routine consequence of logical compactness. For those familiar with ultraproducts, it can also be seen as follows: Let $(\mathcal{M}_n)$ be a sequence of $\mathcal{L}$-structures and let α be a nonprincipal ultrafilter on $\mathbb{N}$. Then the ultraproduct $\prod_{n\to\alpha} \mathcal{M}_n$ (in the notation of [2]) is a rich $\mathcal{L}$-structure. Taking $\mathcal{M}_n = \mathcal{M}$ for all n, this ultraproduct is elementarily equivalent to $\mathcal{M}$. If all the $\mathcal{M}_n$ are finite K-approximate groups (for the same K), then their ultraproduct is a pseudofinite K-approximate group.

Recall that we have a finite language $\mathcal{L}$ for pseudofinite approximate groups and a set Σ_K of $\mathcal{L}$-sentences whose models are the pseudofinite K-approximate groups. Here is how this is relevant for getting the bound L in Theorem 1.1. Let K be given. The existence, for every rich pseudofinite K-approximate group, of some pseudofinite Y and definable subgroups H_i of $\langle Y \rangle^*$ as in Theorem 7.2 means that every rich pseudofinite K-approximate group—and therefore, every pseudofinite K-approximate group—satisfies a certain infinite disjunction $\bigvee_n \sigma_n$ of $\mathcal{L}$-sentences σ_n, that is, every pseudofinite K-approximate group makes one of the σ_n true. Then by logical compactness, some $\sigma_0 \vee \cdots \vee \sigma_n$ (depending on K) is true in all pseudofinite K-approximate groups, and thus in all finite K-approximate groups.

The standard part

Let $\mathcal{M}$ be a rich structure, and let $\mathbb{R}^*$ be an ordered field definable in $\mathcal{M}$, that is, its underlying set is a definable set of $\mathcal{M}$ and the ordering, addition, and multiplication of $\mathbb{R}^*$ are definable in $\mathcal{M}$. We identify $\mathbb{Q}$ with the prime field of $\mathbb{R}^*$. Since $\mathcal{M}$ is rich, there will be elements $r > \mathbb{Q}$ in $\mathbb{R}^*$. Let $\mathcal{O}$ be the bounded part of $\mathbb{R}^*$, that is,

$$\mathcal{O} := \{r \in \mathbb{R}^* : |r| \leq n \text{ for some } n\}.$$

So $\mathcal{O}$ is a convex subring of $\mathbb{R}^*$, with maximal ideal

$$o := \{r \in \mathbb{R}^* : |r| \leq 1/n \text{ for all } n \geq 1\}.^{(9)}$$

9. The symbols $\mathcal{O}$ and o are to remind the reader of Landau's big O and small o.

It may help to think of the elements of $\mathcal{o}$ as the *infinitesimals* of $\mathbb{R}^*$. Note that $\mathcal{O}$ is Σ-definable and $\mathcal{o}$ is Π-definable. Define the *standard part map*

$$\mathrm{st} \ : \ \mathcal{O} \to \mathbb{R}$$

to be the unique ring morphism that respects $\leq$: it sends each $r \in \mathcal{O}$ to the nearest real number s, that is, $s \in \mathbb{R}$ and for all rational q_1, q_2, if $q_1 \leq r \leq q_2$, then $q_1 \leq s \leq q_2$. From the richness of $\mathcal{M}$ it follows that st is surjective. Its kernel is $\mathcal{o}$, and thus $\mathcal{O}/\mathcal{o} \cong \mathbb{R}$.

Bounded quotients

Rich structures may be artifacts, but they breed "natural" objects in the form of (non-definable) quotients; example: the above isomorphism $\mathcal{O}/\mathcal{o} \cong \mathbb{R}$. We use this in Section 5 to construct a locally compact group from a pseudofinite approximate group. The bounded quotients below appear in [18, 14, 13] and other places, and I only add here an interpretation in terms of uniform spaces.

Assume $\mathcal{M} = \big((M_i); \dots\big)$ is rich. Let an (ambient) definable set $D \subseteq M_{\boldsymbol{i}}$ be given and a Π-definable equivalence relation $\mathbf{E}$ on D. A routine logical compactness argument yields definable binary relations E_n on D such that

$$\mathbf{E} = \bigcap_n E_n, \quad \text{and } E_{n+1} \subseteq E_n^{-1}, \quad E_{n+1} \circ E_{n+1} \subseteq E_n \text{ for all } n.$$

Below we fix such a sequence (E_n). Then (E_n) is a base of entourages for a uniform structure on D which is independent of the choice of (E_n). Taking D as a uniform space in this way, each point $a \in D$ has neighborhood base $\{E_n(a) : n = 0, 1, \dots\}$, and if $X \subseteq D$ is open or closed, then X is $\mathbf{E}$-saturated, that is, $X = \mathbf{E}(X)$. The quotient space $D/\mathbf{E}$ is hausdorff.

Next we fix an $\mathbf{E}$-saturated Σ-definable set $\mathbf{S} \subseteq D$. Then $\mathbf{S}$ is open in D, and we consider $\mathbf{S}$ as a subspace of the uniform space D. Call $\mathbf{E}$ *bounded on* $\mathbf{S}$ if for each n countably many sets $E_n(a)$ with $a \in \mathbf{S}$ cover $\mathbf{S}$. This is again independent of (E_n). For $\mathbf{S} = \bigcup_m S_m$ where each S_m is definable we have: $\mathbf{E}$ is bounded on $\mathbf{S}$ iff each S_m is totally bounded, that is, S_m is covered for each n by finitely many sets $E_n(a)$ with $a \in S_m$.

Assume below that $\mathbf{E}$ is bounded on $\mathbf{S}$. For each n, pick points $a_{m,n} \in \mathbf{S}$ such that $\mathbf{S} \subseteq \bigcup_{m=0}^{\infty} E_n(a_{m,n})$. Take a countable $A \subseteq (M_i)$ such that all E_n and all $\{a_{m,n}\}$ are A-definable. The topology of the space $\mathbf{S}$, the A-topology on $M_{\boldsymbol{i}}$, and the logical notion of Σ-definability are closely related:

Lemma 4.1. — *The interiors in* $\mathbf{S}$ *of the sets* $E_n(a_{m,n}) \subseteq \mathbf{S}$ *form a countable base for the topology of* $\mathbf{S}$. *For* $X \subseteq \mathbf{S}$, *the following are equivalent:*

(i) X *is open in the space* $\mathbf{S}$;

(ii) X *is* $\mathbf{E}$*-saturated and* A*-open;*

(iii) X *is* $\mathbf{E}$*-saturated and* Σ*-definable.*

Proof. — Let X be open in $\mathbf{S}$. Given $x \in X$, take m, n such that $E_n(x) \subseteq X$, and $(x, a_{m,n+1}) \in E_{n+1}$, so $x \in E_{n+1}(a_{m,n+1}) \subseteq E_n(x) \subseteq X$. Thus X is a union of A-definable sets $E_{n+1}(a_{m,n+1})$. This proves (i)⇒(ii) and the countable base claim. The direction (ii)⇒(iii) is obvious. Assume (iii), and let $x \in X$. From $\mathbf{E}(x) = \bigcap_n E_n(x) \subseteq X$ we get n with $E_n(x) \subseteq X$. □

To keep notation simple, let $\mathbf{S}/\mathbf{E}$ denote the image of $\mathbf{S}$ in the (hausdorff) quotient space $D/\mathbf{E}$. We consider $\mathbf{S}/\mathbf{E}$ as a subspace of $D/\mathbf{E}$ and we let $\pi : \mathbf{S} \to \mathbf{S}/\mathbf{E}$ be the canonical map. Then the topology of $\mathbf{S}/\mathbf{E}$ is also the quotient topology induced by π: a set $Y \subseteq \mathbf{S}/\mathbf{E}$ is open iff $\pi^{-1}(Y)$ is open in $\mathbf{S}$, iff $\pi^{-1}(Y)$ is Σ-definable (by Lemma 4.1). Because of this relation to the logical notion of Σ-definability, this topology on $\mathbf{S}/\mathbf{E}$ is called the *logic topology*. Nevertheless, this topology is also "classical":

Corollary 4.2. — *The space $\mathbf{S}/\mathbf{E}$ is locally compact and second countable. (The latter means there is a countable base for the topology.) For $Y \subseteq \mathbf{S}/\mathbf{E}$,*

$$Y \text{ is compact} \iff \pi^{-1}(Y) \subseteq M_{\boldsymbol{i}} \text{ is } \Pi\text{-definable.}$$

If $X \subseteq \mathbf{S}$ is Π-definable as subset of $M_{\boldsymbol{i}}$, then $\pi(X)$ is compact.

The proof is an exercise in point set topology, using the lemma above, and the quasi-compactness of the A-topology on $M_{\boldsymbol{i}}$.

Bounded quotient groups

Let $\mathcal{M}$ be rich, and assume the group G is definable in $\mathcal{M}$, that is, the underlying set of G and the graph of its product operation are definable in $\mathcal{M}$. In addition, let G^Σ be a subgroup of G whose underlying set is Σ-definable, and let G^Π be a subgroup of G^Σ whose underlying set is Π-definable. Thus G^Σ is $\mathbf{E}$-saturated, where $\mathbf{E}$ is the Π-definable equivalence relation on G given by $x\mathbf{E}y \Leftrightarrow x \in yG^\Pi$. With $D := G$ and $\mathbf{S} := G^\Sigma$ this puts us in the situation of the previous subsection.

Lemma 4.3 (cf. [13], 3.3). — *The following are equivalent:*

(1) $\mathbf{E}$ *is bounded on G^Σ;*

(2) *for all definable $X, Y \subseteq G^\Sigma$ with $X \supseteq G^\Pi$, finitely many left cosets of X cover Y.*

We leave the proof of Lemma 4.3 and that of the next result as an exercise in point set topology. Let us call G^Σ/G^Π a *bounded quotient* if the equivalent conditions (1) and (2) of Lemma 4.3 are satisfied.

Corollary 4.4. — *Let G^Π be normal in G^Σ with bounded quotient $\mathcal{G} = G^\Sigma/G^\Pi$. Then $\mathcal{G}$ with the logic topology is a locally compact topological group. Let $\pi : G^\Sigma \to \mathcal{G}$ be the canonical map. Then every definable $X \subseteq G$ with $G^\Pi \subseteq X$ contains $\pi^{-1}(U)$ for some neighborhood U of the identity in $\mathcal{G}$.*

5. HRUSHOVSKI'S LIE MODEL THEOREM

We now fix a rich pseudofinite K-approximate group $X \subseteq G$. As specified in Section 4, this is officially a 2-sorted structure $\mathcal{M}$ with underlying sets G and $\mathbb{R}^*$ and among its primitives the set $X \subseteq G$, a subring $\mathbb{Z}^*$ of $\mathbb{R}^*$ and a bijection $X \to [N]$ with $N \in \mathbb{N}^* := (\mathbb{Z}^*)^{\geq 0}$.

The subgroup $\langle X \rangle = \bigcup_n X^n$ of G is Σ-definable, and if $Y \in \mathrm{Def}(\langle X \rangle)$, then $Y \subseteq X^n$ for some n, and so Y is pseudofinite with $|Y| \leq K^n|X|$ for such n. Normalizing the above pseudocounting and taking standard parts yields therefore a finitely additive *real valued* measure $\boldsymbol{\mu}$ on $\mathrm{Def}(\langle X \rangle)$:

$$\boldsymbol{\mu}(Y) \ := \ \mathrm{st}\left(|Y|/|X|\right) \qquad \text{with } \boldsymbol{\mu}(X) = 1.$$

Note that $\boldsymbol{\mu}$ is left-and-right invariant: $\boldsymbol{\mu}(aY) = \boldsymbol{\mu}(Y) = \boldsymbol{\mu}(Ya)$ for all $a \in \langle X \rangle$ and all $Y \in \mathrm{Def}(\langle X \rangle)$. Also, $\boldsymbol{\mu}(X^4) \leq K^3$.

The "measure zero" ideal $\{Y \in \mathrm{Def}(\langle X \rangle) : \ \boldsymbol{\mu}(Y) = 0\}$ is an invariant S1-ideal, in Hrushovski's terminology. See his paper [13] for what this means. In view of Theorem 3.5 and Corollary 3.6 from [13] we may conclude:

THEOREM 5.1. — *There is a Π-definable normal subgroup $\mathcal{o}(X)$ of $\langle X \rangle$ such that $\mathcal{o}(X) \subseteq X^4$, and the quotient $\mathcal{G} := \langle X \rangle / \mathcal{o}(X)$ is bounded.*

As the notation suggests, we think of the elements of $\mathcal{o}(X)$ as *infinitesimals*. Let $\pi : \langle X \rangle \to \mathcal{G}$ be the canonical map, so $\ker(\pi) = \mathcal{o}(X) \subseteq X^4$. We make $\mathcal{G}$ into a locally compact group by giving it the logic topology: a set $S \subseteq \mathcal{G}$ is open iff $\pi^{-1}(S) \subseteq G$ is Σ-definable. It follows from Corollaries 4.2 and 4.4 that $\pi(X^4)$ is a compact neighborhood of the identity in $\mathcal{G}$.

In the space available I cannot give an adequate account of the very general [10] Theorem 3.5 in [13] of which the above Theorem 5.1 is a special case. I can present, however, another proof of the latter that was found later by Breuillard, Green, Tao [2] and is based on Corollary 2.13 above:

LEMMA 5.2. — *There is a descending sequence*

$$X^4 \ = \ X_0 \ \supseteq \ X_1 \ \supseteq \ X_2 \ \supseteq \ \cdots \ \supseteq X_n \ \supseteq \ \ldots$$

of definable symmetric subsets of G such that for every n,

$$X_{n+1}^2 \ \subseteq \ X_n, \qquad X_{n+1}^X \ \subseteq \ X_n$$

and X^4 can be covered by finitely many left cosets of X_n.

10. Hrushovski has further generalized this to "approximate equivalence relations."

Proof. — Suppose $Y \subseteq G$ is a definable symmetric set such that $Y^4 \subseteq X^4$ and X^4 can be covered by finitely many left cosets of Y. Applying the pseudofinite version of Lemma 2.13 to X^4 in the role of X gives a definable symmetric $S \subseteq Y^4$ such that $(S^{16})^X \subseteq Y^4$ and $|S| \geq \varepsilon |X^4|$ for some rational $\varepsilon > 0$. Note that X^4 can be covered by finitely many left cosets of $Z := S^2$, by the pseudofinite version of Corollary 2.3. Moreover,

$$Z^4 \subseteq Y^4, \qquad (Z^4)^2 = S^{16} \subseteq Y^4, \qquad (Z^4)^X \subseteq Y^4.$$

Applying this construction of Z from Y recursively gives a sequence

$$X = Y_0, Y_1, Y_2, \ldots$$

of definable approximate groups in G such that with $X_n := Y_n^4$ we have for all n: $X_{n+1} \subseteq X_n$, $X_{n+1}^2 \subseteq X_n$, $X_{n+1}^X \subseteq X_n$, and $X_0 = X^4$ is covered by finitely many left cosets of Y_n and thus of X_n. □

Given a descending sequence $X_0 \supseteq X_1 \supseteq X_2 \supseteq \cdots$ as in Lemma 5.2, it follows easily that $\mathcal{o}(X) := \bigcap X_n$ is as in Theorem 5.1: it is a Π-definable normal subgroup of $\langle X \rangle$ contained in X^4, with bounded quotient $\langle X \rangle / \mathcal{o}(X)$.

In the rest of this section we fix $\mathcal{o}(X)$ as in Theorem 5.1, we equip $\mathcal{G} := \langle X \rangle / \mathcal{o}(X)$ with its logic topology, and let $\pi : \langle X \rangle \to \mathcal{G}$ be the canonical map. It is worth mentioning that $\pi(X^2)$ is a neighborhood of the identity in $\mathcal{G}$. This can be seen as follows. The neighborhood $\pi(X^4)$ of $1 \in \mathcal{G}$ is covered by K^3 left cosets of the compact set $\pi(X)$, so $\pi(X)$ has nonempty interior, and thus $\pi(X^2) = \pi(X)\pi(X)^{-1}$ is a neighborhood of $1 \in \mathcal{G}$. (We will not use this, but it is also of interest to note that $\mathcal{G}$ is necessarily unimodular: the above finitely additive measure $\boldsymbol{\mu}$ on $\mathrm{Def}(\langle X \rangle)$ induces a left-and-right-invariant Haar measure on $\mathcal{G}$.)

Good models

In the *sketch* in Section 1 our morphism $\pi : \langle X \rangle \to \mathcal{G}$ got modified repeatedly. Following [2] we formalize the properties to be preserved under these modifications in the notion of a good model. [11] Let H be a group definable in some ambient rich structure $\mathcal{N}$: the underlying set of H and the graph of its product operation are definable in $\mathcal{N}$. Let $Y \subseteq H$ be definable and symmetric, so $\langle Y \rangle = \bigcup_n Y^n \subseteq H$ is Σ-definable. A *good model* of $Y \subseteq H$ is a surjective group morphism $\rho : \langle Y \rangle \to \mathcal{H}$ onto a second countable locally compact group $\mathcal{H}$ such that:

(g1) $\rho^{-1}(U) \subseteq Y$ for some neighborhood U of 1 in $\mathcal{H}$; so $\ker(\rho) \subseteq Y$;

(g2) the closure of $\rho(Y)$ in $\mathcal{H}$ is compact;

(g3) for all compact $C \subseteq \mathcal{H}$ and open $U \subseteq \mathcal{H}$ with $C \subseteq U$, we have $\rho^{-1}(C) \subseteq D \subseteq \rho^{-1}(U)$ for some definable $D \subseteq H$.

11. This use of "model" is in the spirit of the Freiman models in additive combinatorics, and does not correspond to the use of this term in model theory.

(It follows from (g1) and (g2) that Y is an approximate group in H.) Note that our $\pi : \langle X \rangle = \langle X^4 \rangle \to \mathcal{G}$ is a good model of $X^4 \subseteq G$.

LEMMA 5.3. — *Let $\rho : \langle Y \rangle \to \mathcal{H}$ be a good model of Y. Then $\ker(\rho)$ is Π-definable, the quotient $\langle Y \rangle / \ker(\rho)$ is bounded, and the map*

$$a \ker(\rho) \mapsto \rho(a) \ : \ \langle Y \rangle / \ker(\rho) \to \mathcal{H} \qquad (a \in \langle Y \rangle)$$

is a topological group isomorphism, with the logic topology on $\langle Y \rangle / \ker(\rho)$.

Proof. — If $C \subseteq \mathcal{H}$ is compact, then C is a countable intersection of open sets in $\mathcal{H}$, so $\rho^{-1}(C)$ is Π-definable by (g3). Thus $\ker(\rho)$ is Π-definable. Using (g2) it follows that the quotient $\langle Y \rangle / \ker(\rho)$ is bounded. Also, if $U \subseteq \mathcal{H}$ is open, then U is a countable union of compact subsets, so $\rho^{-1}(U)$ is Σ-definable. Thus the bijection of the lemma is continuous, and as $\langle Y \rangle / \ker(\rho)$ is σ-compact, it is a homeomorphism. □

See [2] for instructive examples of good models. Good models are robust:

LEMMA 5.4. — *Let $\rho : \langle Y \rangle \to \mathcal{H}$ be a good model of Y. Then:*

(i) *For any open subgroup $\mathcal{H}'$ of $\mathcal{H}$, the set $Y \cap \rho^{-1}(\mathcal{H}') \subseteq Y$ is definable, symmetric, and contains $\ker(\rho)$.*

(ii) *If $Y' \subseteq Y$ is definable, symmetric, and contains $\ker(\rho)$, then $\rho(\langle Y' \rangle)$ is open in $\mathcal{H}$, and $\rho|_{\langle Y' \rangle} : \langle Y' \rangle \to \rho(\langle Y' \rangle)$ is a good model of Y'.*

(iii) *If N is a compact normal subgroup of $\mathcal{H}$ and $\rho^{-1}(N) \subseteq Y$, then the composition $\langle Y \rangle \to \mathcal{H} \to \mathcal{H}/N$ is a good model of Y.*

Proof. — Suppose $\mathcal{H}'$ is an open subgroup of $\mathcal{H}$. Then the set $Y \cap \rho^{-1}(\mathcal{H}')$ is Σ-definable, and its complement in Y is also Σ-definable since $\mathcal{H}'$ is closed in $\mathcal{H}$. This gives (i). Use Corollary 4.4 for the rest. □

Lie models

We recall here a version of Yamabe's Theorem: *For any locally compact group $\mathcal{G}$ and any neighborhood U of the identity in $\mathcal{G}$, there is an open subgroup $\mathcal{G}'$ of $\mathcal{G}$ and a compact normal subgroup $N \subseteq U$ of $\mathcal{G}'$ such that $\mathcal{G}'/N$ is a connected Lie group.* This allows us to upgrade our original good model $\pi : \langle X \rangle \to \mathcal{G}$ of X^4 to a Lie model:

THEOREM 5.5. — *There exists a definable K^6-approximate group $Y \subseteq G$ such that $\ker(\pi) \subseteq Y \subseteq X^4$, and such that there is a good model $\rho : \langle Y \rangle \to \mathcal{H}$ of Y onto a connected Lie group $\mathcal{H}$.*

Proof. — Let U be an open neigborhood of the identity in $\mathcal{G}$ with $\pi^{-1}(U) \subseteq X^4$, and take an open subgroup $\mathcal{G}'$ of $\mathcal{G}$ and a compact normal subgroup $N \subseteq U$ of $\mathcal{G}'$ such that $\mathcal{G}'/N$ is a connected Lie group. Set $Y := X^4 \cap \pi^{-1}(\mathcal{G}')$. Since X^2 is a K^2-approximate group, Y is a K^6-approximate group by Lemma 2.9. Applying Lemma 5.4 to the good model π of X^4 we see that $Y \subseteq G$ is definable, $\pi(\langle Y \rangle)$ is open in $\mathcal{G}$, and π restricts to a good model $\langle Y \rangle \to \pi(\langle Y \rangle)$ of Y. But $\pi(\langle Y \rangle) = \mathcal{G}'$, since $\pi^{-1}(N) \subseteq Y$, so $N \subseteq \pi(Y)$ and

$\pi(\langle Y\rangle)/N$ is open in $\mathcal{G}'/N$, which is connected. Composing this good model $\langle Y\rangle \to \mathcal{G}'$ with the canonical map $\mathcal{G}' \to \mathcal{G}'/N$ gives a good model $\langle Y\rangle \to \mathcal{G}'/N$ of Y onto a connected Lie group, by (3) of Lemma 5.4. □

Except for the K^6-bound, this is part of Hrushovski's Theorem 4.2 in [13] (which has other parts). We now turn to an application.

The case of finite exponent

Let $e \in \mathbb{N}^{\geq 1}$. A subset S of a group is said to have *exponent* e if $s^e = 1$ for all $s \in S$. The next result is Corollary 4.18 in [13], and Theorem 6.15 in [2], except that the hypothesis there is a bit stronger, namely that the ambient group has exponent e. In the additive (abelian) setting, the result is due to Ruzsa [19].

COROLLARY 5.6. — *Let S be a finite K-approximate group such that S^2 has exponent e. Then S^4 contains a subgroup H of $\langle S\rangle$ such that S can be covered by L left cosets of H, with L depending only on K, e.*

Proof. — Our rich pseudofinite K-approximate group $X \subseteq G$ is arbitrary, so by logical compactness it suffices to show: if X^2 has exponent e, then X^4 contains a definable subgroup $H \supseteq \ker(\pi)$ of G (and thus finitely many left cosets of H cover X.)

So assume X^2 has exponent e. Since $\pi(X^2)$ is a neighborhood of the identity in $\mathcal{G}$, we have an open neigborhood U of the identity in $\mathcal{G}$ such that $\pi^{-1}(U) \subseteq X^4$ and U has exponent e. Then the proof of Theorem 5.5 yields a symmetric definable $Y \subseteq G$ with $\ker(\pi) \subseteq Y \subseteq X^4$ and a good model $\rho : \langle Y\rangle \to \mathcal{H}$ of Y onto a connected Lie group $\mathcal{H}$ with a neighborhood of its identity of exponent e. So $\mathcal{H}$ is trivial, and thus $H := \ker(\rho) = Y = \langle Y\rangle$ has the desired properties. □

The proof gives an algorithm that on any input $K, e \in \mathbb{N}^{\geq 1}$ finds an $L \in \mathbb{N}^{\geq 1}$ as in Corollary 5.6. Breuillard has an example with $K = 4$, $e = 2$ showing that "S^2 has exponent e" cannot be weakened to "S has exponent e".

6. THE EXIT NORM

We give here an account of Sections 7, 8 in [2], with different or more explicit constants in some places. The main result is Theorem 6.4. But first we relate the "no small subgroups" property of Lie groups to a "trapping property" as it is called in [2] (with easy proof left to the reader):

LEMMA 6.1. — *Let $\mathcal{G}$ be a locally compact group and V a compact neighborhood of $1 \in \mathcal{G}$ that contains no subgroup of $\mathcal{G}$ other than $\{1\}$. Then there is for each neighborhood U of $1 \in \mathcal{G}$ an $n = n(U) \geq 1$ such that for all $a \in \mathcal{G}$,*

$$a^i \in V \text{ for } i = 1, \ldots, n \implies a \in U.$$

For the rest of this section we fix an ambient structure $\mathcal{M}$ with an ordered field $\mathbb{R}^*$, definable in $\mathcal{M}$ and definably complete, and a discrete definable subring $\mathbb{Z}^*$ of $\mathbb{R}^*$. This gives sense to "pseudofinite". (At this point we do allow $\mathbb{R}^* = \mathbb{R}$ and $\mathbb{Z}^* = \mathbb{Z}$, in which case *pseudofinite* just means *finite*. Later we add richness of $\mathcal{M}$ as extra assumption.) Let ν range over $\mathbb{N}^*$.

We also fix a definable group G in $\mathcal{M}$. We need to consider "products" $g_1 \cdots g_\nu$ where $g_1, \ldots, g_\nu$ is a definable sequence in G (that is, the graph of the map $i \mapsto g_i : [\nu] \to G$ is definable in $\mathcal{M}$). Such a sequence gives rise to a unique definable sequence $h_1, \ldots, h_\nu$ in G of "partial products" with $h_1 = g_1$ and $h_{i+1} = h_i g_{i+1}$ for all $i \in [\nu]$ with $i < \nu$. We set $g_1 \cdots g_\nu := h_\nu$, with the convention that this equals $1 \in G$ if $\nu = 0$. If all g_i equal the same element $g \in G$, we denote $g_1 \cdots g_\nu$ by g^ν. Set $g^{-\nu} := (g^\nu)^{-1}$. For definable symmetric $X \subseteq G$, let $\langle X \rangle^*$ be the set of all products $g_1 \cdots g_\nu$ with $g_1, \ldots, g_\nu$ a definable sequence in X. Then $\langle X \rangle^*$ is the smallest definable subgroup of G that contains X. Of course, $\langle X \rangle \subseteq \langle X \rangle^*$. Also, $g^{\mathbb{Z}^*} := \{g^k : \ k \in \mathbb{Z}^*\}$ is the smallest definable subgroup of G containing $g \in G$.

Below g, h range over G. Let also a definable symmetric $X \subseteq G$ be given. The *exit norm* $|g|_X$ of g with respect to X is the element of $\mathbb{R}^*$ defined by

$$|g|_X \ := \ \inf\{\frac{1}{\nu+1} : \ g^i \in X \text{ for } i = 1, \ldots, \nu\}, \text{ so } \ 0 \le |g|_X \le 1,$$
$$|g|_X \le 1/2 \iff g \in X, \qquad |g|_X = 0 \iff g^\nu \in X \text{ for all } \nu.$$

Here the infimum is taken in $\mathbb{R}^*$. So the longer it takes for the powers of g to exit X, the smaller $|g|_X$. Note that $g \mapsto |g|_X : G \to \mathbb{R}^*$ is definable, and we have the norm-like properties $|1|_X = 0$ and $|g^{-1}|_X = |g|_X$. To get something more useful, we assume a trapping condition:

Lemma 6.2. — *Assume that for all g, if $g^i \in X^4$ for $i = 1, \ldots, 8$, then $g \in X$. Then for all g, all $x \in X$, and all $\nu \ge 1$,*

(i) $|g|_X \ \le \ 4|g|_{X^2} \ \le \ 8|g|_{X^4}$;

(ii) $|g|_X \ < \ \frac{1}{\nu} \implies |g^\nu|_X \ \ge \ \frac{\nu}{4}|g|_X$ (escape property);

(iii) $|g^x|_X \ \le \ 8|g|_X$ (conjugacy bound).

This is straightforward, although (ii) requires some care. When X is clear from the context, we drop the subscript X in $|g|_X$. Let us say that X has *strong exit norm* if there is a constant $C \in \mathbb{N}^{\ge 1}$ such that for every definable sequence $g_1, \ldots, g_\nu$ in G, and all g, h,

(s1) $|g_1 \cdots g_\nu| \ \le \ C \cdot \big(|g_1| + \cdots + |g_\nu|\big)$;

(s2) $h \in X \implies |g^h| \ \le \ C \cdot |g|$;

(s3) $g, h \in X \implies |[g, h]| \ \le \ C \cdot |g| \cdot |h|$.

From (s3) we get: if $0 < |g|, |h| < 1/C$, then $|[g,h]| < \min\big(|g|, |h|\big)$ (shrinking commutators). Having strong exit norm is rather robust, in contrast to the sensitivity of trapping properties; see Lemmas 6.6 and 6.7. Here is an immediate consequence of having strong exit norm:

LEMMA 6.3. — *Suppose X has strong exit norm. Then*

$$G_1 \ := \ \{g : \ |g| = 0\} \ = \ \{g : \ g^\nu \in X \text{ for all } \nu\}$$

is the largest definable subgroup of G contained in X. Moreover, $G_1^x = G_1$ for all $x \in X$, so G_1 is a normal subgroup of $\langle X\rangle^*$.

We rectify the exit norm $|\cdot|_X$ to a definable function $\|\cdot\|_X : G \to \mathbb{R}^*$ by

$$\|g\| \ = \ \|g\|_X := \ \inf\{\sum_{i=1}^{\nu} |g_i| : \ g = g_1 \cdots g_\nu\},$$

with $g_1, \ldots, g_\nu$ ranging over definable sequences in G. Thus

$$0 \ \le \ \|g\| \ \le \ |g|, \qquad \|g^{-1}\|_X \ = \ \|g\|, \qquad \|gh\| \ \le \ \|g\| + \|h\|.$$

Note that condition (s1) above translates into $|g| \le C\|g\|$. In the next subsection we indicate how strong exit norms arise from strong trapping conditions that originate in having good Lie models.

Strong trapping conditions

For $K \in \mathbb{Q}^{\ge 1}$ we say that X is *K-strong* if

(t1) X is pseudofinite and $|X^2| \le K|X|$,

(t2) $g^i \in X^4$ for $i = 1, \ldots, 8 \implies g \in X$ (first trapping condition),

and for some definable symmetric $S \subseteq G$:

(t3) $\Big(S^{X^3}\Big)^q \subseteq X$, $q := \lfloor 2^9 K \rfloor$ (S is small compared to X),

(t4) $g^i \in X$ for $i = 1, \ldots, 8q \implies g \in S$ (second trapping condition).

Note that (t1) holds if $X \subseteq G$ is a pseudofinite K-approximate group.

THEOREM 6.4. — *Suppose X is K-strong.*[12] *Then X has strong exit norm:*

$$|g| \ \le \ 2^{16}K^2 \cdot \|g\| \ \text{ for all } g, \quad |[x,y]| \ \le \ 2^{51}K^5 \cdot |x| \cdot |y| \ \text{ for all } x, y \in X.$$

The proof in [2] adapts and finitizes Gleason [6], where G is a locally compact group without small subgroups and X is a compact symmetric neighborhood of $1 \in G$ with $\mu(K^2) \le K\mu(X)$ for a left Haar measure μ on G. In the situation above we have instead of such μ the normalized counting measure $Y \mapsto |Y|/|X|$ (for pseudofinite $Y \subseteq G$), taking values in $\mathbb{R}^*$. We can use this measure to construct definable functions by convolution. Since a proof of Theorem 6.4 would take at least 5 pages, we just give some indications:

12. The hypothesis in [2] includes X being a K-approximate group.

Suppose $\psi : G \to \mathbb{R}^*$ is a definable function such that $\psi(1) \geq 1/2$ and $0 \leq \psi(g) \leq K$ for all g, and $\psi = 0$ outside X^4. Define the "derivative" $\partial_g\psi : G \to \mathbb{R}^*$ by $\partial_g\psi(x) = \psi(g^{-1}x) - \psi(x)$. Then by (t2),

$$|g| \ \leq \ 16\|\partial_g\psi\|_\infty, \ \text{ where } \|\partial_g\psi\|_\infty \ := \ \sup\{|\partial_g\psi(x)| : \ x \in G\},$$

as is easily checked. Thus the first bound in Theorem 6.4 would follow from having $\|\partial_g\psi\|_\infty \leq 2^{12}K^2\|g\|$ for all g. To get ψ with this last bound seems too ambitious, but for any $\varepsilon > 0$ in $\mathbb{R}^*$ we can construct a definable function $\psi_\varepsilon : G \to \mathbb{R}^*$ with $\psi_\varepsilon(1) \geq 1/2$, $0 \leq \psi_\varepsilon(g) \leq K$ for all g, and $\psi_\varepsilon = 0$ outside X^4 such that $\|\partial_g\psi_\varepsilon\|_\infty \leq 2^{12}K^2\|g\|_\varepsilon$ for all g, where $|g|_\varepsilon := |g| + \varepsilon$ and

$$\|g\|_\varepsilon \ := \ \inf\{\sum_{i=1}^{\nu} |g_i|_\varepsilon : \ \nu \geq 1, \ g = g_1 \cdots g_\nu\}.$$

By letting ε tend to 0 we get the first bound in Theorem 6.4 as before. For the second bound we construct again a suitable function ψ, and then use the (easy) upper bound $\|\partial_{[x,y]}\psi\|_\infty \leq \|\partial_x\partial_y\psi\|_\infty + \|\partial_y\partial_x\psi\|_\infty$.

In the rest of this section we assume that $\mathcal{M}$ is rich, that X is pseudofinite, and that we are given a good model

$$\pi : \langle X\rangle \to \mathcal{G}$$

of X onto a connected Lie group $\mathcal{G}$. A set $Y \subseteq \langle X\rangle$ is said to be π-*thick* if Y is definable, symmetric, and $Y \supseteq \ker(\pi)$. By Lemma 5.4, Corollary 4.4, and the connectedness of $\mathcal{G}$, if $Y \subseteq \langle X\rangle$ is π-thick, then $\langle X\rangle = \langle Y\rangle$ and $\pi : \langle Y\rangle \to \mathcal{G}$ is a good model of Y. Since X is an approximate group, we have $|X^2| \leq K|X|$ for some $K \in \mathbb{Q}^{\geq 1}$. Call X *strong* if it is K-strong for some $K \in \mathbb{Q}^{\geq 1}$. This notion applies also to any π-thick subset of $\langle X\rangle$. What we need for later is that X has a strong π-thick subset. This follows from the next more precise result, in which $\exp : \mathfrak{g} \to \mathcal{G}$ is the exponential map of the Lie algebra $\mathfrak{g}$ of $\mathcal{G}$:

Lemma 6.5. — *Let B be an open ball centered at the origin in $\mathfrak{g}$, with respect to some vector space norm on $\mathfrak{g}$. Then for all sufficiently small $r > 0$, any symmetric definable set $Y \subseteq \langle X\rangle$ with*

$$\pi^{-1}\Big(\exp(rB)\Big) \ \subseteq Y \ \subseteq \ \pi^{-1}\Big(\exp(\frac{3}{2}rB)\Big)$$

is a strong π-thick subset of X.

Proof (Sketch). — Let $r > 0$ and let the symmetric definable $Y \subseteq \langle X\rangle$ satisfy the inclusions of the lemma. Then Y is a π-thick subset of X. Basic properties of the exponential map imply that Y satisfies the first trapping condition for small enough $r > 0$. Take $K \in \mathbb{N}^{\geq 1}$ such that Y is a K-approximate group, set $q := 2^9K$, and take a symmetric definable $S \subseteq G$ such that

$$\pi^{-1}\Big(\exp(\frac{r}{4q}B)\Big) \ \subseteq \ S \ \subseteq \ \pi^{-1}\Big(\exp(\frac{r}{2q}B)\Big).$$

It is routine to check that if $r > 0$ is small enough, then $\left(S^{Y^3}\right)^q \subseteq Y$ and the second trapping condition holds. □

LEMMA 6.6. — *Suppose $\pi(X)$ contains no subgroup of $\mathcal{G}$ other than $\{1\}$. Let Y be a π-thick subset of X. Then there is $n \geq 1$ such that for all g,*

$$\frac{1}{n}|g|_Y \;\leq\; |g|_X \;\leq\; |g|_Y.$$

In particular, if X has strong exit norm, then so does Y.

Proof. — Take a neighborhood U of $1 \in \mathcal{G}$ with $\pi^{-1}(U) \subseteq Y$. By Lemma 6.1 we can take $n \geq 1$ such that if $a \in \mathcal{G}$ and $a^i \in \pi(X)$ for $i = 1, \ldots, n$, then $a \in U$. It follows that if g is such that $g^i \in X$ for $i = 1, \ldots, n$, then $g \in Y$. Therefore, $|g|_Y \;\leq\; n|g|_X$ for all g. □

LEMMA 6.7. — *Suppose $\pi(X^2)$ contains no subgroup of $\mathcal{G}$ other than $\{1\}$. Then there is $n \geq 1$ such that for all g,*

$$\frac{1}{n}|g|_X \;\leq\; |g|_{X^2} \;\leq\; |g|_X.$$

In particular, if X has strong exit norm, then so does X^2.

Proof. — Let U be a neighborhood of the identity in $\mathcal{G}$ such that $\pi^{-1}(U) \subseteq X$. By Lemma 6.1 we can take $n \geq 1$ such that if $a \in \mathcal{G}$ is such that $a^i \in \pi(X^2)$ for $i = 1, \ldots, n$, then $a \in U$. It follows that if g is such that $g^i \in X^2$ for $i = 1, \ldots, n$, then $g \in X$. Thus $|g|_X \;\leq\; n|g|_{X^2}$ for all g. □

From the nonstandard approach [12] to Hilbert's 5th problem we borrow the generation of one-parameter subgroups by infinitesimals:

LEMMA 6.8. — *Suppose $\pi(X)$ contains no subgroup of $\mathcal{G}$ other than $\{1\}$. Let $u \in X$ be such that $0 \neq |u| \in \mathcal{o}$, so $|u| = \frac{1}{N}$ with $N \in \mathbb{N}^*,\ N > \mathbb{N}$. Then:*

(i) *the map $t \mapsto \pi(u^{\lfloor tN \rfloor}) \;:\; \mathbb{R} \to \mathcal{G}$ is a continuous group morphism;*

(ii) *the map in* (i) *is injective on $[0, 1]$.*

Proof. — We have $\pi(u^{\mathbb{Z}}) \subseteq \pi(X)$, so $\pi(u) = 1$. Also for $s, t \in \mathbb{R}$ we have $\lfloor (s+t)N \rfloor \;=\; \lfloor sN \rfloor + \lfloor tN \rfloor + k$ with $k \in \{-1, 0, 1\}$, so the map in (i) is a group morphism. For it to be continuous, it is enough that it is continuous at 0. Let U be an open neighborhood of $1 \in \mathcal{G}$, and take $n \geq 1$ so large that for all $a \in \mathcal{G}$, if $a^i \in \pi(X)$ for $i = 1, \ldots, n$, then $a \in U$. It is then easy to verify that if $|t| < \frac{1}{2n}$, then $\pi(u^{\lfloor tN \rfloor}) \in U$. □

7. THE MAIN THEOREM

We prove here Theorem 1.1. To substantiate the *sketch* in Section 1, we need the ability to pass to certain definable quotient groups on the "rich" side. More generally, in order to study the situation arising from the Lie Model Theorem 5.5, we fix till further notice an ambient rich structure $\mathcal{M}$ and a definable group G in $\mathcal{M}$ with a definable symmetric set $Y \subseteq G$, and a good model $\rho : \langle Y\rangle \to \mathcal{H}$ of Y.

Passing to definable quotient groups

Let D be a definable normal subgroup of G, with quotient map $\eta : G \to G/D$. Then $\eta Y \subseteq G/D$ is symmetric and $\langle \eta Y\rangle = \eta\langle Y\rangle$. The closure $\mathcal{D}$ of $\rho(D \cap \langle Y\rangle)$ in $\mathcal{H}$ is a normal subgroup of $\mathcal{H}$. Thus ρ induces a surjective group morphism

$$\overline{\rho} \ : \ \langle \eta Y\rangle \to \overline{\mathcal{H}} \ := \ \mathcal{H}/\mathcal{D}, \qquad \overline{\rho}(\eta(x)) \ := \ \rho(x)\mathcal{D} \ \text{ for } x \in \langle Y\rangle.$$

Take a countable parameter set A of $\mathcal{M}$ such that G and D are A-definable. Then G and D are 0-definable in $\mathcal{M}_A$, so $\eta G = G/D$ is a definable group in the expansion $\mathcal{M}_{A,\eta}$ of $\mathcal{M}$, and $\eta Y \subseteq \eta G$ is a definable set of this expansion. With $\mathcal{M}_{A,\eta}$ now as the ambient rich structure, we have:

LEMMA 7.1. — *The map $\overline{\rho}$ is a good model of ηY with* $\ker(\overline{\rho}) = \eta(\ker(\rho))$.

The main steps in the proof are as follows. Take a descending sequence of definable sets $Z_n \subseteq \langle Y\rangle$ such that $\ker(\rho) = \bigcap_n Z_n$, and note that then $\eta(\ker(\rho)) = \bigcap_n \eta Z_n$. Thus $\eta(\ker(\rho)) \subseteq G/D$ is Π-definable and a normal subgroup of $\langle \eta Y\rangle$. The quotient $\langle \eta Y\rangle/\eta(\ker(\rho))$ is bounded since $\langle Y\rangle/\ker(\rho)$ is. Use $\eta(\ker(\rho)) = \bigcap_n \eta Z_n$ (and Corollary 4.4) to get $\ker(\overline{\rho}) = \eta(\ker(\rho))$. Thus $\overline{\rho}$ induces a group isomorphism

$$\langle \eta Y\rangle/\eta(\ker(\rho)) \ \to \ \overline{\mathcal{H}},$$

which is easily verified to be continuous, with the logic topology on the bounded quotient $\langle \eta Y\rangle/\eta(\ker(\rho))$. As this quotient is σ-compact, this group isomorphism is also a homeomorphism.

The main induction

We have our good model $\rho : \langle Y\rangle \to \mathcal{H}$. We now assume also that our ambient rich $\mathcal{M}$ has a distinguished ordered field $\mathbb{R}^*$, definable in $\mathcal{M}$ and definably complete, with a discrete definable subring $\mathbb{Z}^*$ of $\mathbb{R}^*$. This gives sense to "pseudofinite".

THEOREM 7.2. — *Assume Y is pseudofinite and $\mathcal{H}$ is a connected Lie group. Set $H := \langle Y\rangle^*$ and $d := \dim \mathcal{H}$. Then there is an increasing sequence*

$$\{1\} \ = \ H_0 \ \subseteq \ H_1 \ \subseteq \ \cdots \ \subseteq \ H_{2d+1} \ = \ H$$

of definable normal subgroups of H with the following properties:

(1) *if $i \leq 2d$ is even, then H_{i+1}/H_i is pseudofinite,*

(2) *if $i \leq 2d$ is odd, then $H_{i+1} = u^{\mathbb{Z}^*} H_i$ for some $u \in H_{i+1}$, and H_{i+1}/H_i is central in H/H_i.*

Moreover, the Lie group $\mathcal{H}$ is nilpotent.

Proof. — By induction on d. If $d = 0$, then $\mathcal{H} = \{1\}$, so $\ker(\rho) = Y = \langle Y \rangle = H$, and we are done. Assume $d > 0$ below. In view of Lemmas 6.5 and 6.6 we can shrink Y without changing $\langle Y \rangle$ and arrange:

(i) $\rho(Y^2)$ contains no subgroup of $\mathcal{H}$ other than $\{1\}$;

(ii) the exit norm of Y is strong.

By (ii) we have a largest definable subgroup $H_1 = \{h \in H : \ |h|_Y = 0\}$ of H contained in Y. Then $\rho(H_1)$ is a subgroup of $\mathcal{H}$ contained in $\rho(Y)$, so $H_1 \subseteq \ker \rho$. Since $H_1^y = H_1$ for all $y \in Y$, we have $H_1 \trianglelefteq H$. The image $\overline{Y}$ of Y in the definable quotient group H/H_1 is pseudofinite, the group $\langle \overline{Y} \rangle$ it generates equals the image of $\langle Y \rangle$ in H/H_1, and ρ induces a group morphism $\overline{\rho} : \langle \overline{Y} \rangle \to \mathcal{H}$, which is easily checked to be a good model of $\overline{Y}$. Using $YH_1 \subseteq Y^2$, we have for all $h \in H$,

$$|h|_{Y^2} \ \leq \ |hH_1|_{\overline{Y}} \ \leq \ |h|_Y.$$

Then Lemma 6.7 gives that $\overline{Y}$ has still strong exit norm and every non-identity element of H/H_1 has nonzero exit norm with respect to $\overline{Y}$. Thus, replacing H and Y by H/H_1 and $\overline{Y}$, respectively, and renaming, we maintain (i) and (ii) above and reduce to the case that in addition

(iii) $|h|_Y \neq 0$ for all $h \neq 1$ in H.

Thus Y contains no definable subgroup of H other than $\{1\}$. Take a neighborhood U of the identity in $\mathcal{H}$ such that $\rho^{-1}(U) \subseteq Y$. Then, given any $n \geq 1$ there are elements $a \neq 1$ in U such that $a^i \in U$ for $i = 1, \ldots, n$, and so there are elements $g \neq 1$ in $\rho^{-1}(U)$ with $|g|_Y < 1/n$. As Y is pseudofinite, we can take $u \neq 1$ in Y for which $|u|_Y$ is minimal. Thus $|u|_Y$ is infinitesimal. Take $C \geq 1$ in $\mathbb{N}$ such that

$$|[g, h]|_Y \ \leq \ C \cdot |g|_Y \cdot |h|_Y \ \text{ for all } g, h \in Y.$$

It follows that for all $g \in Y$ with $|g|_Y < 1/C$,

$$|[g, u]|_Y \ < |u|_Y$$

which by the minimality of $|u|_Y$ gives $[g, u] = 1$. Hence u commutes with all $g \in Y$ such that $|g|_Y < 1/C$. Consider the ρ-thick subset Z of Y,

$$Z \ := \ \{g \in Y : \ |g|_Y < 1/C\}.$$

Then $\langle Z \rangle = \langle Y \rangle$, so u lies in the center of H. Let $D := u^{\mathbb{Z}^*}$ be the smallest definable subgroup of H containing u. Then $D \trianglelefteq H$, and with the notations of Lemma 7.1 we have a pseudofinite approximate group ηY in the quotient $\eta H = H/D$. Let $\mathcal{D}$ be the closure of $\rho(D \cap \langle Y \rangle)$ in $\mathcal{H}$, so $\mathcal{D}$ is a closed central subgroup of $\mathcal{H}$ and ρ induces a good model

$$\overline{\rho} : \langle \eta Y \rangle \to \mathcal{H}/\mathcal{D}.$$

Note that $\mathcal{D}$ is infinite, by Lemma 6.8. Hence $\mathcal{H}/\mathcal{D}$ is a connected Lie group of lower dimension than $\mathcal{H}$, and we are done by induction. □

Back to reality

From the artificially rich environment of $\mathcal{M}$ we now return to the real world, and so G can be any group in this subsection. Theorem 5.5 and Theorem 7.2 with its proof, together with logical compactness, yield:

COROLLARY 7.3. — *Let $X \subseteq G$ be a finite K-approximate group. Then there is a K^6-approximate group $Y \subseteq X^4$ such that:*

(1) *X can be covered by L left cosets of Y, with L depending only on K;*
(2) *$\langle Y \rangle$ has normal subgroups $\{1\} = H_0 \subseteq H_1 \subseteq \cdots \subseteq H_{2d+1} = \langle Y \rangle$, with $d \in \mathbb{N}$ depending only on K, such that if $i \leq 2d$ is even, then $H_{i+1} \subseteq YH_i$, and so H_{i+1}/H_i is finite, and if $i \leq 2d$ is odd, then $H_{i+1} = u^{\mathbb{Z}}H_i$ for some $u \in Y$, and H_{i+1}/H_i is central in $\langle Y \rangle/H_i$.*

The finite quotients H_{i+1}/H_i for even i can be replaced by just one at the end. This follows from the next group theoretic lemma and its corollary.

LEMMA 7.4. — *Let $G_1 \subseteq G_2$ be normal subgroups of G such that G_1 is finite and G_2/G_1 is cyclic and central in G/G_1. Then G_2 has a cyclic subgroup C such that $C \subseteq \mathrm{Z}(G)$ and $[G_2 : C] < \infty$.*

Proof. — If G_2 is finite, then we can take $C = \{1\}$. Assume G_2 is infinite. Take $u \in G_2$ such that $G_2 = u^{\mathbb{Z}}G_1$. Then $u^{\mathbb{Z}}$ is an infinite cyclic subgroup of G_2 and $[G_2 : u^{\mathbb{Z}}] < \infty$. Each $g \in G$ gives a subgroup $gu^{\mathbb{Z}}g^{-1}$ of G_2 of the same finite index in G_2 as $u^{\mathbb{Z}}$. As G_2 is finitely generated, there are only finitely many subgroups of G_2 of that index, so we have $n \geq 1$ such that $\bigcap_{g \in G} gu^{\mathbb{Z}}g^{-1} = u^{n\mathbb{Z}}$ is an infinite cyclic subgroup of G_2, of finite index in G_2, and normal in G. Note that $u^{n\mathbb{Z}} \cap G_1 = \{1\}$. We claim that $u^{n\mathbb{Z}} \subseteq \mathrm{Z}(G)$. Otherwise we have $g \in G$ with $gu^n \neq u^n g$. Then $gu^n g^{-1} = u^{-n}$. As G_2/G_1 is central in G/G_1, we also have $gu^n = g_1 u^n g$ with $g_1 \in G_1$. Then $u^{-n} = g_1 u^n$, so $u^{2n} \in G_1$, a contradiction. □

By induction on d this lemma and its proof give the following:

COROLLARY 7.5. — *Let $\{1\} = G_0 \subseteq G_1 \subseteq \cdots \subseteq G_{2d+1} = G$ be normal subgroups of G such that for $i \leq 2d$, if i is even, then G_{i+1}/G_i is finite, and if i is odd, then G_{i+1}/G_i is cyclic and central in G/G_i. Let $u_i \in G_{2i}$ be such that $G_{2i} = G_{2i-1}u_i^{\mathbb{Z}}$, for $i = 1, \ldots, d$. Then G has a d-nilpotent normal subgroup of finite index, with nilpotent base $u_1^{n_1}, \ldots, u_d^{n_d}$ for some $n_1, \ldots, n_d \in \mathbb{N}$.*

From Corollaries 7.3 and 7.5 we obtain:

COROLLARY 7.6. — *Let $X \subseteq G$ be a finite K-approximate group. Then there is a K^6-approximate group $Y \subseteq X^4$, such that:*

(1) *X can be covered by L left cosets of Y, where L depends only on K;*

(2) *$\langle Y\rangle$ has a normal subgroup N of finite index such that N is d-nilpotent with $d \in \mathbb{N}$ depending only on K;*

(3) *N has a nilpotent base $y_1^{n_1}, \ldots, y_d^{n_d}$ with $y_1, \ldots, y_d \in Y$.*

If we drop (3) in Corollary 7.6 we can get $d \leq 3\log_2 K$ in (2), which is Theorem 1.1. Following [13] and [2] we prove this in the next subsection.

A logarithmic bound

We are going to use that a connected nilpotent Lie group $\mathcal{H}$ is unimodular and has a largest compact subgroup T. This T is central, and $\mathcal{H}/T$ is simply connected; see [25], p. 192. Moreover:

LEMMA 7.7. — *Let $\mathcal{H}$ be a simply connected nilpotent Lie group with Haar measure μ, and let $C \subseteq \mathcal{H}$ be compact. Then $\mu(C^2) \geq 2^{\dim \mathcal{H}}\mu(C)$.*

Proof. — Let $d := \dim \mathcal{H}$, and let $\mathfrak{h}$ be the Lie algebra of $\mathcal{H}$. The assumption on $\mathcal{H}$ guarantees that the exponential map $\exp : \mathfrak{h} \to \mathcal{H}$ is a homeomorphism, and that the Lebesgue measure on $\mathfrak{h}$ induced by a linear isomorphism $\mathbb{R}^d \cong \mathfrak{h}$ corresponds under exp to a Haar measure μ on $\mathcal{H}$. Now

$$\log C^2 \ \supseteq\ \{2\log x :\ x \in C\},$$

so $\mu(C^2) \geq 2^d\mu(C)$. □

Let $X \subseteq G$ be a pseudofinite K-approximate group and $\pi : \langle X\rangle \to \mathcal{G}$ be good model of X^4, as we know exists by Section 5. Take an open neigborhood U of the identity in $\mathcal{G}$ with $\pi^{-1}(U) \subseteq X^4$. Next, take an open subgroup $\mathcal{G}'$ of $\mathcal{G}$ and a compact normal subgroup $N \subseteq U$ of $\mathcal{G}'$ such that $\mathcal{H} := \mathcal{G}'/N$ is a connected Lie group, and set $Y := X^4 \cap \pi^{-1}(\mathcal{G}')$. The proof of the Lie Model Theorem 5.5 shows that $Y \subseteq G$ is a definable K^6-approximate group and yields a good model $\rho : \langle Y\rangle \to \mathcal{H}$ of Y. Then $\mathcal{H}$ is nilpotent by Theorem 7.2. Let T be the largest compact subgroup of $\mathcal{H}$. We now have:

COROLLARY 7.8. — $\dim \mathcal{H}/T \leq 3\log_2 K$. (*The bound in* [2] *is* $6\log_2 K$.)

Proof. — Since $\pi(X)$ is a K-approximate group in $\mathcal{G}$, Lemma 2.9 says that $\pi(X)^2 \cap \mathcal{G}'$ is a K^3-approximate group in $\mathcal{G}'$. Now $\pi(X)^2 \cap \mathcal{G}'$ is a compact neighborhood of the identity in $\mathcal{G}'$, and so is its image Z in $\mathcal{H} = \mathcal{G}'/N$. Hence the image C of Z in $\mathcal{H}/T$ is a compact neighborhood of the identity in $\mathcal{H}/T$ and also a K^3-approximate group. With μ a Haar measure on the simply connected nilpotent Lie group $\mathcal{H}/T$ and $d = \dim \mathcal{H}/T$ we have $\mu(C^2) \geq 2^d\mu(C)$ by Lemma 7.7. In view of $\mu(C^2) \leq K^3\mu(C)$ and $\mu(C) > 0$ we get $d \leq 3\log K$. □

Lemma 5.3 gives $m \geq 1$ such that $\rho^{-1}(T) \subseteq Y^m$, and so composing ρ with the canonical map $\mathcal{H} \to \mathcal{H}/T$ gives a good model $\langle Y \rangle \to \mathcal{H}/T$ of $Y^m \subseteq X^{4m}$. Now $\langle Y^m \rangle = \langle Y \rangle$, and so Theorem 7.2 applied to this good model of Y^m and logical compactness give the variant of Corollary 7.3 where in (2) we have $d \leq 3 \log_2 K$, and $H_{i+1} \subseteq Y^m H_i$ for even i (instead of $H_{i+1} \subseteq Y H_i$), and $u \in Y^m$ for odd i (instead of $u \in Y$), with $m \geq 1$ depending only on K. In combination with Corollary 7.5 we obtain Theorem 1.1, as promised.

REFERENCES

[1] L. Bieberbach, Über einen Satz des Hrn. C. Jordan in der Theorie der endlichen Gruppen linearer Substitutionen, Sitzungsber. Preuss. Akad. Wiss. (1911).

[2] E. Breuillard, B. Green and T. Tao, The structure of approximate groups, Publ. Math. i.h.é.s. **116** (2012), 115–221.

[3] G. Courtois, Lemme de Margulis à courbure de Ricci minorée, Séminaire Bourbaki (2013-2014), Exp. n° 1075, novembre 2013.

[4] E. Croot and O. Sisask, A probabilistic technique for finding almost-periods of convolutions, Geom. Funct. Anal. **20** (2010), 1367–1396.

[5] A. Fornasiero and P. Hieronymi, A fundamental dichotomy for definably complete expansions of ordered fields, available at `arXiv:1305.4767`, to appear in J. Symb. Logic.

[6] A. Gleason, Groups without small subgroups, Ann. of Math. **56** (1952), 193–212.

[7] I. Goldbring, Hilbert's fifth problem for local groups, Ann. of Math. **172** (2010), 1269–1314.

[8] M. Gromov, Groups of polynomial growth and expanding maps, Publ. Math. i.h.é.s. **53** (1981), 53–73.

[9] H. Helfgott, Growth and generation in $\mathrm{SL}_2(\mathbb{Z}/p\mathbb{Z})$, Ann. of Math. **167** (2008), 601–623.

[10] H. Helfgott, Growth in $\mathrm{SL}_3(\mathbb{Z}/p\mathbb{Z})$, J. Eur. Math. Soc. **13** (2011), 761–851.

[11] H. Helfgott, Growth in groups: ideas and perspectives, available at `arXiv:1303.023gv4`.

[12] J. Hirschfeld, The nonstandard treatment of Hilbert's fifth problem, Trans. AMS **321** (1990), 379–400.

[13] E. Hrushovski, Stable group theory and approximate subgroups, J. Amer. Math. Soc. **25** (2012), 189–243.

[14] E. Hrushovski, Y. Peterzil, and A. Pillay, Groups, measures, and the NIP, J. Amer. Math. Soc. **21** (2008), 563–596.

[15] B. Kleiner, A new proof of Gromov's theorem on groups of polynomial growth, J. Amer. Math. Soc. **23** (2010), 815–829.

[16] D. Montgomery and L. Zippin, *Topological Transformation Groups*, Interscience Publishers, New-York-London, 1955.

[17] G. Petridis, New proofs of Plünnecke-type estimates for product sets in groups, Combinatorica **32** (2012), 712–733.

[18] A. Pillay, Type-definability, compact Lie groups, and o-minimality, J. of Math. Logic **4** (2004), 147–162.

[19] I. Ruzsa, An analog of Freiman's theorem in groups, Astérisque **258** (1999), 323–326.

[20] I. Ruzsa, Towards a noncommutative Plünnecke-type inequality, In: Bolyai Society Mathematical Studies, vol. 21, Springer, 2010.

[21] T. Sanders, On a non-abelian Balog-Szemeredi-type lemma, J. Aust. Math. Soc. **89** (2010), 127–132.

[22] Y. Shalom and T. Tao, A finitary version of Gromov's polynomial growth theorem, Geom. Funct. Anal. **20** (2010), 1502–1547.

[23] T. Tao, Product set estimates for non-commutative groups, Combinatorica **28** (2008), 547–594.

[24] T. Tao and Van H. Vu, *Additive combinatorics*, Cambridge U. Press, 2006.

[25] J. Tits, *Liesche Gruppen und Algebren*, Springer, 1983.

[26] H. Weyl, Über die Definitionen der mathematischen Grundbegriffe, in *Gesammelte Abhandlungen, Band I*, pp. 298–304.

[27] H. Yamabe, A generalization of a theorem of Gleason, Ann. of Math. **58** (1953), 351–365.

Lou van den DRIES

University of Illinois at Urbana-Champaign
Department of Mathematics
1409 W. Green Street
Urbana, IL 61801-2975, U.S.A.
E-mail : vddries@math.uiuc.edu