CAMBRIDGE
UNIVERSITY PRESS

ARTICLE

On finite sets of small tripling or small alternation in arbitrary groups

Gabriel Conant

Department of Pure Mathematics & Mathematical Statistics, University of Cambridge, Cambridge CB3 0WB, UK
Email: gconant@maths.cam.ac.uk

Abstract

We prove Bogolyubov–Ruzsa-type results for finite subsets of groups with small tripling, $|A^3| \leqslant O(|A|)$, or small alternation, $|AA^{-1}A| \leqslant O(|A|)$. As applications, we obtain a qualitative analogue of Bogolyubov's lemma for dense sets in arbitrary finite groups, as well as a quantitative arithmetic regularity lemma for sets of bounded VC-dimension in finite groups of bounded exponent. The latter result generalizes the abelian case, due to Alon, Fox and Zhao, and gives a quantitative version of previous work of the author, Pillay and Terry.

1. Introduction

Freiman's theorem (see [13], [12]) is a combinatorial result in additive number theory which states that if A is a finite subset of a torsion-free abelian group G, and $|A + A| \leqslant k|A|$ (*i.e. A has small doubling*), then A is contained in an n-dimensional arithmetic progression of length $c|A|$, where c and n depend only on k. Ruzsa [30] gave a new proof of this result, and a similar strategy was later used by Green and Ruzsa [16] to prove a generalization of Freiman's theorem, involving coset progressions, for arbitrary abelian groups. A key part of this work is that a set of small doubling in an abelian group can be 'Freiman-isomorphically' mapped to a dense set in a finite abelian group (see [16, Proposition 1.2]). This allows one to apply the following result, which Ruzsa [30] adapted from Bogolyubov [5]. For comparison to our work, we state this result in two cases.

Theorem 1.1 (Bogolyubov's lemma). *Fix $r \in \mathbb{Z}^+$ and $\alpha \in \mathbb{R}^+$.*

(a) (Bounded exponent case) *Suppose G is a finite abelian group of exponent at most r, and $A \subseteq G$ is such that $|A| \geqslant \alpha|G|$. Then there is a subgroup $H \leqslant G$ such that $[G:H] \leqslant r^{\alpha^{-2}}$ and $H \subseteq 2A - 2A$.*

(b) (General case) *Suppose G is a finite abelian group, and $A \subseteq G$ is such that $|A| \geqslant \alpha|G|$. Then there is a $(1/4, n)$-Bohr neighbourhood B in G such that $n < \alpha^{-2}$ and $B \subseteq 2A - 2A$.*

Bohr neighbourhoods (see Definition 2.1) are certain kinds of well-structured sets which, in the abelian case, contain large coset progressions (preserved by Freiman isomorphism). This yields

CrossMark

the 'Bogolyubov–Ruzsa lemma[1] for finite abelian groups': if G is abelian and $A \subseteq G$ is finite, with $|A + A| \leqslant k|A|$, then $2A - 2A$ contains an n-dimensional coset progression of size $c|A|$, where c and n depend only on k. The conclusion of Freiman's theorem (exchanging the arithmetic progression for one containing A) then follows after a little more work (see [16, Proposition 5.1]).

For an abelian group G, Freiman's theorem also yields a classification of k-*approximate subgroups* of G, *i.e.* finite symmetric subsets $A \subseteq G$ such that $A + A$ can be covered by k translates of A. Approximate subgroups of arbitrary groups have been studied by many authors, culminating in the work of Breuillard, Green and Tao [7].

The goal of the present paper is to give generalizations of Bogolyubov's lemma to arbitrary finite groups, as well as similar statements about finite subsets of arbitrary groups whose product set growth can be controlled. For this, we focus on sets of small *tripling*, which satisfy Plünnecke–Ruzsa inequalities for product sets (as observed by Helfgott [18]: see Proposition 3.2(a)). Motivated by the work of Hrushovski [19] on approximate groups, we also consider finite sets A of *small alternation*, *i.e.* $|AA^{-1}A| \leqslant k|A|$ for some fixed k (see Remark 2.2).

Our main results, Theorems 2.3 and 2.4, are versions of the Bogolyubov–Ruzsa lemma for finite sets of small tripling or alternation in arbitrary groups (with some further constraints). Our first application of these results is the following qualitative analogue of Bogolyubov's lemma (Theorem 1.1) for arbitrary groups.

Theorem 1.2. *Fix a positive integer r and a positive real number α.*

(a) *(Bounded exponent case) Suppose G is a finite group of exponent at most r, and $A \subseteq G$ is such that $|A| \geqslant \alpha|G|$. Then there is a normal subgroup $H \leqslant G$ such that $[G : H] \leqslant O_{r,\alpha}(1)$ and $H \subseteq (AA^{-1})^2 \cap A^2A^{-2} \cap (A^{-1}A)^2 \cap A^{-2}A^2$.*

(b) *(General case) Suppose G is a finite group, and $A \subseteq G$ is such that $|A| \geqslant \alpha|G|$. Then there is a normal subgroup $H \leqslant G$ and a (δ, n)-Bohr neighbourhood B in H, such that $[G : H], \delta^{-1}, n \leqslant O_{\alpha}(1)$ and $B \subseteq (AA^{-1})^2 \cap A^2A^{-2} \cap (A^{-1}A)^2 \cap A^{-2}A^2$.*

This is proved in Section 2. For later applications, and also to illustrate the use of Bohr neighbourhoods in the non-abelian setting, we prove (in Section 2) the following easy corollary of Theorem 1.2(b). Call a (non-trivial) group G *purely non-abelian* if no normal subgroup $H \leqslant G$ has a non-trivial abelian quotient (*i.e.* $[H, H] = H$ for all normal $H \leqslant G$). The class of purely non-abelian groups contains all non-abelian simple groups, and is closed under direct product by Goursat's lemma.

Corollary 1.3. *Fix a positive real number α. Suppose G is a purely non-abelian finite group and $A \subseteq G$ is such that $|A| \geqslant \alpha|G|$. Then there is a normal subgroup $H \leqslant G$ such that $[G : H] \leqslant O_{\alpha}(1)$ and $H \subseteq (AA^{-1})^2 \cap A^2A^{-2} \cap (A^{-1}A)^2 \cap A^{-2}A^2$.*

Before continuing to the next application, we state the following consequences for symmetric subsets of groups of bounded exponent (part (a) follows immediately from Theorem 2.3 and part (b) is a special case of Theorem 1.2(a); see Section 2).

Corollary 1.4.

(a) *Fix positive integers k and r. Suppose G is a group of exponent r and $A \subseteq G$ is finite and symmetric, with $|A^3| \leqslant k|A|$. Then there is a subgroup $H \leqslant G$ such that A is covered by $O_{k,r}(1)$ translates of H and $H \subseteq A^4$.*

(b) *Fix a positive integer r and a positive real number α. Suppose G is a finite group of exponent r and A ⊆ G is symmetric, with $|A| \geqslant \alpha|G|$. Then there is a normal subgroup $H \leqslant G$, of index $O_{\alpha,r}(1)$, such that $H \subseteq A^4$.*

Remark 1.5. We do not know if H can also be made normal in part (a) of the previous corollary. In addition to improving Theorem 1.6 below (see Remark 1.7), such a result could be quite interesting, depending on the methods used. For example, together with the Feit–Thompson theorem and the Brauer–Fowler theorem, this strengthening of Corollary 1.4(a) would imply that for any positive integer r, there are only finitely many finite simple groups of exponent r. This is a known fact, but its proof requires the classification of finite simple groups (*e.g.* [4, Theorem 5.4]).

Our final applications are in the subject of arithmetic regularity (developed by Green [15] for abelian groups). There has been recent interest in strengthened arithmetic regularity lemmas for subsets of groups satisfying special tameness assumptions. This was initiated by the work of Terry and Wolf [38] on 'stable arithmetic regularity' in $\mathbb{F}_p^n$, which is continued in [9] and [39]. Arithmetic regularity in the setting of bounded VC-dimension is considered in [3], [10] and [34]. Given a group G and $A \subseteq G$, define the *VC-dimension of A* to be the VC-dimension of the collection of left translates of A, *i.e.* the supremum of all integers d such that, for some d-element set $X \subseteq G$, we have $\mathcal{P}(X) = \{X \cap gA : g \in G\}$. Alon, Fox and Zhao [3] have shown that if G is a finite *abelian* group of exponent at most r, and $A \subseteq G$ has VC-dimension at most d, then there is a subgroup $H \leqslant G$ of index $(1/\varepsilon)^{d+o_{r,d}(1)}$, and a subset $D \subseteq G$ which is a (possibly empty) union of cosets of H, such that $|A \triangle D| \leqslant \varepsilon|G|$. A main tool in their proof is Theorem 1.1(a), and we will use Corollary 1.4(a) to give the following generalization to arbitrary groups.

Theorem 1.6. *Fix positive integers r and d. Suppose G is a finite group of exponent at most r, and $A \subseteq G$ has VC-dimension at most d. Then, for any $\varepsilon, v > 0$, there is a subgroup H of G, of index $O_{r,d,v}((1/\varepsilon)^{d+v})$, which satisfies the following properties.*

   (i) *(Structure) There is a set $D \subseteq G$, which is a union of right cosets of H, such that $|A \triangle D| \leqslant \varepsilon|G|$.*
   (ii) *(Regularity) There is a set $Z \subseteq G$, with $|Z| < \frac{1}{2}\varepsilon^{1/2}|G|$, such that for any $x \in G \backslash Z$, either $|Hx \cap A| \leqslant \varepsilon^{1/4}|H|$ or $|Hx \backslash A| \leqslant \varepsilon^{1/4}|H|$.*

Remark 1.7. There are several comments to be made about Theorem 1.6.

   (1) In [3], Alon, Fox and Zhao conjecture that condition (i) of Theorem 1.6 holds for a *normal* subgroup H of index $\varepsilon^{-O_{r,d}(1)}$. This would follow from the proof if one could show that Corollary 1.4(a) holds with H being normal (see Remark 1.5). However, it does follow from the proof that one can replace the subgroup H with the intersection of its conjugates to obtain a normal subgroup of index $2^{\varepsilon^{-O_{r,d}(1)}}$ satisfying conditions (i) and (ii) (see Remark 8.3).
   (2) In [10] (joint with Pillay and Terry), we gave a version of Theorem 1.6 in which H is also normal, but without effective bounds on its index. One could instead use Corollary 1.4(b) to deduce this, yielding a very different proof compared to what is done in [10] (see Remark 8.3).
   (3) The 'regularity' statement in condition (ii) is not made explicit in [3], but follows implicitly from their methods (see Lemma 8.2).[2]
   (4) The $O_{r,d,v}$ constant in the statement of the theorem comes from Corollary 1.4(a) and so, unlike the abelian case, is not explicit (see Section 9.1).

---

[2]This was first observed by C. Terry.

In Section 8 we also show that a qualitative version of Theorem 1.6 holds for the class of purely non-abelian finite groups (see Theorem 8.4), which yields an interesting divergence between sets of bounded VC-dimension in non-abelian finite simple groups, compared to the abelian setting (see Corollary 8.6).

We end this Introduction with some discussion of our proofs. The results above involving groups of bounded exponent will be derived from Theorem 2.3. By the work of Hrushovski [19] and Breuillard, Green and Tao [7], approximate subgroups of groups with bounded exponent are close to genuine subgroups (see Theorem 5.1). Moreover, in any group, finite sets of small tripling are close to approximate subgroups by a result of Tao [35]. Together, these two facts imply a weaker version of Theorem 2.3 (see Section 9). To prove our result, we will sharpen what is essentially the first step of the work in [7], which is a theorem about sets of small doubling (proved independently by Croot and Sisask [11] and Sanders [32]). Namely, if G is a group, $A \subseteq G$ is finite, and $|A^2| \leqslant k|A|$, then A^2A^{-2} contains S^n, for some symmetric $S \subseteq G$ of size $\Omega_{k,n}(|A|)$. In Section 4 we re-prove this result using the same techniques, but for sets of small tripling or small alternation, which leads to stronger conclusions. We also work in the setting of measures (similar to Massicot and Wagner [24]), so that this analysis can be applied later to pseudofinite subsets of ultraproducts of groups. We then prove Theorem 2.3 in Section 5.

For Theorem 2.4, we will need to delve a bit deeper into the underlying methods of [7] and [19], in particular, the ultraproduct construction. To prove the theorem, we will first prove a pseudo-finite analogue, and then deduce the finitary version using an 'ultraproduct of counterexamples'. To simplify this discussion, and illustrate the leverage obtained by working with pseudofinite sets, we focus on the case of symmetric sets of small tripling. In this case, the pseudofinite analogue of Theorem 2.4 deals with a group G and a *pseudofinite* (symmetric) subset $A \subseteq G$. In other words, G is an ultraproduct of groups, and A is an ultraproduct of finite (symmetric) subsets of those groups. We also assume $\langle A \rangle = A^m$ for some fixed m (which holds, for example, if A is an ultra-product of uniformly dense subsets of finite groups). If A has small tripling (formulated using a pseudofinite counting measure), then the Sanders–Croot–Sisask analysis from Section 4 yields a symmetric set S such that $S^8 \subseteq A^4$. Moreover, S itself has small tripling (in fact it is an approximate subgroup), allowing us to iterate the process. After infinitely many iterations, we obtain a decreasing sequence of symmetric subsets of A^4, whose intersection is a demonstrably 'large' subgroup of $\langle A \rangle$ contained in A^4.

We now reach an obstacle, in that although G is an ultraproduct of groups, the subgroup constructed above need not be an ultraproduct of subgroups of those groups. In order to salvage this, we move to a saturated elementary extension G_* of G (in a suitable first-order language). We then find a normal subgroup Γ of $\langle A_* \rangle$ of small index (where A_* is the interpretation of A in G_*), which is contained in A_*^4 and is an intersection of countably many definable sets. By standard facts, $\langle A_* \rangle / \Gamma$ is a compact Hausdorff group under a certain topology controlled by definable sets in G_*. By a result of Pillay [28], the connected component of $\langle A_* \rangle / \Gamma$ is a compact connected *abelian* group, and thus an inverse limit of tori, supplying us with Bohr neighbourhoods in $\langle A_* \rangle$. Finally, in order to transfer these Bohr neighbourhoods through the ultraproduct, we will use an approximation method from [10], and a result about approximate homomorphisms from [1]. This will yield Bohr neighbourhoods in the original groups, and allow us to prove Theorem 2.4.

2. Definitions, main theorems and corollaries

Before stating the main theorems, we set some notation and definitions (used throughout the paper). Let G be a group. Given $n \geqslant 1$, let $G^{\times n} = G \times .^n. \times G$. Given $X, Y \subseteq G$, let $XY = \{xy : x \in X, \ y \in Y\}$. Set $X^0 = \{1\}$ and inductively define $X^{n+1} = X^nX$. Let $X^{-1} = \{x^{-1} : x \in X\}$. We say that $X \subseteq G$ is *symmetric* if $1 \in X$ and $X = X^{-1}$. A *Y-translate of X* is a set of the form aX where $a \in Y$. Given a set $X \subseteq G$, we let $\langle X \rangle$ denote the subgroup of G generated by X, and we use the notation $\bar{X}$ for the set $X \cup X^{-1} \cup \{1\}$.

Let $\mathbb{T}^n$ denote the n-dimensional torus $\mathbb{R}/\mathbb{Z} \times .^n. \times \mathbb{R}/\mathbb{Z}$, considered as an additive group with identity 0. Let d_n denote the invariant metric on $\mathbb{T}^n$ induced by the product of the arclength metric on $\mathbb{R}/\mathbb{Z}$ (identified with S^1).

Definition 2.1. Given a group G, a positive integer n and a positive real number δ, we say that $B \subseteq G$ is a (δ, n)-*Bohr neighbourhood in* G if there is a homomorphism $\tau \colon G \to \mathbb{T}^n$ such that

$$B = B_{\tau,\delta}^n := \{x \in G \colon d_n(0, \tau(x)) < \delta\}.$$

In the setting of abelian groups, Bohr neighbourhoods are often used as replacements for subgroups in cases where few subgroups are available (*e.g.* in $\mathbb{Z}/p\mathbb{Z}$). In general, if G is a group and $B = B_{\tau,\delta}^n$ is a (δ, n)-Bohr neighbourhood in G, then B is symmetric, closed under conjugation, and contains the kernel N of a homomorphism from G to some $\mathbb{T}^n$ (so G/N is abelian). While B may not be closed under the group operation, one can obtain control in pairs by allowing the radius δ to vary. For instance, $B^2 \subseteq B_{\tau,2\delta}^n$ by the triangle inequality. A more sophisticated manifestation of this idea can be found in the work of Bourgain [6]. Finally, if G is finite then Bohr neighbourhoods are 'large', for instance $|B_{\tau,\delta}^n| \geqslant \delta^n |G|$ (see [37, Lemma 4.20] or [10, Proposition 4.5]).

Recall from the Introduction that we are interested in finite subsets A of some group G, which either have *small tripling*, *i.e.* $|A^3| \leqslant k|A|$ for some fixed constant k, or have *small alternation*, *i.e.* $|AA^{-1}A| \leqslant k|A|$ for some fixed k.

Remark 2.2. The notion of small alternation is motivated by Hrushovski's definition of a *near-subgroup* from [19]. Our terminology is explained by Proposition 3.2(b), which shows that small alternation for a finite set A in a group G implies 'very small' tripling for AA^{-1} (see Section 9.2, and especially Remark 9.4, for discussion on the relationship to approximate subgroups). Note that small tripling implies small doubling, and also small alternation due to the general Plünnecke–Ruzsa inequalities observed by Helfgott (see Proposition 3.2(a)). For abelian groups, small alternation clearly implies small doubling, and it is well known that small doubling implies small tripling (see [29]), making the three notions equivalent. However, in non-abelian groups, there are no general implications between small doubling and small alternation. For example, let G be the free product $H * F_2$, where H is some finite group and F_2 is the free group on two generators, say a and b. Set $A = H \cup \{a\}$ and $B = aHb$. Then A satisfies small doubling but not small alternation, and B satisfies small alternation but not small doubling.

We now state our two main theorems, which are Bogolyubov–Ruzsa-type statements for finite sets of small alternation or small tripling. Each statement involves two crucial assumptions, the first being either small tripling or small alternation for some finite set, and the second being one of the following options: (1) bounded exponent of a certain subgroup, (2) bounded generation of a certain subgroup, or (3) both. Altogether, this yields six statements, which we have divided into two theorems, one for the bounded exponent case and the other for the general case. The two results are proved in Sections 5 and 7, respectively.

Theorem 2.3 (bounded exponent case)**.** *Fix positive integers k, m and r. Let G be a group, and fix a non-empty finite subset $A \subseteq G$.*

(1) (Small alternation) *Suppose $|AA^{-1}A| \leqslant k|A|$ and $\langle AA^{-1} \rangle$ has exponent r.*
   (a) *There is a subgroup $H \leqslant \langle AA^{-1} \rangle$ such that:*
      (i) $(AA^{-1})^m$ *is covered by $O_{k,m,r}(1)$ left cosets of H, and*
      (ii) $H \subseteq (AA^{-1})^2$.
   (b) *Assume $\langle AA^{-1} \rangle = (AA^{-1})^m$. Then there is a normal subgroup $H \leqslant \langle AA^{-1} \rangle$, of index $O_{k,m,r}(1)$, such that $H \subseteq (AA^{-1})^2$.*

similar to Freiman's theorem, this result for abelian groups reduces to Theorem 1.1(b), using the fact that Bohr neighbourhoods in abelian groups can be approximated by coset progressions (see [37, Lemma 4.22]), and that finite sets of small doubling in abelian groups have 'good models' as dense sets in finite abelian groups (see [16, Proposition 2.1]). Thus, for the sake of completeness, we will explain in Remark 7.4 how to obtain $G = H$ in Theorem 1.2(b) when G is abelian.

The rest of this section is devoted to proving the theorems and corollaries in Section 1 (except for Theorem 1.6, which is proved in Section 8). We first consider Corollary 1.4 since it is immediate from the theorems above.

Proof of Corollary 1.4. Part (a) is immediate from Theorem 2.3. Part (b) is immediate from Theorem 1.2(a). □

Proof of Theorem 1.2. Part (a). Fix a positive integer r and a positive real number α. Suppose G is a finite group of exponent r and $A \subseteq G$ is such that $|A| \geqslant \alpha|G|$. It is straightforward to show that $\langle A \rangle = \bar{A}^m$ for some $m \leqslant \lceil 3\alpha + 1 \rceil$ (see *e.g.* [24, Remark 4]). So we can apply Theorem 2.3(2b), with $k = \lceil \alpha^{-1} \rceil$ and $m = \lceil 3\alpha + 1 \rceil$, to obtain a subgroup $K \leqslant \langle A \rangle$, of index $n = n(\alpha, r)$, such that $K \subseteq (AA^{-1})^2 \cap A^2A^{-2} \cap (A^{-1}A)^2 \cap A^{-2}A^2$. Note also that $[G \colon \langle A \rangle] \leqslant \lceil \alpha^{-1} \rceil$, and so $[G \colon K] \leqslant n\lceil \alpha^{-1} \rceil$. Now, if $H = \bigcap_{g \in G} gKg^{-1}$, then $[G \colon H] \leqslant [G \colon K]! \leqslant O_{\alpha,r}(1)$, H is normal in G, and $H \subseteq (AA^{-1})^2 \cap A^2A^{-2} \cap (A^{-1}A)^2 \cap A^{-2}A^2$.

Part (b). Fix a positive real number α. Suppose G is a finite group and $A \subseteq G$ is such that $|A| \geqslant \alpha|G|$. In analogy to part (a), Theorem 2.4(2) provides a subgroup $K \leqslant \langle A \rangle$ and a Bohr neighbourhood $B_{\tau,\delta}^n \subseteq K$ such that $[G \colon K]$, δ^{-1} and n are bounded above in terms of α, and $B_{\tau,\delta}^n \subseteq (AA^{-1})^2 \cap A^2A^{-2} \cap (A^{-1}A)^2 \cap A^{-2}A^2$. If $H = \bigcap_{g \in G} gKg^{-1}$ and $B = B_{\tau,\delta}^n \cap H$, then H is normal in G, $[G \colon H] \leqslant O_\alpha(1)$ and $B = B_{\tau \upharpoonright H,\delta}^n$. So B and H are as desired. □

Remark 2.7. It is worth pointing out that part (a) of Theorem 1.2 also follows easily from part (b), due to the fact that a (δ, n)-Bohr neighbourhood in a group of exponent $r > \delta^{-1}$ is a subgroup. We leave details to the reader. In a similar way, parts (1b) and (2b) of Theorem 2.3 follow from parts (1) and (2) of Theorem 2.4, respectively. On the other hand, the proof given below of Theorem 2.3 is more direct, and does not require the model-theoretic methods employed here, nor the work from [10] on approximate Bohr neighbourhoods.

Proof of Corollary 1.3. Fix $\alpha > 0$. Suppose G is a purely non-abelian finite group and $A \subseteq G$ is such that $|A| \geqslant \alpha|G|$. By Theorem 1.2(b), there is a normal subgroup $H \leqslant G$ and a (δ, n)-Bohr neighbourhood $B \subseteq H$, such that $[G \colon H] \leqslant O_\alpha(1)$ and $B \subseteq (AA^{-1})^2 \cap (A^{-1}A)^2 \cap A^2A^{-2} \cap A^{-2}A^2$. Since B contains $\ker(\tau)$ for some homomorphism $\tau \colon H \to \mathbb{T}^n$, and G is purely non-abelian, we must have $B = H$. □

Remark 2.8. Corollary 1.3 implies that for any $\alpha > 0$, if G is a non-abelian finite simple group of size at least $\Omega_\alpha(1)$, and $A \subseteq G$ is such that $|A| \geqslant \alpha|G|$, then $G = (AA^{-1})^2 = (A^{-1}A)^2 = A^2A^{-2} = A^{-2}A^2$. Applied to the case of alternating groups A_n, this implies that the least upper bound on the index of H in Theorem 1.2(b) must be greater than $\frac{1}{2}\lfloor \alpha^{-1} \rfloor!$ (at least for $\alpha \leqslant \frac{1}{5}$). However, it should be noted that stronger results about dense sets in non-abelian finite simple groups are already known. In particular, if G is a non-abelian finite simple group with $\log|G| \geqslant \Omega(\alpha^{-6})$, and $A, B, C \subseteq G$ are such that $|A|, |B|, |C| \geqslant \alpha|G|$, then $G = ABC$.[3] This follows from work of Gowers [14] on quasirandom groups (as observed by Nikolov and Pyber [25]; see also [25, Corollary 1],

---

[3]By [14], the implied constant in $\Omega(\alpha^{-6})$ is no more than $25^{\log(25)}$. Using the classification of finite simple groups, the overall bound can be improved to $|G| > (\lceil \alpha^{-3} \rceil + 1)!$ (see [8]).

[14, Theorem 3.3] and [14, Theorem 4.7]). Similar results are shown by Hrushovski in [19] (*e.g.* [19, Corollary 1.4]).

3. Ultraproducts of groups

In this section we review the ultraproduct construction in the case of groups. The reader only interested in Theorem 2.4 (the bounded exponent case) can skip this section. Throughout this section, let $(G_s)_{s \in \mathbb{N}}$ be a fixed sequence of groups, and fix a non-principal ultrafilter $\mathcal{U}$ on $\mathbb{N}$. Let $G = \prod_{\mathcal{U}} G_s$ be the ultraproduct of the sequence $(G_s)_{s \in \mathbb{N}}$ with respect to $\mathcal{U}$. Explicitly, $G = (\prod_s G_{s \in \mathbb{N}})/\sim$, where $(a_s) \sim (b_s)$ if and only if $\{s : a_s = b_s\} \in \mathcal{U}$. Recall that G is a group under the (well-defined) operation $[(a_s)] \cdot [(b_s)] = [(a_s \cdot b_s)]$. A subset $X \subseteq G$ is *internal* if there is a sequence $(X_s)_{s \in \mathbb{N}}$, with $X_s \subseteq G_s$, such that

$$X = \prod_{\mathcal{U}} X_s := \left(\prod_{s \in \mathbb{N}} X_s \right) \Big/ \sim .$$

The collection of internal subsets of G forms a Boolean algebra.

We also assume that G is infinite, *i.e.* $\{s \in \mathbb{N} : |G_s| > n\} \in \mathcal{U}$ for all $n \in \mathbb{N}$. As a result, we obtain the following saturation property of G.

Fact 3.1 (Keisler [21]). *Suppose $(X_i)_{i=0}^{\infty}$ is a sequence of internal subsets of $G^{\times n}$ such that $\bigcap_{i=0}^{k} X_i \neq \emptyset$ for all $k \in \mathbb{N}$. Then $\bigcap_{i=0}^{\infty} X_i \neq \emptyset$.*

Finally, we fix a distinguished internal set $A \subseteq G$ (so $A = \prod_{\mathcal{U}} A_s$ for some $A_s \subseteq G_s$), and we assume that A is non-empty and *pseudofinite* (*i.e.* A_s is non-empty and finite for all $s \in \mathbb{N}$). With A fixed, we define the $|A|$-*normalized pseudofinite counting measure* μ on the Boolean algebra of internal subsets of G. Specifically, given an internal set $X = \prod_{\mathcal{U}} X_s$, define

$$\mu(X) = \lim_{\mathcal{U}} \frac{|X_s|}{|A_s|} \in \mathbb{R}_{\geqslant 0} \cup \{\infty\},$$

where $\lim_{\mathcal{U}} x_s = y$ if and only if, for all $\varepsilon > 0$, $\{s : |x_s - y| < \varepsilon\} \in \mathcal{U}$. Note that μ is a left and right invariant finitely additive measure on the internal subsets of G.

Properties of finite subsets of groups such as small alternation or small tripling can be formulated using μ. For example, $\mu(A^3) < \infty$ if and only if, for some fixed $k > 0$, $\{s : |A_s^3| \leqslant k|A_s|\} \in \mathcal{U}$. The fact that μ is controlled by discrete counting measures allows us to transfer the following Plünnecke–Ruzsa inequalities to G.

Proposition 3.2. *Fix an internal set $X \subseteq G$, with $0 < \mu(X) < \infty$.*

(a) *Suppose $\mu(X^3) \leqslant k\mu(X)$ for some $k > 0$. Then, for any $n \geqslant 1$ and $\varepsilon_1, \ldots, \varepsilon_n \in \{-1, 1\}$,*

$$\mu(X^{\varepsilon_1} \cdot \ldots \cdot X^{\varepsilon_n}) \leqslant k^{O_n(1)}\mu(X).$$

(b) *Suppose $\mu(XX^{-1}X) \leqslant k\mu(X)$ for some $k > 0$. Then $\mu((XX^{-1})^n) \leqslant k^{O_n(1)}\mu(X)$ for any $n \geqslant 1$.*

Proof. It suffices to fix $s \in \mathbb{N}$ and prove the claims for G_s with μ replaced by the usual counting measure. In this setting, part (a) is precisely the 'discrete case' of [35, Lemma 3.4] (first observed by Helfgott [18, Lemma 2.2]). So we only need to show part (b). The proof is similar to that of [35, Lemma 3.4] and relies on the triangle inequality for the Ruzsa distance. In particular, given non-empty finite $X, Y \subseteq G_s$, the Ruzsa distance between X and Y is defined as

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

Then, for non-empty finite $X, Y, Z \subseteq G_s$, we have $d(X, Z) \leqslant d(X, Y) + d(Y, Z)$ (this is due to Ruzsa [31] in the commutative setting; see also [35, Lemma 3.2]).

Now fix a non-empty finite set $X \subseteq G_s$, and assume $|XX^{-1}X| \leqslant k|X|$. By part (a) (in G_s) it is enough to show $|(XX^{-1})^3| \leqslant k^{O(1)}|X|$. For this, first note that $d(XX^{-1}, X^{-1}) \leqslant \log k$. So $d(XX^{-1}, XX^{-1}) \leqslant \log k^2$ by the triangle inequality, and thus $|(XX^{-1})^2| \leqslant k^2|XX^{-1}|$. Then $d(XX^{-1}X, X) \leqslant \log k^2$. By the triangle inequality, $d(XX^{-1}X, XX^{-1}X) \leqslant \log k^4$, and thus $|(XX^{-1})^3| \leqslant k^4|XX^{-1}X| \leqslant k^5|X|$. □

4. Sanders–Croot–Sisask analysis

In this section we prove Lemma 4.1, which is the main technical lemma of the paper. It is essentially a modification of a result of Croot and Sisask [11] and Sanders [32], which was later adapted by Breuillard, Green and Tao [7, Section 5] for their results on the structure of approximate groups. In the model-theoretic setting, these same techniques were used by Massicot and Wagner [24] in their work on 'definably amenable' approximate groups, and also by Krupiński and Pillay [22]. Part (a) of Lemma 4.1, which deals with sets of small alternation, is similar to some of Hrushovski's work with near-subgroups, especially [19, Corollary 3.11]. Our proof follows Sanders [32] (as do [7], [22] and [24]), and makes the modifications necessary to work with sets of small tripling or small alternation, and also to account for working with internal sets in the case of ultraproducts.

In this section we work with a fixed group G, a fixed subset $A \subseteq G$ and a finitely additive measure μ, defined on a certain Boolean algebra of subsets of G and taking values in $\mathbb{R}_{\geqslant 0} \cup \{\infty\}$. While one could formulate a precise axiomatic framework to allow for a more general setting, it will suffice for our purposes to further assume that one of the following two cases holds.

Discrete case. A is a non-empty finite subset of G and μ is the $|A|$-normalized counting measure: $\mu(X) = |X|/|A|$, defined for any $X \subseteq G$.

Pseudofinite case. G, A and μ are as in Section 3.

The reader only interested in Theorem 2.4 can assume the discrete case and ignore the pseudofinite case. We call a set $X \subseteq G$ *measurable* if $\mu(X)$ is defined. Note that Proposition 3.2 makes sense in the discrete case if we remove the word 'internal' and the statement remains true (this is what was shown in the proof).

Lemma 4.1. *Fix $m, n \geqslant 1$ and a measurable set $X \subseteq G$, with $0 < \mu(X) < \infty$.*

(a) *Suppose $\mu(XX^{-1}X) \leqslant k\mu(X)$ for some $k \geqslant 1$. Then there is a measurable symmetric set $Y \subseteq G$ such that $Y^n \subseteq (XX^{-1})^2$ and $(XX^{-1})^m$ is covered by $O_{k,m,n}(1)$ $(XX^{-1})^m$-translates of Y.*

(b) *Suppose $\mu(X^3) \leqslant k\mu(X)$ for some $k \geqslant 1$. Then there is a measurable symmetric set $Y \subseteq G$ such that $Y^n \subseteq (XX^{-1})^2 \cap X^2X^{-2} \cap (X^{-1}X)^2 \cap X^{-2}X^2$ and $\bar{X}^m$ is covered by $O_{k,m,n}(1)$ $\bar{X}^m$-translates of Y.*

Proof. We will prove the two statements in parallel, as the arguments are similar. Let $X \subseteq G$ be a fixed internal set, with $0 < \mu(X) < \infty$. Fix an integer $k \geqslant 1$. By 'case (a)' we mean the assumption that $\mu(XX^{-1}X) \leqslant k\mu(X)$, and by 'case (b)' we mean the assumption that $\mu(X^3) \leqslant k\mu(X)$.

Before starting the argument, we first simplify case (b). Define

$$\Pi_1(X) = (XX^{-1})^2, \quad \Pi_2(X) = X^2X^{-2}, \quad \Pi_3(X) = (X^{-1}X)^2 \quad \text{and} \quad \Pi_4(X) = X^{-2}X^2.$$

We claim that it suffices to find, for each individual $c \in \{1, 2, 3, 4\}$, a set Y_c as described but only with $Y_c^n \subseteq \Pi_c(X)$. To see this, we apply some elementary tools from [7] (which transfer

to the pseudofinite setting). So fix $m, n \geqslant 1$ and suppose that, for $c \in \{1, 2, 3, 4\}$, we have measurable symmetric $Y_c \subseteq G$ such that $Y_c^{4n} \subseteq \Pi_c(X)$ and $\bar{X}^m$ is covered by $O_{k,m,n}(1)$ $\bar{X}^m$-translates of Y_c. Let $S = \bar{X}^{\max\{m,2\}}$. Then S is an $O_{k,m}(1)$-approximate subgroup by Proposition 3.2(a) and [7, Corollary 5.2]. Setting $Y_* = \bigcap_{c=1}^4 Y_c^2$, we have $\mu(\bar{X}^m) \leqslant \mu(S) \leqslant O_{k,m,n}(\mu(Y_*))$ by [7, Corollary 5.9]. By [7, Lemma 5.1], there is a finite set $F \subseteq \bar{X}^m$ such that $|F| \leqslant \mu(\bar{X}^m Y_*)/\mu(Y_*)$ and $\bar{X}^m \subseteq FY_*^2$. Since $Y_* \subseteq \bar{X}^2$, we have $\mu(\bar{X}^m Y_*) \leqslant O_{k,m}(\mu(\bar{X}^m))$ by Proposition 3.2(a), which implies $|F| \leqslant O_{k,m,n}(1)$. So if we set $Y = Y_*^2$, then Y is a measurable symmetric set, $\bar{X}^m$ is covered by $O_{k,m,n}(1)$ $\bar{X}^m$-translates of Y, and

$$Y^n = Y_*^{2n} \subseteq \bigcap_{c=1}^4 Y_c^{4n} \subseteq \bigcap_{c=1}^4 \Pi_c(X),$$

as desired.

So now in case (b) we fix $c \in \{1, 2, 3, 4\}$ and find $Y = Y_c$ with $Y^n \subseteq \Pi_c(X)$. Since $\Pi_1(X^{-1}) = \Pi_3(X)$ and $\Pi_2(X^{-1}) = \Pi_4(X)$, it suffices to assume $c \in \{1, 2\}$. Set

$$V = \begin{cases} (XX^{-1})^m & \text{in case (a)} \\ \bar{X}^m & \text{in case (b)} \end{cases} \quad \text{and} \quad Z = \begin{cases} X^{-1} & \text{in case (a), or case (b) with } c = 1, \\ X & \text{in case (b) with } c = 2. \end{cases}$$

Note that V is symmetric. We now closely follow Sanders [32]. For $t \in (0, 1]$, define

$$\mathcal{B}_t = \{B \subseteq X : B \text{ is internal and } \mu(B) \geqslant t\mu(X)\}.$$

Then $X \in \mathcal{B}_t$ for all $t \in (0, 1]$, and so we may define a function $f : (0, 1] \to [1, \infty)$ such that $f(t) = \inf\{\mu(BZ)/\mu(X) : B \in \mathcal{B}_t\}$.

By Proposition 3.2, we may fix $\ell \geqslant 1$ such that $\ell \leqslant k^{O_m(1)}$ and $\mu(VX) \leqslant \ell\mu(X)$ (in case (a) use $\mu((XX^{-1})^m X) \leqslant \mu((XX^{-1})^{m+1})$). By [24, Lemma 11] (taken from [32]), we may choose $t \in (0, 1]$ such that $t^{-1} \leqslant O_{k,m,n}(1)$ and $f(t^2/2\ell) \geqslant ((2n-1)/2n)f(t)$. Choose $B \in \mathcal{B}_t$ such that $\mu(BZ)/\mu(X) \leqslant ((2n+1)/2n)f(t)$.

Define

$$Y_* = \{g \in V^2 : \mu(gB \cap B) \geqslant (t^2/2\ell)\mu(X)\},$$

and note that $1 \in Y_*$, since $B \in \mathcal{B}_t$ and $t > t^2/2\ell$.

Claim 1. *V is covered by $O_{k,m,n}(1)$ V-translates of Y_*.*

Proof. Let $w = \lfloor 2\ell/t \rfloor$ and note that $w \leqslant O_{k,m,n}(1)$. Suppose, for a contradiction, that V is not covered by w V-translates of Y_*. Then we may construct a sequence $(g_i)_{i=0}^w$ from V such that, for all $i \leqslant w$, $g_i \notin \bigcup_{j<i} g_j Y_*$. For any $0 \leqslant i < j \leqslant w$, $g_i^{-1}g_j \in V^2 \setminus Y_*$, and so we have $\mu(g_iB \cap g_jB) < (t^2/2\ell)\mu(X)$. We also have $g_iB \subseteq VX$ for any $0 \leqslant i \leqslant w$. Now we obtain a contradiction:

$$\ell\mu(X) \geqslant \mu(VX)$$

$$\geqslant \mu\left(\bigcup_{i \leqslant w} g_iB\right)$$

$$\geqslant (w+1)\mu(B) - \sum_{i<j\leqslant w} \mu(g_iB \cap g_jB)$$

$$> (w+1)t\mu(X) - \frac{w(w+1)t^2\mu(X)}{4\ell}$$

$$= (w+1)\left(1 - \frac{wt}{4\ell}\right)t\mu(X)$$

$$> \ell\mu(X),$$

where the last inequality uses $w \leqslant 2\ell/t < w + 1$.   □

Set

$$W = \begin{cases} (XX^{-1})^2 & \text{in case (a), or case (b) with } c = 1, \\ X^2X^{-2} & \text{in case (b) with } c = 2. \end{cases}$$

Claim 2. $Y_*^n \subseteq W$.

Proof. We first show that $\mu(gBZ \triangle BZ) < 2\mu(BZ)/n$ for any $g \in Y_*$. To see this, note that if $g \in Y_*$ then $gB \cap B \in \mathcal{B}_{t^2/2\ell}$, and so

$$\mu(gBZ \cap BZ) \geqslant \mu((gB \cap B)Z) \geqslant f(t^2/2\ell)\mu(X) \geqslant \frac{2n-1}{2n}f(t)\mu(X) \geqslant \frac{2n-1}{2n+1}\mu(BZ).$$

So, for any $g \in Y_*$,

$$\mu(gBZ \triangle BZ) = 2\mu(BZ) - 2\mu(gBZ \cap BZ) \leqslant \frac{4}{2n+1}\mu(BZ) < \frac{2}{n}\mu(BZ).$$

Now fix $g_1, \ldots, g_n \in Y_*$ and, for $0 \leqslant i \leqslant n$, let $h_i = \prod_{j \leqslant i} g_j$ (so $h_0 = 1$). Then

$$\mu(h_n BZ \triangle BZ) \leqslant \mu\left(\bigcup_{i=0}^{n-1} h_i(g_{i+1}BZ \triangle BZ)\right) \leqslant \sum_{i=0}^{n-1} \mu(g_{i+1}BZ \triangle BZ) < 2\mu(BZ).$$

It follows that $h_n BZ \cap BZ \neq \emptyset$, which implies $h_n \in BZZ^{-1}B^{-1} \subseteq W$.   □

Now, in the discrete case, we may take $Y = Y_*$ and the proof is finished. In the pseudofinite case, we must address the fact that Y_* may not be internal. So suppose we are in the pseudofinite case.

Claim 3. $Y_* = \bigcap_{i=0}^{\infty} Y_i$ where, for each $i \in \mathbb{N}$, Y_i is symmetric and internal, and contains Y_{i+1}.

Proof. Let $\beta = (t^2/2\ell)\mu(X)$, and so $Y_* = \{g \in V^2 : \mu(gB \cap B) \geqslant \beta\}$. By assumption, X and B are internal and so we may choose sets $X_s, B_s \subseteq G_s$, for $s \in \mathbb{N}$, such that $X = \prod_{\mathcal{U}} X_s$ and $B = \prod_{\mathcal{U}} B_s$. Given $s \in \mathbb{N}$, set

$$V_s = \begin{cases} (X_sX_s^{-1})^n & \text{in case (a),} \\ \bar{X}_s^n & \text{in case (b).} \end{cases}$$

Note that each V_s is symmetric, and $V = \prod_{\mathcal{U}} V_s$. Given $i \in \mathbb{N}$ and $s \in \mathbb{N}$, define

$$Y_{i,s} = \left\{g \in V_s^2 : \frac{|gB_s \cap B_s|}{|G_s|} > \beta - \frac{1}{i+1}\right\}.$$

Note that $Y_{i,s}^{-1} = Y_{i,s}$ for all $i, s \in \mathbb{N}$. Given $i \in \mathbb{N}$, let $Y_i = \prod_{\mathcal{U}} Y_{i,s}$. Then, for any $i \in \mathbb{N}$, $Y_i^{-1} = Y_i$ and $Y_{i+1} \subseteq Y_i$. Moreover, $Y_* = \bigcap_{i=0}^{\infty} Y_i$.   □

Fix $(Y_i)_{i=0}^{\infty}$ as in Claim 3. To finish the proof of the lemma in the pseudofinite case, it suffices to show that $Y_i^n \subseteq W$ for some $i \in \mathbb{N}$. Toward this end, we first show $Y_*^n = \bigcap_{i=0}^{\infty} Y_i^n$. We clearly have $Y_*^n \subseteq \bigcap_{i=0}^{\infty} Y_i^n$. For the other direction, fix $a \in \bigcap_{i=0}^{\infty} Y_i^n$. For $i \in \mathbb{N}$, define

$$D_i = \{(g_1, \ldots, g_n) \in G^{\times n} : g_j \in Y_i \text{ for } 1 \leqslant j \leqslant i, \text{ and } a = g_1 \cdot \ldots \cdot g_n\}.$$

Then, for all $i \in \mathbb{N}$, D_i is non-empty, internal, and $D_{i+1} \subseteq D_i$. By Fact 3.1, there is $(g_1, \ldots, g_n) \in \bigcap_{i=0}^{\infty} D_i$, and so $a = g_1 \cdot \ldots \cdot g_n \in Y_*^n$.

Finally, since $\bigcap_{i=0}^{\infty} Y_i^n = Y_*^n \subseteq W$, it follows from Fact 3.1 that $Y_i^n \subseteq W$ for some $i \in \mathbb{N}$.   □

At this point, we have all the necessary tools to proceed with the proof of Theorem 2.3 (see Section 5). For Theorem 2.4 we will need the following corollary of the previous lemma, which is only meaningful in the pseudofinite case.

Corollary 4.2.

(a) *Suppose $\mu(AA^{-1}A) \leqslant k < \infty$. Then there is a sequence $(X_n)_{n=0}^{\infty}$ of symmetric, internal subsets of G such that $X_0 \subseteq (AA^{-1})^2$ and, for any $n \in \mathbb{N}$, $X_{n+1}^2 \subseteq X_n$ and $(AA^{-1})^4$ is covered by $O_{k,n}(1) \langle AA^{-1} \rangle$-translates of X_n.*

(b) *Suppose $\mu(A^3) \leqslant k < \infty$. Then there is a sequence $(X_n)_{n=0}^{\infty}$ of symmetric, internal subsets of G such that $X_0 \subseteq (AA^{-1})^2 \cap A^2A^{-2} \cap (A^{-1}A)^2 \cap A^{-2}A^2$ and, for any $n \in \mathbb{N}$, $X_{n+1}^2 \subseteq X_n$ and $\bar{A}^8$ is covered by $O_{k,n}(1) \langle A \rangle$-translates of X_n.*

Proof. As in Lemma 4.1, we prove the two statements in parallel. Set

$$(V, W, \Sigma) = \begin{cases} ((AA^{-1})^4, (AA^{-1})^2, \langle AA^{-1} \rangle) & \text{in case (a),} \\ (\bar{A}^8, (AA^{-1})^2 \cap A^2A^{-2} \cap (A^{-1}A)^2 \cap A^{-2}A^2, \langle A \rangle) & \text{in case (b).} \end{cases}$$

We construct a sequence $(Y_n)_{n=0}^{\infty}$ of symmetric, internal subsets of G, such that $Y_0^8 \subseteq W$ and, for all $n \in \mathbb{N}$, $Y_{n+1}^8 \subseteq Y_n^4$ and V is covered by $O_{K,n}(1)$ Σ-translates of Y_n. Given this, the result follows by setting $X_n = Y_n^4$.

By Lemma 4.1, there is a symmetric, internal set $Y_0 \subseteq G$ such that $Y_0^8 \subseteq W$ and V is covered by $O_k(1)$ V-translates of Y_0. Suppose we have constructed $Y_0, \ldots, Y_n$ satisfying the desired properties. Note that $\mu(W) < \infty$ by Proposition 3.2, and thus $\mu(Y_n) < \infty$ since $Y_n \subseteq Y_n^8 \subseteq W$. Since V is covered by $O_{k,n}(1)$ translates of Y_n, we have $0 < \mu(V) \leqslant O_{k,n}(1)\mu(Y_n)$, and thus $\mu(Y_n) > 0$. Since $Y_n^3 \subseteq W$, we also have $\mu(Y_n^3) \leqslant O_{k,n}(1)\mu(Y_n)$. By Lemma 4.1(b), there is a symmetric, internal $Y_{n+1} \subseteq G$ such that $Y_{n+1}^8 \subseteq Y_n^4$ and Y_n^8 is covered by $O_{k,n+1}(1)$ Y_n^8-translates of Y_{n+1}. Since $Y_n \subseteq Y_n^8 \subseteq W$, it follows that Y_n is covered by $O_{k,n+1}(1)$ W-translates of Y_{n+1}. Since V is covered by $O_{k,n}(1)$ Σ-translates of Y_n, it follows that V is covered by $O_{k,n+1}(1)$ Σ-translates of Y_{n+1}.   □

5. Proof of Theorem 2.3

The following theorem is [7, Theorem 6.15]. It can also be deduced from [19, Corollary 4.18].

Theorem 5.1 ([7])**.** *Let G be a group of exponent r, and suppose $X \subseteq G$ is a k-approximate subgroup. Then X^4 contains a subgroup $H \leqslant G$ such that X is covered by $O_{k,r}(1)$ left cosets of H.*

We now give the proof of Theorem 2.3.

Proof of Theorem 2.3. We prove parts (1) and (2) of the theorem in two parallel cases. Fix positive integers k, m and r. Let G be a group and fix $A \subseteq G$ non-empty and finite, with $|AA^{-1}A| \leqslant k|A|$ in case (1) and $|A^3| \leqslant k|A|$ in case (2). Set

$$(V, W, \Sigma) = \begin{cases} (AA^{-1}, (AA^{-1})^2, \langle AA^{-1} \rangle) & \text{in case (1),} \\ (\bar{A}, (AA^{-1})^2 \cap A^2A^{-2} \cap (A^{-1}A)^2 \cap A^{-2}A^2, \langle A \rangle) & \text{in case (2).} \end{cases}$$

By increasing m if necessary, we may assume $W \subseteq V^m$ without loss of generality.

Assume Σ has exponent r. By Lemma 4.1, there is a non-empty finite symmetric set $Y \subseteq G$ such that $Y^4 \subseteq W$ and V^m is covered by $O_{k,m}(1)$ left translates of Y. Since $Y^2 \subseteq Y^4 \subseteq W \subseteq V^m$, it follows that Y is an $O_{k,m}(1)$-approximate group. By Theorem 5.1, Y^4 contains a subgroup $H \leqslant G$ such that Y is covered by $O_{k,m,r}(1)$ left cosets of H. Then $H \subseteq W$ and V^m is covered by $O_{k,m,r}(1)$ left cosets of H. This proves part (a) in both cases (1) and (2).

For part (b), suppose $\Sigma = V^m$. By part (a) there is a subgroup $K \leqslant \Sigma$, of index $O_{k,m,r}(1)$, such that $K \subseteq W$. If $H = \bigcap_{g \in \Sigma} gKg^{-1}$, then H is normal in Σ, $H \subseteq W$ and $[G\colon H] \leqslant [G\colon K]! \leqslant O_{k,m,r}(1)$. $\qquad\square$

6. Saturated extensions and approximate Bohr neighbourhoods

Throughout this section, let G be an ultraproduct constructed as in Section 3. We will now endow G with a first-order structure, and then take a sufficiently saturated elementary extension G_*. Specifically, we define the *internal language of G*, denoted $\mathcal{L}$, to be the group language together with a unary relation R_X for any internal $X \subseteq G$. We view G as an $\mathcal{L}$-structure by interpreting each R_X as X. We also view each G_s as an $\mathcal{L}$-structure by interpreting R_X as some set $X(G_S) \subseteq G_s$, so that $X = \prod_{\mathcal{U}} X(G_s)$. In particular, G is also the ultraproduct of the sequence of $\mathcal{L}$-structures $(G_s)_{s \in \mathbb{N}}$.

Now let G_* be a sufficiently saturated elementary extension of G with respect to the language $\mathcal{L}$.[4] When we say $X \subseteq G_*$ (resp. $X \subseteq G$) is *definable*, we mean definable in the language $\mathcal{L}$ using parameters from G_* (resp. from G). If we want to specify that X is definable using parameters from some set C, we will say *C-definable*. Let A_* be the interpretation in G_* of the predicate in $\mathcal{L}$ naming A.

Note that the measure μ naturally extends to G-definable subsets of G_*. In particular, given a G-definable set $X \subseteq G_*$, the interpretation $X(G)$ of X in G is internal, so we let $\mu(X) = \mu(X(G))$. We say that a G-definable set $X \subseteq G_*$ is *pseudofinite* if $X(G)$ is an ultraproduct of finite sets.

Remark 6.1. Although it will not be necessary for our work, we recall that μ can be extended (not necessarily uniquely) to all definable subsets of G_*. For example, one can add a sort for $[0, 1]$ and a function f_ϕ for each formula $\phi(x; \bar{y})$, from the home sort to $[0, 1]$, which is interpreted as $f_\phi(\bar{b}) = \mu(\phi(G; \bar{b}))$. Then take G_* to be a saturated extension in this larger language. See [20, Section 2].

A cardinal is *bounded* if it is strictly smaller than the saturation of G_*. A set $X \subseteq G_*$ is *type-definable* (resp. *countably type-definable*) if it is an intersection of a bounded (resp. countable) number of definable subsets of G_*.

Now suppose Σ is a definable subgroup of G_*, and Γ is a type-definable normal subgroup of Σ such that $[\Sigma\colon \Gamma]$ is bounded. Call a set $X \subseteq \Sigma/\Gamma$ *closed* if $\pi^{-1}(X)$ is type-definable, where π is the canonical projection from Σ. It is a standard fact that this defines a topology on Σ/Γ, called the *logic topology*, under which Σ/Γ is a compact (Hausdorff) topological group. If Γ is countably type-definable, then Σ/Γ is second countable. See [27, Section 2] for details.

The rest of this section summarizes some tools from [10] concerning Bohr neighbourhoods in G_* and issues regarding their transfer to G and the groups G_s.

Given a compact space $\mathcal{X}$, we say that a map $f\colon G \to \mathcal{X}$ is *definable* if $f^{-1}(C)$ is type-definable for any closed $C \subseteq \mathcal{X}$. The next proposition is a special case of [10, Proposition 5.1], and crucially relies on the result of Pillay [28] that the connected component of a definable compactification of a pseudofinite group is abelian.

---

[4]'Sufficiently saturated' typically means κ-saturated and strongly κ-homogeneous for some very large (*e.g.* strongly inaccessible) cardinal κ.

Proposition 6.2. *Suppose Σ is a G-definable pseudofinite subgroup of G_*, and $\Gamma \leqslant \Sigma$ is a countably type-definable bounded-index normal subgroup of Σ. Then there is a decreasing sequence $(X_i)_{i=0}^\infty$ of definable subsets of Σ such that $\Gamma = \bigcap_{i=0}^\infty X_i$ and, for all $i \in \mathbb{N}$, there are:*

- *a definable finite-index normal subgroup $H_i \leqslant \Sigma$, and*
- *a definable homomorphism $\pi_i \colon H_i \to \mathbb{T}^{n_i}$, for some $n_i \in \mathbb{N}$,*

such that $\Gamma \subseteq \ker \pi_i \subseteq X_i \subseteq H_i$. If, moreover, Σ / Γ is abelian, then we may assume $H_i = \Sigma$ for all $i \in \mathbb{N}$.

In the setting of the previous proposition, the fact that X_i is definable and contains $\ker \pi_i$ implies that it contains a Bohr neighbourhood $B_{\pi_i, \varepsilon_i}^{n_i}$ for sufficiently small $\varepsilon_i > 0$. However, these Bohr neighbourhoods are not necessarily definable, and so we will need to approximate them by definable objects.

Definition 6.3. Fix a group H and an integer $n \in \mathbb{N}$.

(1) Given $\delta > 0$, we say that a function $f \colon H \to \mathbb{T}^n$ is a *δ-homomorphism* if $f(1) = 0$ and, for all $x, y \in H$, $d_n(f(xy), f(x) + f(y)) < \delta$.
(2) Given $\delta, \varepsilon > 0$, we say that $Y \subseteq H$ is a *δ-approximate (ε, n)-Bohr neighbourhood in H* if there is a δ-homomorphism $f \colon H \to \mathbb{T}^n$ such that $Y = \{x \in H \colon d_n(f(x), 0) < \varepsilon\}$.
(3) Assume H is a definable subgroup of G_*, and $\pi \colon H \to \mathbb{T}^n$ is a definable homomorphism. Given an integer $t \geqslant 1$, we say that a decreasing sequence $(Y_i)_{i=0}^\infty$ of subsets of H is a *definable (t, π)-approximate Bohr chain in H* if $\bigcap_{i=0}^\infty Y_i = \ker \pi$ and there is a decreasing sequence $(\delta_i)_{i=0}^\infty$ in $\mathbb{R}_{>0}$ converging to 0 such that, for all $i \geqslant 0$,

$$Y_i = \{x \in H \colon d_n(f_i(x), 0) < t\delta_i\}$$

for some definable δ_i-homomorphism $f_i \colon H \to \mathbb{T}^n$ with finite image.

Note that if $(Y_i)_{i=0}^\infty$ is a definable (t, π)-approximate Bohr chain in H, then each Y_i is a δ_i-approximate $(t\delta_i, n)$-Bohr set in H. It is also worth emphasizing that each Y_i is indeed a *definable* subset of H (see [10, Proposition 5.3]). The next result is a special case of [10, Lemma 5.4].

Lemma 6.4. *Suppose $H \leqslant G_*$ is definable and $\pi \colon H \to \mathbb{T}^n$ is a definable homomorphism for some $n \in \mathbb{N}$. Then, for any integer $t \geqslant 1$, there is a definable (t, π)-approximate Bohr chain $(Y_i)_{i=0}^\infty$ in H.*

Finally, we state a special case of [10, Corollary 4.4], which is an immediate consequence of [1, Theorem 5.13].

Proposition 6.5. *There is a real number $\theta > 0$ such that if H is a finite group, $n \in \mathbb{N}$, and $0 < \delta < \theta$, then every δ-approximate $(3\delta, n)$-Bohr neighbourhood in H contains a (δ, n)-Bohr neighbourhood in H.*

7. Proof of Theorem 2.4

7.1 Transfer to G_*

Throughout this subsection, let G be an ultraproduct constructed as in Section 3, and let G_* be the saturated extension from Section 6. The goal of this subsection is to transfer the analysis in Section 4 to the saturated group G_*. The main idea is that the decreasing sequence $(X_n)_{n=0}^\infty$ of internal sets constructed in Corollary 4.2 converges to a subgroup of G, which is 'large' in a certain

sense. By transferring the sequence first to G_*, we will have more precise control over exactly what this means, and it will be easier to find normal subgroups.

Recall that A_* is the interpretation in G_* of our distinguished internal set $A \subseteq G$.

Lemma 7.1.

(a) *Suppose $\mu(A_* A_*^{-1} A_*) < \infty$. Then there is a countably type-definable subgroup $\Gamma \leqslant G_*$ such that $\Gamma \subseteq (A_* A_*^{-1})^2$ and Γ has index at most $2^{\aleph_0}$ in $\langle A_* A_*^{-1} \rangle$.*
(b) *Suppose $\mu(A_*^3) < \infty$. Then there is a countably type-definable subgroup $\Gamma \leqslant G_*$ such that $\Gamma \subseteq (A_* A_*^{-1})^2 \cap A_*^2 A_*^{-2} \cap (A_*^{-1} A_*)^2 \cap A_*^{-2} A_*^2$ and Γ has index at most $2^{\aleph_0}$ in $\langle A_* \rangle$.*

Proof. Set

$$(V, W, \Sigma) = \begin{cases} ((AA^{-1})^2, (AA^{-1})^2, \langle AA^{-1} \rangle) & \text{in case (a)}, \\ (\bar{A}^4, (AA^{-1})^2 \cap A^2 A^{-2} \cap (A^{-1}A)^2 \cap A^{-2}A^2, \langle A \rangle) & \text{in case (b)}. \end{cases}$$

Let (V_*, W_*, Σ_*) be defined similarly, but with A_* in place of A.

Working first in G, we apply Corollary 4.2 to find a sequence $(Y_n)_{n=0}^\infty$ of symmetric, internal subsets of G such that $Y_0 \subseteq W$ and, for any $n \in \mathbb{N}$, $Y_{n+1}^2 \subseteq Y_n$ and V^2 is covered by finitely many Σ-translates of Y_n. So, for any $n \in \mathbb{N}$, there is some $k_n \in \mathbb{N}$ such that V^2 is covered by finitely many V^{k_n}-translates of Y_n.

Now in G_*, let X_n be the $\varnothing$-definable set given by the interpretation of the unary relation R_{Y_n}. By elementarity, $X_0 \subseteq W_*$ and, for any $n \in \mathbb{N}$, X_n is symmetric and internal, $X_{n+1}^2 \subseteq X_n$, and V_*^2 is covered by finitely many $V_*^{k_n}$-translates of X_n.

Fix $n \in \mathbb{N}$, and let $F \subseteq \Sigma_*$ be finite such that $V_*^2 \subseteq FX_n$. By induction on $k \geqslant 1$, we show that $V_*^k \subseteq F^k X_n$. The base case is given, so assume the result for $k \geqslant 1$. Then $V_*^{k+1} = V_*^k V_* \subseteq F^k X_n V_* \subseteq F^k W_* V_* \subseteq F^k V_*^2 \subseteq F^{k+1} X_n$.

We have shown that, for any $n \in \mathbb{N}$, there is a countable set $F_n \subseteq \Sigma_*$ such that $\Sigma_* = FX_n$. Let $\Gamma = \bigcap_{n=0}^\infty X_n$, and note that Γ is a countably type-definable subgroup of G_*, which is contained in W_*. Since Σ_* is covered by countably many Σ_*-translates of X_n for all $n \geqslant 1$, it follows that Γ has index at most $2^{\aleph_0}$ in Σ_*. $\qquad\square$

Corollary 7.2.

(a) *Suppose $\mu(A_* A_*^{-1} A_*) < \infty$. Then there is a countably type-definable subgroup $\Gamma \leqslant G_*$ such that:*
  (i) *$\Gamma \subseteq (A_* A_*^{-1})^2$,*
  (ii) *Γ is normal in $\langle A_* A_*^{-1} \rangle$, and*
  (iii) *Γ has index at most $2^{\aleph_0}$ in $\langle A_* A_*^{-1} \rangle$.*
(b) *Suppose $\mu(A_*^3) < \infty$. Then there is a countably type-definable subgroup $\Gamma \leqslant G_*$ such that:*
  (i) *$\Gamma \subseteq (A_* A_*^{-1})^2 \cap A_*^2 A_*^{-2} \cap (A_*^{-1} A_*)^2 \cap A_*^{-2} A_*^2$,*
  (ii) *Γ is normal in $\langle A_* \rangle$, and*
  (iii) *Γ has index at most $2^{\aleph_0}$ in $\langle A_* \rangle$.*

Proof. Set

$$(V, W, \Sigma) = \begin{cases} (A_* A_*^{-1}, (A_* A_*^{-1})^2, \langle A_* A_*^{-1} \rangle) & \text{in case (a)}, \\ (\bar{A}_*, (A_* A_*^{-1})^2 \cap A_*^2 A_*^{-2} \cap (A_*^{-1} A_*)^2 \cap A_*^{-2} A_*^2, \langle A_* \rangle) & \text{in case (b)}. \end{cases}$$

By Lemma 7.1, we have a countably type-definable subgroup $\Gamma_0 \leqslant G_*$ such that $\Gamma_0 \subseteq W$ and Γ_0 has index at most $2^{\aleph_0}$ in Σ. Let $\Gamma = \bigcap_{g \in \Sigma} g\Gamma_0 g^{-1}$. Then Γ is an intersection of at most $2^{\aleph_0}$ conjugates $g\Gamma_0 g^{-1}$ with $g \in \Sigma$. So Γ is a type-definable subgroup of G_*, which is normal in Σ and has index at most $2^{2^{\aleph_0}}$ in Σ.

We now show that Γ is countably type-definable of index at most $2^{\aleph_0}$ in Σ. For the first part, let $\mathcal{L}_0 \subseteq \mathcal{L}$ be a countable language containing the language of groups, and unary predicates defining A and Y_n for $n \in \mathbb{N}$, where Y_n are the predicates used to obtain Γ_0 (via the proof of Lemma 7.1). Then Γ is $\mathcal{L}_0$-type-definable. Moreover, Γ_0 is $\mathcal{L}_0$-type-definable over $\emptyset$, and so $\sigma(\Gamma_0) = \Gamma_0$ for any $\sigma \in \operatorname{Aut}_{\mathcal{L}_0}(G_*)$. Since Σ is $\operatorname{Aut}_{\mathcal{L}_0}(G_*)$-invariant, $\sigma(\Gamma) = \Gamma$ for any $\sigma \in \operatorname{Aut}_{\mathcal{L}_0}(G_*)$, and so Γ is $\mathcal{L}_0$-type-definable over $\emptyset$. Since $\mathcal{L}_0$ is countable, Γ is countably type-definable.

Finally, let $\Gamma = \bigcap_{n=0}^{\infty} D_n$, where each D_n is definable and (without loss of generality) contained in Σ. Since Γ has bounded index in Σ, we may fix some bounded set $C \subset \Sigma$ such that $\Sigma = C\Gamma$. Fix $m, n \in \mathbb{N}$. Then $V^m \subseteq \Sigma = C\Gamma = CD_n$. By saturation of G_*, it follows that there is some finite $C_{n,m} \subseteq C$ such that $V^m \subseteq C_{n,m}D_n$. So, if $C_n = \bigcup_{m \in \mathbb{N}} C_{n,m}$, then C_n is countable and $\Sigma = C_n D_n$. Once again, this implies that Γ has index at most $2^{\aleph_0}$ in Σ. $\square$

Corollary 7.2 is a non-standard Bogolyubov–Ruzsa-type statement about pseudofinite sets of small alternation or small tripling. However, since the subgroup Γ is not necessarily definable, it cannot be directly transferred to statements about internal subsets of G (which are needed in order to transfer to the finite groups G_s). For this, we need the material in Section 6 on approximate Bohr neighbourhoods.

7.2 Ultraproduct argument

We now prove parts (1) and (2) of Theorem 2.4 simultaneously. Given a group G and a set $A \subseteq G$, let

$$\Sigma(A) = \begin{cases} \langle AA^{-1} \rangle \\ \langle A \rangle \end{cases} \qquad U(A) = \begin{cases} AA^{-1}A \\ A^3 \end{cases} \qquad V(A) = \begin{cases} AA^{-1} & \text{in part (1)} \\ \bar{A} & \text{in part (2)} \end{cases}$$

$$\text{and} \quad W(A) = \begin{cases} (AA^{-1})^2 & \text{in part (1),} \\ (AA^{-1})^2 \cap A^2 A^{-2} \cap (A^{-1}A)^2 \cap A^{-2}A^2 & \text{in part (2).} \end{cases}$$

The ambient group G is suppressed from the notation, but this should cause no confusion in the following proof.

The next result is a restatement of Theorem 2.4, which we will prove by taking an ultraproduct of counterexamples, and using the material in Section 6 in order to transfer Bohr neighbourhoods through ultraproducts and saturated extensions.

Theorem 7.3. *For any positive integers k and m, there is an integer $s = s(k, m)$ such that the following holds. Suppose G is a group and $A \subseteq G$ is finite such that $|U(A)| \leqslant k|A|$ and $\Sigma(A) = V(A)^m$. Then there are:*

- *a normal subgroup $H \leqslant \Sigma(A)$, of index at most s, and*
- *a (δ, n)-Bohr neighbourhood B in H, with $\delta^{-1}, n \leqslant s$,*

such that $B \subseteq W(A)$. Moreover, if $\Sigma(A)$ is abelian then we may assume $H = \Sigma(A)$.

Proof. Suppose not. Then for any $s \in \mathbb{N}$, we may fix a group G_s and a finite set $A_s \subseteq G_s$ such that $|U(A_s)| \leqslant k|A_s|$, $\Sigma(A_s) = V(A_s)^m$, and there does not exist a normal subgroup $H \leqslant \Sigma(A_s)$ and a

(δ, n)-Bohr neighbourhood B in H such that $[\Sigma(A_s)\colon H], \delta^{-1}, n \leqslant s$ and $H \subseteq W(A_s)$. Note that $|\Sigma(A_s)| > \max\{s, m\}$, since otherwise we could take $H = \{1\}$.

Let $\mathcal{U}$ be a non-principal ultrafilter on $\mathbb{N}$ and set $G = \prod_{\mathcal{U}} G_s$. Let $A = \prod_{\mathcal{U}} A_s$, and note that A is an internal subset of G. We also have

$$U(A) = \prod_{\mathcal{U}} U(A_s), \quad V(A) = \prod_{\mathcal{U}} V(A_s), \quad W(A) = \prod_{\mathcal{U}} W(A_s),$$

and

$$\Sigma(A) = \bigcup_{n \in \mathbb{N}} V(A)^n = \bigcup_{n \in \mathbb{N}} \prod_{\mathcal{U}} V(A_s)^n = \prod_{\mathcal{U}} V(A_s)^m = V(A)^m$$

Note, in particular, that $\Sigma(A) = \prod_{\mathcal{U}} \Sigma(A_s)$ is infinite, and so G is infinite. Let μ be the $|A|$-normalized pseudofinite counting measure on internal subsets of G. By Łoś's theorem, $\mu(U(A)) \leqslant k < \infty$.

Let G_* be a saturated elementary extension of G in the internal language of G (see Section 6), and let A_* be the interpretation in G_* of the predicate defining A in G. By Lemma 7.1, there is a countably type-definable subgroup $\Gamma \leqslant G_*$ such that $\Gamma \subseteq W(A_*)$ and Γ has index at most $2^{\aleph_0}$ in $\Sigma(A_*)$. Note also that $\Sigma(A_*) = V(A_*)^m$. In particular, $\Sigma(A_*)$ is G-definable and pseudofinite.

By Proposition 6.2 and saturation of G_*, there is a definable finite-index normal subgroup $H \leqslant \Sigma(A_*)$ and a definable homomorphism $\pi\colon H \to \mathbb{T}^n$, for some $n \in \mathbb{N}$, such that $\Gamma \subseteq \ker \pi \subseteq H \cap W(A_*)$. By Lemma 6.4, there is a definable $(3, \pi)$-approximate Bohr chain $(Y_i)_{i=0}^{\infty}$ in H. By saturation, $Y_i \subseteq W(A_*)$ for sufficiently large $i \in \mathbb{N}$. So we may fix $\delta < \theta$, where θ is as in Proposition 6.5, and a definable δ-homomorphism $f\colon H \to \mathbb{T}^n$ such that $Y := \{x \in H\colon d(f(x), 0) < 3\delta\} \subseteq W(A^*)$.

Let $\Lambda = f(H)$, and note that Λ is finite (see Definition 6.3(3)). Given $\lambda \in \Lambda$, let $F(\lambda) = f^{-1}(\lambda) \subseteq H$. Then each $F(\lambda)$ is definable. Set $r = [\Sigma(A_*)\colon H] < \infty$.

Fix $\mathcal{L}$-formulas $\phi(x; \bar{y})$, $\psi(x; \bar{z})$ and $\zeta_\lambda(x; \bar{u}_\lambda)$ for $\lambda \in \Lambda$, such that H is defined by an instance of $\phi(x; \bar{y})$, Y is defined by an instance of $\psi(x; \bar{z})$, and, for $\lambda \in \Lambda$, $F(\lambda)$ is defined by an instance of $\zeta_\lambda(x; \bar{u}_\lambda)$. Let $I \subseteq \mathbb{N}$ be the set of $s \in \mathbb{N}$ such that, for some tuples $\bar{a}_s$, $\bar{b}_s$ and $\bar{c}_{\lambda,s}$ (for $\lambda \in \Lambda$) from G_s, we have:

(i) $\phi(x; \bar{a}_s)$ defines a normal subgroup H_s of $\Sigma(A_s) = V(A_s)^m$ of index r,
(ii) for all $\lambda \in \Lambda$, $\zeta_\lambda(x; \bar{c}_{\lambda,s})$ defines a subset $F_s(\lambda)$ of H_s,
(iii) if $f_s\colon H_s \to \Lambda$ is defined so that $f_s(x) = \lambda$ if and only if $x \in F_s(\lambda)$, then f_s is a well-defined δ-homomorphism (from H_s to $\mathbb{T}^n$),
(iv) $\psi(x; \bar{b}_s)$ defines a subset Y_s of H_s, and $Y_s = \{x \in H_s\colon d(f_s(x), 0) < 3\delta\}$,
(v) $Y_s \subseteq W(A_s)$.

Then $I \in \mathcal{U}$ by Łoś's theorem and elementarity (checking that (i)–(v) are first-order expressible is somewhat cumbersome, but fairly routine; see the proof of [10, Lemma 5.6]). So we may fix some $s \in I$ such that $r, n, \delta^{-1} \leqslant s$. For this s, Y_s is a δ-approximate $(3\delta, n)$-Bohr set in H_s. By Proposition 6.5, there is a (δ, n)-Bohr set $B \subseteq Y_s$. So $B \subseteq W(A_s)$, which contradicts the choice of G_s and A_s.

Finally, if we assume $\Sigma(A)$ is abelian then, in the above proof, we may take $H = \Sigma(A_*)$ by Proposition 6.2, and thus assume $H_s = \Sigma(A_s)$ for all $s \in \mathbb{N}$. $\qquad\square$

Remark 7.4. Suppose that in Theorem 7.3 we further assume G is abelian and $|A| \geqslant c|G|$ for some fixed $c > 0$. Then we have $\langle A \rangle = \bar{A}^m$, where $m \leqslant \lceil 3c^{-1} + 1 \rceil$, and $[G\colon \langle A \rangle] \leqslant \lceil c^{-1} \rceil$. Therefore, in the proof of the theorem, $\Sigma(A_*)$ has finite index in G_*, and so Γ has index at most $2^{\aleph_0}$ in G_*. So we can carry out the rest of the proof with G_* in place of $\Sigma(A_*)$, obtaining $H_s = G_s$ in the conclusion. Consequently, in Theorem 1.2(b), if G is abelian then we may take $H = G$.

8. Arithmetic regularity and VC-dimension

The goal of this section is to prove Theorem 1.6. As indicated in the Introduction, the only ingredient in the work of Alon, Fox and Zhao [3] requiring abelian groups is Theorem 1.1(a). The (qualitative) non-abelian version of this result for sets of small *tripling*, provided by Corollary 1.4(a), will be sufficient to essentially carry out the same proof as in [3] (see also Remark 9.3). The only extra work is in specifying the numerics and clarifying the 'regularity' aspect of the result (*i.e.* condition (ii) of Theorem 1.6). We also make some similar (but mostly qualitative) statements for purely non-abelian finite groups, and finite simple groups.

Definition 8.1. Let G be a finite group. Given a subset $A \subseteq G$ and some $\varepsilon > 0$, define the ε-*stabilizer of A* to be the set $\mathrm{Stab}_\varepsilon(A) := \{x \in G \colon |xA \triangle A| \leqslant \varepsilon|G|\}$.

The following lemma, which we have extracted from the counting techniques done in [3], makes explicit the connection between ε-stabilizers and strong arithmetic regularity involving subgroups.

Lemma 8.2. *Let G be a finite group and fix a subset $A \subseteq G$ and some $\varepsilon > 0$. Suppose H is a subgroup of G contained in $\mathrm{Stab}_\varepsilon(A)$.*

(a) *There is $D \subseteq G$, which is a union of right cosets of H, such that $|A \triangle D| \leqslant \varepsilon|G|$.*
(b) *There is $Z \subseteq G$, with $|Z| < \frac{1}{2}\varepsilon^{1/2}|G|$, such that, for any $x \in G \backslash Z$, either $|Hx \cap A| \leqslant \varepsilon^{1/4}|H|$ or $|Hx \backslash A| \leqslant \varepsilon^{1/4}|H|$.*

Proof. Let $\mathcal{C}$ be the set of right cosets of H in G. Given $C \in \mathcal{C}$, define $P_C = (C \cap A) \times (C \backslash A)$. Let

$$P = \bigcup_{C \in \mathcal{C}} P_C = \{(a, g) \in A \times G \backslash A \colon ga^{-1} \in H\},$$

and note that $P_C \cap P_{C'} = \emptyset$ for distinct $C, C' \in \mathcal{C}$. From the proof of [3, Lemma 2.4], we obtain

$$2 \sum_{C \in \mathcal{C}} |P_C| = 2|P| = \sum_{x \in H} |xA \triangle A| \leqslant \varepsilon|G||H|. \tag{\dag}$$

For part (a), we continue to follow [3]. Let

$$D = \bigcup \{C \in \mathcal{C} \colon |C \cap A| \geqslant |H|/2\}.$$

Then, by ($\dag$),

$$|A \triangle D| = \sum_{C \in \mathcal{C}} \min\{|C \cap A|, |C| - |C \cap A|\} \leqslant \sum_{C \in \mathcal{C}} \frac{2}{|H|}|P_C| \leqslant \varepsilon|G|.$$

For part (b), let

$$\mathcal{Z} = \{C \in \mathcal{C} \colon |P_C| > \varepsilon^{1/2}|H|^2\}.$$

By ($\dag$),

$$\frac{1}{2}\varepsilon|G||H| \geqslant \sum_{C \in \mathcal{C}} |P_C| > \varepsilon^{1/2}|H|^2|\mathcal{Z}|.$$

So $|\mathcal{Z}| < \frac{1}{2}\varepsilon^{1/2}|G|/|H|$. Now set $Z = \bigcup_{C \in \mathcal{Z}} C$. Then $|Z| < \frac{1}{2}\varepsilon^{1/2}|G|$. Moreover, if $x \in G \backslash Z$ then $Hx \notin \mathcal{Z}$, and thus $|P_{Hx}| \leqslant \varepsilon^{1/2}|H|^2$, which implies $|Hx \cap A| \leqslant \varepsilon^{1/4}|H|$ or $|Hx \backslash A| \leqslant \varepsilon^{1/4}|H|$. $\square$

We can now prove Theorem 1.6, following the same steps as in [3].

Proof of Theorem 1.6. Fix positive integers r and d, and real numbers $\varepsilon, \nu > 0$. Suppose G is a finite group of exponent at most r, and $A \subseteq G$ has VC-dimension at most d. Let $S = \text{Stab}_\delta (A)$, where $\delta = (\varepsilon/4)^{(d+\nu)/d}/30^{\nu/d}$. Note that S is symmetric. Set $k = (30/\delta)^d$ and $p = d(d+\nu)/\nu$. It is an immediate consequence of *Haussler's Packing Lemma* [17], for set systems of finite VC-dimension, that $|S| \geqslant |G|/k$ (see Lemmas 2.1 and 2.2 of [3]). Therefore we cannot have $|S^{3^{i+1}}| > 3^p|S^{3^i}|$ for all $i \leqslant \log_{3^p}(k)$. So we may fix some $t \leqslant \log_{3^p}(k)$ such that, setting $B = S^{3^t}$, we have $|B^3| \leqslant 3^p|B|$. By Corollary 1.4(a), there is a subgroup $H \leqslant G$ such that $H \subseteq B^4$ and B is covered by $O_{r,d,\nu}(1)$ left translates of H. Since $|G| \leqslant k|B|$, we see that H has index at most $O_{r,d,\nu}(k) = O_{r,d,\nu}((1/\varepsilon)^{d+\nu})$.

To finish the proof, it suffices by Lemma 8.2 to show that $H \subseteq \text{Stab}_\varepsilon (A)$. We have $|xA \triangle A| \leqslant \delta|G|$ for all $x \in S$, and $H \subseteq B^4 = S^{4 \cdot 3^t}$. So, for any $x \in H$,

$$|xA \triangle A| \leqslant 4 \cdot 3^t \delta|G| \leqslant 4k^{1/p}\delta|G| = \varepsilon|G|. \qquad \square$$

Remark 8.3. We make some comments to follow up on Remark 1.7.

(1) Note that, in the proof of Theorem 1.6, if $K = \bigcap_{g \in G} gHg^{-1}$ then K is normal of index at most $[G{:}H]!$ and $K \subseteq \text{Stab}_\varepsilon (A)$. So, if $[G{:}H] \leqslant O_{r,d,\nu}((1/\varepsilon)^{d+\nu})$, for some chosen $\varepsilon, \nu > 0$, then $\log[G{:}K] \leqslant O_{r,d,\nu}(\varepsilon^{-(d+\nu)}\log(\varepsilon^{-1}))$. Altogether, we have a statement identical to Theorem 1.6, but with a *normal* subgroup of index $2^{O_{r,d,\nu}((1/\varepsilon)^{d+\nu})}$. One reason a normal subgroup is desirable in this situation is that it implies a very strong graph regularity conclusion for the bipartite graph $xy \in A$ on G, in which the pieces of the regular partition are the cosets of H (see [10, Corollary 3.3]).

(2) A non-effective version of Theorem 1.6, with a normal subgroup, can also be proved by applying Corollary 1.4(b) directly to $\text{Stab}_\varepsilon (A)$. Together with Haussler's Packing Lemma, this would directly yield a normal subgroup of index $O_{r,d,\varepsilon}(1)$ contained in $\text{Stab}_\varepsilon (A)$. It is interesting to note that a qualitative version of Theorem 1.6, with a normal subgroup, was already shown in [10] using fairly different techniques (although there are some aspects of the work in [10] which are not recovered here, including definability of the subgroup H and a stronger regularity statement).

Finally, we prove similar results about purely non-abelian finite groups and finite simple groups. To motivate our interest in this setting, we recall some of the previous work on arithmetic regularity for subsets of finite groups satisfying extra tameness properties. One example of such a property is bounded VC-dimension, which we have already discussed. Another important example is that of a *d-stable* subset A of a group G, for some integer $d \geqslant 1$, which means there do not exist $a_1, \ldots, a_d, b_1, \ldots, b_d \in G$ such that $a_ib_j \in A$ if and only if $i \leqslant j$. Note that a d-stable set has VC-dimension at most $d-1$. Both of these properties were previously studied in the setting of Szemerédi regularity for graphs (see [2], [23]).

In [9] (joint with Pillay and Terry), we showed that, given $d \geqslant 1$ and $\varepsilon > 0$, if G is a finite group and $A \subseteq G$ is d-stable, then there is a normal subgroup $H \leqslant G$, of index $O_{d,\varepsilon}(1)$, and a union D of cosets of G such that $|A \triangle D| \leqslant \varepsilon|H|$. Informally, stable subsets of finite groups are structurally approximated by cosets of a bounded-index normal subgroup. In the setting of finite groups, this phenomenon was first investigated by Terry and Wolf [38], who proved a similar result for $G = \mathbb{F}_p^n$ with strong quantitative bounds, but with the approximation $|A \triangle D| \leqslant \varepsilon|G|$. (This was recently generalized to arbitrary finite abelian groups in [39].)

In contrast, easy examples show that subgroups are not sufficient to control sets of bounded VC-dimension. For example, as noted in [3], if $p \geqslant 3$ is prime and $G = \mathbb{Z}/p\mathbb{Z}$ and $A = \{1, \ldots, (p-1)/2\}$, then A has VC-dimension 2, but A cannot be approximated by cosets of a non-trivial subgroup of G. This is one reason to use Bohr neighbourhoods in the formulation of arithmetic regularity for sets of bounded VC-dimension, which was done by Sisask in the abelian setting [34], and independently in [10] for general finite groups. As we have seen above, if one introduces a uniform bound on the exponent of the groups, then subgroups can be used to approximate sets of bounded VC-dimension. So this motivates the following result that purely non-abelian groups (see Corollary 1.3) also exhibit this behaviour.

Theorem 8.4. *Fix a positive integer d. Suppose G is a purely non-abelian finite group, and $A \subseteq G$ has VC-dimension at most d. Then, for any $\varepsilon > 0$, there is a normal subgroup H of G, of index $O_{d,\varepsilon}(1)$, which satisfies the following properties.*

(i) *(Structure) There is a set $D \subseteq G$, which is a union of cosets of H, such that $|A \triangle D| \leqslant \varepsilon |G|$.*
(ii) *(Regularity) There is a set $Z \subseteq G$, with $|Z| < \frac{1}{2}\varepsilon^{1/2}|G|$, such that for any $x \in G \backslash Z$, either $|xH \cap A| \leqslant \varepsilon^{1/4}|H|$ or $|xH \backslash A| \leqslant \varepsilon^{1/4}|H|$.*

Proof. Fix a purely non-abelian finite group G, a subset $A \subseteq G$ of VC-dimension at most d, and $\varepsilon > 0$. As in Theorem 1.6, if $S = \mathrm{Stab}_{\varepsilon/4}(A)$ then $|S| \geqslant (\varepsilon/120)^d|G|$. By Corollary 1.3 there is a normal subgroup $H \leqslant G$, of index $O_{d,\varepsilon}(1)$, such that $H \subseteq S^4$. So $|xA \triangle A| \leqslant \varepsilon|G|$ for any $x \in H$. Now apply Lemma 8.2. $\square$

Remark 8.5. The previous theorem can also be deduced from [10, Theorem 5.7], yielding further information as discussed in Remark 8.3(2). On the other hand, the proof here seems more direct, and certainly uses a more acute application of VC-theory. (Both proofs involve identical uses of [1] and [28]).

The work in [9] on stable regularity implies that, for any $d \geqslant 1$ and $\varepsilon > 0$, if G is a finite simple group of size $\Omega_{d,\varepsilon}(1)$ and $A \subseteq G$ is d-stable, then $|A| \leqslant \varepsilon|G|$ or $|A| \geqslant (1-\varepsilon)|G|$.[5] For the abelian case (*i.e.* $G = \mathbb{Z}/p\mathbb{Z}$), a quantitative lower bound on $p = |G|$, in terms of d and ε, could be deduced from [39]. On the other hand, the example above, which shows that subgroups are not sufficient to approximate sets of bounded VC-dimension, takes place in *abelian* finite simple groups. This motivates the following corollary of Theorem 8.4.

Corollary 8.6. *For any integer d and any $\varepsilon > 0$, there is an integer $n = n(d, \varepsilon)$ such that, if G is a non-abelian finite simple group of size greater than n, and $A \subseteq G$ has VC-dimension at most d, then $|A| \leqslant \varepsilon|G|$ or $|A| \geqslant (1-\varepsilon)|G|$.*

Remark 8.7. Using a similar strategy, we can give a direct proof of the previous corollary, which yields $\log(n(d, \varepsilon)) \leqslant O((\varepsilon/90)^{-6d})$ as an explicit bound. Namely, by the work of Gowers [14] discussed in Remark 2.8, there is some $c > 0$ such that if G is a non-abelian finite simple group with $\log|G| \geqslant c(\varepsilon/90)^{-6d}$, and $S \subseteq G$ is such that $|S| \geqslant (\varepsilon/90)^{-d}|G|$, then $G = S^3$. So fix such a G, and suppose $A \subseteq G$ is of VC-dimension at most d. By Haussler's Packing Lemma, and choice of c, we have $G = (\mathrm{Stab}_{\varepsilon/3}(A))^3 = \mathrm{Stab}_\varepsilon(A)$. Now apply Lemma 8.2.[6]

---

[5]This is a finitary analogue of the older fact that any definable subset of an (infinite) definably-connected stable group has measure 0 or 1 with respect to the unique Keisler measure.
[6]As in Remark 2.8, the work in [8] implies $n(d, \varepsilon) \leqslant (\lceil(\varepsilon/90)^{-3d}\rceil + 1)!$ in Corollary 8.6.

9. Final remarks

9.1 Quantitative bounds

An obvious question at this point is on effective bounds for Theorems 1.2, 2.3 and 2.4. Our proof of Theorem 2.4 used an ultraproduct construction, and did not give explicit bounds of any kind. While ultraproducts do not appear explicitly in our proof of Theorem 2.3, they are used similarly in previous work on approximate groups (in both [7] and [19]).

It is sometimes possible, with enough work, to reverse-engineer effective bounds from arguments with ultraproducts, but these bounds are usually very bad (see *e.g.* [36, Chapter 7] for some discussion of this topic). Altogether, it seems that in order to obtain efficient bounds for the above results, one would need efficient bounds for results on approximate groups, or a different proof strategy altogether.

9.2 Small tripling versus approximate groups

For the sake of completeness, we note that weaker versions of our main results can be obtained without the revised Sanders–Croot–Sisask analysis in Section 4. This is because of the following result of Tao, which follows from the proof of [35, Theorem 3.9] (or see [7, Corollary 5.2]).

Theorem 9.1 ([35])**.** *Suppose A is a non-empty finite subset of a group G. If $|A^3| \leqslant k|A|$ then $\bar{A}^2$ is an $O(k^{O(1)})$-approximate group containing A.*

Together with Theorem 5.1, one obtains a weaker version of Theorem 2.3.

Corollary 9.2 ([7], [35])**.** *Fix positive integers k and r. Let G be a group of exponent r, and fix a finite subset $A \subseteq G$. Suppose $|A^3| \leqslant k|A|$. Then there is $H \leqslant \langle A \rangle$ such that $\bar{A}^2$ is covered by $O_{k,r}(1)$ left cosets of H and $H \leqslant \bar{A}^8$.*

Remark 9.3. Corollary 9.2 could be used instead of Corollary 1.4(a) in the proof of Theorem 1.6.

Remark 9.4. Recall that if $A \subseteq G$ is finite and non-empty, with $|AA^{-1}A| \leqslant k|A|$, then $|(A^{-1}A)^3| \leqslant k^{O(1)}|A^{-1}A|$ by Proposition 3.2(b), and so $(AA^{-1})^2$ is an $O(k^{O(1)})$-approximate group by Theorem 9.1. Altogether, this is essentially the 'discrete case' of [19, Corollary 3.11].

A weaker version of Theorem 2.4 can also be formulated using Theorem 9.1, but the proof would still require our work with saturated extensions and approximate Bohr neighbourhoods, and so we will not go into it any further. On the other hand, the following statement about sets of small tripling in arbitrary groups follows by combining Theorem 9.1 with the main structure theorems for approximate groups from Breuillard, Green and Tao [7].

Theorem 9.5 ([7, 35])**.** *Fix a positive integer k. Suppose G is a group and $A \subseteq G$ is finite and non-empty, with $|A^3| \leqslant k|A|$. Then there is a subgroup H of G and a finite normal subgroup N of H with the following properties:*

  (i) *A is covered by $O_k(1)$ left cosets of H,*
  (ii) *H/N is nilpotent and finitely generated of rank and step $O_k(1)$,*
 (iii) *$\bar{A}^8$ contains N and a generating set for H.*

Moreover, there is a coset nilprogression $P \subseteq \bar{A}^8$ of rank and step $O_k(1)$ such that A is covered by $O_k(1)$ left translates of P.

The final clause of the previous theorem is a non-abelian analogue of the *Bogolyubov–Ruzsa lemma* for finite abelian groups, which was stated after Theorem 1.1. However, we have the qualitative discrepancy between $\bar{A}^8$ in condition (iii) and $2A - 2A$ in the abelian case. Given our earlier results, one naturally wonders if $\bar{A}^8$ can be replaced by $(AA^{-1})^2 \cap A^2A^{-2} \cap (A^{-1}A)^2 \cap A^{-2}A^2$. This also raises a similar question about small alternation. So we observe that these issues can be addressed simply by combining the results in [7] with Lemma 4.1.

Theorem 9.6. *Fix a positive integer k. Suppose G is a group and $A \subseteq G$ is finite and non-empty. Furthermore,*

(a) *assume $|AA^{-1}A| \leqslant k|A|$ and set $V = AA^{-1}$ and $W = (AA^{-1})^2$, or*
(b) *assume $|A^3| \leqslant k|A|$ and set $V = \bar{A}$ and $W = (AA^{-1})^2 \cap A^2A^{-2} \cap (A^{-1}A)^2 \cap A^{-2}A^2$.*

Then there is a subgroup H of G and a finite normal subgroup N of H with the following properties:

(i) *for all $m \geqslant 1$, V^m is covered by $O_{k,m}(1)$ left cosets of H,*
(ii) *H/N is nilpotent and finitely generated of rank and step $O_k(1)$,*
(iii) *W contains N and a generating set for H.*

Moreover, there is a coset nilprogression $P \subseteq W$ of rank and step $O_k(1)$ such that, for all $m \geqslant 1$, V^m is covered by $O_{k,m}(1)$ left translates of P.

Proof. By Lemma 4.1, there is a symmetric set $Y \subseteq G$ such that $Y^8 \subseteq W$ and V^3 is covered by $O_k(1)$ left translates of Y. Since $Y \subseteq V^2$, it follows that for all $m \geqslant 1$, V^m is covered by $O_{k,m}(1)$ left translates of Y. Note also that Y is an $O_k(1)$-approximate group. So by [7, Theorem 1.6], there are $N \trianglelefteq H \leqslant G$ such that Y is covered by $O_k(1)$ left cosets of H, H/N is nilpotent and finitely generated of rank and step $O_k(1)$, and Y^4 contains N and a generating set for H. Moreover, by [7, Theorem 2.10], there is a coset nilprogression $P \subseteq Y^8$ of rank and step $O_k(1)$ such that $|Y| \leqslant O_k(|P|)$ and Y is covered by $O_k(1)$ left translates of P. Altogether, the result follows by choice of Y. $\square$

Acknowledgements

References

[1] Alekseev, M. A., Glebskiĭ, L. Y. and Gordon, E. I. (1999) On approximations of groups, group actions and Hopf algebras. *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI)* **256** 224–262.

[2] Alon, N., Fischer, E. and Newman, I. (2007) Efficient testing of bipartite graphs for forbidden induced subgraphs. *SIAM J. Comput.* **37** 959–976.

[3] Alon, N., Fox, J. and Zhao, Y. (2019) Efficient arithmetic regularity and removal lemmas for induced bipartite patterns. *Discrete Anal.*, Paper no. 3, 14 MR3943117.

[4] Babai, L., Goodman, A. J. and Pyber, L. (1997) Groups without faithful transitive permutation representations of small degree. *J. Algebra* **195** 1–29.

[5] Bogolioùboff, N. (1939) Sur quelques propriétés arithmétiques des presque-périodes. *Ann. Chaire Phys. Math. Kiev* **4** 185–205.

[6] Bourgain, J. (1999) On triples in arithmetic progression. *Geom. Funct. Anal.* **9** 968–984.

[7] Breuillard, E., Green, B. and Tao, T. (2012) The structure of approximate groups. *Publ. Math. Inst. Hautes Études Sci.* **116** 115–221.

[8] Collins, M. J. (2007) On Jordan's theorem for complex linear groups. *J. Group Theory* **10** 411–423.

[9] Conant, G., Pillay, A. and Terry, C. (2020) A group version of stable regularity. *Math. Proc. Cambridge Philos. Soc.*, (2) **168** 405–413. MR4064112.

[10] Conant, G., Pillay, A. and Terry, C. (2018) Structure and regularity for subsets of groups with finite VC-dimension. arXiv:1802.04246

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

[12] Freĭman, G. A. (1973) *Foundations of a Structural Theory of Set Addition*, Vol. 37 of Translations of Mathematical Monographs, AMS.

[13] Freiman, G. A. (1987) What is the structure of K if $K + K$ is small? In *Number Theory (New York, 1984–1985)*, Vol. 1240 of Lecture Notes in Mathematics, pp. 109–134, Springer.

[14] Gowers, W. T. (2008) Quasirandom groups. *Combin. Probab. Comput.* **17** 363–387.

[15] Green, B. (2005) A Szemerédi-type regularity lemma in abelian groups, with applications. *Geom. Funct. Anal.* **15** 340–376.

[16] Green, B. and Ruzsa, I. Z. (2007) Freiman's theorem in an arbitrary abelian group. *J. London Math. Soc.* (2) **75** 163–175.

[17] Haussler, D. (1995) Sphere packing numbers for subsets of the Boolean n-cube with bounded Vapnik–Chervonenkis dimension. *J. Combin. Theory Ser. A* **69** 217–232.

[18] Helfgott, H. A. (2008) Growth and generation in $SL_2(\mathbb{Z}/p\mathbb{Z})$. *Ann. of Math.* (2) **167** 601–623.

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

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

[21] Keisler, H. J. (1964) Ultraproducts and saturated models. *Nederl. Akad. Wetensch. Proc. Ser. A* **67** = *Indag. Math.* **26** 178–186.

[22] Krupiński, K. and Pillay, A. (2016) Amenability, definable groups, and automorphism groups. *Adv. Math.* **345** 1253–1299. MR3904280.

[23] Malliaris, M. and Shelah, S. (2014) Regularity lemmas for stable graphs. *Trans. Amer. Math. Soc.* **366** 1551–1585.

[24] Massicot, J.-C. and Wagner, F. O. (2015) Approximate subgroups. *J. École Polytech. Math.* **2** 55–64.

[25] Nikolov, N. and Pyber, L. (2011) Product decompositions of quasirandom groups and a Jordan type theorem. *J. Eur. Math. Soc.* **13** 1063–1077.

[26] Nikolov, N., Schneider, J. and Thom, A. (2018) Some remarks on finitarily approximable groups. *J. École Polytech. Math.* **5** 239–258.

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

[28] Pillay, A. (2017) Remarks on compactifications of pseudofinite groups. *Fund. Math.* **236** 193–200.

[29] Plünnecke, H. (1969) *Eigenschaften und Abschätzungen von Wirkungsfunktionen*, BMwF-GMD-22, Gesellschaft für Mathematik und Datenverarbeitung.

[30] Ruzsa, I. Z. (1994) Generalized arithmetical progressions and sumsets. *Acta Math. Hungar.* **65** 379–388.

[31] Ruzsa, I. Z. (1996) Sums of finite sets. In *Number Theory (New York, 1991–1995)*, pp. 281–293, Springer.

[32] Sanders, T. (2010) On a nonabelian Balog–Szemerédi-type lemma. *J. Aust. Math. Soc.* **89** 127–132.

[33] Sanders, T. (2012) On the Bogolyubov–Ruzsa lemma. *Anal. PDE* **5** 627–655.

[34] Sisask, O. (2018) Convolutions of sets with bounded VC-dimension are uniformly continuous. arXiv:1802.02836

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

[36] Tao, T. (2014) *Hilbert's Fifth Problem and Related Topics*, Vol. 153 of Graduate Studies in Mathematics, AMS.

[37] Tao, T. and Vu, V. (2006) *Additive Combinatorics*, Vol. 105 of Cambridge Studies in Advanced Mathematics, Cambridge University Press.

[38] Terry, C. and Wolf, J. (2019) Stable arithmetic regularity in the finite field model. *Bull. Lond. Math. Soc.* **51** 70–88.

[39] Terry, C. and Wolf, J. (2018) Quantitative structure of stable sets in finite abelian groups. *Trans. Amer. Math. Soc.* accepted.