SOLVING $xz = y^2$ IN CERTAIN SUBSETS OF FINITE GROUPS

TOM SANDERS

Abstract. Suppose that G is a finite group and $A \subset G$ has no non-trivial solutions to $xz = y^2$. We shall show that $|A| = |G|/(\log\log|G|)^{\Omega(1)}$.

1. Introduction

Suppose that G is a group and $A \subset G$. We say that A contains a solution to $xz = y^2$ if there is a triple $(x, y, z) \in A^3$ such that $xz = y^2$; we say the solution is non-trivial if $x \neq y$. This paper is concerned with the following result of Bergelson, McCutcheon and Zhang.

Theorem 1.1 ([BMZ97, Corollary 6.5])**.** *Suppose that G is a finite group and $A \subset G$ contains no non-trivial solutions to $xz = y^2$. Then $|A| = o(|G|)$.*

Note that if G is Abelian then $xz = y^2$ and $x \neq y$ if and only if $(x, y, z) = (x, x+d, x+2d)$ for some $d \neq 0$ *i.e.* if any only if (x, y, z) is a non-trivial three-term arithmetic progression. With this observation in hand it is a short argument to get Roth's celebrated theorem on three-term arithmetic progressions [Rot52, Rot53] from Theorem 1.1 in the special case when G is a cyclic group. This should should give some idea of why Theorem 1.1 might be of interest.

Theorem 1.1 was proved in the case when G is Abelian (and of odd order) in [FGR87, Theorem 1], in which paper the authors attribute the result to [BB82]. The odd order condition here is a technical convenience which can be ignored on a first reading. Frankl, Graham, and Rödl's proof goes by the triangle removal lemma of Ruzsa and Szemerédi [RS78], and it was noted by Král, Serra, and Vena, in the course of a wider generalisation, that the removal lemma can also be used to prove Theorem 1.1 [KSV09, Corollary 3].

The removal lemma suffers from notorious poor dependencies and the reader is directed to [CF12] for a discussion of this and related matters. For our purposes it suffices to know that the best are due to Fox [Fox11] and inserting his work into the proof of [KSV09, Corollary 3] would give

$$|A| = \frac{|G|}{\exp(\Omega(\log_* |G|))} \tag{1.1}$$

when A is a set satisfying the hypotheses of Theorem 1.1 (and G is a group of odd order). Here for $R \in \mathbb{N}$ we define $\log_* R$ to be the minimal $n \in \mathbb{N}_0$ such that

$$\overbrace{\log_2(\log_2 \dots (\log_2}^{n \text{ times}} R)) \leqslant 1.$$

This function grows more slowly than any finite composition of logarithms. It is the purpose of this paper to improve on the bound (1.1) by showing the following.

Theorem 1.2. *Suppose that G is a finite group and $A \subset G$ contains no non-trivial solutions to $xz = y^2$. Then*

$$|A| = \frac{|G|}{(\log\log|G|)^{\Omega(1)}}.$$

For various classes of Abelian groups there has been considerable work improving the bounds in Theorem 1.2. The best known arguments are due to Bloom [Blo14], although he concentrates on the case of initial segments of the integers, he uses the framework of Bourgain [Bou99] and so the argument easily extends to finite Abelian groups to show the following.

Theorem 1.3. *Suppose that G is a finite Abelian group and $A \subset G$ contains no non-trivial solutions to $x + z = 2y$. Then*

$$|A| = \frac{|G|}{\log^{1-o(1)}|G|}.$$

A small amount of additional care is needed to avoid restricting attention to groups of odd order, but it will be useful to avoid this restriction in the remaining discussion.

It is worth noting that when G is non-Abelian there are two possible notions of three-term arithmetic progression. The first is triples (x, y, z) such that $xz = y^2$; the second is triples (x, y, z) such that $z = yx^{-1}y$. This second notion is, perhaps, more natural since it is left (and so by symmetry right) invariant meaning that if (x, y, z) has $z = yx^{-1}y$ then (ax, ay, az) has $az = ay(ax)^{-1}ay$. The first notion is not, in general, left or right invariant. (This does not, however, lead to the problems that non-translation invariant equations have in Abelian groups. For example, if $G = \mathbb{Z}/2N\mathbb{Z}$ then the odd numbers form a subset of G of density $\frac{1}{2}$ not containing any solutions to $x + y = z$.) Nevertheless our proof is iterative and the lack of translation invariance does present issues. These are discussed in more detail at the start of §4.

In the same way as before we say that a solution to $z = yx^{-1}y$ is non-trivial if $x \neq y$. The next result is a slight variant of [Sol13, Theorem 2.5] also communicated to the author personally by Ernie Croot, and is included for comparison with Theorem 1.2.

Theorem 1.4. *Suppose that G is a finite group and $A \subset G$ contains no non-trivial solutions to $z = yx^{-1}y$. Then*

$$|A| = \frac{|G|}{\log^{\frac{1}{2}-o(1)}|G|}.$$

Proof. By a result of Pyber [Pyb97] every finite group G contains an Abelian subgroup H of size $\exp(\Omega(\sqrt{\log|G|}))$. (This is essentially best possible.) Averaging, there is some $t \in G$ such that

$$|tH \cap A| \geqslant \mathbb{E}_{t\in G}|tH \cap A| = \sum_{h\in H}\mathbb{E}_{t\in G}1_A(th) = \sum_{h\in H}\frac{|A|}{|G|} = \frac{|A||H|}{|G|}.$$

Since solutions to $z = yx^{-1}y$ are left invariant we conclude that $H \cap t^{-1}A$ contains no non-trivial solutions to $z = yx^{-1}y$. However, this set is a subset of H which is an Abelian group and so we can apply Bloom's result (Theorem 1.3) to see that

$$\frac{|A||H|}{|G|} \leqslant |tH \cap A| = |H \cap t^{-1}A| = \frac{|H|}{\log^{1-o(1)}|H|} = \frac{|H|}{\log^{\frac{1}{2}-o(1)}|G|},$$

which can be rearranged to give the claimed bound. □

We shall prove Theorem 1.2 by proving the following.

Theorem 1.5. *Suppose that G is a finite group, and $A \subset G$ has size $\alpha|G|$ and distinct squares* i.e. *$a^2 \neq b^2$ if $a, b \in A$ are distinct. Then the number of triples $(x, y, z) \in A^3$ such that $xz = y^2$ is $\exp(-\exp(\alpha^{-O(1)}))|G|^2$.*

Theorem 1.2 follows immediately from this.

Proof of Theorem 1.2. If $a, b \in A$ have $a^2 = b^2$ and $a \neq b$, then (a, b, a) is a triple with $aa = b^2$ and $a \neq b$ and we are done. It follows that we may assume A has distinct squares and so we have a lower bound on the number of triples $(x, y, z) \in A^3$ such that $xz = y^2$. By hypothesis we know that in this case $x = y$ whence $x = z$ and so the number of such triples is $|A|$ and hence

$$\alpha|G| = |A| \geqslant \exp(-\exp(\alpha^{-O(1)}))|G|^2$$

which can be rearranged to give the result. □

The remainder of the paper is concerned with proving Theorem 1.5.

2. Notation

Given a finite set Z we write $M(Z)$ for the set of complex-valued measures on Z and put

$$\|\mu\| := \int d|\mu| \text{ for all } \mu \in M(Z).$$

Suppose that μ is a non-negative measure supported on Z. We write $L_p(\mu)$ for the space of functions $f : Z \to \mathbb{C}$ endowed with the (semi-)norm

$$\|f\|_{L_p(\mu)} := \left(\int |f(z)|^p d\mu(z)\right)^{1/p},$$

with the usual convention when $p = \infty$. Of course $L_2(\mu)$ is a Hilbert space we denote the inner product inducing the norm $\|\cdot\|_{L_2(\mu)}$ by

$$\langle f, g\rangle_{L_2(\mu)} = \int f(x)\overline{g(x)}d\mu(x) \text{ for all } f, g \in L_2(\mu).$$

We shall often take μ to be the uniform probability measure supported on Z which we denote μ_Z.

Throughout the paper we work with sets in some finite group G. The group structure will be encoded by the left and right regular representations defined on functions $f : G \to \mathbb{C}$ by

$$\lambda_x(f)(y) := f(x^{-1}y) \text{ for all } y \in G;$$

and

$$\rho_x(f)(y) := f(yx) \text{ for all } y \in G.$$

This extends to (complex-valued) measures, μ on G, where for each $x \in G$ we write $\rho_x(\mu)$ for the measure induced by

$$\int f d\rho_x(\mu) = \int \rho_{x^{-1}}(f) d\mu \text{ for all } f : G \to \mathbb{C},$$

and similarly for $\lambda_x(\mu)$. Thus,

$$\rho_x(\mu)(A) = \mu(Ax) \text{ and } \lambda_x(\mu)(A) = \mu(xA) \text{ for all } x \in G, A \subset G.$$

Given a (complex-valued) measure μ on G, and a function $f : G \to \mathbb{C}$ we define

$$\langle f, \mu\rangle := \int f d\overline{\mu} \text{ and } \langle \mu, f\rangle := \int \overline{f} d\mu;$$

similarly define the **convolution** of f and μ to be

$$f * \mu(x) := \int f(xy^{-1}) d\mu(y) \text{ and } \mu * f(x) = \int f(y^{-1}x) d\mu(y) \text{ for all } x \in G.$$

It is worth noting that if $A \subset G$ is non-empty then

$$f * \mu_A(x) = \mathbb{E}_{a\in A} f(xa^{-1}) = \mathbb{E}_{z\in xA^{-1}} f(z);$$

i.e. $f * \mu_A(x)$ is the average value of f on xA^{-1}. Similarly, $\mu_A * f(x)$ is the average value of f on $A^{-1}x$.

We shall also look to convolve two measures: suppose that μ and ν are such. Then we define their convolution to be the measure induced by

$$\int f(z) d(\mu * \nu)(z) = \int f(xy) d\mu(x) d\mu(y) \text{ for all } f : G \to \mathbb{C}.$$

Here it is worth noting that if $A, A' \subset G$ are non-empty then

$$\operatorname{supp} \mu_A * \mu_{A'} = AA' := \{aa' : a \in A, a' \in A'\}.$$

Finally, for $f : G \to \mathbb{C}$ define the

$$\tilde{f}(x) := \overline{f(x^{-1})} \text{ for all } x \in G,$$

and similarly for measures. Note that if $A \subset G$ is non-empty then $\widetilde{\mu_A} = \mu_{A^{-1}}$ where $A^{-1} := \{a^{-1} : a \in A\}$.

We introduced the last piece of notation because it captures the adjoint operation. In particular, the adjoint of $\mu \mapsto f * \mu$ is $\nu \mapsto \tilde{f} * \nu$ *i.e.*

$$\langle f * \mu, \nu\rangle = \langle \mu, \tilde{f} * \nu\rangle \text{ for all measures } \mu, \nu,$$

and, again, similarly for convolution with measures. One can also use this notation to capture convolution:

$$f * \mu(x) = \langle \rho_x(\mu), \tilde{f} \rangle \text{ and } \mu * f(x) = \langle \rho_x(f), \tilde{\mu} \rangle \text{ for all } x \in G,$$

and, similarly for λ.

3. Multiplicative systems

Since the work of Roth [Rot52, Rot53] the standard approach to problems of the type considered in this paper has been inductive. As often happens with inductive arguments they become possible when we enlarge the class we are working over. In this case we shall prove our result not just for large subsets of groups, but for large subsets of certain group-like objects which we shall call multiplicative systems.

There is nothing particularly novel about the multiplicative systems presented in this paper and there are numerous essentially equivalent definitions extracting the key properties of a group which we require. In the Abelian setting this has been explored extensively since the pioneering work of Bourgain [Bou99]. It may be worth noting that Gowers and Wolf use some nice notation in [GW11] and Bloom [Blo14] takes as basic a structure which is pretty close to ours.

Two of the axioms a group satisfies are particularly easy to guarantee for subsets of a group: we say that a set $A \subset G$ is a **symmetric neighbourhood of the identity** if it contains the identity and is closed under taking inverses *i.e.* $1_G \in A$ and $x \in A$ implies $x^{-1} \in A$. What is harder to capture is closure under multiplication and, indeed, we have to make do with a sort of 'approximate' closure. Given $r \in \mathbb{N}$ and $\epsilon \in [0, 1]$ we say that

$$\mathcal{B} = (B_{0+}, B_0, B_{0-}; B_{1+}, B_1, B_{1-}; \dots; B_{r+}, B_r, B_{r-}; B_{r+1})$$

is an $(r+1)$**-step ϵ-closed multiplicative system** if

(i) *(Symmetric neighbourhoods)* B_{r+1}, and B_{i+}, B_i, and B_{i-}, are symmetric neighbourhoods of the identity for all $0 \leqslant i \leqslant r$;

(ii) *(Nesting)* we have

$$B_{0+} \supset B_0 \supset B_{0-} \supset B_{1+} \supset B_1 \supset B_{1-} \supset \cdots \supset B_{r+} \supset B_r \supset B_{r-} \supset B_{r+1};$$

(iii) *(Closure)* we have

$$B_{i-} \subset yB_ix \subset B_{i+} \text{ for all } x, y \in B_{i+1} \text{ for all } 0 \leqslant i \leqslant r,$$

and the estimates

$$\frac{1}{1+\epsilon}|B_{i+}| \leqslant |B_i| \leqslant (1+\epsilon)|B_{i-}| \text{ for all } 0 \leqslant i \leqslant r.$$

Note here that since the B_is and $B_{i\pm}$s contain the identity, a number of the inclusions in nesting follow from closure.

The idea here is that an r-step system in enough to multiply group elements together about 'r times'. The parameter ϵ captures the error each time we do this. While we have set these systems up quite generally we shall only make use of systems with r small, typically 1 or 2.

When we need more than one multiplicative system they will be denoted $\mathcal{B}'$, $\mathcal{B}''$ *etc.* with the obvious convention that

$$\mathcal{B}' = (B'_{0+}, B'_0, B'_{0-}; B'_{1+}, B'_1, B'_{1-}; \dots; B'_{r+}, B'_r, B'_{r-}; B'_{r+1}).$$

This is the reason that we have not chosen the easier-to-read notation B_i^+ and B_i^- for B_{i+} and B_{i-}.

The model we have in mind, and perhaps the simplest example, is given by groups.

Example 3.1 (Groups)**.** Suppose that $H_{r+1} \leqslant H_r \leqslant \dots \leqslant H_1 \leqslant H_0 \leqslant G$. Then

$$(H_0, H_0, H_0; H_1, H_1, H_1; \dots; H_r, H_r, H_r; H_{r+1})$$

is an $(r+1)$-step 0-closed multiplicative system.

This example, and in fact the special case when $H_0 = H_{r+1}$ is a very useful example to have in mind on a first reading of many of the results below.

In a 1-step multiplicative system $\mathcal{B}$ we think of B_1 as 'acting on' B_0, and there are many examples of multiplicative systems resulting from trivial action sets.

Example 3.2 (Trivial action set)**.** For any symmetric neighbourhood of the identity A, we have that $(A, A, A; \{1_G\})$ is a 1-step 0-closed system.

The point of this example is just to emphasise that we shall be interested in the case when B_{i+1} is not too much smaller than B_i.

To get a sense of how the various parameters behave it is useful to record some basic properties of multiplicative systems.

Lemma 3.3 (Basic properties of multiplicative systems)**.** *Suppose that $\mathcal{B}$ and $\mathcal{B}'$ are $(r+1)$-step (resp. $(r'+1)$-step) ϵ-closed multiplicative systems. Then*

(i) (Monotonicity) *$\mathcal{B}$ is an $(r+1)$-step ϵ''-closed multiplicative system for all $\epsilon'' \geqslant \epsilon$;*

(ii) (Truncation) *for $0 \leqslant l \leqslant m \leqslant r$ and any symmetric neighbourhood of the identity $B_* \subset B_{m+1}$,*

$$(B_{l+}, B_l, B_{l-}; \dots; B_{m+}, B_m, B_{m-}; B_*)$$

is an $(m-l+1)$-step ϵ-closed multiplicative system;

(iii) (Gluing) *if $B'_{0+} \subset B_{r+1}$ then*

$$(B_{0+}, B_0, B_{0-}; \dots; B_{r+}, B_r, B_{r-}; B'_{0+}, B'_0, B'_{0-}; \dots; B'_{r'+}, B'_{r'}, B'_{r'-}; B_{(r'+1)})$$

is an $(r+r'+1)$-step ϵ-closed multiplicative system.

At this stage we have not given any examples of multiplicative systems that are not either trivial or endowed with the far stronger structure of being a nested sequence of subgroups. In Abelian groups we have a rich set of examples provided by Bohr sets.

Example 3.4 (Bohr sets)**.** Throughout this example suppose that G is an Abelian group and use additive rather than multiplicative notation for the group operation. Suppose that Γ is a set of d characters on G, and $\delta \in (0, 2]$. We define the **Bohr set with frequency set Γ and width δ** to be the set

$$\mathrm{Bohr}(\Gamma, \delta) := \{x \in G : |\gamma(x) - 1| \leqslant \delta \text{ for all } \gamma \in \Gamma\}.$$

Some fairly straight-forward arguments which can be found in *e.g.* [TV06, Lemma 4.19] show that for $l \in \mathbb{N}$ we have

$$|\operatorname{Bohr}(\Gamma, \delta)| \leqslant l^{O(d)} |\operatorname{Bohr}(\Gamma, \delta/l)|$$

and an easy application of the triangle inequality shows us that

$$l \operatorname{Bohr}(\Gamma, \delta/l) := \overbrace{\operatorname{Bohr}(\Gamma, \delta/l) + \cdots + \operatorname{Bohr}(\Gamma, \delta/l)}^{l \text{ times.}} \subset \operatorname{Bohr}(\Gamma, \delta).$$

The ability to *dilate* Bohr sets lets us use them to produce multiplicative systems. In particular, we have

$$l \operatorname{Bohr}(\Gamma, \delta/l) + \operatorname{Bohr}(\Gamma, \delta) \subset \operatorname{Bohr}(\Gamma, 2\delta)$$

and so by the pigeonhole principle there is some $j < l$ such that

$$|\operatorname{Bohr}(\Gamma, \delta/l) + (j \operatorname{Bohr}(\Gamma, \delta/l) + \operatorname{Bohr}(\Gamma, \delta))| \leqslant \exp(O(d/l)) |j \operatorname{Bohr}(\Gamma, \delta/l) + \operatorname{Bohr}(\Gamma, \delta)|$$

For any $\epsilon \in (0, 1]$ (the closure parameter of the system) we can pick $l = O(d\epsilon^{-1})$ such that $\exp(O(d/l)) \leqslant 1 + \epsilon$, and so we have two sets

$$B_0 := j \operatorname{Bohr}(\Gamma, \delta/l) + \operatorname{Bohr}(\Gamma, \delta) \text{ and } B_1 := \operatorname{Bohr}(\Gamma, \delta/l)$$

such that $|B_1 + B_0| \leqslant (1 + \epsilon)|B_0|$. It is a short step from here to defining a 2-step ϵ-closed multiplicative system.[1] Crucially this multiplicative system satisfies

$$|B_1| = \Omega(1/l)^{O(d)} |B_0|.$$

In fact in the case of Bohr sets a structure somewhat stronger than a multiplicative system can be constructed: we can arrange for a nested sequence of so-called regular Bohr sets. This was originally done in [Bou99]; an exposition may be found around [TV06, Definition 4.23].

There are natural analogues of Bohr sets in general finite groups, but they tend to describe only normal subsets of a group and that is not rich enough for our purposes. We shall take a different approach to find a supply of multiplicative systems motivated by a result of Bogolioúboff [Bog39].

Bogolioúboff showed that if A is a symmetric subset of an Abelian group of density α then $4A := A+A+A+A$ contains a large Bohr set. Turning this around, in a general group we shall look for our multiplicative systems inside four-fold product set of large subsets of G. Bogolioúboff's lemma was improved by Chang in [Cha02], and recently Croot and Sisask discovered a generalisation of Chang's argument to non-Abelian groups.

The following result is essentially [CS10, Corollary 1.4] extended to functions. We only need the version for sets here as it happens, but the proof for functions is no harder. (We shall have to examine this later in Lemma 3.14 when we establish a version of the Croot-Sisask Lemma for multiplicative systems.)

[1] We do not do it because, while B_{0+} can just be defined to be $B_0 + B_1$, there is a small technical obstacle to defining B_{0-}. This is easy to resolve but detracts from the example.

Lemma 3.5 (The Croot-Sisask Lemma)**.** *Suppose that $f \in L_p(\mu_G)$ for some $p \in [2,\infty)$, $X \subset G$ has density $\delta := \mu_G(X) > 0$, and $\eta \in (0,1]$ is a parameter. Then the set of x such that*

$$\|\rho_{x^{-1}}(f * \mu_X) - f * \mu_X\|_{L_p(\mu_G)} \leqslant \eta \|f\|_{L_p(\mu_G)}$$

is a symmetric neighbourhood of the identity and has density $\exp(-O(\eta^{-2}p\log 2\delta^{-1}))$.

We shall now use this to establish a Bogolioúboff-type lemma in general groups. Before diving in, we should say that this result is just a variant of [CS10, Theorem 1.6].

Lemma 3.6. *Suppose that $X \subset G$ is a symmetric neighbourhood of the identity of density $\delta := \mu_G(X) > 0$, and $k \in \mathbb{N}$ is a parameter. Then there is a symmetric neighbourhood of the identity S such that*

$$S^k \subset X^4 \text{ and } \mu_G(S) \geqslant \exp(-O(k^2 \log^2 2\delta^{-1})).$$

Proof. Let $p \geqslant 2$ and $\eta \in (0,1]$ be parameters to be chosen later and apply Lemma 3.5 with $f = 1_{X^2}$ to get a symmetric neighbourhood of the identity S such that

$$\|\rho_{x^{-1}}(1_{X^2} * \mu_X) - 1_{X^2} * \mu_X\|_{L_p(\mu_G)} \leqslant \eta \|1_{X^2}\|_{L_p(\mu_G)}.$$

By the triangle inequality we have that

$$\|\rho_{x^{-1}}(1_{X^2} * \mu_X) - 1_{X^2} * \mu_X\|_{L_p(\mu_G)} \leqslant k\eta \|1_{X^2}\|_{L_p(\mu_G)},$$

for all $x \in S^k$, and so by Hölder's inequality we see that

$$|\langle \rho_{x^{-1}}(1_{X^2} * \mu_X), 1_X\rangle_{L_2(\mu_G)} - \langle 1_{X^2} * \mu_X, 1_X\rangle_{L_2(\mu_G)}| \leqslant k\eta |X|^{1-1/p}|X^2|^{1/p} \leqslant k\eta\delta^{-1/p}|X|$$

since $X^2 \subset G$. On the other hand

$$\langle 1_{X^2} * \mu_X, 1_X\rangle_{L_2(\mu_G)} = |X|,$$

so we can take $p = O(\log 2\delta^{-1})$ and $\eta = \Omega(1/k)$ such that

$$\langle \rho_{x^{-1}}(1_{X^2} * \mu_X), 1_X\rangle_{L_2(\mu_G)} > |X|/2 \text{ for all } x \in S^k.$$

But then the result follows since

$$\langle \rho_{x^{-1}}(1_{X^2} * \mu_X), 1_X\rangle_{L_2(\mu_G)} = 1_X * 1_{X^2} * \mu_X(x^{-1}),$$

and $\operatorname{supp} 1_X * 1_{X^2} * \mu_X \subset X^4$. □

If G were Abelian, and X a Bohr set then it would be possible to take S with

$$|S| \geqslant \exp(-O((\log 2k)(\log 2\delta^{-1}))),$$

and so the δ-dependence in the above is not too bad even though the k dependence is rather poor. Fortunately in our applications k will tend to be fixed, while δ will decrease.

Finally we can use this lemma to produce some multiplicative systems. The argument is essentially the same pigeon-hole as we did with Bohr sets in Example 3.4, replacing the nice properties of dilates of Bohr sets by applications of Lemma 3.6.

Corollary 3.7. *Suppose that X is a symmetric neighbourhood of the identity in G with $\delta := \mu_G(X) > 0$, and $r \in \mathbb{N}_0$ and $\epsilon \in (0, 1]$ are parameters. Then there is an $(r+1)$-step ϵ-closed multiplicative system $\mathcal{B}$ and some symmetric neighbourhood of the identity S such that*

$$B_{0+} \subset X^4, S^4 \subset B_{r+1} \text{ and } \mu_G(S) \geqslant \exp(-O((\epsilon^{-2}\log 2\delta^{-1})^{4^{r+1}})).$$

Proof. First apply Lemma 3.6 to get a symmetric neighbourhood of the identity S_0 such that

$$S_0^9 \subset X^4 \text{ and } \mu_G(S_0) \geqslant \exp(-O(\log^2 2\delta^{-1})).$$

We shall now proceed inductively to construct sequences $((B_{l+}, B_l, B_{l-}))_{l=0}^{r}$ and $(S_l)_{l=0}^{r+1}$ with

$$S_{i+1}^9 \subset B_{i-}, \subset B_{i+} \subset S_i^9 \text{ and } B_{i-} \subset xB_iy \subset B_{i+} \text{ for all } x, y \in S_{i+1}^9,$$

and

$$\frac{1}{1+\epsilon}\mu_G(B_{i+}) \leqslant \mu_G(B_i) \leqslant (1+\epsilon)\mu_G(B_{i-}),$$

and, writing $\delta_i := \mu_G(S_i)$, such that

$$\delta_{i+1} \geqslant \exp(-O(\epsilon^{-2}\log^4 2\delta_i^{-1})).$$

Suppose we have constructed S_i, but not B_{i+}, B_i, B_{i-} or S_{i+1}. Apply Lemma 3.6 to the set S_i with a natural l_i to be chosen later to get a symmetric neighbourhood of the identity S_{i+1} such that

$$S_{i+1}^{l_i} \subset S_i^4 \text{ and } \delta_{i+1} \geqslant \exp(-O(l_i^2 \log^2 2\delta_i^{-1})).$$

Now

$$\mu_G(S_{i+1}^{l_i} S_i S_{i+1}^{l_i}) \leqslant 1 = \delta_i^{-1}\mu_G(S_i),$$

and so, by telescoping products,

$$\prod_{j=1}^{\lfloor l_i/18 \rfloor - 1} \frac{\mu_G(S_{i+1}^{18(j+1)} S_i S_{i+1}^{18(j+1)})}{\mu_G(S_{i+1}^{18j} S_i S_{i+1}^{18j})} \leqslant \delta_i^{-1}.$$

Thus we can take $l_i = O(\epsilon^{-1}\log 2\delta_i^{-1})$ so that there is some $0 < j < \lfloor l_i/18 \rfloor$ with

$$\mu_G(S_{i+1}^{18}(S_{i+1}^{18j} S_i S_{i+1}^{18j}) S_{i+1}^{18}) \leqslant (1+\epsilon)\mu_G(S_{i+1}^{18j} S_i S_{i+1}^{18j}).$$

We put

$$B_{i-} := S_{i+1}^{18j} S_i S_{i+1}^{18j}, B_i := S_{i+1}^9 B_{i-} S_{i+1}^9, \text{ and } B_{i+} := S_{i+1}^9 B_i S_{i+1}^9$$

and note that B_{i+}, B_i and B_{i-} are all symmetric neighbourhoods of the identity and

$$B_{i-} \subset xB_iy \subset B_{i+} \text{ for all } x, y \in S_{i+1}^9.$$

Furthermore we have $\mu_G(B_{i+}) \leqslant (1+\epsilon)\mu_G(B_{i-})$ and so

$$\frac{1}{1+\epsilon}\mu_G(B_{i+}) \leqslant \mu_G(B_i) \leqslant (1+\epsilon)\mu_G(B_{i-}).$$

Finally

$$S_{i+1}^9 \subset B_{i-} \text{ and } B_{i+} \subset S_{i+1}^{18j+18} S_i S_{i+1}^{18j+18} \subset S_{i+1}^{l_i} S_i S_{i+1}^{l_i} \subset S_i^9,$$

and

$$\delta_{i+1} \geqslant \exp(-O(\epsilon^{-2}\log^4 2\delta_i^{-1})).$$

We complete the construction by putting $S := S_{r+1}$ and $B_{r+1} := S^4$, and initialise the construction with S_0 as described above. The result follows on working out the bounds for δ_{r+1}. □

It is possible to relate X to a much more rigid structure called a coset nilprogression. These are somewhat complicated to define and the interested reader is directed to the paper [BGT10] Breuillard, Green and Tao, where this and related matters are addressed although at the expense of bounds.

At this stage we have established the results we shall need for the generation of suitable multiplicative systems and we can turn to tools for using them. First note the simple observation that conjugation preserves multiplicative systems.

Lemma 3.8 (Conjugation of multiplicative systems)**.** *Suppose that $\mathcal{B}$ is an $(r+1)$-step ϵ-closed multiplicative system and $g \in G$. Then*

$$(gB_{0+}g^{-1}, gB_0g^{-1}, gB_{0-}g^{-1}; \dots; gB_{r+}g^{-1}, gB_rg^{-1}, gB_{r-}g^{-1}; gB_{r+1}g^{-1})$$

is also an $(r+1)$-step ϵ-closed multiplicative system.

As will be clear from what we have already written, we shall find our multiplicative systems inside four-fold product sets. The following lemma will help us combine this with conjugation.

Lemma 3.9. *Suppose that $S \subset G$ is a symmetric set with density $\sigma := \mu_G(S) > 0$ and $g, h \in G$. Then there is a symmetric neighbourhood of the identity, X, such that*

$$X^4 \subset (gS^4g^{-1}) \cap (hS^4h^{-1}) \text{ and } \mu_G(X) \geqslant \exp(-O(\log^2 2\sigma^{-1})).$$

Proof. We apply Lemma 3.6 to get a symmetric neighbourhood of the identity R, such that $R^8 \subset S^4$, and $\mu_G(R) \geqslant \exp(-O(\log^2 2\sigma^{-1}))$. Now note that

$$\begin{aligned}
\mu_G((gR^2g^{-1}) \cap (hR^2h^{-1}))\mu_G(R)^2 &\geqslant \langle 1_{gR} * \widetilde{1_{gR}}, 1_{hR} * \widetilde{1_{hR}}\rangle_{L_2(\mu_G)} \\
&= \|\widetilde{1_{hR}} * 1_{gR}\|^2_{L_2(\mu_G)} \\
&\geqslant \|\widetilde{1_{hR}} * 1_{gR}\|^2_{L_1(\mu_G)} \\
&= (\mu_G(hR)\mu_G(gR))^2 = \mu_G(R)^4,
\end{aligned}$$

and so

$$\mu_G((gR^2g^{-1}) \cap (hR^2h^{-1})) \geqslant \mu_G(R)^2 = \exp(-O(\log^2 2\sigma^{-1})).$$

On the other hand

$$(gS^4g^{-1}) \cap (hS^4h^{-1}) \supset (gR^8g^{-1}) \cap (hR^8h^{-1}) \supset (gR^2g^{-1} \cap hR^2h^{-1})^4,$$

and $(gR^2g^{-1}) \cap (hR^2h^{-1})$ is also a symmetric neighbourhood of the identity. The result follows on letting X be this set. □

At this point we turn to analysis on multiplicative systems. Just as analysis on finite groups begins with Haar measure – the unique translation invariant probability measure on G – analysis on multiplicative systems begins with an approximate version of this.

Lemma 3.10 (Approximate right invariant Haar measure)**.** *Suppose that $\mathcal{B}$ is an $(r+1)$-step ϵ-closed multiplicative system. Then*

$$\|\rho_{x^{-1}}(\mu_{B_i}) - \mu_{B_i}\| = O(\epsilon) \text{ for all } x \in B_{i+1},$$

where $0 \leqslant i \leqslant r$.

Proof. Just note that

$$\begin{aligned}\|\rho_{x^{-1}}(\mu_{B_i}) - \mu_{B_i}\| &= \mu_{B_i}(B_i \backslash B_i x) + \mu_{B_i x}(B_i x \backslash B_i) \\ &\leqslant \mu_{B_i}(B_i \backslash x B_{i-}) + \mu_{B_i}(B_i \backslash B_{i-} x^{-1}) \\ &\leqslant 2 - 2\mu_{B_i}(B_{i-}) = O(\epsilon).\end{aligned}$$

The result is proved. □

We shall need a number of results in both a left and right hand form. Typically we shall prove the right hand version and simply state the left hand version after it, the proof being essentially the same.

Lemma 3.11 (Approximate left invariant Haar measure)**.** *Suppose that $\mathcal{B}$ is an $(r+1)$-step ϵ-closed multiplicative system. Then*

$$\|\lambda_x(\mu_{B_i}) - \mu_{B_i}\| = O(\epsilon) \text{ for all } x \in B_{i+1}$$

where $0 \leqslant i \leqslant r$.

Our arguments will involve finding (smaller and smaller) multiplicative systems on which our original set has larger and larger density. There are various different ways of doing this in Abelian groups and many of them have analogues in our setting. We shall use the energy increment technique developed by Heath-Brown and Szemerédi in [HB87] and [Sze90] and record an appropriate version now.

Lemma 3.12. *Suppose that $\mathcal{B}$ is a 1-step ϵ-closed multiplicative system; $A \subset Z := gB_0$ has density $\alpha := \mu_Z(A) > 0$; and*

$$\|1_A * \mu_{B_1} - \alpha 1_Z\|^2_{L_2(\mu_Z)} \geqslant \eta\alpha^2.$$

Then there is some $z \in Z$ such that

$$\mu_{zB_1}(A) = 1_A * \mu_{B_1}(z) \geqslant \alpha(1+\eta) - O(\epsilon).$$

Proof. This is just a calculation. First note that

$$\begin{aligned}\|1_A * \mu_{B_1}\|^2_{L_2(\mu_Z)} &= \|1_A * \mu_{B_1} - \alpha 1_Z\|^2_{L_2(\mu_Z)} \\ &\quad + \alpha\langle 1_A * \mu_{B_1}, 1_Z\rangle_{L_2(\mu_Z)} + \alpha\langle 1_Z, 1_A * \mu_{B_1}\rangle_{L_2(\mu_Z)} - \alpha^2. \qquad (3.1)\end{aligned}$$

Now Lemma 3.10 and the integral triangle inequality tell us that

$$\begin{aligned}\|\mu_Z * \mu_{B_1} - \mu_Z\| &= \|\mu_{B_0} * \mu_{B_1} - \mu_{B_0}\| \\ &\leqslant \int \|\rho_{x^{-1}}(\mu_{B_0}) - \mu_{B_0}\| d\mu_{B_1}(x) = O(\epsilon).\end{aligned}$$

Hence

$$\langle 1_A * \mu_{B_1}, 1_Z\rangle_{L_2(\mu_Z)} = \langle 1_A * \mu_{B_1}, \mu_Z\rangle = \langle 1_A, \mu_Z * \mu_{B_1}\rangle = \alpha + O(\epsilon),$$

and similarly

$$\langle 1_Z, 1_A * \mu_{B_1}\rangle_{L_2(\mu_Z)} = \alpha + O(\epsilon).$$

Inserting these into (3.1) we get

$$\|1_A * \mu_{B_1}\|^2_{L_2(\mu_Z)} = \|1_A * \mu_{B_1} - \alpha 1_Z\|^2_{L_2(\mu_Z)} + \alpha^2 + O(\epsilon\alpha) \geqslant (1+\eta)\alpha^2 - O(\epsilon\alpha)$$

On the other hand we have

$$\|1_A * \mu_{B_1}\|^2_{L_2(\mu_Z)} \leqslant \|1_A * \mu_{B_1}\|_{L_\infty(\mu_Z)}\alpha,$$

by the triangle inequality and the result follows. □

In exactly the same way we have a left hand version of the above.

Lemma 3.13. *Suppose that $\mathcal{B}$ is a 1-step ϵ-closed multiplicative system; $A \subset Z := B_0 h^{-1}$ has density $\alpha := \mu_Z(A) > 0$; and*

$$\|\mu_{B_1} * 1_A - \alpha 1_Z\|^2_{L_2(\mu_Z)} \geqslant \eta\alpha^2.$$

Then there is some $z \in Z$ such that

$$\mu_{B_1 z}(A) = \mu_{B_1} * 1_A(z) \geqslant \alpha(1+\eta) - O(\epsilon).$$

We shall ultimately be interested in examining some fairly thin subsets of multiplicative systems. In Abelian groups the tool for doing this is Chang's lemma [Cha02]; as we mentioned before, in non-Abelian groups the tool is the Croot-Sisask Lemma. We shall actually need a version for functions on multiplicative systems, where the proof is a minor variant on that of Lemma 3.5 with a few technicalities resulting from the fact that ρ is not an isometry when restricted to multiplicative systems.

Lemma 3.14 (The (right hand) Croot-Sisask Lemma for multiplicative systems)**.** *Suppose that $\mathcal{B}$ is a 2-step ϵ-closed multiplicative system, X is a symmetric neighbourhood of the identity and $X^8 \subset B_2$, $f \in L_\infty(\mu_{B_{0+}})$, $A \subset B_{0-}$ has density $\alpha := \mu_{B_0}(A) > 0$, and $\eta \in (0,1]$ and $p \in [2,\infty)$ are parameters. Then there is a symmetric neighbourhood of the identity $T \subset X^2$ with*

$$\mu_G(T) \geqslant \exp(-O(\eta^{-2} p \log 2\alpha^{-1}))\mu_G(X)$$

such that

$$\|\rho_{t^{-1}}(f * \mu_A) - f * \mu_A\|_{L_p(\mu_{B_1})} \leqslant \eta(1 + O(\epsilon/p))\|f\|_{L_\infty(\mu_{B_{0+}})}$$

for all $t \in T^4$.

Proof. Let $z_1, \dots, z_k$ be independent uniformly distributed A-valued random variables, and for each $y \in B_{1+}$ define $Z_i(y) := \rho_{z_i^{-1}}(f)(y) - f * \mu_A(y)$. For fixed y, the variables $Z_i(y)$ are independent and have mean 0, so it follows by the Marcinkiewicz-Zygmund inequality and then Hölder's inequality that

$$\begin{aligned} \|\sum_{i=1}^k Z_i(y)\|^p_{L^p(\mu_A^k)} &= O(p)^{p/2} \int \left(\sum_{i=1}^k |Z_i(y)|^2\right)^{p/2} d\mu_A^k \\ &= O(p)^{p/2} k^{p/2-1} \sum_{i=1}^k \int |Z_i(y)|^p d\mu_A^k. \end{aligned}$$

Integrating over $y \in B_{1+}$ and interchanging the order of summation we get

$$\int \|\sum_{i=1}^k Z_i(y)\|^p_{L^p(\mu_A^k)} d\mu_{B_{1+}}(y) = O(p)^{p/2} k^{p/2-1} \int \sum_{i=1}^k \int |Z_i(y)|^p d\mu_{B_{1+}}(y) d\mu_A^k. \tag{3.2}$$

On the other hand,

$$\begin{aligned} \left(\int |Z_i(y)|^p d\mu_{B_{1+}}(y)\right)^{1/p} &= \|Z_i\|_{L_p(\mu_{B_{1+}})} \\ &\leqslant \|\rho_{z_i^{-1}}(f)\|_{L_p(\mu_{B_{1+}})} + \|f * \mu_A\|_{L_p(\mu_{B_{1+}})} \leqslant 2\|f\|_{L_\infty(\mu_{B_{0+}})} \end{aligned}$$

by the triangle inequality and the fact that $A \subset B_{0-}$. Dividing (3.2) by k^p and inserting the above and the expression for the Z_is we get that

$$\int\int \left|\frac{1}{k}\sum_{i=1}^k \rho_{z_i^{-1}}(f)(y) - f * \mu_A(y)\right|^p d\mu_{B_{1+}}(y) d\mu_A^k(z) = O(pk^{-1}\|f\|^2_{L_\infty(\mu_{B_{0+}})})^{p/2}.$$

Pick $k = O(\eta^{-2}p)$ such that the right hand side is at most $(\eta\|f\|_{L_\infty(\mu_{B_{0+}})}/16)^p$ and write $\mathcal{L}$ for the set of $x \in A \times \cdots \times A$ (where the product is k-fold) for which the integrand above is at most $(\eta\|f\|_{L_\infty(\mu_{B_{0+}})}/8)^p$. By averaging $\mu_A^k(\mathcal{L}^c) \leqslant 2^{-p}$ and so $\mu_A^k(\mathcal{L}) \geqslant 1 - 2^{-p} \geqslant 3/4$.

Now, put $\Delta := \{(x, \dots, x) : x \in X\}$ and note (since $A \subset B_{0-}$ implies $\mathcal{L} \subset B_{0-} \times \cdots \times B_{0-}$ and $\mathcal{L} + \Delta \subset B_0 \times \cdots \times B_0$) that

$$\begin{aligned} \langle \widetilde{1_\mathcal{L}} * \mu_\mathcal{L}, \mu_\Delta * \widetilde{\mu_\Delta}\rangle &= \langle \mu_\mathcal{L} * \mu_\Delta, 1_\mathcal{L} * \mu_\Delta\rangle \\ &= \|1_\mathcal{L} * \mu_\Delta\|_{L_2(\mu^k_{B_0})} \mu^k_{B_0}(\mathcal{L})^{-1} \\ &\geqslant \|1_\mathcal{L} * \mu_\Delta\|^2_{L_1(\mu^k_{B_0})} \mu^k_{B_0}(\mathcal{L})^{-1} = \mu^k_{B_0}(\mathcal{L}) \geqslant \frac{3}{4}\alpha^k. \end{aligned}$$

We let $T \subset X^2$ be the set of t such that $\widetilde{1_\mathcal{L}} * \mu_\mathcal{L}(t, \dots, t) > 0$. Then

$$\mu_\Delta * \widetilde{\mu_\Delta}(\{(t, \dots, t) : t \in T\}) \geqslant \frac{3}{4}\mu_{B_0}(A)^k,$$

and so $\mu_G(T) \geqslant \frac{3}{4}\alpha^k \mu_G(X)$.

Now suppose that $t_1,t_2,t_3,t_4 \in T$. Then for each $1 \leqslant j \leqslant 4$ there are elements $z(t_j), y(t_j) \in \mathcal{L}$ such that $y(t_j)_i = z(t_j)_i t_j$ for all $1 \leqslant i \leqslant k$. By the triangle inequality

$$\begin{aligned}(3.3)\|\rho_{(t_1t_2t_3t_4)^{-1}}(f*\mu_A)-f*\mu_A\|_{L_p(\mu_{B_1})} &\leqslant \|\rho_{(t_2t_3t_4)^{-1}}(\rho_{t_1^{-1}}(f*\mu_A)-f*\mu_A)\|_{L_p(\mu_{B_1})}\\ &+ \|\rho_{(t_3t_4)^{-1}}(\rho_{t_2^{-1}}(f*\mu_A)-f*\mu_A)\|_{L_p(\mu_{B_1})}\\ &+ \|\rho_{t_4^{-1}}(\rho_{t_3^{-1}}(f*\mu_A)-f*\mu_A)\|_{L_p(\mu_{B_1})}\\ &+ \|\rho_{t_4^{-1}}(f*\mu_A)-f*\mu_A\|_{L_p(\mu_{B_1})}\\ &\leqslant \sup_{x\in T^3,t\in T}\|\rho_{t^{-1}}(f*\mu_A)-f*\mu_A\|_{L_p(\rho_x(\mu_{B_1}))}\end{aligned}$$

Now, suppose that $x \in T^3$ and $t \in T$. Then

$$\begin{aligned}\|\rho_{t^{-1}}(f*\mu_A)-f*\mu_A\|_{L_p(\rho_x(\mu_{B_1}))} &\leqslant \|\rho_{t^{-1}}\left(\frac{1}{k}\sum_{i=1}^k \rho_{z(t)_i^{-1}}(f)\right)-f*\mu_A\|_{L_p(\rho_x(\mu_{B_1}))}\\ &\quad+\|\rho_{t^{-1}}\left(\frac{1}{k}\sum_{i=1}^k \rho_{z(t)_i^{-1}}(f)-f*\mu_A\right)\|_{L_p(\rho_x(\mu_{B_1}))}\\ &= \|\frac{1}{k}\sum_{i=1}^k \rho_{y(t)_i^{-1}}(f)-f*\mu_A\|_{L_p(\rho_x(\mu_{B_1}))}\\ &\quad+\|\frac{1}{k}\sum_{i=1}^k \rho_{z(t)_i^{-1}}(f)-f*\mu_A\|_{L_p(\rho_{tx}(\mu_{B_1}))}.\end{aligned}$$

However, since $t \in T$ and $x \in T^3$ we have $x, tx \in T^4 \subset X^8 \subset B_2$, hence $\rho_x(\mu_{B_1}) \leqslant (1+O(\epsilon))\mu_{B_{1+}}$ and so

$$\begin{aligned}\|\rho_{t^{-1}}(f*\mu_A)-f*\mu_A\|_{L_p(\rho_x(\mu_{B_1}))} &\leqslant (1+O(\epsilon/p))\left(\|\frac{1}{k}\sum_{i=1}^k \rho_{y(t)_i^{-1}}(f)-f*\mu_A\|_{L_p(\mu_{B_{1+}})}\right.\\ &\quad\left.+\|\frac{1}{k}\sum_{i=1}^k \rho_{z(t)_i^{-1}}(f)-f*\mu_A\|_{L_p(\mu_{B_{1+}})}\right)\\ &\leqslant (1+O(\epsilon/p))\eta\|f\|_{L_\infty(\mu_{B_{0+}})}/4.\end{aligned}$$

The last inequality is from the definition of $\mathcal{L}$. Inserting this bound into (3.3) gives us the required result. □

Lemma 3.15 (The (left hand) Croot-Sisask Lemma for multiplicative systems)**.** *Suppose that $\mathcal{B}$ is a 2-step ϵ-closed multiplicative system, X is a symmetric neighbourhood of the identity and $X^8 \subset B_2$, $f \in L_\infty(\mu_{B_{0+}})$, $A \subset B_{0-}$ has density $\alpha := \mu_{B_0}(A) > 0$, and $\eta \in (0,1]$ and $p \in [2,\infty)$ are parameters. Then there is a symmetric neighbourhood of the identity $T \subset X^2$ with*

$$\mu_G(T) \geqslant \exp(-O(\eta^{-2}p\log 2\alpha^{-1}))\mu_G(X)$$

such that

$$\|\lambda_t(\mu_A * f) - \mu_A * f\|_{L_p(\mu_{B_1})} \leqslant \eta(1 + O(\epsilon/p))\|f\|_{L_\infty(\mu_{B_{0+}})}$$

for all $t \in T^4$.

4. The iteration lemmas

Having set up the basic machinery in the previous section, this section is devoted to establishing the key iteration lemmas that are specific to the problem we are considering here. All of the results will have the form of a dichotomy: either we shall find some sort of density increment on a new multiplicative system; or we shall be able to ensure some good behaviour.

There are two key results, Proposition 4.1 and Proposition 4.5. These results correspond roughly to making some relative versions of the U^1-norm and U^2-norm small respectively. The first of these is rather easy and in the Abelian case is essentially [Bou99, §5]. This (in the application of Lemma 4.4 inside the proof of Proposition 4.1) is where the requirement that the elements of A have distinct squares comes from.

To understand the argument it may be useful to consider a model example. Meshulam's proof for Roth's Theorem in $\mathbb{F}_3^n$ [Mes95] can be viewed (somewhat anachronistically) as a model version of Bourgain's proof of Roth's Theorem in $\mathbb{Z}$ [Bou99] in which all the Bohr sets are assumed to be subgroups. The same modelling assumption is useful here.

Suppose that $A \subset gHk^{-1}$ for some $H \leqslant G$ and elements $g, k \in G$ has size $\alpha|H|$. Write $K := gHg^{-1} \cap kHk^{-1}$. It is possible to show by averaging that (either we have a density increment or) the set

$$S := \{a \in A : \mu_K * 1_{Agk^{-1}}(akg^{-1}) \approx \alpha \text{ and } 1_{gk^{-1}A} * \mu_K(gk^{-1}a) \approx \alpha\},$$

has $|S| \approx \alpha|H|$. This is the application of Lemmas 4.2 and 4.3 below.

By a slightly more complicated averaging argument (Lemma 4.4) there is some $a \in S$ such that

$$sK = aK \text{ and } Ks = Ka \text{ for at least } \frac{|K|^2}{|KS||SK|}|S| \text{ elements } s \in S.$$

Since $|KS| \leqslant |H|$ and $|SK| \leqslant |H|$ it follows that

$$|\{s^2 : s \in S\} \cap aKa\}| \gtrsim \alpha \left(\frac{|K|}{|H|}\right)|K|,$$

and then we see that

$$A_1 := Aa^{-1} \cap K, A_2 := a^{-1}A \cap K \text{ and } T := a^{-1}\{s^2 : s \in A\}a^{-1} \cap K$$

have

$$|A_1| \approx \alpha|K|, |A_2| \approx \alpha|K| \text{ and } |T| \gtrsim \alpha \left(\frac{|K|}{|H|}\right),$$

and we have an injection from triples $(a_1, a_2, t) \in A_1 \times A_2 \times T$ with $a_2a_1 = t$ to triples $(a_1a, aa_2, ata) \in A \times A \times \{s^2 : s \in A\}$ with $(aa_2)(a_1a) = ata$.

Counting triples $(a,b,c) \in A \times B \times C$ such that $ab = c$ is actually rather well understood in simple groups as was shown by Gowers in [Gow08], but in Abelian groups it leads to a dichotomy: either the count of triples is about right or else there is a structure (for us a multiplicative system) on which at least one of the sets has increased density. This dichotomy also holds for general groups and is our Proposition 4.5.

The tool for proving the analogue of Proposition 4.5 in the Abelian setting is the Fourier transform (and Chang's theorem) and for us here we require the Croot-Sisask Lemma (as mentioned in §3).

Proposition 4.1. *Suppose that $\mathcal{B}$ and $\mathcal{B}'$ are 1-step ϵ-closed multiplicative systems with $B_0' \subset gB_1g^{-1} \cap hB_1h^{-1}$; $A \subset Z := gB_0h^{-1}$ has $\alpha := \mu_Z(A) > 0$ (and distinct squares); X is a symmetric neighbourhood of the identity with $\delta := \mu_G(X) > 0$ such that $X^4 \subset B_1'$; and $\eta \in (0,1]$ is a parameter. Then*

(i) either there is some $a \in gB_0h^{-1}$ such that

$$\mu_{aB_0'}(A), \mu_{B_0'a}(A) \geqslant \alpha(1-\eta)$$

and

$$\mu_{aB_1'a}(\{s^2 : s \in A\}) = \Omega(\alpha\delta^2);$$

(ii) or there is some $z \in gB_0h^{-1}$ such that

$$\mu_{B_1'} * 1_A(z) \geqslant \alpha(1+\Omega(\eta^2)) - O(\epsilon\alpha^{-1});$$

(iii) or there is some $z \in gB_0h^{-1}$ such that

$$1_A * \mu_{B_1'}(z) \geqslant \alpha(1+\Omega(\eta^2)) - O(\epsilon\alpha^{-1}).$$

First we have two lemmas (which are left and right versions of each other) which will be used in producing the second and third outcomes above.

Lemma 4.2. *Suppose $\mathcal{B}$ and $\mathcal{B}'$ are 1-step ϵ-closed multiplicative systems with $B_0' \subset hB_1h^{-1}$; and $A \subset Z := gB_0h^{-1}$ has density $\alpha := \mu_Z(A) > 0$; and $\eta \in (0,1]$ is a parameter. Then*

(i) either we have

$$\|1_A * \mu_{B_0'} - \alpha 1_Z\|_{L_1(\mu_A)} \leqslant \eta\alpha;$$

(ii) or there is some $z \in Z$ such that

$$1_A * \mu_{B_1'}(z) \geqslant \alpha(1+\Omega(\eta^2)) - O(\epsilon\alpha^{-1}).$$

Proof. Write

$$f(z) := \begin{cases} |1_A * \mu_{B_0'}(z) - \alpha 1_Z(z)| & \text{for all } z \in Z = gB_0h^{-1} \\ 0 & \text{otherwise.} \end{cases}$$

Since

$$\|\rho_{x^{-1}}(1_A * \mu_{B_0'}) - 1_A * \mu_{B_0'}\|_{L_\infty(\mu_G)} = O(\epsilon) \text{ for all } x \in B_1',$$

by the closure properties of $\mathcal{B}'$, and $\rho_{x^{-1}}(1_Z)(z) = 1_Z(z)$ for all $x \in B_1'$ and $z \in gB_{0-}h^{-1}$ (since $B_1' \subset B_0' \subset hB_1h^{-1}$), we have

$$|\rho_{x^{-1}}(f)(z) - f(z)| = O(\epsilon) \text{ for all } z \in gB_{0-}h^{-1} \text{ and } x \in B_1'.$$

It follows that

$$|\langle f, 1_A\rangle_{L_2(\mu_Z)} - \langle \rho_{x^{-1}}(f), 1_A\rangle_{L_2(\mu_Z)}| = O(\epsilon\alpha) + O(\mu_Z(Z\backslash gB_{0-}h^{-1})) = O(\epsilon).$$

We conclude that if we are not in the first case of the lemma then we have

$$\langle f, 1_A * \mu_{B_1'}\rangle_{L_2(\mu_Z)} = \langle f * \mu_{B_1'}, 1_A\rangle_{L_2(\mu_Z)} > \eta\alpha^2 - O(\epsilon),$$

and hence that

$$\langle f, 1_A * \mu_{B_1'} - \alpha 1_Z\rangle_{L_2(\mu_Z)} + \alpha\|f\|_{L_1(\mu_Z)} > \eta\alpha^2 - O(\epsilon).$$

Applying the Cauchy-Schwarz inequality to the first inner product, and nesting of norms to $\|f\|_{L_1(\mu_Z)}$ this then tells us that

$$\|1_A * \mu_{B_0'} - \alpha 1_Z\|_{L_2(\mu_Z)}(\|1_A * \mu_{B_1'} - \alpha 1_Z\|_{L_2(\mu_Z)} + \alpha) > \eta\alpha^2 - O(\epsilon). \tag{4.1}$$

If

$$\|1_A * \mu_{B_1'} - \alpha 1_Z\|_{L_2(\mu_Z)} + \alpha \geqslant 2\alpha, \tag{4.2}$$

then

$$\|1_A * \mu_{B_1'} - \alpha 1_Z\|^2_{L_2(\mu_Z)} \geqslant \alpha^2,$$

and we are done by Lemma 3.12 applied to the 1-step ϵ-closed system

$$(hB_{0+}h^{-1}, hB_0h^{-1}, hB_{0-}h^{-1}; B_1')$$

and $A \subset gh^{-1}(hB_0h^{-1})$. Thus we may suppose that (4.2) does not hold and so by (4.1) we have

$$\|1_A * \mu_{B_0'} - \alpha 1_Z\|_{L_2(\mu_Z)} > \eta\alpha/2 - O(\epsilon\alpha^{-1}),$$

and so

$$\|1_A * \mu_{B_0'} - \alpha 1_Z\|^2_{L_2(\mu_Z)} = \Omega(\eta^2\alpha^2) - O(\epsilon + \epsilon^2\alpha^{-2}).$$

Now, $\|\mu_{B_1'} * \mu_{B_0'} - \mu_{B_0'}\| = O(\epsilon)$ and so

$$\|1_A * \mu_{B_1'} * \mu_{B_0'} - \alpha 1_Z\|^2_{L_2(\mu_Z)} = \|1_A * \mu_{B_0'} - \alpha 1_Z\|^2_{L_2(\mu_Z)} + O(\epsilon\alpha).$$

On the other hand by the convexity of $\|\cdot\|_{L_2(\mu_Z)}$ (and the fact that $B_0' \subset hB_1h^{-1}$ and $Z = gB_0h^{-1}$ so $|ZB_0'| \leqslant (1+\epsilon)|Z|$) we have

$$\begin{aligned}
\|1_A * \mu_{B_1'} * \mu_{B_0'} - \alpha 1_Z\|^2_{L_2(\mu_Z)} &\leqslant \int \|\rho_{x^{-1}}(1_A * \mu_{B_1'}) - \alpha 1_Z\|^2_{L_2(\mu_Z)} d\mu_{B_0'}(x) \\
&= \int \|\rho_{x^{-1}}(1_A * \mu_{B_1'} - \alpha 1_Z)\|^2_{L_2(\mu_Z)} d\mu_{B_0'}(x) + O(\epsilon\alpha) \\
&= \int \|1_A * \mu_{B_1'} - \alpha 1_Z\|^2_{L_2(\mu_Z)} d\mu_{B_0'}(x) + O(\epsilon) \\
&= \|1_A * \mu_{B_1'} - \alpha 1_Z\|^2_{L_2(\mu_Z)} + O(\epsilon),
\end{aligned}$$

from which it follows that

$$\|1_A * \mu_{B_0'} - \alpha 1_Z\|_{L_2(\mu_Z)}^2 \leqslant \|1_A * \mu_{B_1'} - \alpha 1_Z\|_{L_2(\mu_Z)}^2 + O(\epsilon).$$

But our lower bound on the left hand side then tells us that

$$\|1_A * \mu_{B_1'} - \alpha 1_Z\|_{L_2(\mu_Z)}^2 = \Omega(\eta^2\alpha^2) - O(\epsilon + \epsilon^2\alpha^{-2}).$$

The result follows again on application of Lemma 3.12. □

We also have the left analogue of the above.

Lemma 4.3. *Suppose $\mathcal{B}$ and $\mathcal{B}'$ are 1-step ϵ-closed multiplicative systems with $B_0' \subset gB_1g^{-1}$; and $A \subset Z := gB_0h^{-1}$ has density $\alpha := \mu_Z(A) > 0$; and $\eta \in (0,1]$ is a parameter. Then*

(i) either we have

$$\|\mu_{B_0'} * 1_A - \alpha 1_Z\|_{L_1(\mu_A)} \leqslant \eta\alpha;$$

(ii) or there is some $z \in Z$ such that

$$\mu_{B_1'} * 1_A(z) \geqslant \alpha(1 + \Omega(\eta^2)) - O(\epsilon\alpha^{-1}).$$

The previous two lemmas will provide us with a left and right translate of some suitable multiplicative system on which A has the 'right' density. We now turn to ensuring that the squares of the elements in A have not-too-small density. This is the lemma for which we need A to have distinct squares.

Lemma 4.4. *Suppose that $\mathcal{B}$ is a 1-step, ϵ-closed multiplicative system, $S \subset gB_0h^{-1}$ has distinct squares, and $X \subset (gB_1g^{-1}) \cap (hB_1h^{-1})$ is a symmetric neighbourhood of the identity. Then there is some $s' \in S$ such that*

$$|\{s^2 : s \in S\} \cap s'X^4s'| \geqslant \frac{|X|^2}{(1+\epsilon)^2|B_0|^2}|S|.$$

Proof. We consider the sum

$$\begin{aligned}
\sum_{s,s'\in S} |sX \cap s'X||Xs \cap Xs'| &= \sum_{z,z'\in G}\sum_{s,s'\in S} 1_X(s^{-1}z)1_X((s')^{-1}z)1_X(z's^{-1})1_X(z'(s')^{-1}) \\
&= \sum_{z\in SX,z'\in XS}\sum_{s,s'\in S} 1_X(s^{-1}z)1_X((s')^{-1}z)1_X(z's^{-1})1_X(z'(s')^{-1}) \\
&= \sum_{z\in SX,z'\in XS}\left(\sum_{s\in S} 1_X(s^{-1}z)1_X(z's^{-1})\right)^2 \\
&\geqslant \frac{1}{|SX||XS|}\left(\sum_{z\in SX,z'\in XS}\sum_{s\in S} 1_X(s^{-1}z)1_X(z's^{-1})\right)^2 \\
&= \frac{|X|^4|S|^2}{|SX||XS|}.
\end{aligned}$$

The summands on the left hand side are at most $|X|^2$, and so by averaging there is some $s' \in S$ such that for at least $\frac{|X|^2}{|SX||XS|}|S|$ elements $s \in S$ we have

$$sX \cap s'X \neq \emptyset \text{ and } Xs \cap Xs' \neq \emptyset.$$

It follows that for each such s we have $s \in s'XX^{-1}$ and $s \in X^{-1}Xs'$ and hence $s^2 \in s'XX^{-2}Xs' = s'X^4s'$. It remains to note that since $X \subset (gB_1g^{-1}) \cap (hB_1h^{-1})$ and $S \subset gB_0h^{-1}$ we have that $XS \subset gB_1B_0h^{-1} \subset gB_{0+}h^{-1}$ and $SX \subset gB_0B_1h^{-1} \subset gB_{0+}h^{-1}$ from which we get the result. □

Proof of Proposition 4.1. Apply Lemmas 4.2 and 4.3 with parameter $\eta/4$ and the set B_0, to see that either we are in the second or third cases of the proposition (and we are done) or else

$$\|1_A * \mu_{B_0'} - \alpha 1_{gB_0h^{-1}}\|_{L_1(\mu_A)} \leqslant \eta\alpha/4 \text{ and } \|\mu_{B_0'} * 1_A - \alpha 1_{gB_0h^{-1}}\|_{L_1(\mu_A)} \leqslant \eta\alpha/4.$$

Let

$$S := \{a \in A : |1_A * \mu_{B_0'}(a) - \alpha| \leqslant \eta\alpha \text{ and } |\mu_{B_0'} * 1_A(a) - \alpha| \leqslant \eta\alpha\},$$

so that

$$\mu_A(A\backslash S)\eta\alpha \leqslant \|1_A * \mu_{B_0'} - \alpha 1_{gB_0h^{-1}}\|_{L_1(\mu_A)} + \|\mu_{B_0'} * 1_A - \alpha 1_{gB_0h^{-1}}\|_{L_1(\mu_A)} \leqslant \eta\alpha/2.$$

It follows that $\mu_A(S) \geqslant 1/2$. Now apply Lemma 4.4 to the set $S \subset gB_0h^{-1}$ and

$$X \subset X^4 \subset B_1' \subset B_0' \subset gB_1g^{-1} \cap hB_1h^{-1}$$

to get that there is some $a \in S$ such that

$$|\{s^2 : s \in S\} \cap aX^4a| \geqslant \frac{|X|^2}{(1+\epsilon)^2|B_0|^2}|S|.$$

Since $a \in S$ we have that

$$1_A * \mu_{B_0'}(a) \geqslant \alpha(1-\eta) \text{ and } \mu_{B_0'} * 1_A(a) \geqslant \alpha(1-\eta),$$

and since $S \subset A$ and $X^4 \subset B_1'$ we have

$$\mu_{aB_1'a}(\{s^2 : s \in A\}) = \Omega(\alpha\delta^2),$$

and the result is proved. □

We now turn to the companion result to Proposition 4.1 which explains how to use the output in the first case of that result.

Proposition 4.5. *Suppose that $\mathcal{B}$ is a 2-step ϵ-closed multiplicative system; that there is a symmetric neighbourhood of the identity X of density $\delta := \mu_G(X) > 0$ such that $X^4 \subset B_2$; and that $U, V \subset B_{0-}$ have $\mu_{B_0}(U) = \mu_{B_0}(V) = \alpha > 0$ and $W \subset B_{1-}$ has $\mu_{B_1}(W) = \omega > 0$. Then*

(i) either

$$\langle 1_U * \mu_V, 1_W \rangle_{L_2(\mu_{B_1})} \geqslant \frac{1}{2}\mu_{B_0}(U)\mu_{B_0}(V)\mu_{B_1}(W);$$

(ii) or there is a 1-step ϵ-closed multiplicative system $\mathcal{B}^L$ and a symmetric neighbourhood of the identity S_L such that $S_L^4 \subset B_1^L$ and

$$\mu_G(S_L) \geqslant \exp(-O(\epsilon^{-1}\alpha^{-1}\log 2\omega^{-1}\log 2\delta^{-1})^{O(1)})$$

and some $z \in B_0$ with

$$\mu_{B_0^L z}(U) \geqslant \mu_{B_0}(U)(1+\Omega(1)) - O(\epsilon\alpha^{-1});$$

(iii) or there is a 1-step ϵ-closed multiplicative system $\mathcal{B}^R$ and a symmetric neighbourhood of the identity S_R such that $S_R^4 \subset B_1^R$ and

$$\mu_G(S_R) \geqslant \exp(-O(\epsilon^{-1}\alpha^{-1}\log 2\omega^{-1}\log 2\delta^{-1})^{O(1)})$$

and some $z \in B_0$ with

$$\mu_{zB_0^R}(V) \geqslant \mu_{B_0}(V)(1+\Omega(1)) - O(\epsilon\alpha^{-1}).$$

Proof. Begin by applying Lemma 3.6 to get a symmetric neighbourhood of the identity T such that $T^8 \subset X^4$ and $\mu_G(T) \geqslant \exp(-O(\log^2 2\delta^{-1}))$.

Let $\eta := \epsilon\alpha \in (0,1]$ and $p \in [2,\infty)$ be a parameter to be chosen later (it will be apparent that it only depends on data available at this stage of the proof) and apply Lemma 3.14 to get that there is a symmetric neighbourhood of the identity $R \subset T^2$ with

$$\mu_G(R) \geqslant \exp(-O(\eta^{-2}p\log 2\mu_G(V)^{-1} + \log^2 2\delta^{-1}))$$

such that for all $t \in R^4$ we have

$$\|\rho_{t^{-1}}(1_U * \mu_V) - 1_U * \mu_V\|_{L_p(\mu_{B_1})} \leqslant \eta(1+O(\epsilon/p))\|1_U\|_{L_\infty(\mu_{B_{0+}})} = O(\eta).$$

Apply Corollary 3.7 to get a 1-step ϵ-closed multiplicative system $\mathcal{B}^R$ and symmetric neighbourhood of the identity S_R such that

$$B_0^R \subset R^4 \text{ and } S_R^4 \subset B_1^R$$

and

$$\mu_G(S_R) \geqslant \exp(-O(\epsilon^{-1}\log 2\mu_G(R)^{-1})^{O(1)}).$$

By integrating we have

$$\|1_U * \mu_V * \mu_{B_0^R} - 1_U * \mu_V\|_{L_p(\mu_{B_1})} = O(\eta).$$

It follows from linearity and Hölder's inequality (writing p' for the conjugate index of p) that

$$|\langle 1_U * \mu_V * \mu_{B_0^R}, 1_W\rangle_{L_2(\mu_{B_1})} - \langle 1_U * \mu_V, 1_W\rangle_{L_2(\mu_{B_1})}| = O(\eta\|1_W\|_{L_{p'}(\mu_{B_1})}).$$

Now, $\|1_W\|_{L_{p'}(\mu_{B_1})} = \mu_{B_1}(W)^{1/p'} = \mu_{B_1}(W)^{1-1/p}$ and so taking $p = 2 + \log\mu_{B_1}(W)^{-1}$ we see that

$$\langle 1_U * \mu_V, 1_W\rangle_{L_2(\mu_{B_1})} = \langle 1_U * \mu_V * \mu_{B_0^R}, 1_W\rangle_{L_2(\mu_{B_1})} + O(\eta\mu_{B_1}(W)).$$

Since $U \subset B_{0-}$ we see that

$$1_U * \mu_{B_0}(x) = \mu_{B_0}(U) \text{ for all } x \in B_1$$

and so

$$\begin{aligned}\langle 1_U * \mu_V, 1_W\rangle_{L_2(\mu_{B_1})} &= \langle 1_U * (\mu_V * \mu_{B_0^R} - \mu_{B_0}), 1_W\rangle_{L_2(\mu_{B_1})} \\ &\quad + \mu_{B_0}(U)\langle 1_{B_1}, 1_W\rangle_{L_2(\mu_{B_1})} + O(\eta\mu_{B_1}(W)).\end{aligned} \tag{4.3}$$

On the other hand, if we write $f := 1_V * \mu_{B_0^R} - \mu_{B_0}(V)1_{B_0}$ then

$$\langle 1_U * (\mu_V * \mu_{B_0^R} - \mu_{B_0}(V)\mu_{B_0}), 1_W\rangle_{L_2(\mu_{B_1})} = \langle \mu_U * f, 1_W\rangle_{L_2(\mu_{B_1})}.$$

Now we can apply Lemma 3.15 (with the same parameters as before) to get that there is a symmetric neighbourhood of the identity $L \subset T^2$ with

$$\mu_G(L) \geqslant \exp(-O(\eta^{-2}p\log 2\mu_G(U)^{-1} + \log^2 2\delta^{-1}))$$

such that for all $t \in L^4$ we have

$$\|\lambda_t(\mu_U * f) - \mu_U * f\|_{L_p(\mu_{B_1})} = O(\eta\|f\|_{L_\infty(\mu_{B_{0^+}})}) = O(\eta),$$

and argue as before. Apply Corollary 3.7 to get a 1-step ϵ-closed multiplicative system $\mathcal{B}^L$ and symmetric neighbourhood of the identity S_L such that

$$B_0^L \subset L^4 \text{ and } S_L^4 \subset B_1^L$$

and

$$\mu_G(S_L) \geqslant \exp(-O(\epsilon^{-1}\log 2\mu_G(L)^{-1})^{O(1)}).$$

By integrating we have

$$\|\mu_{B_0^L} * \mu_U * f - \mu_U * f\|_{L_p(\mu_{B_1})} = O(\eta).$$

Our choice of p ensures that

$$\langle \mu_U * f, 1_W\rangle_{L_2(\mu_{B_1})} = \langle \mu_{B_0^L} * \mu_U * f, 1_W\rangle_{L_2(\mu_{B_1})} + O(\eta\mu_{B_1}(W)).$$

Since $V \subset B_0^-$ we see that for all $x \in B_1$ we have

$$\begin{aligned}\mu_{B_0} * f(x) &\leqslant \mu_{B_0} * 1_V * \mu_{B_0^R}(x) - \mu_{B_0}(V)\mu_{B_0} * 1_{B_0}(x) \\ &\leqslant \mu_{B_0}(V) - (1 - O(\epsilon))\mu_{B_0}(V) = O(\epsilon\mu_{B_0}(V)),\end{aligned}$$

and so

$$\begin{aligned}\langle \mu_U * f, 1_W\rangle_{L_2(\mu_{B_1})} &= \langle (\mu_{B_0^L} * \mu_U - \mu_{B_0}) * f, 1_W\rangle_{L_2(\mu_{B_1})} \\ &\quad + O(\epsilon\mu_{B_0}(V))\mu_{B_1}(W) + O(\eta\mu_{B_1}(W)).\end{aligned} \tag{4.4}$$

Put

$$g := \mu_{B_0^L} * 1_U - \mu_{B_0}(U)1_{B_0} \text{ and } \nu := \mu_{B_0}(U)(\mu_{B_0^L} * \mu_U - \mu_{B_0})$$

so that

$$\nu * h(x) = \langle \rho_x(h), \tilde{g}\rangle_{L_2(\mu_{B_0})} \text{ for all } h : G \to \mathbb{C}.$$

With these definitions, (4.4) and (4.3) give

$$\begin{aligned}\langle \nu * f, 1_W\rangle_{L_2(\mu_{B_1})} &= \mu_{B_0}(U)\langle \mu_U * f, 1_W\rangle_{L_2(\mu_{B_1})} - O((\epsilon\mu_{B_0}(V) + \eta)\mu_{B_1}(W)\mu_{B_0}(U)) \\ &= \mu_{B_0}(U)\mu_{B_0}(V)\mu_{B_1}(W) - O((\epsilon\mu_{B_0}(V) + \eta)\mu_{B_1}(W)\mu_{B_0}(U)).\end{aligned}$$

We conclude that either we are in the first case of the lemma or else there is some $x \in B_1$ such that

$$|\nu * f(x)| \geqslant \frac{1}{2}\mu_{B_0}(U)\mu_{B_0}(V)(1-O(\epsilon))$$

(since $\eta = \epsilon\alpha$). Now, by the Cauchy-Schwarz inequality we have

$$\begin{aligned} |\nu * f(x)| &= |\langle \rho_x(f), \tilde{g}\rangle_{L_2(\mu_{B_0})}| \\ &\leqslant \|\rho_x(f)\|_{L_2(\mu_{B_0})}\|\tilde{g}\|_{L_2(\mu_{B_0})} = (\|f\|_{L_2(\mu_{B_0})} + O(\epsilon))\|g\|_{L_2(\mu_{B_0})}. \end{aligned}$$

It follows that either

$$\|g\|^2_{L_2(\mu_{B_0})} \geqslant \frac{1}{2}\mu_{B_0}(U)^2 - O(\epsilon)$$

or

$$\|f\|^2_{L_2(\mu_{B_0})} \geqslant \frac{1}{2}\mu_{B_0}(V)^2 - O(\epsilon). \tag{4.5}$$

In the first instance we apply Lemma 3.13 to the 1-step ϵ-closed multiplicative system

$$(B_{0+}, B_0, B_{0-}; B_0^L)$$

and the set $U \subset B_0$ to get that there is some $z \in B_0$ such that

$$\mu_{B_0^L z}(U) \geqslant \mu_{B_0}(U)(1+\Omega(1)) - O(\epsilon\alpha^{-1}).$$

We are in the second case of the proposition with the system $\mathcal{B}^L$. In the second instance above (4.5) we apply Lemma 3.12 and are in the third case of the proposition. The result is proved. □

5. Proof of the main result

We are now in a position to prove out main theorem.

Proof of Theorem 1.5. We fix two parameters $\epsilon = c'\alpha^2$ and $\eta = c$ for some absolute constants $c, c' > 0$ whose precise value will become clear later in the argument. All of the multiplicative systems we consider will be ϵ-closed.

We shall proceed iteratively. At stage i of the iteration we suppose that we have the following data:

(i) $\mathcal{B}^{(i)}$ a 1-step (ϵ-closed) multiplicative system;
(ii) X_i, a symmetric neighbourhood of the identity with density $\delta_i := \mu_G(X_i) > 0$ such that $X_i^4 \subset B_1^{(i)}$;
(iii) $g_i, h_i \in G$ such that $A_i := A \cap g_i B_0^{(i)} h_i^{-1}$ has $\alpha_i := \mu_{g_i B_0^{(i)} h_i^{-1}}(A_i) > 0$.

We initialise with the 1-step ϵ-closed multiplicative system $\mathcal{B}^{(0)} := (G, G, G; G)$, $X_0 := G$, $g_0, h_0 = 1_G$ and $A_0 := A$ so that

$$\alpha_0 = \alpha > 0 \text{ and } \delta_0 = 1 > 0.$$

At some stage i_0 the iteration will terminate but for all $0 \leqslant i < i_0$ we shall have

$$\alpha_i \geqslant \alpha(1+\Omega(1))^i \text{ and } \delta_i \geqslant \exp(-O(\alpha^{-1})^{\exp(O(i))}).$$

At stage i of our iteration apply Lemma 3.9 to the set X_i and elements g_i and h_i to get a symmetric neighbourhood of the identity Y_i such that

$$Y_i^4 \subset (g_i X_i^4 g_i^{-1}) \cap (h_i X_i^4 h_i^{-1}) \text{ and } \mu_G(Y_i) \geqslant \exp(-O(\log^2 2\delta_i^{-1})).$$

Now apply Corollary 3.7 to Y_i to get a 2-step multiplicative system $\mathcal{B}^{(i)'}$ and a symmetric neighbourhood of the identity S_i such that

$$B_0^{(i)'} \subset Y_i^4, S_i^4 \subset B_3^{(i)'} \text{ and } \mu_G(S_i) \geqslant \exp(-O(\epsilon^{-1} \log 2\delta_i^{-1})^{O(1)}).$$

Apply Proposition 4.1 to $\mathcal{B}^{(i)}$ and the 1-step ϵ-closed system

$$(B_{0+}^{(i)'}, B_0^{(i)'}, B_{0-}^{(i)'}; B_{1-}^{(i)'}),$$

which have

$$B_0^{(i)'} \subset Y_i^4 \subset (g_i X_i^4 g_i^{-1}) \cap (h_i X_i^4 h_i^{-1}) \subset g_i B_1^{(i)} g_i^{-1} \cap h_i B_1^{(i)} h_i^{-1}$$

by design, and $A_i \subset g_i B_0^{(i)} h_i^{-1}$ and $S_i^4 \subset B_2^{(i)'} \subset B_{1-}^{(i)'}$.

If we are in the second two cases of Proposition 4.1 then we have some $z_i \in g_i B_0^{(i)} h_i^{-1}$ such that either

$$\begin{aligned} \mu_{B_1^{(i)'} z_i}(A_i) &\geqslant \mu_{B_{1-}^{(i)'} z_i}(A_i)(1 - O(\epsilon)) \\ &= \mu_{B_{1-}^{(i)'}} * 1_{A_i}(z_i)(1 - O(\epsilon)) \geqslant \alpha_i(1 + \Omega(\eta^2)) - O(\epsilon \alpha_i^{-1}) \end{aligned} \tag{5.1}$$

or

$$\begin{aligned} \mu_{z_i B_1^{(i)'}}(A_i) &\geqslant \mu_{z_i B_{1-}^{(i)'}}(A_i)(1 - O(\epsilon)) \\ &= 1_{A_i} * \mu_{B_{1-}^{(i)'}}(z_i)(1 - O(\epsilon)) \geqslant \alpha_i(1 + \Omega(\eta^2)) - O(\epsilon \alpha_i^{-1}). \end{aligned} \tag{5.2}$$

Set

$$\mathcal{B}^{(i+1)} := (B_{1+}^{(i)'}, B_1^{(i)'}, B_{1-}^{(i)'}; B_2^{(i)'})$$

and $X_{i+1} := S_i$ so that we have

$$X_{i+1}^4 = S_i^4 \subset B_2^{(i)'} = B_1^{(i+1)} \text{ and } \delta_{i+1} \geqslant \exp(-O(\epsilon^{-1} \log 2\delta_i^{-1})^{O(1)}).$$

It follows by the inductive hypothesis and the value of ϵ that

$$\delta_{i+1} = \exp(-O(\alpha^{-1})^{\exp(O((i+1)))}).$$

If (5.1) holds then take $g_{i+1} = 1_G$, $h_{i+1} = z_i^{-1}$, so

$$\begin{aligned} \alpha_{i+1} &= \mu_{g_{i+1} B_0^{(i+1)} h_{i+1}^{-1}}(A \cap g_{i+1} B_0^{(i+1)} h_{i+1}^{-1}) \\ &= \mu_{B_1^{(i)'} z_i}(A) \geqslant \mu_{B_1^{(i)'} z_i}(A_i) \\ &\geqslant \alpha_i(1 + \Omega(\eta^2)) - O(\epsilon \alpha_i^{-1}) = \alpha_i(1 + \Omega(c^2) - O(c')) = \alpha_i(1 + \Omega(1)) \end{aligned}$$

provided c' is sufficiently small compared with c^2. On the other hand, if (5.2) holds then take $g_{i+1}=z_i$ and $h_{i+1}=1_G$, so

$$\begin{aligned}\alpha_{i+1}&=\mu_{g_{i+1}B_0^{(i+1)}h_{i+1}^{-1}}(A\cap g_{i+1}B_0^{(i+1)}h_{i+1}^{-1})\\&=\mu_{z_iB_1^{(i)'}}(A)\geqslant\mu_{z_iB_1^{(i)'}}(A_i)\\&\geqslant\alpha_i(1+\Omega(\eta^2))-O(\epsilon\alpha_i^{-1})=\alpha_i(1+\Omega(c^2)-O(c'))=\alpha_i(1+\Omega(1)),\end{aligned}$$

again provided c' is sufficiently small compared with c^2. In either case

$$\alpha_{i+1}\geqslant\alpha_i(1+\Omega(1))\geqslant\alpha(1+\Omega(1))^{i+1}.$$

With the easier cases dealt with, suppose that we are in the first case of Proposition 4.1 and there is some a_i such that

$$\mu_{a_iB_0^{(i)'}}(A_i),\mu_{B_0^{(i)'}a_i}(A_i)\geqslant\alpha_i(1-\eta)$$

and

$$\mu_{a_iB_{1-}^{(i)'}a_i}(\{s^2:s\in A_i\})=\Omega(\alpha_i\mu_G(S_i)^2).$$

Now let

$$\tilde{U}_i:=(a_i^{-1}A_i)\cap B_{0-}^{(i)'}\text{ and }\tilde{V}_i:=(A_ia_i^{-1})\cap B_{0-}^{(i)'},$$

and note that

$$\min\left\{\mu_{B_0^{(i)'}}(\tilde{U}_i),\mu_{B_0^{(i)'}}(\tilde{V}_i)\right\}\geqslant\alpha_i(1-\eta)-O(\epsilon).$$

Thus we can pick $U_i\subset\tilde{U}_i$ and $V_i\subset\tilde{V}_i$ such that

$$\mu_{B_0^{(i)'}}(U_i)=\mu_{B_0^{(i)'}}(V_i)=\min\left\{\mu_{B_0^{(i)'}}(\tilde{U}_i),\mu_{B_0^{(i)'}}(\tilde{V}_i)\right\}\geqslant\alpha_i(1-\eta)-O(\epsilon).$$

Let

$$W_i:=\{a_i^{-1}s^2a_i^{-1}:s\in A_i\}\cap B_{1-}^{(i)'}.$$

so that

$$\mu_{B_{1-}^{(i)'}}(W_i)=\Omega(\alpha_i\mu_G(S_i)^2)=\Omega(\alpha_i\exp(-O(\epsilon^{-1}\log 2\delta_i^{-1})^{O(1)})).$$

We apply Proposition 4.5 to the system $\mathcal{B}^{(i)'}$, the symmetric neighbourhood of the identity S_i which has $S_i^4\subset B_2^{(i)'}$, the sets $U_i,V_i\subset B_{0-}^{(i)'}$ which have

$$\mu_{B_0^{(i)'}}(U_i)=\mu_{B_0^{(i)'}}(V_i)\geqslant\alpha_i(1-\eta)-O(\epsilon),$$

and the set $W_i\subset B_{1-}^{(i)'}$. If we are in the first case we terminate our iteration with

$$\langle 1_{U_i}*\mu_{V_i},1_{W_i}\rangle_{L_2\left(\mu_{B_1^{(i)'}}\right)}\geqslant\frac{1}{2}\mu_{B_0^{(i)'}}(U_i)\mu_{B_0^{(i)'}}(V_i)\mu_{B_1^{(i)'}}(W_i).\tag{5.3}$$

If we are in the second case of Proposition 4.5 then let $\mathcal{B}^{(i+1)}$ be the multiplicative system, and X_{i+1} the given symmetric neighbourhood of the identity so that

$$\begin{aligned}\delta_{i+1}&\geqslant\exp(-O(\epsilon^{-1}(\alpha_i(1-\eta)-O(\epsilon))^{-1}\log 2\mu_{B_1}(W)^{-1}\log 2\delta_i^{-1})^{O(1)})\\&=\exp(-O(\epsilon^{-1}\alpha^{-1}\log 2\delta_i)^{O(1)})=\exp(-O(\alpha^{-1})^{\exp(O((i+1)))}).\end{aligned}$$

We are given an h_i such that

$$\mu_{B_0^{(i+1)}h_i^{-1}}(U_i) \geqslant \mu_{B_0^{(i)'}}(U_i)(1+\Omega(1)) - O(\epsilon\alpha_i^{-1}).$$

It follows that putting $g_i := a_i$ we have

$$\begin{aligned}\alpha_{i+1} \geqslant \mu_{g_iB_0^{(i+1)}h_i^{-1}}(A_i \cap g_iB_0^{(i+1)}h_i^{-1}) &\geqslant \mu_{B_0^{(i)'}}(U)(1+\Omega(1)) - O(\epsilon\alpha_i^{-1})\\ &\geqslant \alpha_i(1+\Omega(1)-\eta) - O(\epsilon\alpha^{-1})\\ &= \alpha_i(1+\Omega(1)-c-O(c')) = \alpha_i(1+\Omega(1))\end{aligned}$$

provided c and c' are sufficiently small absolute constants. If we are in the third case of Proposition 4.5 then let $\mathcal{B}^{(i+1)}$ be the multiplicative system, and X_{i+1} the given symmetric neighbourhood of the identity so that

$$\begin{aligned}\delta_{i+1} &\geqslant \exp(-O(\epsilon^{-1}(\alpha_i(1-\eta)-O(\epsilon))^{-1}\log 2\mu_{B_1}(W)^{-1}\log 2\delta_i^{-1})^{O(1)})\\ &= \exp(-O(\epsilon^{-1}\alpha^{-1}\log 2\delta_i)^{O(1)}) = \exp(-O(\alpha^{-1})^{\exp(O((i+1)))}).\end{aligned}$$

We are given an g_i such that

$$\mu_{g_iB_0^{(i+1)}}(V_i) \geqslant \mu_{B_0^{(i)'}}(V_i)(1+\Omega(1)) - O(\epsilon\alpha_i^{-1}).$$

It follows that putting $h_i := a_i^{-1}$ we have

$$\begin{aligned}\alpha_{i+1} \geqslant \mu_{g_iB_0^{(i+1)}h_i^{-1}}(A_i \cap g_iB_0^{(i+1)}h_i^{-1}) &\geqslant \mu_{B_0^{(i)'}}(V)(1+\Omega(1)) - O(\epsilon\alpha_i^{-1})\\ &\geqslant \alpha_i(1+\Omega(1)-\eta) - O(\epsilon\alpha^{-1})\\ &= \alpha_i(1+\Omega(1)-c-O(c')) = \alpha_i(1+\Omega(1)),\end{aligned}$$

again provided c and c' are sufficiently small absolute constants. In either of the above cases

$$\alpha_{i+1} \geqslant (1+\Omega(1))\alpha_i \geqslant \alpha(1+\Omega(1))^{i+1}.$$

Since $\alpha_i \leqslant 1$ for all i it follows that the iteration terminates at some step i_0 with

$$1 \geqslant \alpha(1+\Omega(1))^{i_0},$$

i.e. at some stage $i_0 = O(\log 2\alpha^{-1})$. It follows from this that

$$\delta_{i_0} \geqslant \exp(-\exp(\alpha^{-O(1)})).$$

When we terminate the iteration (5.3) holds and so there are at least

$$\begin{aligned}|B_0^{(i)'}||B_1^{(i)'}| \cdot \frac{1}{2}\mu_{B_0^{(i)'}}(U_i)\mu_{B_0^{(i)'}}(V_i)\mu_{B_1^{(i)'}}(W_i) &= \Omega(\alpha_i^3\mu_G(S_i)^4)|G|^2\\ &= \Omega(\alpha^3\exp(-O(\alpha^{-1}\log 2\delta_{i_0}^{-1})^{O(1)}))|G|^2\end{aligned}$$

triples $(r,s,t) \in U_i \times W_i \times V_i$ such that $rt = s$. The mapping taking (r,s,t) to (a_ir, a_isa_i, ta_i) is an injection into

$$a_iU_i \times a_iW_ia_i \times V_ia_i \subset A \times \{a^2 : a \in A\} \times A,$$

and $(a_ir)(ta_i) = a_i(rt)a_i = a_isa_i$. It follows that there are at least

$$\Omega(\alpha^3\exp(-O(\alpha^{-1}\log 2\delta_{i_0}^{-1})^{O(1)}))|G|^2 = \exp(-\exp(\alpha^{-O(1)}))|G|^2$$

triples in $(x,y,z)\in A\times A\times A$ with $xz=y^2$ as claimed. □

Acknowledgement

We should like to thank the referee for a careful reading of the paper which identified numerous errors.

References

[BB82] T. C. Brown and J. P. Buhler. A density version of a geometric Ramsey theorem. *J. Combin. Theory Ser. A*, 32(1):20–34, 1982.
[BGT10] E. Breuillard, B. Green, and T. Tao. Approximate subgroups of linear groups. *ArXiv e-prints*, May 2010, 1005.1881.
[Blo14] T. F. Bloom. A quantitative improvement for Roth's theorem on arithmetic progressions. 2014, arXiv:1405.5800.
[BMZ97] V. Bergelson, R. McCutcheon, and Q. Zhang. A Roth theorem for amenable groups. *Amer. J. Math.*, 119(6):1173–1211, 1997.
[Bog39] N. Bogolioùboff. Sur quelques propriétés arithmétiques des presque-périodes. *Ann. Chaire Phys. Math. Kiev*, 4:185–205, 1939.
[Bou99] J. Bourgain. On triples in arithmetic progression. *Geom. Funct. Anal.*, 9(5):968–984, 1999.
[CF12] D Conlon and J. Fox. Graph removal lemmas. 2012, arXiv:1211.3487.
[Cha02] M.-C. Chang. A polynomial bound in Freĭman's theorem. *Duke Math. J.*, 113(3):399–419, 2002.
[CS10] E. S. Croot and O. Sisask. A probabilistic technique for finding almost-periods of convolutions. *Geom. Funct. Anal.*, 20(6):1367–1396, 2010.
[FGR87] P. Frankl, R. L. Graham, and V. Rödl. On subsets of abelian groups with no 3-term arithmetic progression. *J. Combin. Theory Ser. A*, 45(1):157–161, 1987.
[Fox11] J. Fox. A new proof of the graph removal lemma. *Ann. of Math. (2)*, 174(1):561–579, 2011, arXiv:1006.1300.
[Gow08] W. T. Gowers. Quasirandom groups. *Comb. Probab. Comput.*, 17(3):363–387, 2008.
[GW11] W. T. Gowers and J. Wolf. Linear forms and quadratic uniformity for functions on $\mathbb{Z}_N$. *J. Anal. Math.*, 115:121–186, 2011, arXiv:1002.2210.
[HB87] D. R. Heath-Brown. Integer sets containing no arithmetic progressions. *J. London Math. Soc. (2)*, 35(3):385–394, 1987.
[KSV09] D. Král, O. Serra, and L. Vena. A combinatorial proof of the removal lemma for groups. *J. Combin. Theory Ser. A*, 116(4):971–978, 2009.
[Mes95] R. Meshulam. On subsets of finite abelian groups with no 3-term arithmetic progressions. *J. Combin. Theory Ser. A*, 71(1):168–172, 1995.
[Pyb97] L. Pyber. How abelian is a finite group? In *The mathematics of Paul Erdős, I*, volume 13 of *Algorithms Combin.*, pages 372–384. Springer, Berlin, 1997.
[Rot52] K. F. Roth. Sur quelques ensembles d'entiers. *C. R. Acad. Sci. Paris*, 234:388–390, 1952.
[Rot53] K. F. Roth. On certain sets of integers. *J. London Math. Soc.*, 28:104–109, 1953.
[RS78] I. Z. Ruzsa and E. Szemerédi. Triple systems with no six points carrying three triangles. In *Combinatorics (Proc. Fifth Hungarian Colloq., Keszthely, 1976), Vol. II*, volume 18 of *Colloq. Math. Soc. János Bolyai*, pages 939–945. North-Holland, Amsterdam, 1978.
[Sol13] J. Solymosi. Roth-type theorems in finite groups. *European J. Combin.*, 34(8):1454–1458, 2013.
[Sze90] E. Szemerédi. Integer sets containing no arithmetic progressions. *Acta Math. Hungar.*, 56(1-2):155–158, 1990.

[TV06] T. C. Tao and H. V. Vu. *Additive combinatorics*, volume 105 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2006.

Mathematical Institute, University of Oxford, Radcliffe Observatory Quarter, Woodstock Road, Oxford OX2 6GG, United Kingdom

E-mail address: tom.sanders@maths.ox.ac.uk