SETS AVOIDING THREE-TERM ARITHMETIC PROGRESSIONS

TRISTAN SHIN

Abstract. In this expository essay, we provide an account of the recent breakthrough of Kelley and Meka on the size of subsets of $\{1, \dots, N\}$ with no nontrivial three-term arithmetic progressions. With an improvement by Bloom and Sisask, the size of such a set must be at most $N/\exp(C(\log N)^{1/9})$ for some constant $C > 0$. This upper bound matches the shape of the size of the largest known such sets, up to the power of the $\log N$ term in the exponent. We first discuss the corresponding problem over a finite field vector space, which provides all of the key ideas, before presenting the details of the full proof over the integers. We organize the arguments for both settings in the same manner as to draw attention to the parallels between the two proofs.

Contents

Centre for Mathematical Sciences, Wilberforce Road, Cambridge CB3 0WB, United Kingdom
E-mail address: ts926@cam.ac.uk.

1. Introduction

In 1936, Erdős and Turán [ET36] posed a conjecture: every set of positive integers with positive upper density[1] contains a nontrivial **three-term arithmetic progression** (3-AP)—that is, a set of the form $\{x, x+d, x+2d\}$ with $d \neq 0$. In 1952, Roth [Rot52] proved this conjecture, which is now known as Roth's theorem. A qualitative equivalent is to consider subsets of $\{1, \dots, N\}$ for some integer $N \geq 1$ and determine a density threshold for which a subset must have a nontrivial three-term arithmetic progression. Let $r_3(N)$ denote the maximum size of a subset of $\{1, \dots, N\}$ with no nontrivial three-term arithmetic progressions. Then Roth's theorem states the following[2].

Theorem 1.1 (Roth)**.** *We have that* $r_3(N) = o(N)$.

A natural follow-up question is to ask for a more precise asymptotic on $r_3(N)$. The proof by Roth [Rot53] using Fourier-analytic methods established a bound of $r_3(N) \lesssim N/\log\log N$. Szemerédi (unpublished; presented in a seminar in 1985) improved this to $r_3(N) \lesssim_c N/(\log\log N)^c$ for any $c > 0$; Balog optimised the approach to improve to $r_3(N) \lesssim N/\exp(\Omega((\log\log N)^{1/2}))$.

Heath-Brown [HB87] applied the circle method and large sieve from analytic number theory to improve this to $r_3(N) \lesssim N/(\log N)^c$ for some effective constant $c > 0$; Szemerédi [Sze90] demonstrated this with $c = 1/4$. Bourgain [Bou99] used Bohr sets to improve this to $r_3(N) \lesssim N(\log\log N)^{O(1)}/(\log N)^c$ for $c = 1/2$ and later to $c = 2/3$ [Bou08]; Sanders [San12a] improved this to $c = 3/4$ and later $c = 1$ with an $O(1)$ term of 6 [San11]. For $c = 1$, Bloom [Blo16] reduced the exponent of the $\log\log N$ term to 4; Schoen [Sch21] improved it further to $3 + o(1)$.

In 2020, Bloom and Sisask [BS20] broke the logarithmic barrier by proving an upper bound of $r_3(N) \lesssim N/(\log N)^{1+c}$ for some effective constant $c > 0$. Their proof was a difficult adaptation of the method used by Bateman and Katz [BK12] to prove a similar result over $\mathbb{F}_3^n$ for integers $n \geq 1$. This upper bound on $r_3(N)$ also implies that if the sum of the reciprocals of a subset of the positive integers diverges, then the set contains infinitely many nontrivial three-term arithmetic progressions. This provides the first nontrivial case of a conjecture of Erdős that such sets contain arbitrarily long arithmetic progressions.

In a major breakthrough in 2023, Kelley and Meka [KM23] broke the quasi-polynomial barrier in the upper bound by proving that $r_3(N) \lesssim N/\exp(\Omega((\log N)^c))$ for $c = 1/12$. A few months later, Bloom and Sisask [BS23a] increased this to $c = 1/9$ by improving one step of the Kelley–Meka method, also claiming that $c = 5/41$ was doable with additional technical work.

Theorem 1.2. *We have that*

$$r_3(N) \leq N/\exp(\Omega((\log N)^c))$$

for $c = 1/9$.

It is also natural to ask how tight the bounds on $r_3(N)$ are, i.e. provide lower bounds. Salem and Spencer [SS42] used a digital construction to prove that $r_3(N) \gtrsim_\epsilon N/N^{(\log 2+\epsilon)/\log\log N}$ for any $\epsilon > 0$. Behrend [Beh46] used the fact that spheres in any dimension avoid three-term arithmetic progressions to improve the lower bound to $r_3(N) \gtrsim N(\log N)^{-1/4}/\exp(c\sqrt{\log N})$ for $c = 2\sqrt{2\log 2} \approx 2.35$. Ever since Behrend's construction in 1946, the lower bound has retained the form of $N/\exp(O((\log N)^{1/2}))$, with improvements coming only in lower-order terms. Elkin [Elk11] used a thin annulus instead of a sphere to replace the $(\log N)^{-1/4}$ with a $(\log N)^{1/4}$. Hunter [Hun24] applied techniques of Elsholtz, Proske, and Sauermann [EPS24] to improve the lower bound to $r_3(N) \gtrsim_c N(\log N)^{-1}/\exp(c\sqrt{\log N})$ for any $c > 2\sqrt{\log(32/9)} \approx 2.25$, in the first quasi-polynomial improvement to Behrend's construction.

The Kelley–Meka result thus provided the first upper bound that came in the same quasi-polynomial shape as the Behrend lower bound, closing a significant amount of the gap between

[1]The **upper density** of a subset A of the positive integers is $\limsup_{N\to\infty} \frac{|A \cap \{1, \dots, N\}|}{N}$.

[2]See Section 2.1.5 for asymptotic notation.

the upper and lower bounds for $r_3(N)$. It is still an open problem to fully close the gap on the power of the $\log N$ term in the exponent.

The Kelley–Meka argument utilises the density increment method of Roth, which is central to many proofs of quantitative bounds for Roth's theorem and other problems in extremal additive combinatorics problems. The general idea is straightforward:

(1) Start by assuming that a set A has no nontrivial three-term arithmetic progressions.
(2) Then the number of three-term arithmetic progressions, which can be counted by a sum of indicators, must be far from what a "random" set with the same density would have.
(3) With some control on the count, this leads to some large quantitative structure—for example, in Roth's original proof, this comes in the form of a large Fourier coefficient. This is a case of what is often referred to as the **dichotomy of structure versus pseudorandomness** [Tao07, Zha23] which prevails across many subfields of combinatorics.
(4) Then this quantitative structure is used to exhibit some structured subset on which A has sufficiently larger density. In vector spaces, such a subset would take the form of an affine subspace (i.e. a coset of a subgroup).
(5) Finally, this passing to structured subsets is iterated until no longer possible, in which case there must be many progressions. The number of iterations is controlled by the size of the density increase, while the size of the original set A is controlled by the gap to the subset that we pass to as well as the number of iterations.

For a clear example of this broken down in Roth's original argument, see [Zha23, Chapter 6]. The details and quantitative bounds of the mechanisms in this method determine how effective the final bounds on $|A|$ are. For much of the history of this problem, most of the structural work was done through Fourier-analytic techniques. One surprise of the novel Kelley–Meka method is that the primary improvement comes on the physical side.

In this essay, we present the Bloom–Sisask simplification [BS23b] of the Kelley–Meka method with the improvement to $c = 1/9$. In Section 2, we provide preliminaries for the method. In Section 3, we present the argument as applied to $\mathbb{F}_q^n$ for an odd prime q and integer $n \geq 1$. Over finite field vector spaces, the argument contains the same main ideas but is technically simpler. Then in Section 4, we provide the full argument in the integer setting to prove Theorem 1.2. In both Sections 3 and 4, we first provide an outline of the proof with statements of key lemmas before proving each step thoroughly. In Section 5, we discuss several related problems which have been tackled using similar techniques.

1.1. **Acknowledgements.** The author would like to thank W. T. Gowers for providing useful guidance and discussion throughout the process of writing this essay, as well as suggesting the topic. This essay was written in partial fulfillment of the requirements for the degree of Master of Advanced Study in Mathematics at the University of Cambridge.

2. Preliminaries

2.1. **Notation.** In this essay, G will refer to a finite abelian group of odd order[3] with addition as its group operation. There are only two types of groups G that will show up:

- $G = \mathbb{F}_q^n$ for some odd prime q and integer $n \geq 1$. In this case, it is useful to think of q as fixed and n as growing to infinity.
- $G = \mathbb{Z}/N\mathbb{Z}$ for some odd integer $N \geq 1$. Again, it is useful to think of N as growing to infinity. This N will not necessarily be the same as the N in the statement of Theorem 1.2, but it will grow in the same manner.

Similarly, V will refer to a finite-dimensional vector space over $\mathbb{F}_q$ for some odd prime q. We withhold from just writing $\mathbb{F}_q^n$ because we will often change which dimension we are working in by taking subspaces.

2.1.1. *Sets.* Let S be a finite set. The **indicator function** of S is defined by

$$\mathbb{1}_S(x) \coloneqq \begin{cases} 1 & \text{if } x \in S \\ 0 & \text{otherwise.} \end{cases}$$

For any function f whose domain contains S, let

$$\mathop{\mathbb{E}}_{x \in S} f(x) = \frac{1}{|S|} \sum_{x \in S} f(x).$$

For shorthand, let $\mathbb{E}$ be the functional that sends a function to its expectation over G, i.e.

$$\mathbb{E}\, f = \mathop{\mathbb{E}}_{x \in G} f(x).$$

For $A \subseteq G$, the **density** of A is $\frac{|A|}{|G|}$. If $A \subseteq B$, the **relative density** of A (with respect to B) is $\frac{|A|}{|B|}$. The **normalised indicator function** of A is $\mu_A \coloneqq \alpha^{-1}\, \mathbb{1}_A$, where α is the density of A. Note that

$$\mathop{\mathbb{E}}_{x \in A} f(x) = \mathop{\mathbb{E}}_{x \in G} \mu_A(x) f(x)$$

for all $f \colon G \to \mathbb{C}$.

For sets A and B, define

$$\begin{aligned} -A &= \{-a \,:\, a \in A\} \\ A + B &= \{a + b \,:\, a \in A, b \in B\} \\ A - B &= \{a - b \,:\, a \in A, b \in B\}. \end{aligned}$$

For shorthand, let $kA = \underbrace{A + \cdots + A}_{k \text{ times}}$ for a positive integer k. This is not to be confused with

$$k \cdot A = \{ka \,:\, a \in A\},$$

where $ka = \underbrace{a + \cdots + a}_{k \text{ times}}$.

For a function f with codomain $\mathbb{R}$, let $\{f > c\}$ be shorthand for the set of values x in the domain of f for which $f(x) > c$, and similarly with other inequalities.

2.1.2. *Functions.* We will be working with the Hilbert space of functions from G to $\mathbb{C}$ with "normalised counting measure". For $f, g \colon G \to \mathbb{C}$, define the following:

- **inner product**:

$$\langle f, g \rangle \coloneqq \mathop{\mathbb{E}}_{x \in G} f(x) \overline{g(x)}$$

[3] Most results still hold when G is just a finite abelian group, but it will be easier to assume that no element has order 2—otherwise three-term arithmetic progressions can behave unusually.

- L^p**-norm** for $p \geq 1$:

$$\|f\|_p \coloneqq \left(\mathop{\mathbb{E}}_{x \in G} |f(x)|^p \right)^{\frac{1}{p}}$$

- **convolution**:

$$(f * g)(x) \coloneqq \mathop{\mathbb{E}}_{y \in G} f(y) g(x - y)$$

- **cross-correlation**[4]:

$$(f \star g)(x) \coloneqq \mathop{\mathbb{E}}_{y \in G} \overline{f(y)} g(x + y)$$

We also define the L^∞-norm in the usual way over a finite domain:

$$\|f\|_\infty \coloneqq \max_{x \in G} |f(x)| .$$

Then L^p-norms are monotonically increasing in p for $p \in [1, \infty]$ by convexity. Note that this is the *opposite* direction as for L^p-norms over $\mathbb{C}$, but is the *same* direction as over probability spaces (which the normalised counting measure provides).

Note that $\langle f, g \rangle = \overline{(f \star g)(0)}$. If $f, g \colon G \to \mathbb{R}$, then we can drop all the conjugations.

It is straightforward to check that convolution is associative and commutative, but cross-correlation is neither. For shorthand, let $f^{*k} = \underbrace{f * \cdots * f}_{k \text{ times}}$ for any integer $k \geq 1$.

For $t \in G$, let τ_t denote the **translation operator** by t, defined by $\tau_t f(x) = f(x + t)$ for all $x \in G$. Let $\mathcal{N}$ denote the **negation operator**, defined by $\mathcal{N} f(x) = f(-x)$ for all $x \in G$. Note that $\mathcal{N}$ is an involution, and that inner products and norms are preserved under translation and negation.

One can check the following for $f, g, h \colon G \to \mathbb{C}$.

$$\begin{aligned}
g \star f &= \overline{\mathcal{N}(f \star g)} \\
f \star g &= \overline{\mathcal{N} f} * g = \overline{\mathcal{N}(f * \overline{\mathcal{N} g})} \\
f * g &= \overline{\mathcal{N} f} \star g = \overline{\mathcal{N}(f \star \overline{\mathcal{N} g})} \\
(f \star g) \star h &= f * (g \star h) \\
f \star (g \star h) &= (f * g) \star h \\
(f \star g) * h &= f \star (g * h) \\
\langle f, g \star h \rangle &= \langle f * g, h \rangle = \overline{\langle f \star h, g \rangle} \\
\langle f, g * h \rangle &= \langle g \star f, h \rangle = \langle h \star f, g \rangle .
\end{aligned}$$

As before, if $f, g, h \colon G \to \mathbb{R}$, then we can drop all the conjugations.

2.1.3. *Probability measures.* We will also need to work with nonuniform measures on G. In a bit of nonstandard notation, we say that $\mu \colon G \to \mathbb{R}_{\geq 0}$ is a **probability measure** on G if $\mathbb{E}\, \mu = 1$. Then we can define the inner product and norm under this measure:

- μ**-inner product**:

$$\langle f, g \rangle_\mu \coloneqq \mathop{\mathbb{E}}_{x \in G} \mu(x) f(x) \overline{g(x)}$$

- $L^p(\mu)$**-norm**:

$$\|f\|_{L^p(\mu)} \coloneqq \left(\mathop{\mathbb{E}}_{x \in G} \mu(x) \, |f(x)|^p \right)^{\frac{1}{p}}$$

[4]Bloom and Sisask [BS23b] refer to this as the **difference convolution** and notate it as $f \circ g$. The notation we use here is the same as that of Kelley and Meka, which is borrowed from signal processing.

Similarly, we have the $L^\infty(\mu)$-norm:

$$\|f\|_{L^\infty(\mu)} \coloneqq \max_{x \in \operatorname{supp} \mu} |f(x)|\,.$$

Again, the $L^p(\mu)$-norms are monotonically increasing in p.

In an abuse of notation, for a set $S \subseteq G$, let $\mu(S) = \|\mathbb{1}_S\|_{L^1(\mu)}$ denote the density of S under the weighting μ. This makes $(G, 2^G, \mu)$ a probability triple in the standard measure-theoretic sense. Note that $\mu(\{x\}) \neq \mu(x)$ as the former has the normalisation factor of $\frac{1}{|G|}$.

Note that μ_S is a probability measure for any $S \subseteq G$, corresponding to the uniform distribution on S. Also, for any probability measures μ_1 and μ_2, the convolution $\mu_1 * \mu_2$ and cross-correlation $\mu_1 \star \mu_2$ are both also probability measures. These correspond to the distributions of $X_1 + X_2$ and $X_2 - X_1$, respectively, where $X_1 \sim \mu_1$ and $X_2 \sim \mu_2$. In particular, expressions such as $(\mu_{B''} \star \mu_{B''}) * (\mu_{B'''} \star \mu_{B'''})$ for $B'', B''' \subseteq G$ are probability measures.

2.1.4. *Dual group.* We will also use the **(Pontryagin) dual group** of G, defined as

$$\widehat{G} \coloneqq \operatorname{Hom}(G, \mathbb{T}) = \{\text{homomorphisms from } G \text{ to } \mathbb{T}\}\,,$$

where $\mathbb{T} = \{z \in \mathbb{C} : |z| = 1\} \cong \mathbb{R}/\mathbb{Z}$. Note that $\widehat{G}$ has multiplication as its group operation. We will write χ_0 for the **trivial character** which evaluates to 1 on all of G. Then χ_0 is the identity element of $\widehat{G}$.

It is not difficult to show that $G \cong \widehat{G}$ for any finite abelian group G. In the groups that we care about, the dual groups can be shown to be the following:

- Let $G = \mathbb{Z}/N\mathbb{Z}$. For $r \in G$, let $\chi_r(x) = \zeta^{rx}$, where $\zeta = e^{2\pi i/N}$. Then
$$\widehat{G} = \{\chi_r : r \in G\}\,.$$
- More generally, this construction generalises for $G = (\mathbb{Z}/N\mathbb{Z})^m$ (and thus $\mathbb{F}_q^n$). For $\mathbf{r} \in G^m$, let $\chi_{\mathbf{r}}(\mathbf{x}) = \zeta^{\mathbf{r}\cdot\mathbf{x}}$, where $\cdot$ denotes the standard dot product. Then
$$\widehat{G} = \{\chi_{\mathbf{r}} : \mathbf{r} \in G^m\}\,.$$
In fact, by changing the root of unity, this generalises to all finite abelian groups upon appealing to the structure theorem.

We will work with the Hilbert space of functions from $\widehat{G}$ to $\mathbb{C}$ with (unnormalised) counting measure, which means that expectations in the definitions of inner product, norm, convolution, and cross-correlation are replaced by sums. Note that this means that L^p-norms are monotonically *decreasing* instead.

2.1.5. *Asymptotic notation.* We use $f \lesssim g$, $f = O(g)$, and $g = \Omega(f)$ to denote $|f| \leq Cg$ for some constant $C > 0$. We use $f = o(g)$ to denote that $f/g \to 0$ as the argument tends to ∞. Subscripts denote that the hidden constant may depend on these parameters. For example, $f = O_\epsilon(g)$ means that for every ϵ, we have that $|f| \leq C_\epsilon g$ for some constant $C_\epsilon > 0$.

2.2. **Fourier transform.** For a function $f\colon G \to \mathbb{C}$, its **Fourier transform** $\hat{f}\colon \widehat{G} \to \mathbb{C}$ is given by

$$\hat{f}(\chi) \coloneqq \langle f, \chi\rangle = \mathop{\mathbb{E}}_{x \in G} f(x)\chi(-x).$$

Note that the Fourier transform is a linear operator. The Fourier transform translates nicely between multiplication, convolution, and cross-correlation.

Proposition 2.1. *Let $f, g\colon G \to \mathbb{C}$. Then $\widehat{f * g} = \hat{f} \cdot \hat{g}$, $\widehat{f \star g} = \overline{\hat{f}} \cdot \hat{g}$, and $\widehat{f \cdot g} = \hat{f} * \hat{g}$.*

Corollary 2.2. *Let $f\colon G \to \mathbb{C}$. Then:*

- *$\widehat{f * f} = \hat{f}^2$ and $\widehat{f \star f} = |\hat{f}|^2$; and*
- *$\widehat{f^{*k}} = \hat{f}^k$ and $\widehat{f^k} = \hat{f}^{*k}$ for any integer $k \geq 1$.*

The Fourier transform also acts as the coefficients when expanding functions in the orthonormal (up to uniform scaling) basis of characters $\chi \in \widehat{G}$. This gives the following two fundamental facts about the Fourier transform.

Proposition 2.3 (Fourier inversion)**.** *Let $f\colon G \to \mathbb{C}$. Then*

$$f = \sum_{\chi \in \widehat{G}} \hat{f}(\chi)\chi.$$

Proposition 2.4 (Parseval)**.** *Let $f, g\colon G \to \mathbb{C}$. Then*

$$\langle f, g\rangle = \langle \hat{f}, \hat{g}\rangle.$$

In particular,

$$\|f\|_2 = \|\hat{f}\|_2.$$

Remark 2.5. Recall that the inner product, norm, and convolution over the dual group is defined using the counting measure rather than the normalised counting measure, so for example $\widehat{f \cdot g} = \hat{f} * \hat{g}$ means that

$$\mathop{\mathbb{E}}_{x \in G} f(x)g(x)\chi(-x) = \sum_{\gamma \in \widehat{G}} \hat{f}(\gamma)\hat{g}(\chi/\gamma).$$

Each of these facts can be proven by expanding and possibly applying the facts that

$$\mathop{\mathbb{E}}_{x \in G} \chi(x) = \begin{cases} 1 & \text{if } \chi = \chi_0 \\ 0 & \text{otherwise} \end{cases}$$

and

$$\sum_{\chi \in \widehat{G}} \chi(x) = \begin{cases} |G| & \text{if } x = 0 \\ 0 & \text{otherwise.} \end{cases}$$

The Fourier inversion formula also allows us to cleanly read off how the Fourier transform interacts with other operators. For example, for all $x, t \in G$, we have that

$$\tau_t f(x) = \sum_{\chi \in \widehat{G}} \hat{f}(\chi)\chi(x+t) = \sum_{\chi \in \widehat{G}} \chi(t)\hat{f}(\chi)\chi(x),$$

so $\widehat{\tau_t f} = \chi(t)\hat{f}$.

If a function has nonnegative Fourier transform everywhere on $\widehat{G}$, we say that it is **spectrally nonnegative**.

For $0 < \lambda \leq 1$ and $f\colon G \to \mathbb{C}$, the **λ-large spectrum** of f is

$$\operatorname{Spec}_\lambda(f) \coloneqq \left\{\chi \in \widehat{G} : |\hat{f}(\chi)| \geq \lambda \|f\|_1\right\}.$$

Since translating a function by t multiplies its Fourier transform by $\chi(t)$, it follows that $\operatorname{Spec}_\lambda(\tau_t f) = \operatorname{Spec}_\lambda(f)$ for all $t \in G$.

2.3. **Bohr sets.** For nonempty $\Gamma \subseteq \widehat{G}$ and $\rho > 0$, define the **Bohr set** $B = \operatorname{Bohr}(\Gamma, \rho)$ to be

$$\operatorname{Bohr}(\Gamma, \rho) \coloneqq \{x \in G : |1 - \chi(x)| \leq \rho \text{ for all } \chi \in \Gamma\}.$$

We call Γ the **frequency set** of B, $\operatorname{rank}(B) \coloneqq |\Gamma|$ the **rank** of B, and $\operatorname{radius}(B) \coloneqq \rho$ the **radius** of B.[5]

Bohr sets are commonly used in additive combinatorics as a structured subset of an arbitrary finite abelian group. In vector spaces, we have subspaces. But in groups such as $\mathbb{Z}/N\mathbb{Z}$, it is possible to have very few subgroups. So we need some other notion of structure. The original proof of Roth's theorem used arithmetic progressions contained in $\{1, \dots, N\}$, which is somewhat of a predecessor to the concept of Bohr sets in additive combinatorics. Bohr sets are

[5]Tao and Vu (and some other sources) define Bohr sets a bit differently by using the angle along the unit circle rather than the distance. This affects the radius up to an irrelevant constant factor, so statements of results here may appear different compared to other sources.

not original to combinatorics; Harald Bohr[6] used them in the early twentieth century to study almost-periodic functions on $\mathbb{R}$ using Fourier analysis.

The following results are standard facts about Bohr sets and can be found in either the book of Tao and Vu [TV06, Section 4.4] or the prior work of Bloom and Sisask [BS20, Section 4].

Remark 2.6. Bohr sets can be viewed as a generalization of subspaces. Indeed, if $V \leq \mathbb{F}_q^n$, then $V = \mathrm{Bohr}(\Gamma, \rho)$, where:

- Γ is a set of characters that correspond to a basis of $V^\perp$ under an isomorphism between V and its dual group; and
- $\rho < 2 \sin \frac{\pi}{q}$.

Note that $\mathrm{rank}(V) = \mathrm{codim}\, V$ under this interpretation.

Since characters have magnitude 1, it is easy to verify that Bohr sets are symmetric. Similarly, one can check that

$$\mathrm{Bohr}(\Gamma, \rho_1) + \mathrm{Bohr}(\Gamma, \rho_2) \subseteq \mathrm{Bohr}(\Gamma, \rho_1 + \rho_2)$$

for all $\rho_1, \rho_2 > 0$. Slightly less obvious is that if k is a positive integer relatively prime to $|G|$ and B is a Bohr set, then $k \cdot B$ is also a Bohr set of the same rank and radius—one can adjust the frequency set accordingly.

For $\lambda > 0$, the **dilate** of a Bohr set $B = \mathrm{Bohr}(\Gamma, \rho)$ is

$$B_\lambda = \mathrm{Bohr}(\Gamma, \lambda\rho).$$

Note that the B_λ are increasing in λ.

By a pigeonhole argument, one can establish that Bohr sets have considerable size.

Lemma 2.7. *Let $B \subseteq G$ be a Bohr set and $0 < \lambda \leq 1$. Then $B_\lambda \subseteq B$ has relative density at least $(\lambda/4)^{\mathrm{rank}(B)}$. In particular, B has density at least $(\mathrm{radius}(B)/8)^{\mathrm{rank}(B)}$.*

See [BS20, Lemma 4.4] for a full proof (also [TV06, Lemma 4.20] for the bound on the density of B with the constant 8 replaced with 2π).

Bohr sets in general may exhibit bad additive structure, but it turns out that all Bohr sets are somewhat close to one that has good structure. We say that a Bohr set B of rank r is **regular** if

$$(1 - 100r\,|\kappa|)\,|B| \leq |B_{1+\kappa}| \leq (1 + 100r\,|\kappa|)\,|B|$$

for all $-\frac{1}{100r} \leq \kappa \leq \frac{1}{100r}$. This essentially means that adding a small dilate of B to B does not expand B by too much if B is regular.

Lemma 2.8. *Let $B \subseteq G$ be a Bohr set. There exists some $\frac{1}{2} \leq \lambda \leq 1$ so that B_λ is regular.*

See [TV06, Lemma 4.25] for a full proof.

It is also straightforward to check that if B is a regular Bohr set, then $k \cdot B$ is also regular. As one might expect, additional additive structure can be compelled of regular Bohr sets. The following is an effective version of [BS20, Lemma 4.5].

Lemma 2.9. *Let $B \subseteq G$ be a regular Bohr set of rank $r \geq 1$. Let μ be a probability measure such that $\mathrm{supp}\,\mu \subseteq B_\lambda$ for some $0 < \lambda < 1$. Then*

$$\|\mu_B * \mu - \mu_B\|_1 \leq 200\lambda r.$$

Proof. If $\lambda \geq \frac{1}{100r}$, then

$$\|\mu_B * \mu - \mu_B\|_1 \leq \|\mu_B * \mu\|_1 + \|\mu_B\|_1 = 2 \leq 200\lambda r.$$

[6] Harald Bohr's older brother Niels Bohr also lends his name to a notion of "Bohr radius" that is perhaps more famous.

Assume otherwise. We have that

$$\begin{aligned}\|\mu_B * \mu - \mu_B\|_1 &= \mathop{\mathbb{E}}_{x\in G}\left|\mathop{\mathbb{E}}_{y\in G}\mu_(y)\mu_B(x-y) - \mu_B(x)\right| \\ &\le \mathop{\mathbb{E}}_{x\in G}\mathop{\mathbb{E}}_{y\in G}\mu(y)\left|\mu_B(x-y) - \mu_B(x)\right| \\ &= \left\|\mathop{\mathbb{E}}_{x\in G}\left|\mu_B(x-\,\cdot\,) - \mu_B(x)\right|\right\|_{L^1(\mu)} \\ &= \left\|\mathop{\mathbb{E}}_{x\in G}\left|\mu_B(x-\,\cdot\,) - \mu_B(x)\right|\right\|_{L^\infty(\mu)}.\end{aligned}$$

Observe that

$$\mathop{\mathbb{E}}_{x\in G}\left|\mu_B(x-\,\cdot\,) - \mu_B(x)\right| = \frac{1}{|B|}\sum_{x\in G}\left|\mathbb{1}_B(x-y) - \mathbb{1}_B(x)\right| = \frac{|(B+y)\setminus B| + |B\setminus(B+y)|}{|B|}.$$

Let $y \in \operatorname{supp}\mu \subseteq B_\lambda$. Then $B + y \subseteq B + B_\lambda \subseteq B_{1+\lambda}$, so by regularity, we have that

$$|(B+y)\setminus B| \le |B_{1+\lambda}\setminus B| = |B_{1+\lambda}| - |B| \overset{\text{reg.}}{\le} 100r\lambda\,|B|\,.$$

Similarly, $B_{1-\lambda} - y \subseteq B_{1-\lambda} + B_\lambda \subseteq B$, so

$$|B\setminus(B+y)| \le |B\setminus B_{1-\lambda}| = |B| - |B_{1-\lambda}| \overset{\text{reg.}}{\le} 100r\lambda\,|B|\,.$$

The conclusion follows. □

One can use this additive structure to produce good translated structure when looking at any sets. The following is a "narrowing" trick due to Bourgain, as stated in [BS20, Lemma 12.1].

Lemma 2.10. *Let $0 < \epsilon < 1$. Let B be a regular Bohr set of rank $r \ge 1$, and let $A \subseteq B$ have relative density $\alpha > 0$. Let $B_1, B_2 \subseteq B_\lambda$ for some $0 < \lambda \le \frac{\epsilon\alpha}{800r}$. Then one of the following alternatives must hold:*

(1) There exists $d \in B$ such that

$$\frac{|(d-A)\cap B_1|}{|B_1|} \ge (1+\tfrac{1}{2}\epsilon)\alpha \quad \textit{or} \quad \frac{|(d-A)\cap B_2|}{|B_2|} \ge (1+\tfrac{1}{2}\epsilon)\alpha.$$

(2) There exists $d \in B$ such that

$$\frac{|(d-A)\cap B_1|}{|B_1|} \ge (1-\epsilon)\alpha \quad \textit{and} \quad \frac{|(d-A)\cap B_2|}{|B_2|} \ge (1-\epsilon)\alpha.$$

Proof. Let B have density $\beta > 0$. Note that for $d \in G$, we have that

$$(\mu_A * \mu_{B_1})(d) = \mathop{\mathbb{E}}_{y\in G}\mu_A(d-y)\mu_{B_1}(y) = (\alpha\beta)^{-1}\frac{|(d-A)\cap B_1|}{|B_1|}$$

and likewise with B_2.

Using the symmetry of B, we have that

$$\begin{aligned}|\langle \mu_A * \mu_{B_1} - \mu_A, \mu_B\rangle| &= |(\mu_A * \mu_{B_1} * \mu_B - \mu_A * \mu_B)(0)| \\ &= |\langle \mu_A, \mu_{B_1} * \mu_B - \mu_B\rangle| \\ &\overset{\text{Hölder}}{\le} \|\mu_A\|_\infty \|\mu_B * \mu_{B_1} - \mu_B\|_1 \\ &\overset{\text{Lem. 2.9}}{\le} (\alpha\beta)^{-1}\cdot 200\lambda r \\ &\le \tfrac{1}{4}\epsilon\beta^{-1}.\end{aligned}$$

The same result holds for B_2. But $\langle \mu_A, \mu_B \rangle = \|\mu_B\|_{L^1(\mu_A)} = \beta^{-1}$, so

$$\begin{aligned}\|\mu_A * \mu_{B_1} + \mu_A * \mu_{B_2}\|_{L^\infty(\mu_B)} &\geq \|\mu_A * \mu_{B_1} + \mu_A * \mu_{B_2}\|_{L^1(\mu_B)} \\ &= \langle \mu_A * \mu_{B_1}, \mu_B \rangle + \langle \mu_A * \mu_{B_2}, \mu_B \rangle \\ &\geq (2 - \tfrac{1}{2}\epsilon)\beta^{-1}.\end{aligned}$$

Thus there exists some $d \in B$ for which $(\mu_A * \mu_{B_1})(d) + (\mu_A * \mu_{B_2})(d) \geq (2 - \frac{1}{2}\epsilon)\beta^{-1}$.

Suppose that the first alternative fails. Then

$$\begin{aligned}(\mu_A * \mu_{B_1})(d) &\geq (2 - \tfrac{1}{2}\epsilon)\beta^{-1} - (\mu_A * \mu_{B_2})(d) \\ &> (2 - \tfrac{1}{2}\epsilon)\beta^{-1} - (1 + \tfrac{1}{2}\epsilon)\beta^{-1} \\ &= (1 - \epsilon)\beta^{-1}\end{aligned}$$

and similarly with B_2, as desired. □

The following is essentially the same as [BS20, Lemma 4.6].

Lemma 2.11. *Let $B \subseteq G$ be a regular Bohr set of rank $r \geq 1$, $k \geq 1$ be an integer, and $0 < \lambda \leq \frac{1}{100kr}$. If μ is a probability measure such that $\operatorname{supp}\mu \subseteq kB_\lambda$, then*

$$\mu_B(x) \leq 2(\mu_{B_{1+k\lambda}} * \mu)(x)$$

for all $x \in G$.

Proof. If $x \in B$ and $y \in \operatorname{supp}\mu \subseteq kB_\lambda$, then $x - y \in B + kB_\lambda \subseteq B_{1+k\lambda}$. It follows that for all $x \in B$, we have that

$$(\mu_{B_{1+k\lambda}} * \mu)(x) = \mathop{\mathbb{E}}_{y \in G} \mu(y) \cdot \frac{|G|}{|B_{1+k\lambda}|} = \frac{|G|}{|B_{1+k\lambda}|} \overset{\text{reg.}}{\geq} \frac{|G|}{2\,|B|} = \mu_B(x).$$

For $x \notin B$, the result is trivial. □

3. Finite field model

In many problems in additive combinatorics over the integers, the corresponding problem over a finite field vector space is often considerably easier to tackle. For example, the original proof of Roth's theorem was adapted by Meshulam [Mes95] to prove a similar result over $\mathbb{F}_q^n$ where q is an odd prime and $n \geq 1$ is an integer.

In an abuse of notation, let $r_3(\mathbb{F}_q^n)$ denote the maximum size of a subset of $\mathbb{F}_q^n$ with no nontrivial three-term arithmetic progressions. Meshulam proved that $r_3(\mathbb{F}_q^n) \lesssim q^n/n$. The Fourier-analytic proof that Meshulam used is quite straightforward compared to Roth's proof over the integers. The primary reason for this is that the strategy involves passing to a highly structured subset of the original ambient set. Over $\mathbb{F}_q^n$, we have a tremendous subspace structure which works perfectly for the argument. Over the integers, we don't quite have anything as nice.

As a result, it is useful to test out methods in this finite field vector space setting before applying them in the integer setting. This is the **finite field model**. This strategy is expanded upon in great detail in the surveys of Green [Gre05], Wolf [Wol15], and Peluse [Pel23]. For the problem at hand, the finite field model has already been quite successful. Bateman and Katz [BK12] broke the logarithmic barrier over $\mathbb{F}_3^n$ by proving that $r_3(\mathbb{F}_3^n) \lesssim 3^n/n^{1+c}$ for some effective constant $c > 0$. Following this, some ideas from their method were used by Schoen [Sch21] to prove that $r_3(N) \lesssim N(\log\log N)^{3+o(1)}/\log N$. Then, Bloom and Sisask [BS20] fully adapted the method to break the logarithmic barrier over the integers.

The corresponding result that the Kelley–Meka method achieves in the finite field model is the following:

Theorem 3.1. *We have that*

$$r_3(\mathbb{F}_q^n) \leq q^n/\exp(\Omega(n^c))$$

for $c = 1/7$.

The original Kelley–Meka argument gave $c = 1/9$, while this $c = 1/7$ version is due to the Bloom–Sisask improvement. Note that for fixed primes q, this is the same shape as in Theorem 1.2 since N is replaced by q^n and $\log N$ is replaced by n.

We note that in the finite field model, the current best upper bound is much stronger than the quasi-polynomial shape given by the Kelley–Meka method. Indeed, Ellenberg and Gijswijt [EG17] used the polynomial method as prescribed in the breakthrough of Croot, Lev, and Pach [CLP17] to prove that for sufficiently large primes q, $r_3(\mathbb{F}_q^n) \leq (cq)^n$ for $c \approx 0.85$.

For lower bounds in the finite field model, a variant of the Salem–Spencer [SS42] or Behrend [Beh46] construction gives that $r_3(\mathbb{F}_q^n) \geq (cq)^{n-o(n)}$ for $c = 1/2$. Recent work by Elsholtz, Proske, and Sauermann [EPS24] improved this lower bound to $c = \sqrt{7/24}$. For fixed primes q, more improvement can be made by constructing suitable sets in fixed dimension and extending these sets to higher dimensions; the current best lower bound of $r_3(\mathbb{F}_3^n) \geq 2.2202^n$ by Romera-Paredes et al. [RPBN$^+$24] applies large language models (FunSearch) to improve the search for such constructions.

So to summarise, we have the following for sufficiently large primes q:

$$\underbrace{(0.54q)^n}_{\text{EPS}} \lesssim r_3(\mathbb{F}_q^n) \lesssim \underbrace{(0.85q)^n}_{\substack{\text{EG}\\ \text{(CLP method)}}} \lesssim \underbrace{q^n/\exp(\Omega(n^{1/7}))}_{\substack{\text{BS}\\ \text{(KM method)}}} \lesssim \underbrace{q^n/n}_{\text{Meshulam}};$$

and the following for $q = 3$:

$$\underbrace{(2.22)^n}_{\text{FunSearch}} \lesssim r_3(\mathbb{F}_3^n) \lesssim \underbrace{(2.76)^n}_{\substack{\text{EG}\\ \text{(CLP method)}}} \lesssim \underbrace{3^n/\exp(\Omega(n^{1/7}))}_{\substack{\text{BS}\\ \text{(KM method)}}} \lesssim \underbrace{3^n/n}_{\text{Meshulam}}.$$

In order to prove the extremal result of Theorem 3.1, it suffices to prove a counting result.

Theorem 3.2. *Let* $A \subseteq \mathbb{F}_q^n$ *have density* $\alpha > 0$. *Then*

$$\#\{\text{3-APs in } A\} \geq q^{2n - O((1+\log\alpha^{-1})^7)}.$$

Remark 3.3. Such a counting result is akin to **supersaturation** results in extremal graph theory which produce a large number of some structure once the threshold for existence of this structure has been passed. However, it is often the case that graph theoretical supersaturation results are proven separately from extremal results, while here we will prove the extremal result using the counting result.

3.1. **Outline.** The rough idea of the density increment used in the Kelley–Meka method is as follows:

(1) Start with a large deviation from the expected number of progressions, assuming a random set of density α. Use Hölder's inequality to lift to a large value of $\mu_A \star \mu_A - 1$ (on average).
(2) Then, "unbalance" the function to get a large value of $\mu_A \star \mu_A$ (on average).
(3) Apply dependent random choice (a probabilistic technique to find large structured subsets) to correlate $\mu_A \star \mu_A$ with $\mu_{A_1} \star \mu_{A_2}$ for some smaller pieces A_1, A_2.
(4) Apply almost-periodicity to convert this correlation to a density increment.

With the right bounds on each step, iterating this density increment would provide the Kelley–Meka result.

Remark 3.4. To count the number of progressions, we can use $\langle \mu_A * \mu_A, \mu_{2\cdot A} \rangle$. This is because

$$\begin{aligned}\langle \mu_A * \mu_A, \mu_{2\cdot A} \rangle &= \alpha^{-3} \mathop{\mathbb{E}}_{x\in G} \mathop{\mathbb{E}}_{y\in G} \mathbb{1}_A(y)\, \mathbb{1}_A(x-y)\, \mathbb{1}_{2\cdot A}(x) \\ &= \alpha^{-3} \frac{1}{|G|^2} \sum_{x\in 2\cdot A} \#\{a, a' \in A : a + a' = x\} \\ &= \alpha^{-3} \frac{\#\{\text{3-APs in } A\}}{|G|^2}.\end{aligned}$$

Remark 3.5. To measure the size of the density increment to a subspace $U \leq V$, we can use $\|\mu_A * \mu_U\|_\infty$. Indeed, suppose that $\|\mu_A * \mu_U\|_\infty \geq 1 + \epsilon$ for some $\epsilon > 0$. Let $t \in V$ be such that $\mu_A * \mu_U(t) \geq 1 + \epsilon$. Then

$$\frac{|A \cap (t - U)|}{|U|} = \frac{1}{|U|} \sum_{y\in V} \mathbb{1}_A(y)\, \mathbb{1}_U(t - y) \geq (1 + \epsilon)\alpha.$$

But shifting by t and replacing U by $-U$ (since U is a subspace) provides the desired density increment: the set $((A - t) \cap U) \subseteq U$ has relative density at least $(1 + \epsilon)\alpha$.

We now provide an outline of the proof, with statements of the key steps. The sequence of steps roughly follows Bloom and Sisask [BS23b].

Step 1: Lifting. The first step allows us to go from a far-from-random progression count to a large discrepancy between $\mu_A \star \mu_A$ and its expectation. Note that a random set A of density α should have $\#\{\text{3-APs in } A\} \approx \alpha^3 |G|^2$ and thus $\langle \mu_A * \mu_A, \mu_{2\cdot A} \rangle \approx 1$.

Lemma 3.6. *Let $A \subseteq G$ have density $\alpha > 0$. Suppose that*

$$\langle \mu_A * \mu_A, \mu_{2\cdot A} \rangle \leq 1 - \epsilon$$

for some $0 < \epsilon < 1$. Then

$$\|\mu_A \star \mu_A - 1\|_p \geq \tfrac{1}{2}\epsilon$$

for some $p = O(1 + \log \alpha^{-1})$.

Step 2: Unbalancing. We now have an estimate that $\mu_A \star \mu_A - 1$ must be large (on average). It makes more sense to analyse the more symmetric $\mu_A \star \mu_A$. Intuitively, the estimate on $\mu_A \star \mu_A - 1$ suggests that $\mu_A \star \mu_A$ ought to be far from 1. But which direction is not immediately clear—for example, the function -1 is far from 0, but adding 1 results in the 0 function. So we want some kind of "unbalancing" move that results in $\mu_A \star \mu_A$ being large.

Lemma 3.7. *Let $A \subseteq G$. Suppose that*

$$\|\mu_A \star \mu_A - 1\|_p \geq \epsilon$$

for some $0 < \epsilon < 1$ and $p \geq 1$. Then

$$\|\mu_A \star \mu_A\|_{p'} \geq 1 + \tfrac{1}{2}\epsilon$$

for some $p' = O_\epsilon(p)$.

Step 3: Correlation via dependent random choice. With a large average $\mu_A \star \mu_A$, we begin working towards the required density increment by essentially correlating $\mu_A \star \mu_A$ with $\mu_{A_1} \star \mu_{A_2}$ for some much smaller sets A_1, A_2.

Lemma 3.8. *Let $A \subseteq G$ have density $\alpha > 0$. Suppose that*

$$\|\mu_A \star \mu_A\|_p \geq 1 + \epsilon$$

for some $0 < \epsilon < 1$ and integer $p \geq 1$. Then there exist $A_1, A_2 \subseteq G$ of density $\Omega(\alpha^{2p+O_\epsilon(1)})$ such that

$$\langle \mu_{A_1} \star \mu_{A_2}, \mathbb{1}_S \rangle \geq 1 - \tfrac{1}{8}\epsilon,$$

where $S = \{\mu_A \star \mu_A > 1 + \frac{1}{2}\epsilon\}$.

The proof of this step involves the technique of **dependent random choice**, where the sets A_1, A_2 are chosen randomly but with some guidance. This idea stems from several similar arguments often phrased in terms of graphs and common neighbours—see [FS11] for a survey of typical applications. The main motivation for this step, original to the Kelley–Meka method, is to feed into the following almost-periodicity step which has been crucial in recent progress on other additive combinatorics problems.

Remark 3.9. The conclusion of Lemma 3.8 can also be stated as $(\mu_{A_1} \star \mu_{A_2})(S) \geq 1 - \frac{1}{8}\epsilon$ (viewing $\mu_{A_1} \star \mu_{A_2}$ as a probability measure). This allows us to interpret the result as that the $x \in G$ for which $\mu_A \star \mu_A$ is large collectively have many representations of the form $a_2 - a_1$ for $a_1 \in A_1, a_2 \in A_2$.

Step 4: Density increment via almost-periodicity. Several **almost-periodicity** results have been heavily used to improve bounds on related problems in additive combinatorics ever since Croot and Sisask [CS10] introduced the idea. We can use an almost-periodicity result due to Schoen and Sisask [SS16] alongside an improved "bootstrapping" argument due to Bloom and Sisask [BS23a] to establish a density increment. Since we wish to pass to subspaces, we have to be more specific and say that our group G is a vector space now.

Lemma 3.10. *Let V be a finite-dimensional vector space over a field of prime order. Let $A, A_1, A_2, S \subseteq V$ such that A, A_1, A_2 have densities $\alpha, \alpha_1, \alpha_2 > 0$, respectively. Suppose that*

$$\langle \mu_{A_1} \star \mu_{A_2}, \mathbb{1}_S \rangle \geq 1 - \epsilon$$

for some $0 < \epsilon < 1/8$, where $S \subseteq \{\mu_A \star \mu_A \geq 1 + 4\epsilon\}$. Then there exists a subspace $U \leq V$ of codimension

$$O_\epsilon((1 + \log \alpha^{-1})^2 (1 + \log \alpha_1^{-1})(1 + \log \alpha_2^{-1}))$$

such that

$$\|\mu_A * \mu_U\|_\infty \geq 1 + \tfrac{1}{2}\epsilon.$$

The original version of this density increment by Kelley and Meka instead exhibited a subspace of codimension $O_\epsilon((1+\log \alpha_1^{-1})^3(1+\log \alpha_2^{-1}))$ by directly applying the almost-periodicity result without further adjustment. The α_1, α_2 produced from the previous three steps will satisfy $\log \alpha_i^{-1} = O_\epsilon((1 + \log \alpha^{-1})^2)$, so the original version gives a codimension bound of order $(1 + \log \alpha^{-1})^8$, while this improved version gives a codimension bound of order $(1 + \log \alpha^{-1})^6$.

Iterate. We can chain together the four steps into a density increment from the progression count.

Proposition 3.11. *Let V be a finite-dimensional vector space over a field of prime order. Let $A \subseteq V$ have density $\alpha > 0$. Suppose that*

$$\#\{\text{3-APs in } A\} \leq \tfrac{1}{2}\alpha^3 |V|^2 .$$

Then there exist a subspace $U \leq V$ of codimension $O((1+\log \alpha^{-1})^6)$ and a $t \in V$ such that

$$\frac{|(A-t)\cap U|}{|U|} \geq \tfrac{129}{128}\alpha.$$

This proposition is ready to be iterated. As a sketch, the idea is to start with V and apply this proposition repeatedly to generate a sequence of subspaces. After $O(\log \alpha^{-1})$ steps, this must stop since density is bounded by 1. At this point, we must have many 3-APs. The codimension bound gives us an idea of how far we have strayed from V, so we can analyze this final state to deduce the desired count. We provide a full proof in Section 3.3.

3.2. **Proof of density increment steps.** We now prove the steps in the density increment for the finite field model.

Step 1: Lifting. To prove Lemma 3.6, we need the following simple Fourier estimate.

Lemma 3.12. *Let $f\colon G \to \mathbb{R}$ and $p \geq 1$ be an even integer. Then*

$$\left\| f * f - (\mathbb{E}\, f)^2 \right\|_p \leq \left\| f \star f - (\mathbb{E}\, f)^2 \right\|_p .$$

Proof. Let $g \coloneqq f * f - (\mathbb{E}\, f)^2$. Since p is even, we have that

$$\|g\|_p^p = \mathbb{E}\, g^p = \widehat{g^p}(0) = \hat{g}^{*p}(0).$$

Similarly, with $h \coloneqq f \star f - (\mathbb{E}\, f)^2$, we have that $\|h\|_p^p = \hat{h}^{*p}(0)$. But

$$\hat{g} = \widehat{f * f} - (\mathbb{E}\, f)^2 \mathbb{1}_{\{\chi_0\}} = \hat{f}^2 \mathbb{1}_{\widehat{G}\setminus\{\chi_0\}}$$

and similarly $\hat{h} = |\hat{f}|^2 \mathbb{1}_{\widehat{G}\setminus\{\chi_0\}}$, so $\hat{h} = |\hat{g}|$ and the result follows from the triangle inequality. □

Proof of Lemma 3.6. Let $p = 2 + 2\lfloor \log \alpha^{-1} \rfloor$. Then

$$\begin{aligned} \epsilon &\leq |\langle \mu_A * \mu_A, \mu_{2\cdot A}\rangle - 1| \\ &= |\langle \mu_A * \mu_A - 1, \mu_{2\cdot A}\rangle| \\ &\overset{\text{Hölder}}{\leq} \|\mu_A * \mu_A - 1\|_p \, \|\mu_{2\cdot A}\|_{p^*} , \end{aligned}$$

where $\frac{1}{p} + \frac{1}{p^*} = 1$. But

$$\|\mu_{2\cdot A}\|_{p^*} = \alpha^{-1} \left(\mathbb{E}\, \mathbb{1}_{2\cdot A}\right)^{1/p^*} = \alpha^{-1/p},$$

so

$$\|\mu_A * \mu_A - 1\|_p \geq \epsilon \alpha^{1/p} > \tfrac{1}{2}\epsilon$$

by the choice of p. The result follows from Lemma 3.12 applied to μ_A. □

Step 2: Unbalancing. It will be helpful in the integer setting to have the following generalisation of the lemma, from which Lemma 3.7 follows by taking $f = \mu_A \star \mu_A - 1$ and ν to be the uniform measure.

Lemma 3.13. *Let $f\colon G \to \mathbb{R}$ be spectrally nonnegative, and let ν be a spectrally nonnegative probability measure on G. Suppose that*

$$\|f\|_{L^p(\nu)} \geq \epsilon$$

for some $0 < \epsilon < 1$ and $p \geq 1$. Then

$$\|f + 1\|_{L^{p'}(\nu)} \geq 1 + \tfrac{1}{2}\epsilon$$

for some $p' = O_\epsilon(p)$.

Remark 3.14. The lemma is false upon removing the spectral nonnegativity conditions (e.g. with $f \equiv -1$). The choice $f = \mu_A \star \mu_A - 1$ is valid because

$$\hat{f} = \widehat{\mu_A \star \mu_A} - \mathbb{1}_{\{\chi_0\}} = |\widehat{\mu_A}|^2 \, \mathbb{1}_{\widehat{G}\setminus\{\chi_0\}} \geq 0.$$

The first step to prove the strengthened lemma is to see that if p is odd, then f^p should be somewhat correlated with the event that $f \geq c\epsilon$ for $0 < c < 1$.

Lemma 3.15. *Let $f\colon G \to \mathbb{R}$ be spectrally nonnegative, and let ν be a spectrally nonnegative probability measure on G. Suppose that*

$$\|f\|_{L^p(\nu)} \geq \epsilon$$

for some $0 < \epsilon < 1$ and odd integer $p \geq 1$. Then for any $0 < c < 1$,

$$\left\langle \mathbb{1}_{\{f \geq c\epsilon\}}, f^p \right\rangle_\nu \geq (\tfrac{1}{2} - c^p)\epsilon^p.$$

Proof. It suffices to check correlation with $f > 0$ and with $0 < f < c\epsilon$. Observe that

$$\begin{aligned}
\left\langle \mathbb{1}_{\{f>0\}}, f^p \right\rangle_\nu &= \mathbb{E}\left[\nu \, \mathbb{1}_{\{f>0\}} \, f^p\right] \\
&= \mathbb{E}\left[\nu \cdot \frac{f^p + |f| \, f^{p-1}}{2}\right] \\
&= \frac{1}{2}\,\mathbb{E}\left[\nu f^p\right] + \frac{1}{2}\,\mathbb{E}\left[\nu \, |f|^p\right] \\
&= \frac{1}{2}\widehat{\nu f^p}(0) + \frac{1}{2}\,\|f\|^p_{L^p(\nu)} \\
&= \frac{1}{2}\hat{\nu} * \hat{f}^{*p}(0) + \frac{1}{2}\,\|f\|^p_{L^p(\nu)} \\
&\geq 0 + \frac{1}{2}\epsilon^p
\end{aligned}$$

and

$$\begin{aligned}
\left\langle \mathbb{1}_{\{0<f<c\epsilon\}}, f^p \right\rangle_\nu &= \mathbb{E}\left[\nu \, \mathbb{1}_{\{0<f<c\epsilon\}} \, f^p\right] \\
&\leq \nu(\{0 < f < c\epsilon\})(c\epsilon)^p \\
&\leq (c\epsilon)^p,
\end{aligned}$$

so subtracting the two bounds finishes. □

But this correlation is controlled by the average of these objects, which can be unbalanced. A little bit of bounding allows us to finish.

Proof of Lemma 3.13. If $\|f+1\|_{L^{2p}(\nu)} \geq 1 + \frac{1}{2}\epsilon$, then we are done with $p' = 2p$. Assume otherwise.

Since L^p-norms are monotonically increasing in p, without loss of generality we can round p up to an odd integer at least 3. Let $c = 0.51$ and $p' = 1000\epsilon^{-1}(1 + \log \epsilon^{-1})p$. Consider the following three estimates:

(1) Cauchy–Schwarz:

$$\begin{aligned}
\left\langle \mathbb{1}_{\{f \geq c\epsilon\}}, f^p \right\rangle_\nu &\leq \left\| \mathbb{1}_{\{f \geq c\epsilon\}} \right\|_{L^2(\nu)} \|f^p\|_{L^2(\nu)} \\
&= \nu(\{f \geq c\epsilon\})^{1/2} \, \|f\|^p_{L^{2p}(\nu)}
\end{aligned}$$

(2) Triangle inequality:

$$\begin{aligned}
\|f\|_{L^{2p}(\nu)} &\leq 1 + \|f+1\|_{L^{2p}(\nu)} \\
&< 2 + \tfrac{1}{2}\epsilon
\end{aligned}$$

(3) Markov's inequality:

$$\begin{aligned}\nu(\{f \geq c\epsilon\}) &\leq \nu(\{|f+1|^{p'} \geq (1+c\epsilon)^{p'}\}) \\ &\leq \frac{\|f+1\|_{L^{p'}(\nu)}^{p'}}{(1+c\epsilon)^{p'}}\end{aligned}$$

Combining all three bounds along with Lemma 3.15 gives that

$$(\tfrac{1}{2} - c^p)\epsilon^p \leq \left\langle \mathbb{1}_{\{f \geq c\epsilon\}}, f^p \right\rangle_\nu \leq \left(\frac{\|f+1\|_{L^{p'}(\nu)}}{1+c\epsilon}\right)^{\frac{p'}{2}} (2+\tfrac{1}{2}\epsilon)^p.$$

With the choices of c and p' above, it follows that $\|f+1\|_{L^{p'}(\nu)} \geq 1 + \frac{1}{2}\epsilon$. □

Remark 3.16. Though we required both f and ν to both be spectrally nonnegative, the only Fourier-side condition used is that $\hat{\nu} * \hat{f}^{*p}(0) \geq 0$.

Step 3: Correlation. Again, it will be helpful in the integer setting to have the following strengthened version of the lemma, from which Lemma 3.8 follows by taking $B_1 = B_2 = G$ (in which case μ is the uniform measure), replacing ϵ with $\frac{\epsilon/3}{1+\epsilon}$ and δ with $\frac{1}{8}\epsilon$, and loosening the constraint in S to the desired size (which does not affect the conclusion as the final inner product can only increase upon expanding S).

Lemma 3.17. *Let $0 < \epsilon, \delta < 1$ and $p \geq 1$ be an integer. Let $A \subseteq G$ have density $\alpha > 0$ and $B_1, B_2 \subseteq G$ be nonempty. Set $\mu = \mu_{B_1} \star \mu_{B_2}$. Suppose that $\|\mu_A \star \mu_A\|_{L^p(\mu)} > 0$. Then there exist $A_1 \subseteq B_1$ and $A_2 \subseteq B_2$ of relative density $\Omega((\alpha \|\mu_A \star \mu_A\|_{L^p(\mu)})^{2p+O_{\epsilon,\delta}(1)})$ such that*

$$\langle \mu_{A_1} \star \mu_{A_2}, \mathbb{1}_S \rangle \geq 1 - \delta,$$

where $S = \{\mu_A \star \mu_A > (1-\epsilon)\|\mu_A \star \mu_A\|_{L^p(\mu)}\}$.

Since

$$\langle \mu_{A_1} \star \mu_{A_2}, 1 \rangle = \mathbb{E}[\mu_{A_1} \star \mu_{A_2}] = 1,$$

it suffices to find appropriate A_1, A_2 such that

$$\langle \mathbb{1}_{A_1} \star \mathbb{1}_{A_2}, \mathbb{1}_{S^c} \rangle \leq \delta\alpha_1\alpha_2,$$

where S^{c} is the complement of $S \subseteq G$ and α_1, α_2 are the densities of A_1, A_2, respectively.

The key idea is to apply dependent random choice: randomly choose some intersection of B_i with several translates of A. To formalise this argument, for $\mathbf{s} = (s_1, \ldots, s_p) \in G^p$, let $A_i(\mathbf{s}) = B_i \cap (A+s_1) \cap \cdots \cap (A+s_p)$ for $i = 1, 2$. Let $\alpha_i(\mathbf{s})$ and β_i denote the densities of $A_i(s)$ and B_i, respectively. We record a straightforward calculation.

Lemma 3.18. *Let $f\colon G \to \mathbb{R}$. Then*

$$\mathop{\mathbb{E}}_{\mathbf{s}\in G^p} \langle \mathbb{1}_{A_1(\mathbf{s})} \star \mathbb{1}_{A_2(\mathbf{s})}, f \rangle = \beta_1\beta_2 \mathop{\mathbb{E}}_{x\in G} \mu(x)(\mathbb{1}_A \star \mathbb{1}_A)(x)^p f(x).$$

Proof. We have that

$$\begin{aligned}\mathop{\mathbb{E}}_{\mathbf{s}\in G^p} \langle \mathbb{1}_{A_1(\mathbf{s})} \star \mathbb{1}_{A_2(\mathbf{s})}, f \rangle &= \mathop{\mathbb{E}}_{\mathbf{s}\in G^p} \mathop{\mathbb{E}}_{x\in G} \mathop{\mathbb{E}}_{y\in G} \mathbb{1}_{A_1(\mathbf{s})}(y)\, \mathbb{1}_{A_2(\mathbf{s})}(x+y) f(x) \\ &\overset{z=x+y}{=} \mathop{\mathbb{E}}_{\mathbf{s}\in G^p} \mathop{\mathbb{E}}_{z\in G} \mathop{\mathbb{E}}_{y\in G} \mathbb{1}_{A_1(\mathbf{s})}(y)\, \mathbb{1}_{A_2(\mathbf{s})}(z) f(z-y).\end{aligned}$$

For fixed $y, z \in G$, we have that

$$\begin{aligned}
\mathop{\mathbb{E}}_{\mathbf{s}\in G^p} \mathbb{1}_{A_1(\mathbf{s})}(y)\,\mathbb{1}_{A_2(\mathbf{s})}(z) &= \mathbb{1}_{B_1}(y)\,\mathbb{1}_{B_2}(z) \mathop{\mathbb{E}}_{s_1,\dots,s_p\in G} \prod_{j=1}^{p} \mathbb{1}_{A+s_j}(y)\,\mathbb{1}_{A+s_j}(z) \\
&= \mathbb{1}_{B_1}(y)\,\mathbb{1}_{B_2}(z) \left(\mathop{\mathbb{E}}_{t\in G} \mathbb{1}_{A+t}(y)\,\mathbb{1}_{A+t}(z)\right)^p \\
&\overset{u=y-t}{=} \mathbb{1}_{B_1}(y)\,\mathbb{1}_{B_2}(z) \left(\mathop{\mathbb{E}}_{u\in G} \mathbb{1}_{A}(u)\,\mathbb{1}_{A}(u+z-y)\right)^p \\
&= \mathbb{1}_{B_1}(y)\,\mathbb{1}_{B_2}(z)(\mathbb{1}_A \star \mathbb{1}_A)(z-y)^p.
\end{aligned}$$

Plugging this in gives that

$$\begin{aligned}
\mathop{\mathbb{E}}_{\mathbf{s}\in G^p} \left\langle \mathbb{1}_{A_1(\mathbf{s})} \star \mathbb{1}_{A_2(\mathbf{s})}, f\right\rangle &= \mathop{\mathbb{E}}_{y\in G}\mathop{\mathbb{E}}_{z\in G} \mathbb{1}_{B_1}(y)\,\mathbb{1}_{B_2}(z)(\mathbb{1}_A \star \mathbb{1}_A)(z-y)^p f(z-y) \\
&\overset{x=z-y}{=} \beta_1\beta_2 \mathop{\mathbb{E}}_{x\in G}\mathop{\mathbb{E}}_{y\in G} \mu_{B_1}(y)\mu_{B_2}(x+y)(\mathbb{1}_A \star \mathbb{1}_A)(x)^p f(x) \\
&= \beta_1\beta_2 \mathop{\mathbb{E}}_{x\in G} \mu(x)(\mathbb{1}_A \star \mathbb{1}_A)(x)^p f(x)
\end{aligned}$$

as desired. □

It is clear that we want to apply this fact to $f = \mathbb{1}_{S^{\mathsf{c}}}$. It turns out that $f \equiv 1$ also provides useful information.

Corollary 3.19. *We have that*

$$\mathop{\mathbb{E}}_{\mathbf{s}\in G^p} \left\langle \mathbb{1}_{A_1(\mathbf{s})} \star \mathbb{1}_{A_2(\mathbf{s})}, \mathbb{1}_{S^{\mathsf{c}}}\right\rangle \le (1-\epsilon)^p \mathop{\mathbb{E}}_{\mathbf{s}\in G^p} \alpha_1(\mathbf{s})\alpha_2(\mathbf{s})$$

and

$$\mathop{\mathbb{E}}_{\mathbf{s}\in G^p} \alpha_1(\mathbf{s})\alpha_2(\mathbf{s}) = \beta_1\beta_2 \left\|\mathbb{1}_A \star \mathbb{1}_A\right\|_{L^p(\mu)}^p.$$

Proof. Apply Lemma 3.18 to $f = \mathbb{1}_{S^{\mathsf{c}}}$ to get that

$$\mathop{\mathbb{E}}_{\mathbf{s}\in G^p} \left\langle \mathbb{1}_{A_1(\mathbf{s})} \star \mathbb{1}_{A_2(\mathbf{s})}, \mathbb{1}_{S^{\mathsf{c}}}\right\rangle = \beta_1\beta_2 \mathop{\mathbb{E}}_{x\in G} \mu(x)(\mathbb{1}_A \star \mathbb{1}_A)(x)^p\,\mathbb{1}_{S^{\mathsf{c}}}(x).$$

But $S^{\mathsf{c}} = \{\mathbb{1}_A \star \mathbb{1}_A \le (1-\epsilon)\left\|\mathbb{1}_A \star \mathbb{1}_A\right\|_{L^p(\mu)}\}$, so we have the bound

$$\begin{aligned}
\mathop{\mathbb{E}}_{x\in G} \mu(x)(\mathbb{1}_A \star \mathbb{1}_A)(x)^p\,\mathbb{1}_{S^{\mathsf{c}}}(x) &\le \mu(S^{\mathsf{c}})(1-\epsilon)^p \left\|\mathbb{1}_A \star \mathbb{1}_A\right\|_{L^p(\mu)}^p \\
&\le (1-\epsilon)^p \left\|\mathbb{1}_A \star \mathbb{1}_A\right\|_{L^p(\mu)}^p.
\end{aligned}$$

Apply Lemma 3.18 to $f \equiv 1$ to get that

$$\begin{aligned}
\mathop{\mathbb{E}}_{\mathbf{s}\in G^p} \left\langle \mathbb{1}_{A_1(\mathbf{s})} \star \mathbb{1}_{A_2(\mathbf{s})}, 1\right\rangle &= \beta_1\beta_2 \mathop{\mathbb{E}}_{x\in G} \mu(x)(\mathbb{1}_A \star \mathbb{1}_A)(x)^p \\
&= \beta_1\beta_2 \left\|\mathbb{1}_A \star \mathbb{1}_A\right\|_{L^p(\mu)}^p.
\end{aligned}$$

But

$$\begin{aligned}
\mathop{\mathbb{E}}_{\mathbf{s}\in G^p} \left\langle \mathbb{1}_{A_1(\mathbf{s})} \star \mathbb{1}_{A_2(\mathbf{s})}, 1\right\rangle &= \mathop{\mathbb{E}}_{\mathbf{s}\in G^p} \alpha_1(\mathbf{s})\alpha_2(\mathbf{s}) \left\langle \mu_{A_1(\mathbf{s})} \star \mu_{A_2(\mathbf{s})}, 1\right\rangle \\
&= \mathop{\mathbb{E}}_{\mathbf{s}\in G^p} \alpha_1(\mathbf{s})\alpha_2(\mathbf{s}),
\end{aligned}$$

so chaining these together gives the desired result. □

At this point, we would like to use the probabilistic method to say that we are done—that is, we have found some s for which $\langle \mathbb{1}_{A_1} \star \mathbb{1}_{A_2}, \mathbb{1}_{S^c} \rangle \leq (1-\epsilon)^p \alpha_1 \alpha_2$, and we just need p to be large enough (which we can do by monotonicity of L^p-norms). But recall that we also had another aim of keeping the sizes of A_1 and A_2 large enough. To do that, we can tack on an extra factor of $\mathbb{1}_{\{\alpha_1\alpha_2 \geq D\beta_1\beta_2\}}$ for some $D > 0$, so that both $\frac{|A_i|}{|B_i|}$ are at least D. But we have to keep the sum large, so we must control the amount that we lose when we throw away the small $\alpha_1\alpha_2$.

Lemma 3.20. *Let $D > 0$. Then*

$$\mathop{\mathbb{E}}_{\mathbf{s}\in G^p} \alpha_1(\mathbf{s})\alpha_2(\mathbf{s})\, \mathbb{1}_{\{\alpha_1\alpha_2 < D\beta_1\beta_2\}} < D^{1/2}\beta_1\beta_2\alpha^p.$$

Proof. By size bounding and the Cauchy–Schwarz inequality, we have that

$$\begin{aligned}\mathop{\mathbb{E}}_{\mathbf{s}\in G^p} \alpha_1(\mathbf{s})\alpha_2(\mathbf{s})\, \mathbb{1}_{\{\alpha_1\alpha_2 < D\beta_1\beta_2\}} &\leq D^{1/2}\beta_1^{1/2}\beta_2^{1/2} \mathop{\mathbb{E}}_{\mathbf{s}\in G^p} \alpha_1(\mathbf{s})^{1/2}\alpha_2(\mathbf{s})^{1/2} \\ &\overset{\text{CS}}{\leq} D^{1/2}\beta_1^{1/2}\beta_2^{1/2} \left(\mathop{\mathbb{E}}_{\mathbf{s}\in G^p} \alpha_1(\mathbf{s})\right)^{\frac{1}{2}} \left(\mathop{\mathbb{E}}_{\mathbf{s}\in G^p} \alpha_2(\mathbf{s})\right)^{\frac{1}{2}}.\end{aligned}$$

A similar calculation to the proof of Lemma 3.18 gives that

$$\mathop{\mathbb{E}}_{\mathbf{s}\in G^p} \alpha_i(\mathbf{s}) = \mathop{\mathbb{E}}_{\mathbf{s}\in G^p} \mathop{\mathbb{E}}_{x\in G} \mathbb{1}_{A_i(\mathbf{s})}(x) = \mathop{\mathbb{E}}_{x\in G} \mathbb{1}_{B_i}(x) \left(\mathop{\mathbb{E}}_{t\in G} \mathbb{1}_{A+t}(x)\right)^p = \beta_i\alpha^p,$$

so

$$\mathop{\mathbb{E}}_{\mathbf{s}\in G^p} \alpha_1(\mathbf{s})\alpha_2(\mathbf{s})\, \mathbb{1}_{\{\alpha_1\alpha_2 < D\beta_1\beta_2\}} \leq D^{1/2}\beta_1\beta_2\alpha^p.$$

In order for equality to hold, both sides of the very first inequality would have to be 0—if $\alpha_1\alpha_2 > 0$, then either $\alpha_1\alpha_2 \geq D\beta_1\beta_2$ or $D^{1/2}\beta_1^{1/2}\beta_2^{1/2}\alpha_1^{1/2}\alpha_2^{1/2} > \alpha_1\alpha_2$. So then the final upper bound would have to be 0, which it is not. So equality cannot hold. □

We can now tie everything together.

Proof of Lemma 3.17. Let $p' = p + \lceil \epsilon \log \delta^{-1} \rceil$ and $D = 0.01\alpha^{-2p'} \|\mathbb{1}_A \star \mathbb{1}_A\|_{L^{p'}(\mu)}^{2p'} \neq 0$. We have that

$$\begin{aligned}\mathop{\mathbb{E}}_{\mathbf{s}\in G^{p'}} \alpha_1(\mathbf{s})\alpha_2(\mathbf{s})\, \mathbb{1}_{\{\alpha_1\alpha_2 < D\beta_1\beta_2\}} &\overset{\text{Lem. 3.20}}{<} 0.1\beta_1\beta_2 \|\mathbb{1}_A \star \mathbb{1}_A\|_{L^{p'}(\mu)}^{p'} \\ &\overset{\text{Cor. 3.19}}{=} 0.1 \mathop{\mathbb{E}}_{\mathbf{s}\in G^{p'}} \alpha_1(\mathbf{s})\alpha_2(\mathbf{s}),\end{aligned}$$

so

$$\begin{aligned}\mathop{\mathbb{E}}_{\mathbf{s}\in G^{p'}} \alpha_1(\mathbf{s})\alpha_2(\mathbf{s})\, \mathbb{1}_{\{\alpha_1\alpha_2 \geq D\beta_1\beta_2\}} &> 0.9 \mathop{\mathbb{E}}_{\mathbf{s}\in G^{p'}} \alpha_1(\mathbf{s})\alpha_2(\mathbf{s}) \\ &\overset{\text{Cor. 3.19}}{\geq} 0.9(1-\epsilon)^{-p'} \mathop{\mathbb{E}}_{\mathbf{s}\in G^{p'}} \left\langle \mathbb{1}_{A_1(\mathbf{s})} \star \mathbb{1}_{A_2(\mathbf{s})}, \mathbb{1}_{S^c} \right\rangle.\end{aligned}$$

By the probabilistic method, there exists some s for which

$$\alpha_1\alpha_2\, \mathbb{1}_{\{\alpha_1\alpha_2 \geq D\beta_1\beta_2\}} > 0.9(1-\epsilon)^{-p'} \left\langle \mathbb{1}_{A_1} \star \mathbb{1}_{A_2}, \mathbb{1}_{S^c} \right\rangle.$$

Since the right-hand side is (termwise) nonnegative, the left-hand side must be positive, so $\alpha_1\alpha_2 \geq D\beta_1\beta_2$. Thus

$$\begin{aligned}\frac{|A_i|}{|B_i|} \geq \frac{|A_1|\,|A_2|}{|B_1|\,|B_2|} &= \frac{\alpha_1\alpha_2}{\beta_1\beta_2} \\ &\geq D = 0.01(\alpha \|\mu_A \star \mu_A\|_{L^{p'}(\mu)})^{2p'} \\ &\geq 0.01(\alpha \|\mu_A \star \mu_A\|_{L^p(\mu)})^{2p + O_{\epsilon,\delta}(1)}.\end{aligned}$$

Furthermore, since $p' > \log(0.9\delta)/\log(1-\epsilon)$, the bound on the inner product becomes

$$\alpha_1\alpha_2 > \delta^{-1}\left\langle \mathbb{1}_{A_1} \star \mathbb{1}_{A_2}, \mathbb{1}_{S^c}\right\rangle,$$

which gives the desired bound on $\langle \mu_{A_1} \star \mu_{A_2}, \mathbb{1}_S\rangle$ upon rearranging. □

Step 4: Density increment. We will use the following L^∞-almost-periodicity result, which is essentially the same as a statement of Schoen and Sisask [SS16, Theorem 3.2] and is a special case of the result we will use for the integer case.

Proposition 3.21. *Let $0 < \epsilon < 1$ and $k \geq 2$ be an integer. Let $A_1, A_2, S \subseteq G$ be such that A_1, A_2 have density $\alpha_1, \alpha_2 > 0$, respectively. Then there exists a set $T \subseteq G$ of density at least*

$$\exp(-O_\epsilon(k^2(1+\log\alpha_1^{-1})(1+\log\alpha_2^{-1})))$$

such that

$$\|\mu_T^{*k} * (\mu_{A_1} \star \mu_{A_2}) * \mathbb{1}_S - (\mu_{A_1} \star \mu_{A_2}) * \mathbb{1}_S\|_\infty \leq \epsilon.$$

We defer the proof of this result to Appendix A.

This result alone is enough to give Kelley and Meka's original result with $c = 1/9$. The improvement of Bloom and Sisask was to improve the "bootstrap" procedure in which a μ_U factor for some subspace U is introduced in place of the μ_T^{*k} factor. This would give us the structure that we want. In order to do this, we must take advantage of the structure of S as given to us in the result of Step 3; note that the almost-periodicity result alone assumes no structure on S.

Proof of Lemma 3.10. Apply Proposition 3.21 with $k = 2 + 2\lceil\log\epsilon^{-1}\rceil + 2\lceil\log\alpha^{-1}\rceil$, $G = V$, and S replaced by $-S$ to produce a set of almost-periods T. Let

$$U = \left\{x \in V \,:\, \chi(x) = 1 \text{ for all } \chi \in \operatorname{Spec}_{1/2}(\mu_T)\right\},$$

and note that U is a vector space isomorphic to the orthogonal complement of $\operatorname{span}(\operatorname{Spec}_{1/2}(\mathbb{1}_T))$. By Chang's lemma (Corollary A.4), we have that

$$\begin{aligned}\operatorname{codim} U &= \dim \operatorname{span}(\operatorname{Spec}_{1/2}(\mathbb{1}_T)) \\ &\overset{\text{Chang}}{=} O\left(\log\frac{|V|}{|T|}\right) \\ &= O_\epsilon(k^2(1+\log\alpha_1^{-1})(1+\log\alpha_2^{-1})) \\ &= O_\epsilon((1+\log\alpha^{-1})^2(1+\log\alpha_1^{-1})(1+\log\alpha_2^{-1})).\end{aligned}$$

Now, we compute the density increment on U. Writing $f = \mu_T^{*k} * (\mu_{A_1} \star \mu_{A_2}) - \mu_{A_1} \star \mu_{A_2}$ for shorthand, we have that

$$|(f \star \mathbb{1}_S)(0)| \leq \|f \star \mathbb{1}_S\|_\infty = \|f * \mathbb{1}_{-S}\|_\infty \overset{\text{Prop. 3.21}}{\leq} \epsilon,$$

so

$$((\mu_T^{*k} * (\mu_{A_1} \star \mu_{A_2})) \star \mathbb{1}_S)(0) = (f \star \mathbb{1}_S)(0) + \langle \mu_{A_1} \star \mu_{A_2}, \mathbb{1}_S\rangle \geq 1 - 2\epsilon.$$

We have large correlation with $\mathbb{1}_S$, which we would now like to convert into large correlation with $\mu_A \star \mu_A$. By the definition of S, we have that

$$\begin{aligned}((\mu_T^{*k} * (\mu_{A_1} \star \mu_{A_2})) \star (\mu_A \star \mu_A))(0) &= \mathop{\mathbb{E}}_{y \in G} (\mu_T^{*k} * (\mu_{A_1} \star \mu_{A_2})(y) \cdot (\mu_A \star \mu_A)(y) \\ &\geq \mathop{\mathbb{E}}_{y \in G} (\mu_T^{*k} * (\mu_{A_1} \star \mu_{A_2})(y) \cdot (1+4\epsilon)\, \mathbb{1}_S(y) \\ &= (1+4\epsilon)((\mu_T^{*k} * (\mu_{A_1} \star \mu_{A_2})) \star \mathbb{1}_S)(0) \\ &\geq (1+4\epsilon)(1-2\epsilon) \\ &= 1 + 2\epsilon - 8\epsilon^2 \\ &> 1 + \epsilon\end{aligned}$$

as $0 < \epsilon < 1/8$. Let $\mu = (\mu_T^{*k} * (\mu_{A_1} \star \mu_{A_2})) \star (\mu_A \star \mu_A)$ so that $\mu(0) > 1 + \epsilon$.

Next, we would like large cross-correlation on a structured subset—that is, to go from $\mu(0)$ to $\mu(x)$. We have that $|\hat{\mu}| = |\widehat{\mu_T}|^k |\widehat{\mu_{A_1}}||\widehat{\mu_{A_2}}||\widehat{\mu_A}|^2$, so for all $x \in G$, we have that

$$\begin{aligned}|\mu(x) - \mu(0)| &= \left| \sum_{\chi \in \widehat{G}} \hat{\mu}(\chi)(\chi(x) - 1) \right| \\ &\leq \sum_{\chi \in \widehat{G}} |\widehat{\mu_T}(\chi)|^k |\widehat{\mu_{A_1}}(\chi)||\widehat{\mu_{A_2}}(\chi)||\widehat{\mu_A}(\chi)|^2 \, |\chi(x) - 1| \,.\end{aligned}$$

But observe that:

- if $\chi \in \mathrm{Spec}_{1/2}(\mu_T)$ and $x \in U$, then $\chi(x) = 1$;
- if $\chi \notin \mathrm{Spec}_{1/2}(\mu_T)$, then $|\widehat{\mu_T}(\chi)| < 1/2$; and
- $|\widehat{\mu_{A_1}}(\chi)| \leq \|\mu_{A_1}\|_1 \leq 1$, and similarly $|\widehat{\mu_{A_2}}(\chi)| \leq 1$.

It follows that for all $x \in U$, we have that

$$\begin{aligned}|\mu(x) - \mu(0)| &\leq \sum_{\chi \notin \mathrm{Spec}_{1/2}(\mu_T)} |\widehat{\mu_A}(\chi)|^2 \cdot 2^{-k} \cdot 2 \\ &\leq 2^{1-k} \|\widehat{\mu_A}\|_2^2 \\ &\overset{\text{Parseval}}{=} 2^{1-k} \|\mu_A\|_2^2 \\ &= 2^{1-k} \alpha^{-1} \\ &\leq \tfrac{1}{2}\epsilon\end{aligned}$$

by the choice of k.

Finally, we can aggregate this correlation and extract the desired terms. We have that

$$\begin{aligned}|(\mu_U \star \mu)(0) - \mu(0)| &= \left| \mathop{\mathbb{E}}_{y \in G} \mu_U(y)(\mu(y) - \mu(0)) \right| \\ &\leq \|\mu - \mu(0)\|_{L^1(\mu_U)} \\ &\leq \|\mu - \mu(0)\|_{L^\infty(\mu_U)} \\ &\leq \tfrac{1}{2}\epsilon,\end{aligned}$$

so $(\mu_U \star \mu)(0) > 1 + \frac{1}{2}\epsilon$. But one can check that

$$\mu_U \star \mu = (\mu_A * \mu_U) \star ((\mu_T^{*k} * (\mu_{A_1} \star \mu_{A_2})) \star \mu_A),$$

so

$$\begin{aligned}\|\mu_A * \mu_U\|_\infty &= \|\mu_A * \mu_U\|_\infty \|(\mu_T^{*k} * (\mu_{A_1} \star \mu_{A_2})) \star \mu_A\|_1 \\ &\overset{\text{Hölder}}{\geq} \left\langle \mu_A * \mu_U, (\mu_T^{*k} * (\mu_{A_1} \star \mu_{A_2})) \star \mu_A \right\rangle \\ &= (\mu_U \star \mu)(0) \\ &> 1 + \tfrac{1}{2}\epsilon. \end{aligned}$$

□

Remark 3.22. Throughout the proof, we referred to correlation, i.e. $\langle f, g\rangle$, but required expanding it as $(f \star g)(0)$ for the technical computations.

3.3. **Finishing the argument.** Now that we have the steps of the density increment, it suffices to put them together. We restate the aggregated result here for convenience.

Proposition 3.11. *Let V be a finite-dimensional vector space over a field of prime order. Let $A \subseteq V$ have density $\alpha > 0$. Suppose that*

$$\#\{\text{3-APs in } A\} \leq \tfrac{1}{2}\alpha^3 |V|^2 .$$

Then there exist a subspace $U \leq V$ of codimension $O((1 + \log \alpha^{-1})^6)$ and a $t \in V$ such that

$$\frac{|(A - t) \cap U|}{|U|} \geq \tfrac{129}{128}\alpha.$$

Proof. By Remarks 3.4 and 3.5, we can replace the hypothesis with $\langle \mu_A * \mu_A, \mu_{2\cdot A}\rangle \leq \frac{1}{2}$ and the conclusion with $\|\mu_A * \mu_U\|_\infty \geq 1 + \frac{1}{128}$.

Apply Lemma 3.6 with $\epsilon = 1/2$, Lemma 3.7 with $\epsilon = 1/4$, Lemma 3.8 with $p = \lceil p' \rceil$ (where p' is obtained from Lemma 3.7) and $\epsilon = 1/8$, and Lemma 3.10 with $\epsilon = 1/64$. To get the desired codimension bound, it suffices to note that the densities α_i of the A_i produced by Lemma 3.8 satisfy $\log \alpha_i^{-1} = O((1 + \log \alpha^{-1})^2)$. □

We now demonstrate how to prove the Kelley–Meka result in the finite field model with $c = 1/7$ by iterating this density increment. First, we prove the counting result.

Proof of Theorem 3.2. Consider the following process:

(1) Initialise $V_0 = \mathbb{F}_q^n$, $A_0 = A$, $\alpha_i = \alpha$, and $i = 0$.
(2) Assert that $A_i \subseteq V_i \leq \mathbb{F}_q^n$.
(3) Set $\alpha_i = \frac{|A_{i+1}|}{|V_{i+1}|}$.
(4) If $\#\{\text{3-APs in } A_i\} > \frac{1}{2}\alpha_i^3 |V_i|^2$, then set $m = i$ and **STOP**.
(5) Otherwise, by Proposition 3.11, there exists a subspace $U_i \leq V_i$ of codimension $O((1 + \log \alpha_i^{-1})^6)$ and a $t_i \in V_i$ such that
$$\frac{|(A_i - t_i) \cap U_i|}{|U_i|} \geq \tfrac{129}{128}\alpha_i.$$
(6) Set $V_{i+1} = U_i$ and $A_{i+1} = (A_i - t_i) \cap U_i$.
(7) Increment i and go back to step (2).

By construction, we have that:

- $\mathbb{F}_q^n = V_0 \geq \cdots \geq V_m$;
- $\#\{\text{3-APs in } A_m\} > \frac{1}{2}\alpha_m^3 |V_m|^2$;
- the α_i are increasing, so $V_{i+1} \leq V_i$ has codimension $O((1 + \log \alpha^{-1})^6)$ and thus $V_m \leq \mathbb{F}_q^n$ has codimension $O(m(1 + \log \alpha^{-1})^6)$;
- $\alpha_m \geq \left(\frac{129}{128}\right)^m \alpha$, so since $\alpha_m \leq 1$, we have that $m \leq 200 \log \alpha^{-1}$ (in particular, the process terminates); and
- $A_{i+1} \subseteq A_i - t_i$, so $A_m \subseteq A - t$, where $t = t_0 + \cdots + t_{m-1}$.

The bound on m implies that $V_m \leq \mathbb{F}_q^n$ has codimension $O((1+\log \alpha^{-1})^7)$, so

$$|V_m| \geq q^{n-O((1+\log \alpha^{-1})^7)}.$$

Then

$$\begin{aligned}
\#\{\text{3-APs in } A\} &= \#\{\text{3-APs in } A-t\} \\
&\geq \#\{\text{3-APs in } A_m\} \\
&> \tfrac{1}{2}\alpha_m^3 \, |V_m|^2 \\
&\geq q^{-3(1+\log \alpha^{-1})} q^{2n-O((1+\log \alpha^{-1})^7)} \\
&= q^{2n-O((1+\log \alpha^{-1})^7)}.
\end{aligned}$$

□

Now, the extremal result is immediate.

Proof of Kelley–Meka in $\mathbb{F}_q^n$ (Theorem 3.1). Suppose that $A \subseteq \mathbb{F}_q^n$ of density $\alpha > 0$ has no nontrivial three-term arithmetic progressions. Then $\#\{\text{3-APs in } A\} = |A|$. By Theorem 3.2, we have that

$$1 \geq \alpha \geq q^{n-O((1+\log \alpha^{-1})^7)}.$$

Solving for α gives the desired bound. □

We have thus completed the proof of the Kelley–Meka result in the finite field setting.

4. Integer setting

We now turn our attention to the main problem in the integer setting. All of the main ideas will carry over, but as has been true for all previous quantitative bounds on $r_3(N)$ and $r_3(\mathbb{F}_q^n)$, the integer setting is filled with many more technicalities.

We recall the current state of results as stated in Section 1:

$$\underbrace{N/\exp(O((\log N)^{1/2}))}_{\substack{\text{Behrend}\\ \text{(best constants: Hunter)}}} \lesssim r_3(N) \lesssim \underbrace{N/\exp(\Omega((\log N)^{1/9}))}_{\substack{\text{BS}\\ \text{(KM method)}}} \lesssim \underbrace{N/(\log N)^{1+c}}_{\substack{\text{BS}\\ \text{(previous best)}}} \lesssim \underbrace{N/\log\log N}_{\text{Roth}}.$$

As with the finite field model, it suffices to prove a counting result.

Theorem 4.1. *Let $A \subseteq \{1,\dots,N\}$ have size $|A| = \alpha N$ for some $\alpha > 0$. Then*

$$\#\{\text{3-APs in } A\} \geq N^2/\exp(O((1+\log\alpha^{-1})^9)).$$

4.1. **Outline.** The structure of the proof in the integer setting is essentially the same as in the finite field model. The key difference is that we no longer have subspaces to work with in $G = \mathbb{Z}/N\mathbb{Z}$. We instead turn to regular Bohr sets as described in Section 2.3, which provide enough additive structure to be useful in the same sense.

Remark 4.2. Kelley and Meka's original proof passed back and forth between Bohr sets and **generalised arithmetic progressions**, which are sets of the form

$$\{a_0 + j_1 d_1 + \cdots + j_r d_r \,:\, j_i \in \{1,\dots,\ell_i\}\}$$

for some $a_0, d_1, \dots, d_r \in G$ and $\ell_1,\dots,\ell_r \in \mathbb{Z}_{>0}$. One of Bloom and Sisask's contributions in their rephrasing was to clean this up by only staying within the world of Bohr sets, which is possible because Bohr sets and generalised arithmetic progressions both play the role of an "approximate subgroup" in the sense of additive structure.

We now provide an outline of the proof, with parallel steps to the finite field model proof. While the proof follows the methodology of Kelley and Meka [KM23] with the rephrasing of Bloom and Sisask [BS23b, BS23a], we differ from both of their presentations by breaking up the key lemmas in precisely the same way as the finite field model to demonstrate the parallels as clearly as possible.

Step 1: Lifting. The Hölder lifting step is similar to before, but a bit more general by counting solutions to a linear equation in $A \times A \times C$ rather than A^3. It turns out that it will be easier to narrow down to two different subsets of A when performing the density increment.

Lemma 4.3. *Let $0 < \epsilon < 1$. Let $B \subseteq G$ be a regular Bohr set of rank r and density $\beta > 0$, and let $A \subseteq B$ with relative density $\alpha > 0$. Let $B' \subseteq B_{\frac{\epsilon\alpha}{4000r}}$ be a regular Bohr set, and let $C \subseteq B'$ with relative density $\gamma > 0$. Suppose that*

$$\langle \mu_A * \mu_A, \mu_C\rangle \leq (1-\epsilon)\beta^{-1}$$

Then for any regular Bohr sets $B'', B''' \subseteq B'_{\frac{1}{400r}}$, we have that

$$\|(\mu_A - \mu_B) \star (\mu_A - \mu_B)\|_{L^p(\mu)} \geq \tfrac{1}{2}\epsilon\beta^{-1}$$

for some $p = O(1+\log\gamma^{-1})$, where $\mu = (\mu_{B''} \star \mu_{B''}) * (\mu_{B'''} \star \mu_{B'''})$.

Here, $(\mu_A - \mu_B) \star (\mu_A - \mu_B)$ will play the role of $\mu_A \star \mu_A - 1$ from the finite field model. Indeed, we can compute that

$$\mu_A \star \mu_A - 1 = (\mu_A - 1) \star (\mu_A - 1),$$

and we essentially use $B = G$ (so that $\mu_B \equiv 1$) in the finite field model.

The lifting lemma for the finite field model (Lemma 3.6) could also have been stated in such generality, but this was not needed. The general version of that lemma is useful in other problems though, such as in finding large arithmetic progressions in sumsets (see Section 5.3).

Step 2: Unbalancing. Though we are using $(\mu_A - \mu_B) \star (\mu_A - \mu_B)$ instead of $\mu_A \star \mu_A - 1$, the correct function to "unbalance" to is still $\mu_A \star \mu_A$. The result is essentially the same as in the finite field model.

Lemma 4.4. *Let $0 < \epsilon < 1$. Let $B \subseteq G$ be a regular Bohr set of rank r and density $\beta > 0$, and let $A \subseteq B$ with relative density $\alpha > 0$. Let ν be a spectrally nonnegative probability measure such that $\operatorname{supp} \nu \subseteq B_{\frac{\epsilon\alpha}{4000r}}$. Suppose that*

$$\|(\mu_A - \mu_B) \star (\mu_A - \mu_B)\|_{L^p(\nu)} \geq \epsilon\beta^{-1}$$

for some integer $p \geq 2$. Then

$$\|\mu_A \star \mu_A\|_{L^{p'}(\nu)} \geq (1 + \tfrac{1}{4}\epsilon)\beta^{-1}$$

for some $p' = O_\epsilon(p)$.

Step 3: Correlation via dependent random choice. This correlation step requires us to narrow down where A_1 and A_2 may lie. This will be useful when we apply the almost-periodicity result over the integers. Note that the variable B' is omitted from the statement as to be consistent with other statements. The same will be true in Step 4.

Lemma 4.5. *Let $B, B'', B''' \subseteq G$ such that B has density $\beta > 0$ and B'', B''' are symmetric. Let $A \subseteq B$ with relative density $\alpha > 0$. Suppose that*

$$\|\mu_A \star \mu_A\|_{L^p(\mu)} \geq (1 + \epsilon)\beta^{-1}$$

for some $0 < \epsilon < 1$ and integer $p \geq 1$, where $\mu = (\mu_{B''} \star \mu_{B''}) * (\mu_{B'''} \star \mu_{B'''})$. Then there exist $d \in B'' + B'''$ and sets $A_1 \subseteq B''$ and $A_2 \subseteq B''' - d$ of relative density $\Omega(\alpha^{2p+O_\epsilon(1)})$ such that

$$\langle \mu_{A_1} \star \mu_{A_2}, \mathbb{1}_S \rangle \geq 1 - \tfrac{1}{4}\epsilon,$$

where $S = \{\mu_A \star \mu_A > (1 + \frac{1}{2}\epsilon)\beta^{-1}\}$.

Step 4: Density increment via almost-periodicity. In the finite field model, the codimension bound was crucial for keeping the size of the final subspace large. Here, we will need something to keep the size of the final Bohr set large. To do this, we can appeal to the Bohr set size bound (Lemma 2.7), which requires us to control the rank and radius. It is useful to keep in mind that α_1 and α_2 will satisfy $\log \alpha_i^{-1} \lesssim (1 + \log \alpha^{-1})^2$ as before, so the bounds really do turn out to be what we desire.

Lemma 4.6. *Let $r \geq 1$ be an integer. Let $B, B'', B''' \subseteq G$ be regular Bohr sets of rank r such that B has density $\beta > 0$. Let $A \subseteq B$ with relative density $\alpha > 0$. Suppose that there exist $d \in G$ and sets $A_1 \subseteq B''$ of relative density $\alpha_1 > 0$ and $A_2 \subseteq B''' - d$ of relative density $\alpha_2 > 0$ such that*

$$\langle \mu_{A_1} \star \mu_{A_2}, \mathbb{1}_S \rangle \geq 1 - \epsilon$$

for some $0 < \epsilon < 1/16$, where $S \subseteq \{\mu_A \star \mu_A \geq (1 + 2\epsilon)\beta^{-1}\}$ with $|S| \leq 2\,|B''|$. Then there exists a regular Bohr set $B^\dagger \subseteq B'''$ of rank at most

$$r + O_\epsilon((1 + \log \alpha^{-1})^2 (1 + \log \alpha_1^{-1})(1 + \log \alpha_2^{-1}))$$

and radius at least

$$\operatorname{radius}(B''') \exp(-O_\epsilon(1 + \log \alpha^{-1} + \log r + \log(1 + \log \alpha_1^{-1}) + \log(1 + \log \alpha_2^{-1})))$$

such that

$$\|\mu_A * \mu_{B^\dagger}\|_\infty \geq (1 + \tfrac{1}{4}\epsilon)\beta^{-1}.$$

Iterate. Now, all that is left to do is combine the steps of the density increment and iterate.

Proposition 4.7. *Let $r \geq 1$ be an integer. Let $B \subseteq G$ be a regular Bohr set of rank r and density $\beta > 0$, and let $A \subseteq B$ with relative density $\alpha > 0$. Let $B' = B_{\lambda'}$ for some $0 < \lambda' \leq \frac{\alpha}{16000r}$ be a regular Bohr set (of rank r) and $B'' \subseteq B'_{\frac{1}{400r}}$ be a regular Bohr set of rank r. Let $C \subseteq B'$ with relative density $\gamma > 0$. Suppose that*

$$\#\{(a_1, a_2, c) \in A \times A \times C : a_1 + a_2 = c\} \leq \tfrac{3}{4}\alpha^2\beta^{-1}\gamma\,|G|^2\,.$$

Then there exists a regular Bohr set $B^\dagger \subseteq B''$ such that:

- $\operatorname{rank}(B^\dagger) \leq r + O((1 + \log\alpha^{-1})^4(1 + \log\gamma^{-1})^2)$;
- $\operatorname{radius}(B^\dagger) \geq \operatorname{radius}(B'')\exp(-O((1 + \log\alpha^{-1} + \log r + \log(1 + \log\gamma^{-1}))))$; *and*
- *there exists $t \in G$ such that*

$$\frac{\left|(A - t) \cap B^\dagger\right|}{|B^\dagger|} \geq \tfrac{513}{512}\alpha.$$

We warn the reader that the iteration argument is far more technical in the integer setting than in the finite field model, due to the nature of Bohr sets. In Bloom and Sisask's rephrasing [BS23b], only Steps 3 and 4 were iterated, with other lemmas woven in between to complete the argument. In this presentation, we instead choose to iterate the entire density increment argument as to parallel the finite field model as much as possible.

4.2. **Proof of density increment steps.** We now prove the steps in the density increment for the integer setting. Throughout, we take advantage of the fact that Bohr sets are symmetric, and thus

$$(\mu_{B''} \star \mu_{B''}) * (\mu_{B'''} \star \mu_{B'''}) = \mu_{B''} * \mu_{B''} * \mu_{B'''} * \mu_{B'''}$$

for Bohr sets B'', B'''.

Step 1: Lifting. To prove Lemma 4.3, we will once again use an estimate to go from convolution to difference convolution. This time, we must switch measures due to using Bohr sets.

Lemma 4.8. *Let $B \subseteq G$ be a regular Bohr set of rank r, $f\colon G \to \mathbb{R}$, and $p \geq 1$ be an even integer. Then for any symmetric sets $B'', B''' \subseteq B_\lambda$ for some $0 < \lambda \leq \frac{1}{400r}$, we have that*

$$\|f * f\|_{L^p(\mu_B)} \leq 2^{1/p}\,\|f \star f\|_{L^p(\mu)}\,,$$

where $\mu = (\mu_{B''} \star \mu_{B''}) * (\mu_{B'''} \star \mu_{B'''})$.

Proof. We have that $\mu = \mu_{B''} * \mu_{B''} * \mu_{B'''} * \mu_{B'''}$ is supported on $B'' + B'' + B''' + B''' \subseteq 4B_\lambda$. So using the fact that p is even, we have that

$$\begin{aligned}
\|f * f\|^p_{L^p(\mu_B)} &= \mathbb{E}\,\mu_B \cdot (f * f)^p \\
&\overset{\text{Lem. 2.11}}{\leq} 2\,\mathbb{E}(\mu_{B_{1+4\lambda}} * \mu) \cdot (f * f)^p \\
&= 2 \mathop{\mathbb{E}}_{x\in G}\left(\mathop{\mathbb{E}}_{y\in G}\mu_{B_{1+4\lambda}}(y)\mu(x - y)\right)(f * f)(x)^p \\
&= 2 \mathop{\mathbb{E}}_{y\in G}\mu_{B_{1+4\lambda}}(y) \mathop{\mathbb{E}}_{x\in G}\mu(x - y)(f * f)(x)^p \\
&= 2 \mathop{\mathbb{E}}_{y\in B_{1+4\lambda}} \mathop{\mathbb{E}}_{x\in G}\mu(x - y)(f * f)(x)^p.
\end{aligned}$$

Thus there exists $y \in B_{1+4\lambda}$ such that

$$\begin{aligned}\|f*f\|_{L^p(\mu_B)}^p &\le 2 \mathop{\mathbb{E}}_{x\in G} \mu(x-y)(f*f)(x)^p \\ &= 2\langle \tau_{-y}\mu, (f*f)^p\rangle \\ &\overset{\text{Parseval}}{=} 2\left\langle \chi(-y)\hat{\mu}, (\hat{f^2})^{*p}\right\rangle.\end{aligned}$$

But $\hat{\mu} = |\widehat{\mu_{B''}}|^2|\widehat{\mu_{B'''}}|^2 \ge 0$, so the triangle inequality implies that

$$\begin{aligned}\|f*f\|_{L^p(\mu_B)}^p &\le 2\left\langle \hat{\mu}, (|\hat{f}|^2)^{*p}\right\rangle \\ &\overset{\text{Parseval}}{=} 2\langle \mu, (f\star f)^p\rangle \\ &= 2\|f\star f\|_{L^p(\mu)}^p.\end{aligned}$$

□

Proof of Lemma 4.3. Let $f = \mu_A - \mu_B$, and write

$$\langle f*f, \mu_C\rangle = \langle \mu_A*\mu_A, \mu_C\rangle - 2\langle \mu_A*\mu_B, \mu_C\rangle + \langle \mu_B*\mu_B, \mu_C\rangle.$$

Using the fact that B is symmetric so $\mu_B \star \mu_C = \mu_B * \mu_C$, we have that

$$\begin{aligned}\langle \mu_A*\mu_B, \mu_C\rangle &= \langle \mu_A, \mu_B\star\mu_C\rangle \\ &= \langle \mu_A, \mu_B*\mu_C - \mu_B\rangle + \|\mu_B\|_{L^1(\mu_A)} \\ &= \langle \mu_A, \mu_B*\mu_C - \mu_B\rangle + \beta^{-1}\end{aligned}$$

since $A \subseteq B$. By Hölder's inequality, we have that

$$|\langle \mu_A, \mu_B*\mu_C - \mu_B\rangle| \le \|\mu_A\|_\infty \|\mu_B*\mu_C - \mu_B\|_1 = (\alpha\beta)^{-1}\|\mu_B*\mu_C - \mu_B\|_1.$$

Since $C \subseteq B' \subseteq B_\lambda$ with $\lambda = \frac{\epsilon\alpha}{4000r} < 1$, Lemma 2.9 implies that

$$\|\mu_B*\mu_C - \mu_B\|_1 \le 200\lambda r = \tfrac{1}{20}\epsilon\alpha.$$

It follows that

$$\langle \mu_A*\mu_B, \mu_C\rangle \ge \left(1 - \tfrac{1}{20}\epsilon\right)\beta^{-1}.$$

Similarly,

$$\langle \mu_B*\mu_B, \mu_C\rangle \le \left(1 + \tfrac{1}{20}\epsilon\alpha\right)\beta^{-1} \le \left(1 + \tfrac{1}{20}\epsilon\right)\beta^{-1}.$$

Thus

$$\langle f*f, \mu_C\rangle \le (1-\epsilon)\beta^{-1} - 2\left(1 - \tfrac{1}{20}\epsilon\right)\beta^{-1} + \left(1 + \tfrac{1}{20}\epsilon\right)\beta^{-1} = -\tfrac{17}{20}\epsilon\beta^{-1}.$$

Since $C \subseteq B'$ and $\mathbb{1}_{B'}\,\mu_C = \gamma^{-1}\mu_{B'}\,\mathbb{1}_C$, we have that

$$\langle f*f, \mu_C\rangle = \gamma^{-1}\langle f*f, \mathbb{1}_C\rangle_{\mu_{B'}}.$$

Let $p = 2 + 2\lceil \log\gamma^{-1}\rceil$. Then

$$\begin{aligned}\tfrac{17}{20}\gamma\epsilon\beta^{-1} &\le \left|\langle f*f, \mathbb{1}_C\rangle_{\mu_{B'}}\right| \\ &\overset{\text{Hölder}}{\le} \|f*f\|_{L^p(\mu_{B'})}\|\mathbb{1}_C\|_{L^{p^*}(\mu_{B'})} \\ &= \|f*f\|_{L^p(\mu_{B'})}\gamma^{1/p^*},\end{aligned}$$

where $\frac{1}{p} + \frac{1}{p^*} = 1$. It follows that

$$\|f*f\|_{L^p(\mu_{B'})} \ge \tfrac{17}{20}\gamma^{1/p}\epsilon\beta^{-1} \ge 2^{1/p}\cdot\tfrac{1}{2}\epsilon\beta^{-1}$$

by the choice of p. The result follows from Lemma 4.8. □

Step 2: Unbalancing. We have already done the bulk of the work for this step in the finite field case, particularly in Lemma 3.13. The following lemma will help us break up the necessary bound into pieces.

Lemma 4.9. *Let B be a regular Bohr set of density $\beta > 0$, and let $S \subseteq B$ with relative density $\omega > 0$. For all $x \in B_\lambda$ for some $0 < \lambda < 1$, we have that*

$$\left|(\mu_S \star \mu_B)(x) - \beta^{-1}\right| \leq 200(\omega\beta)^{-1}\lambda \operatorname{rank}(B).$$

Proof. Check that

$$(\mu_S \star \mu_B)(0) = \|\mu_B\|_{L^1(\mu_S)} = \beta^{-1},$$

so we have that

$$\begin{aligned}
\left|(\mu_S \star \mu_B)(x) - \beta^{-1}\right| &= |(\mu_S \star \mu_B)(x) - (\mu_S \star \mu_B)(0)| \\
&= |(\mu_S \star \tau_x\mu_B - \mu_S \star \mu_B)(0)| \\
&= |\langle \mu_S, \tau_x\mu_B - \mu_B\rangle| \\
&\overset{\text{Hölder}}{\leq} \|\mu_S\|_\infty \|\tau_x\mu_B - \mu_B\|_1 \\
&= \omega^{-1} \|\tau_x\mu_B - \mu_B\|_1 .
\end{aligned}$$

But $\tau_x\mu_B = \mu_B * \mu_{\{-x\}}$, and $-x \in B_\lambda$ by symmetry of Bohr sets, so we can apply Lemma 2.9 to $\mu_{\{-x\}}$ to finish. □

Proof of Lemma 4.4. Let $f = \mu_A - \mu_B$ and $g = \mu_A \star \mu_A - f \star f = \mu_A \star \mu_B + \mu_B \star \mu_A - \mu_B \star \mu_B$. Apply Lemma 3.13 to $\beta f \star f$, which is spectrally nonnegative because its Fourier transform is $\beta|\hat{f}|^2$. Then

$$\left\|f \star f + \beta^{-1}\right\|_{L^{p'}(\nu)} \geq (1 + \tfrac{1}{2}\epsilon)\beta^{-1}$$

for some $p' = O_\epsilon(p)$. It suffices to show that

$$\left\|g - \beta^{-1}\right\|_{L^{p'}(\nu)} \leq \tfrac{1}{4}\epsilon\beta^{-1}.$$

Let $x \in \operatorname{supp}\nu \subseteq B_\lambda$, where $\lambda = \frac{\epsilon\alpha}{4000r} < 1$. Applying Lemma 4.9 with $S = A$ gives that

$$\left|(\mu_A \star \mu_B)(x) - \beta^{-1}\right| \leq 200(\alpha\beta)^{-1}\lambda r = \tfrac{1}{20}\epsilon\beta^{-1}.$$

Applying Lemma 4.9 with $S = A$ and x replaced by $-x$ gives that

$$\left|(\mu_B \star \mu_A)(x) - \beta^{-1}\right| \leq \tfrac{1}{20}\epsilon\beta^{-1}.$$

Applying Lemma 4.9 with $S = B$ gives that

$$\left|(\mu_B \star \mu_B)(x) - \beta^{-1}\right| \leq \tfrac{1}{20}\alpha\epsilon\beta^{-1} \leq \tfrac{1}{20}\epsilon\beta^{-1}.$$

It follows that

$$\left|g(x) - \beta^{-1}\right| \leq \tfrac{3}{20}\epsilon\beta^{-1}$$

for all $x \in \operatorname{supp}\nu$, so

$$\left\|g - \beta^{-1}\right\|_{L^{p'}(\nu)} \leq \left\|g - \beta^{-1}\right\|_{L^\infty(\nu)} \leq \tfrac{3}{20}\epsilon\beta^{-1} < \tfrac{1}{4}\epsilon\beta^{-1}$$

as desired. □

Step 3: Correlation. Again, we have already done most of the work for this step in the finite field case.

Proof of Lemma 4.5. We have that

$$\begin{aligned}
&\mathop{\mathbb{E}}_{s\in B''}\mathop{\mathbb{E}}_{t\in B'''} \|\mu_A \star \mu_A\|^p_{L^p(\mu_{B''} * \mu_{B'''+s+t})} \\
&= \mathop{\mathbb{E}}_{s\in G}\mathop{\mathbb{E}}_{t\in G} \mu_{B''}(s)\mu_{B'''}(t) \mathop{\mathbb{E}}_{x\in G}\mathop{\mathbb{E}}_{y\in G} \mu_{B''}(y)\mu_{B'''+s+t}(x-y)(\mu_A \star \mu_A)(x)^p \\
&= \mathop{\mathbb{E}}_{x\in G} \mathop{\mathbb{E}}_{(s,t,y)\in G^3} \mu_{B''}(s)\mu_{B'''}(t)\mu_{B''}(y)\mu_{B'''}(x-y-s-t)(\mu_A \star \mu_A)(x)^p.
\end{aligned}$$

Since $\mu = \mu_{B''} * \mu_{B''} * \mu_{B'''} * \mu_{B'''}$, this is just $\|\mu_A \star \mu_A\|_{L^p(\mu)}^p \geq ((1+\epsilon)\beta^{-1})^p$. It follows that there exist $s \in B''$ and $t \in B'''$ such that

$$\|\mu_A \star \mu_A\|_{L^p(\mu_{B''} * \mu_{B'''} + s + t)} \geq (1+\epsilon)\beta^{-1}.$$

Let $d = -s - t \in B'' + B'''$. Apply Lemma 3.17 with $B_1 = B''$, $B_2 = B''' - d$, ϵ replaced by $\frac{\epsilon/3}{1+\epsilon}$, and δ replaced by $\frac{1}{4}\epsilon$; loosen the constraint in S to the desired size. Note that $\mu_{B_1} * \mu_{B_2} = \mu_{B_1} \star \mu_{B_2}$ as $B_1 = B''$ is symmetric. The conclusion follows. □

Step 4: Density increment. For the integer setting, we will use the following strengthened L^∞-almost-periodicity result, which is essentially the same as a statement of Schoen and Sisask [SS16, Theorem 5.1]. Again, this result alone is enough for Kelley and Meka's original proof for $c = 1/12$, but an improved bootstrap by Bloom and Sisask allows us to achieve $c = 1/9$.

Proposition 4.10. *Let $0 < \epsilon < 1$, $\eta > 0$, $K \geq 2$, and $k \geq 1$ be an integer. Let $A_1, A_2, B, S \subseteq G$ be such that $|A_1| = \eta |S|$ and $|A_2 + B| \leq K |A_2|$. There exist $b \in B$ and $T \subseteq B - b$ of relative density at least*

$$\exp(-O_\epsilon(k^2 \max\{\log \eta^{-1}, 1\} \log K))$$

such that

$$\|\mu_T^{*k} * (\mu_{A_1} \star \mu_{A_2}) * \mathbb{1}_S - (\mu_{A_1} \star \mu_{A_2}) * \mathbb{1}_S\|_\infty \leq \epsilon.$$

We defer the proof of this result to Appendix A.

Proof of Lemma 4.6. We will identify such a $B^\dagger$ contained in $B'''_\kappa \subseteq B'''$, where $\kappa = \frac{\lambda}{100r}$ for some $\frac{1}{2} \leq \lambda \leq 1$ chosen so that B'''_κ is regular (by Lemma 2.8).

Apply Proposition 4.10 with ϵ replaced by $\frac{1}{3}\epsilon$, $\eta = \frac{|A_1|}{|S|}$, $k = 4 + 2\lceil \log \epsilon^{-1} \rceil + 2\lceil \log \alpha^{-1} \rceil$, B replaced by B'''_κ, and S replaced by $-S$. We can take $K = 2\alpha_2^{-1}$ because

$$\left|A_2 + B'''_\kappa\right| \leq \left|B''' - d + B'''_\kappa\right| \leq \left|B'''_{1+\kappa}\right| \overset{\text{reg.}}{\leq} (1+\lambda)\left|B'''\right| \leq 2\alpha_2^{-1}|A_2|.$$

Note that $\eta^{-1} \leq \frac{2|B''|}{|A_1|} = 2\alpha_1^{-1}$, so $\log \eta^{-1} \lesssim 1 + \log \alpha_1^{-1}$. Also $\log K \lesssim 1 + \log \alpha_2^{-1}$. Then by Proposition 4.10, there exist $b \in B'''_\kappa$ and $T \subseteq B'''_\kappa - b$ of relative density ω satisfying

$$\log \omega^{-1} \lesssim_\epsilon (1 + \log \alpha^{-1})^2 (1 + \log \alpha_1^{-1})(1 + \log \alpha_2^{-1})$$

such that $\|f * \mathbb{1}_{-S}\|_\infty \leq \frac{1}{3}\epsilon$, where $f = \mu_T^{*k} * (\mu_{A_1} \star \mu_{A_2}) - \mu_{A_1} \star \mu_{A_2}$.

Apply Sanders' local version of Chang's lemma (Lemma A.5) with $\delta = \frac{1}{8}\epsilon\alpha$, λ replaced by $1/2$, B replaced by B'''_κ, and $Y = T + b$. Then there exists a Bohr set $B^\dagger \subseteq B'''_\kappa$ of rank

$$\begin{aligned}\operatorname{rank}(B^\dagger) &\leq \operatorname{rank}(B'''_\kappa) + O(1 + \log \omega^{-1}) \\ &\lesssim r + O_\epsilon((1 + \log \alpha^{-1})^2 (1 + \log \alpha_1^{-1})(1 + \log \alpha_2^{-1}))\end{aligned}$$

and radius satisfying

$$\begin{aligned}\operatorname{radius}(B^\dagger) &\gtrsim \frac{\operatorname{radius}(B'''_\kappa) \cdot \frac{1}{8}\epsilon\alpha}{\operatorname{rank}(B'''_\kappa)^2 (1 + \log \omega^{-1})} \\ &\gtrsim \frac{r^{-1} \operatorname{radius}(B''')\epsilon\alpha}{r^2 (1 + \log \alpha^{-1})^2 (1 + \log \alpha_1^{-1})(1 + \log \alpha_2^{-1})} \\ &\gtrsim \operatorname{radius}(B''') \exp(-O_\epsilon(1 + \log \alpha^{-1} + \log r + \log(1 + \log \alpha_1^{-1}) + \log(1 + \log \alpha_2^{-1})))\end{aligned}$$

such that $|1 - \chi(x)| \leq \frac{1}{8}\epsilon\alpha$ for all $\chi \in \operatorname{Spec}_{1/2}(\mu_{T+b}) = \operatorname{Spec}_{1/2}(\mu_T)$ and $x \in B^\dagger$.

The density increment is computed in essentially the same manner as in the finite field model. We skip computations that are identical to the finite field case. We have that

$$\left((\mu_T^{*k} * (\mu_{A_1} \star \mu_{A_2})) \star (\mathbb{1}_S)\right)(0) \geq \langle \mu_{A_1} \star \mu_{A_2}, \mathbb{1}_S \rangle - \|f * \mathbb{1}_{-S}\| \geq 1 - \tfrac{4}{3}\epsilon.$$

Letting $\mu = (\mu_T^{*k} * (\mu_{A_1} \star \mu_{A_2})) \star (\mu_A \star \mu_A)$, this implies that

$$\mu(0) \geq (1 + 2\epsilon)(1 - \tfrac{4}{3}\epsilon)\beta^{-1} > (1 + \tfrac{1}{2}\epsilon)\beta^{-1}.$$

For all $x \in G$, we have that

$$\begin{aligned}|\mu(x) - \mu(0)| &= \left|\sum_{\chi \in \widehat{G}} \hat{\mu}(\chi)(\chi(x) - 1)\right| \\ &\leq \sum_{\chi \in \widehat{G}} |\widehat{\mu_T}(\chi)|^k |\widehat{\mu_{A_1}}(\chi)||\widehat{\mu_{A_2}}(\chi)||\widehat{\mu_A}(\chi)|^2 \, |\chi(x) - 1|\,.\end{aligned}$$

But observe that:

- if $\chi \in \mathrm{Spec}_{1/2}(\mu_T)$ and $x \in B^\dagger$, then $|1 - \chi(x)| \leq \frac{1}{8}\epsilon\alpha$;
- if $\chi \notin \mathrm{Spec}_{1/2}(\mu_T)$, then $|\widehat{\mu_T}(\chi)| < 1/2$; and
- $|\widehat{\mu_{A_1}}(\chi)| \leq \|\mu_{A_1}\|_1 \leq 1$, and similarly $|\widehat{\mu_{A_2}}(\chi)| \leq 1$ and $|\widehat{\mu_T}(\chi)| \leq 1$.

In particular, $|\widehat{\mu_T}(\chi)|^k \, |\chi(x) - 1| \leq \frac{1}{8}\epsilon\alpha + 2^{-k} \cdot 2$ for all $x \in B^\dagger$. It follows that for all $x \in B^\dagger$, we have that

$$\begin{aligned}|\mu(x) - \mu(0)| &\leq \sum_{\chi \in \widehat{G}} |\widehat{\mu_A}(\chi)|^2 \cdot (\tfrac{1}{8}\epsilon\alpha + 2^{-k} \cdot 2) \\ &\overset{\text{Parseval}}{=} (\tfrac{1}{8}\epsilon\alpha + 2^{1-k})\|\mu_A\|_2^2 \\ &= (\tfrac{1}{8}\epsilon\alpha + 2^{1-k})(\alpha\beta)^{-1} \\ &\leq \tfrac{1}{4}\epsilon\beta^{-1}\end{aligned}$$

by the choice of k. Then we have that

$$(\mu_{B^\dagger} \star \mu)(0) \geq \mu(0) - \|\mu - \mu(0)\|_{L^\infty(\mu_{B^\dagger})} > (1 + \tfrac{1}{2}\epsilon\beta^{-1}) - \tfrac{1}{4}\epsilon\beta^{-1},$$

so

$$\|\mu_A * \mu_{B^\dagger}\|_\infty \overset{\text{Hölder}}{\geq} (\mu_{B^\dagger} \star \mu)(0) > (1 + \tfrac{1}{4}\epsilon\beta^{-1}). \qquad \square$$

4.3. **Finishing the argument.** We must now tie together the steps of the density increment. We restate the boost here for convenience. The proof is more involved than the simple combination that took place in the finite field model.

Proposition 4.7. *Let $r \geq 1$ be an integer. Let $B \subseteq G$ be a regular Bohr set of rank r and density $\beta > 0$, and let $A \subseteq B$ with relative density $\alpha > 0$. Let $B' = B_{\lambda'}$ for some $0 < \lambda' \leq \frac{\alpha}{16000r}$ be a regular Bohr set (of rank r) and $B'' \subseteq B'_{\frac{1}{400r}}$ be a regular Bohr set of rank r. Let $C \subseteq B'$ with relative density $\gamma > 0$. Suppose that*

$$\#\{(a_1, a_2, c) \in A \times A \times C \,:\, a_1 + a_2 = c\} \leq \tfrac{3}{4}\alpha^2\beta^{-1}\gamma\,|G|^2\,.$$

Then there exists a regular Bohr set $B^\dagger \subseteq B''$ such that:

- $\mathrm{rank}(B^\dagger) \leq r + O((1 + \log\alpha^{-1})^4(1 + \log\gamma^{-1})^2)$;
- $\mathrm{radius}(B^\dagger) \geq \mathrm{radius}(B'')\exp(-O((1 + \log\alpha^{-1} + \log r + \log(1 + \log\gamma^{-1}))))$; *and*
- *there exists $t \in G$ such that*

$$\frac{\left|(A - t) \cap B^\dagger\right|}{|B^\dagger|} \geq \tfrac{513}{512}\alpha.$$

Proof. By a similar calculation to Remark 3.4, the hypothesis is equivalent to

$$\langle \mu_A * \mu_A, \mu_C \rangle \leq \tfrac{3}{4}\beta^{-1}.$$

Let $B''' = B''_{\frac{1}{400r}}$ be a regular Bohr set (of rank r). Let $\mu = (\mu_{B''} \star \mu_{B''}) * (\mu_{B'''} \star \mu_{B'''})$.

Step 1: Apply Lemma 4.3 with $\epsilon = 1/4$ to deduce that

$$\|(\mu_A - \mu_B) \star (\mu_A - \mu_B)\|_{L^p(\mu)} \geq \tfrac{1}{8}\beta^{-1}$$

for some $p = O(1 + \log\gamma^{-1})$.

Step 2: We wish to apply Lemma 4.4 with $\nu = \mu$. We have that $\hat{\mu} = |\widehat{\mu_{B''}}|^2|\widehat{\mu_{B'''}}|^2 \geq 0$. In addition,

$$\operatorname{supp}\mu = B'' + B'' + B''' + B''' \subseteq 4B'' \subseteq 4B'_{\lambda''} \subseteq B'_{4\lambda''} = B_{4\lambda'\lambda''}$$

with $4\lambda'\lambda'' \leq \frac{\alpha}{1600000r^2} < \frac{\alpha/8}{4000r}$. So we can apply Lemma 4.4 with $\epsilon = 1/8$ and $\nu = \mu$ to deduce that

$$\|\mu_A \star \mu_A\|_{L^{p'}(\mu)} \geq (1 + \tfrac{1}{32})\beta^{-1}$$

for some $p' = O(1 + \log\gamma^{-1})$.

Step 3: Apply Lemma 4.5 with $\epsilon = 1/32$ and p replaced by $\lceil p' \rceil$. We get that there exist $d \in B'' + B'''$ and sets $A_1 \subseteq B''$ of relative density $\alpha_1 = \Omega(\alpha^{2p+O(1)})$ and $A_2 \subseteq B''' - d$ of relative density $\alpha_2 = \Omega(\alpha^{2p+O(1)})$ such that

$$\langle \mu_{A_1} \star \mu_{A_2}, \mathbb{1}_S \rangle \geq 1 - \tfrac{1}{128},$$

where $S = \{\mu_A \star \mu_A > (1 + \frac{1}{64})\beta^{-1}\}$.

Step 4: We wish to apply Lemma 4.6, but there is a size restriction on S. We can massage our S by taking $\tilde{S} = S \cap (A_2 - A_1)$. Indeed, $\mu_{A_1} \star \mu_{A_2}$ is supported on $A_2 - A_1$, so the correlations $\langle \mu_{A_1} \star \mu_{A_2}, \mathbb{1}_S \rangle$ and $\langle \mu_{A_1} \star \mu_{A_2}, \mathbb{1}_{\tilde{S}} \rangle$ are the same. As for the size restriction, check that

$$|\tilde{S}| \leq |A_2 - A_1| \leq \left|B''' - d - B''\right| = \left|B'' + B'''\right| \leq \left|B''_{1+\frac{1}{100r}}\right| \overset{\text{reg.}}{\leq} 2\left|B''\right|$$

as desired. So we can apply Lemma 4.6 with $\epsilon = 1/128$ and S replaced by $\tilde{S}$ to produce a regular Bohr set $B^\dagger \subseteq B'''$ such that

$$\|\mu_A * \mu_{B^\dagger}\|_\infty \geq (1 + \tfrac{1}{512})\beta^{-1}.$$

Using the fact that $\log\alpha_i^{-1} \lesssim (1 + \log\alpha^{-1})(1 + \log\gamma^{-1})$, this Bohr set $B^\dagger$ has the correct rank. It also has the correct radius with $\operatorname{radius}(B''')$ in place of $\operatorname{radius}(B'')$. But

$$\operatorname{radius}(B''') = \frac{1}{400r}\operatorname{radius}(B'') = \operatorname{radius}(B'')\exp(-O(1 + \log r)),$$

so the radius bound is also correct. And clearly $B^\dagger \subseteq B''' \subseteq B''$.

In the same manner as Remark 3.5, this implies the desired conclusion. □

Given this boost, we demonstrate how to prove the upper bound on $r_3(N)$ by iterating this density increment. First, we prove the counting result that a subset of $\{1, \ldots, N\}$ of size αN contains $N^2/\exp(O((1 + \log\alpha^{-1})^9))$ three-term arithmetic progressions. The iteration is far more technical than in the finite field model.

Proof of Theorem 4.1. Embed A into $G = \mathbb{Z}/(2N+1)\mathbb{Z}$, noting that arithmetic progressions in $\{1, \ldots, N\} \subseteq G$ correspond to those in $\{1, \ldots, N\} \subseteq \mathbb{Z}$. Replace α with the density of A in G, which is fine for the desired bound as it only affects α by a factor of at most 3.

Consider the following process:

(1) Initialise $\mathcal{B}_{(0)} = G$, $A_0 = A$, and $i = 0$.
(2) Assert that $\mathcal{B}_{(i)} \subseteq G$ is a regular Bohr set and $A_i \subseteq \mathcal{B}_{(i)}$.
(3) Set the following:
- $\alpha_i = \frac{|A_i|}{|\mathcal{B}_{(i)}|}$;
- $r_i = \operatorname{rank}(\mathcal{B}_{(i)})$;
- $B_{(i)} = (\mathcal{B}_{(i)})_{\lambda_i}$, where $\lambda_i = \frac{c_i\alpha_i}{r_i}$ for some $\frac{1}{2^{21}} \leq c_i \leq \frac{1}{2^{20}}$ so that $B_{(i)}$ is regular;
- $\beta_i = \frac{|B_{(i)}|}{|G|}$;
- $B'_{(i)} = (B_{(i)})_{\lambda'_i}$, where $\lambda'_i = \frac{c'_i\alpha_i}{r_i}$ for some $\frac{1}{2^{15}} \leq c'_i \leq \frac{1}{2^{14}}$ so that $B'_{(i)}$ is regular; and
- $B''_{(i)} = (B'_{(i)})_{\lambda''_i}$, where $\lambda''_i = \frac{c''_i\alpha_i}{r_i}$ for some $\frac{1}{2^{11}} \leq c''_i \leq \frac{1}{2^{10}}$ so that $B''_{(i)}$ is regular.

(4) If there exists $t_i \in G$ such that

$$\frac{\left|(A_i - t_i) \cap B_{(i)}\right|}{\left|B_{(i)}\right|} \geq (1 + \tfrac{1}{2048})\alpha_i,$$

then set $A_{i+1} = (A_i - t_i) \cap B_{(i)}$ and $\mathcal{B}_{(i+1)} = B_{(i)}$. Increment i and go back to step (2).

(5) If there exists $t_i \in G$ such that

$$\frac{\left|(A_i - t_i) \cap B'_{(i)}\right|}{\left|B'_{(i)}\right|} \geq (1 + \tfrac{1}{2048})\alpha_i,$$

then set $A_{i+1} = (A_i - t_i) \cap B'_{(i)}$ and $\mathcal{B}_{(i+1)} = B'_{(i)}$. Increment i and go back to step (2).

(6) Otherwise, apply Bourgain's narrowing trick (Lemma 2.10) with B replaced by $\mathcal{B}_{(i)}$, A replaced by A_i, ϵ replaced by $\frac{1}{1024}$, B_1 replaced by $B_{(i)}$, and B_2 replaced by $B'_{(i)}$. These parameters are valid by the bounds on the constant c_i. Then we are in the second alternative, so there exists $d_i \in G$ such that

$$\frac{\left|(A_i - d_i) \cap B_{(i)}\right|}{\left|B_{(i)}\right|} \geq (1 - \tfrac{1}{1024})\alpha_i \quad \text{and} \quad \frac{\left|(A_i - d_i) \cap B'_{(i)}\right|}{\left|B'_{(i)}\right|} \geq (1 - \tfrac{1}{1024})\alpha_i.$$

Let $\tilde{A}_i = (A_i - d_i) \cap B_{(i)}$ and $\tilde{A}'_i = (A_i - d_i) \cap B'_{(i)}$ so that $\tilde{A}_i \subseteq B_{(i)}$ and $\tilde{A}'_i \subseteq B'_{(i)}$ both have relative density at least $\frac{1023}{1024}\alpha_i$.

(7) If

$$\#\left\{(x, y, z) \in \tilde{A}_i \times \tilde{A}'_i \times \tilde{A}_i \,:\, x + z = 2y\right\} > \tfrac{3}{4}\beta_i^{-1}\frac{|\tilde{A}_i|^2|\tilde{A}'_i|}{|G|},$$

set $m = i$ and **STOP**.

(8) Otherwise, apply Proposition 4.7 with B replaced by $B_{(i)}$, A replaced by $\tilde{A}_i$, B' replaced by $B'_{(i)}$, B'' replaced by $2 \cdot B''_{(i)}$, and C replaced by $2 \cdot \tilde{A}'_i$. Note that α is replaced by a density that is at least $\frac{1023}{1024}\alpha_i$. These parameters are valid by the bounds on the constants c'_i, c''_i and observing that $2 \cdot B''_{(i)} = 2 \cdot (B'_{(i)})_{\lambda''_i} \subseteq (B'_{(i)})_{2\lambda''_i}$. Then there exists a regular Bohr set $B^\dagger \subseteq 2 \cdot B''_{(i)}$ such that:

- $\operatorname{rank}(B^\dagger) \leq r_i + O((1 + \log \alpha_i^{-1})^6)$;
- $\operatorname{radius}(B^\dagger) \geq \operatorname{radius}(2 \cdot B''_{(i)}) \exp(-O(1 + \log \alpha_i^{-1} + \log r_i))$; and
- there exists $d'_i \in G$ such that

$$\frac{|(\tilde{A}_i - d'_i) \cap B^\dagger|}{|B^\dagger|} \geq \tfrac{513}{512} \cdot \tfrac{1023}{1024}\alpha_i \geq (1 + \tfrac{1}{2048})\alpha_i.$$

(9) Set $A_{i+1} = (\tilde{A}_i - d'_i) \cap B^\dagger$ and $\mathcal{B}_{(i+1)} = B^\dagger$.

(10) Increment i and go back to step (2).

Now, we analyze the procedure. The only way to terminate is at step (7). At this point, we have accumulated m density increments of the form $\alpha_{i+1} \geq (1 + \frac{1}{2048})\alpha_i$, so $m \lesssim \log \alpha^{-1}$. At each step, the rank either stays the same or is increased by $O((1+\log \alpha_i^{-1})^6) \leq O((1+\log \alpha^{-1})^6)$, so

$$\begin{aligned}\operatorname{rank}(\mathcal{B}_{(m)}) &\leq \operatorname{rank}(\mathcal{B}_{(0)}) + O(m(1 + \log \alpha^{-1})^6) \\ &\lesssim (1 + \log \alpha^{-1})^7.\end{aligned}$$

We also have the same bound on r_i for any $i \leq m$.

Recall that $\operatorname{radius}(2 \cdot B''_{(i)}) = \operatorname{radius}(B''_{(i)})$. Note that

$$\operatorname{radius}(B_{(i)}) = \Omega(\alpha_i/r_i) \operatorname{radius}(\mathcal{B}_{(i)}) = \operatorname{radius}(\mathcal{B}_{(i)}) \exp(-O(1 + \log \alpha_i^{-1} + \log r_i))$$

and similarly with the other Bohr sets, so it is always true that

$$\begin{aligned}\operatorname{radius}(\mathcal{B}_{(i+1)}) &\geq \operatorname{radius}(\mathcal{B}_{(i)}) \exp(-O_{\epsilon,\delta}(1 + \log \alpha_i^{-1} + \log r_i)) \\ &\geq \operatorname{radius}(\mathcal{B}_{(i)}) \exp(-O_{\epsilon,\delta}(1 + \log \alpha^{-1})).\end{aligned}$$

Thus

$$\begin{aligned}\operatorname{radius}(\mathcal{B}_{(m)}) &\geq \operatorname{radius}(\mathcal{B}_{(0)}) \exp(-O(m(1 + \log \alpha^{-1}))) \\ &\geq \exp(-O((1 + \log \alpha^{-1})^2)).\end{aligned}$$

It follows by the Bohr set size bound (Lemma 2.7) that

$$\left|\mathcal{B}_{(m)}\right| \geq |G| \exp(-O((1 + \log \alpha^{-1})^9)).$$

Now, we take into account the termination condition in step (7). Since $\tilde{A}_i$ and $\tilde{A}_i'$ are subsets of the same translate of A, the left-hand side is at most the number of solutions to $x + z = 2y$ in A. Thus it suffices to compute the right-hand side. We have

$$|\tilde{A}_m| = \tfrac{|\tilde{A}_m|}{\left|B_{(m)}\right|} \cdot \tfrac{\left|B_{(m)}\right|}{\left|\mathcal{B}_{(m)}\right|} \cdot \left|\mathcal{B}_{(m)}\right|.$$

The first term satisfies

$$\tfrac{|\tilde{A}_m|}{\left|B_{(m)}\right|} \geq \tfrac{1023}{1024}\alpha_m \geq \exp(-O(1 + \log \alpha^{-1})).$$

The second term satisfies

$$\tfrac{\left|B_{(m)}\right|}{\left|\mathcal{B}_{(m)}\right|} \overset{\text{Lem. 2.7}}{\geq} (\lambda_m/4)^{r_m} \geq (\Omega(\alpha/r_m))^{r_m} \geq \exp(-O((1 + \log \alpha^{-1})^8)).$$

This means that the $\left|\mathcal{B}_{(m)}\right|$ term is the main term, so

$$|\tilde{A}_m| \geq |G| \exp(-O((1 + \log \alpha^{-1})^9)).$$

A similar conclusion is reached for A_m'', so

$$\begin{aligned}\#\{\text{3-APs in } A\} &> |G|^2 \exp(-O((1 + \log \alpha^{-1})^9)) \\ &\geq N^2/\exp(O((1 + \log \alpha^{-1})^9)).\end{aligned}$$

□

As with the finite field model, the extremal result is immediate.

Proof of Kelley–Meka (Theorem 1.2). Suppose that $A \subseteq \{1, \dots, N\}$ with size $|A| = \alpha N$ has no nontrivial three-term arithmetic progressions. Then $\#\{\text{3-APs in } A\} = |A|$. By Theorem 4.1, we have that

$$1 \geq \alpha \geq N/\exp(O((1 + \log \alpha^{-1})^9)).$$

Solving for α gives the desired bound. □

This concludes the proof of the current best upper bound on $r_3(N)$, provided by Bloom and Sisask using the Kelley–Meka method.

5. Related problems

There are numerous extensions to the problem posed by Roth's theorem.

5.1. **Szemerédi's theorem.** Perhaps the most obvious extension is to ask to forbid k-term arithmetic progressions for arbitrary fixed $k \geq 3$. To that end, let $r_k(N)$ denote the maximum size of a subset of $\{1, \dots, N\}$ with no nontrivial k-term arithmetic progression. Then Roth's theorem generalises to Szemerédi's theorem [Sze75].

Theorem 5.1 (Szemerédi)**.** *For any fixed integer $k \geq 3$, we have that $r_k(N) = o_k(N)$.*

Szemerédi's original proof relied on his powerful regularity lemma which provides tremendous structural results but has terrible quantitative bounds. For example, Szemerédi's proof gives that $r_3(N) \leq N/(\log^* N)^c$ for some $c > 0$, where $\log^*$ denotes the iterated logarithm—and larger k only get worse along the Ackermann hierarchy.

In a breakthrough, Gowers [Gow01] improved the upper bound to $r_k(N) \leq N/(\log\log N)^{c_k}$ for $c_k = 2^{-2^{k+9}}$ by introducing higher order Fourier analysis in order to parallel the Fourier-analytic proof of Roth's theorem. For arbitrary k, this was not improved until very recently, when Leng, Sah, and Sawhney proved quasi-polynomial bounds on the inverse theorem for the Gowers U^{k+1}-norm [LSS24b] and applied this to prove an upper bound of

$$r_k(N) \leq N/\exp((\log\log N)^{c_k})$$

for some effective constant $c_k > 0$ [LSS24a]. This Leng–Sah–Sawhney upper bound is the current best for any $k \geq 5$. For $k = 4$, Green and Tao [GT17] previously proved that $r_4(N) \lesssim N/(\log N)^c$ for some effective constant $c > 0$.

There is still a substantial gap between the upper and lower bounds for Szemerédi's theorem, as the current best lower bound for $r_k(N)$ is of a similar quasi-polynomial shape as the Behrend lower bound, provided by O'Bryant [O'B11]:

$$r_k(N) \gtrsim N(\log N)^{O_k(1)}/\exp(O_k((\log N)^{O_k(1)})),$$

where each of the O_k terms have effective implicit constants.

As with the three-term arithmetic progression case, breaking the logarithmic barrier in the denominator would provide more cases of the conjecture of Erdős that sets of positive integers with divergent reciprocal sum must contain arbitrarily long arithmetic progressions.

It is worth noting that the extremal problem for k-term arithmetic progressions in the finite field model is still wide open. Recall that $r_3(\mathbb{F}_q^n)$ has been bounded between two nontrivial exponentials for all odd primes q, so the shape of the bounds is correct. Let $r_k(\mathbb{F}_q^n)$ denote the analogous quantity for k-term arithmetic progressions over $\mathbb{F}_q^n$ (for q large enough so that modular arithmetic constraints do not affect such progressions). The current best upper bound for $k = 4$ is due to Green and Tao [GT09, GT12], similar to their bound in the integers: $r_4(\mathbb{F}_q^n) \lesssim q^n/n^c$ for $c = 2^{-22}$.

5.2. **Corners.** A closely related object to arithmetic progressions is the **corner**, i.e. a set of the form $\{(x, y), (x, y + d), (x + d, y)\}$. Let $r_\llcorner(N)$ denote the maximum size of a subset of $\{1, \dots, N\}^2$ with no nontrivial corners. One reason that corners are relevant when discussing Roth's theorem is the relation between avoiding corners and avoiding three-term arithmetic progressions.

Lemma 5.2. *We have that $r_3(N) \leq \frac{r_\llcorner(2N)}{N}$.*

Proof. Suppose $A \subseteq \{1, \dots, N\}$ contains no nontrivial three-term arithmetic progressions. Consider the set $\{(x, y) \in \{1, \dots, 2N\} : x - y \in A\}$. This set has at least $|A|\, N$ elements, namely $(y + a, y)$ for $a \in A, y \in \{1, \dots, N\}$. But it is corner-free: if $(x, y), (x, y + d), (x + d, y)$ were all in the set with $d \neq 0$, then $x - y - d, x - y, x - y + d$ would be a nontrivial three-term arithmetic progression in A. Thus $|A|\, N \leq r_\llcorner(2N)$. Taking the maximum over all such A finishes. □

In a similar result to Theorem 5.1, Ajtai and Szemerédi [AS74] proved the following.

Theorem 5.3. *We have that $r_\llcorner(N) = o(N^2)$.*

The original proof used Szemerédi's theorem as a black box, though a much simpler proof was given by Solymosi [Sol03] by invoking the triangle removal lemma that follows from the regularity lemma. As previously mentioned, quantitative bounds for the regularity lemma and the triangle removal lemma are quite poor, so alternate methods must be used to produce a better upper bound. Indeed, Shkredov [Shk05] applied a density increment argument using the box norm to prove a much better bound of $r_{\llcorner}(N) \lesssim N^2/(\log\log N)^c$ for $c = 1/73$. Lower bounds for $r_{\llcorner}(N)$ also follow Behrend-type quasi-polynomial densities, with the current best constants coming from Green [Gre21].

In the finite field model, there is a similar situation as to k-term arithmetic progressions: the upper bound is polylogarithmic due to Lacey and McClain [LM07], while the lower bound is exponential due to Christandl, Fawzi, Ta, and Zuiddam [CFTZ22]. However, the situation is actually more dire. While the Croot–Lev–Pach polynomial method has some hope of achieving exponential savings in large arithmetic progressions, a result of Christandl, Fawzi, Ta, and Zuiddam [CFTZ22, Theorem 8] shows that existing tensor-based methods (including the slice rank method, which is a reformulation of the Croot–Lev–Pach polynomial method) cannot provide good upper bounds on corner-free sets in $\mathbb{F}_q^n$.

Beyond corners, various multidimensional shapes have similar extremal questions. One particularly interesting shape is the **skew corner**, i.e. a set of the form $\{(x, y), (x, y+d), (x+d, y')\}$. In a quick success of the Kelley–Meka method, Milićević [Mil24] and Jaber, Lovett, and Ostuni [JLO24] independently applied the method to prove a quasi-polynomial shape upper bound on the size of sets with no nontrivial skew corners. As with three-term arithmetic progressions, this matches the best lower bound of Beker [Bek24] up to the power of the $\log N$ term in the exponent. There is good reason to believe that the Kelley–Meka method will be successful in improving upper bounds in many other similar settings.

5.3. **Sumsets containing arithmetic progressions.** In their original paper, Kelley and Meka [KM23] also applied their method to find large structured subspaces in $A + A + A$ for $A \subseteq \mathbb{F}_q^n$. With the reformulation of Bloom and Sisask in terms of only Bohr sets, a similar result [BS23a, Theorem 3] is able to be proven in the integers.

Theorem 5.4. *Let $A \subseteq \{1, \dots, N\}$ have size $|A| = \alpha N$ for some $\alpha > 0$. Then $A + A + A$ contains an arithmetic progression of length at least*

$$\exp(-O((1 + \log \alpha^{-1})^2))N^{\Omega((1+\log \alpha^{-1})^{-7})}.$$

The proof of this theorem follows the exact same steps as the proof of Theorem 1.2, but with different parameters. This narrows the gap on this problem in a similar fashion to the bounds on $r_3(N)$: the previous best lower bound due to Sanders [San08] had exponent $\alpha^{1+o(1)}$, while the current best construction due to Freiman, Halberstam, and Ruzsa [FHR92] gives a lower bound with exponent $O((1 + \log \alpha^{-1})^{-1})$.

5.4. **Back to Roth's theorem.** Turning our attention back to the problem of three-term arithmetic progressions, there is still much work to be done. Closing the gap on the power of the $\log N$ term in the exponent is an enticing open problem. Bloom and Sisask suggest that $c = 1/7$ is the limit of this technique without substantially new ideas—indeed, the rank bound in Lemma 4.6 necessarily carries a $\log \alpha_1^{-1}$ and $\log \alpha_2^{-1}$ term, and as previously discussed, each of those contributes 2 to the exponent of the $\log \alpha^{-1}$ term. Thus the best bound we could hope for is on the order of $(1 + \log \alpha^{-1})^4$, which would give $c = 1/7$.

Whether other methods may be applicable is also unclear—perhaps a revolutionary technique similar to the polynomial method is needed to improve the bounds further. Regardless, the innovative method of Kelley and Meka has proven to be extremely powerful in advancing the state of knowledge in additive combinatorics, centered on three-term arithmetic progressions but sure to be widely applicable elsewhere in the field.

Appendix A. Almost-periodicity

In this appendix, we derive the almost-periodicity result needed for the Kelley–Meka method from the original Croot–Sisask almost-periodicity and provide statements of related results. We will use the following statement of Croot–Sisask almost-periodicity [CS10] as stated and proven by Sanders [San12b, Lemma 4.3].

Theorem A.1 (Croot–Sisask almost-periodicity)**.** *Let $0 < \epsilon < 1$, $p \geq 2$, and $K \geq 2$. Let $A, B \subseteq G$ be such that $|A + B| \leq K|A|$. Let $f\colon G \to \mathbb{C}$. Then there exist $b \in B$ and $T \subseteq B - b$ of relative density at least $\exp(-O(\epsilon^{-2} p \log K))$ such that*

$$\|\tau_t(\mu_A * f) - \mu_A * f\|_p \leq \epsilon \|f\|_p$$

for all $t \in T$.

We can aggregate the results for each of the almost-periods to tack on an additional convolution factor.

Corollary A.2. *Let $0 < \epsilon < 1$, $p \geq 2$, $K \geq 2$, and $k \geq 1$ be an integer. Let $A, B \subseteq G$ be such that $|A + B| \leq K|A|$. Let $f\colon G \to \mathbb{C}$. Then there exist $b \in B$ and $T \subseteq B - b$ of relative density at least $\exp(-O(k^2 \epsilon^{-2} p \log K))$ such that*

$$\|\mu_T^{*k} * \mu_A * f - \mu_A * f\|_p \leq \epsilon \|f\|_p.$$

Proof. Apply Theorem A.1 with ϵ replaced by $\frac{1}{k}\epsilon$. Then there exist $b \in B$ and $T \subseteq B - b$ of relative density at least $\exp(-O(k^2 \epsilon^{-2} p \log K))$ such that

$$\|\tau_t(\mu_A * f) - \mu_A * f\|_p \leq \tfrac{1}{k}\epsilon \|f\|_p$$

for all $t \in T$. Since shifting a function does not change its L^p-norm, the triangle inequality implies that

$$\left\|\tau_{-(t_1+\cdots+t_k)}(\mu_A * f) - \mu_A * f\right\|_p \leq \epsilon \|f\|_p$$

for all $t_1, \ldots, t_k \in T$.

Let $\mu = \mu_T^{*k}$ and $g = \mu_A * f$. Then with $\frac{1}{p} + \frac{1}{p^*} = 1$, we have that

$$\begin{aligned}
\|\mu * g - g\|_p &= \left(\mathop{\mathbb{E}}_{x \in G} \left| \mathop{\mathbb{E}}_{y} \mu(y)(g(x-y) - g(x)) \right|^p \right)^{\frac{1}{p}} \\
&= \left(\mathop{\mathbb{E}}_{x \in G} \left| \langle 1, g(x - \cdot) - g(x) \rangle_\mu \right|^p \right)^{\frac{1}{p}} \\
&\overset{\text{Hölder}}{\leq} \left(\mathop{\mathbb{E}}_{x \in G} \|1\|_{L^{p^*}(\mu)}^p \, \|g(x - \cdot) - g(x)\|_{L^p(\mu)}^p \right)^{\frac{1}{p}} \\
&= \left(\mathop{\mathbb{E}}_{x \in G} \left(\mathop{\mathbb{E}}_{y \in G} \mu(y) \, |g(x-y) - g(x)|^p \right) \right)^{\frac{1}{p}} \\
&= \left(\mathop{\mathbb{E}}_{y \in G} \mu(y) \left(\mathop{\mathbb{E}}_{x \in G} |g(x-y) - g(x)|^p \right) \right)^{\frac{1}{p}} \\
&\leq \max_{y \in \operatorname{supp} \mu} \|\tau_{-y} g - g\|_p .
\end{aligned}$$

But $\operatorname{supp} \mu = kT$, so this is bounded by $\epsilon \|f\|_p$ as desired. □

From this version of L^p-almost-periodicity, the L^∞-almost-periodicity results that are needed for the Kelley–Meka method are immediate. We restate the more general result here for convenience.

Proposition 4.10. *Let $0 < \epsilon < 1$, $\eta > 0$, $K \geq 2$, and $k \geq 1$ be an integer. Let $A_1, A_2, B, S \subseteq G$ be such that $|A_1| = \eta |S|$ and $|A_2 + B| \leq K |A_2|$. There exist $b \in B$ and $T \subseteq B - b$ of relative density at least*

$$\exp(-O_\epsilon(k^2 \max\{\log \eta^{-1}, 1\} \log K))$$

such that

$$\|\mu_T^{*k} * (\mu_{A_1} \star \mu_{A_2}) * \mathbb{1}_S - (\mu_{A_1} \star \mu_{A_2}) * \mathbb{1}_S\|_\infty \leq \epsilon.$$

Proof of Propositions 4.10 and 3.21. Observe that $\mu_{-A_1} * \mu_{A_2} = \mu_{A_1} \star \mu_{A_2}$. Apply Corollary A.2 with ϵ replaced by $\frac{1}{2}\epsilon$, $p = \max\{\log_2(\eta^{-1}), 2\}$, $A = A_2$, and $f = \mathbb{1}_S$. Then for all $t \in T$, Young's convolution inequality with $\frac{1}{p} + \frac{1}{p^*} = 1$ implies that

$$\begin{aligned}
\|\mu_T^{*k} * \mu_{-A_1} * \mu_{A_2} * \mathbb{1}_S - \mu_{-A_1} * \mu_{A_2} * \mathbb{1}_S\|_\infty &\overset{\text{Young}}{\leq} \|\mu_T^{*k} * \mu_{A_2} * \mathbb{1}_S - \mu_{A_2} * \mathbb{1}_S\|_p \|\mu_{-A_1}\|_{p^*} \\
&\overset{\text{Cor. A.2}}{\leq} \tfrac{1}{2}\epsilon \left(\frac{|S|}{|G|}\right)^{\frac{1}{p}} \left(\frac{|R|}{|G|}\right)^{\frac{1}{p^*}-1} \\
&= \tfrac{1}{2}\epsilon \eta^{-1/p} \\
&\leq \epsilon
\end{aligned}$$

by choice of p. This gives Proposition 4.10.

Now Proposition 3.21 follows by taking $\eta = \alpha_1 \frac{|G|}{|S|} \geq \alpha_1$, $K = \max\{\alpha_2^{-1}, 2\}$, and $B = G$. □

A.1. **Chang's lemma.** Chang's lemma is a powerful tool often used in combination with almost-periodicity results to gain additional structure in a bootstrapping procedure. We list some versions of Chang's lemma here without proof.

We say that $S \subseteq G$ is **dissociated** if for all $(\epsilon_s)_{s \in S} \in \{-1, 0, 1\}^S$, we have that

$$\sum_{s \in S} \epsilon_s s = 0 \quad \iff \quad \epsilon_s = 0 \text{ for all } s \in S.$$

By applying probabilistic tools, one can show Chang's lemma [Cha02, Lemma 3.1].

Lemma A.3 (Chang)**.** *Let $A \subseteq G$ have density $\alpha > 0$, and let $0 < \lambda \leq 1$. If $\Lambda \subseteq \operatorname{Spec}_\lambda(\mathbb{1}_A)$ is dissociated, then $|\Lambda| = O(\lambda^{-2} \log \alpha^{-1})$.*

As a corollary, we have the following formulation stated by Tao and Vu [TV06, Lemma 4.36].

Corollary A.4. *Let $A \subseteq G$ have density $\alpha > 0$, and let $0 < \lambda \leq 1$. Then for some $d = O(\lambda^{-2} \log \alpha^{-1})$, there exist $\chi_1, \ldots, \chi_d \in \widehat{G}$ such that*

$$\operatorname{Spec}_\lambda(\mathbb{1}_A) \subseteq \left\{ \sum_{i=1}^d \epsilon_i \chi_i \,:\, \epsilon_i \in \{-1, 0, 1\} \right\}.$$

In particular, if $G = V$ is a finite field vector space, then

$$\dim \operatorname{span}(\operatorname{Spec}_\lambda(\mathbb{1}_A)) = O(\lambda^{-2} \log \alpha^{-1}).$$

In a similar vein, one can prove the following "local version" of Chang's lemma for Bohr sets, due to Sanders [San08, Proposition 4.2].

Lemma A.5 (Chang–Sanders)**.** *Let $0 < \delta, \lambda < 1$. Let $B = \operatorname{Bohr}(\Gamma, \rho) \subseteq G$ be a regular Bohr set, and let $Y \subseteq B$ with relative density ω. There exist $\Lambda \subseteq \widehat{G}$ of size at most $O(\lambda^{-2}(1 + \log \omega^{-1}))$ and $\rho' < \rho$ at least*

$$\Omega\left(\frac{\lambda^2 \rho \delta}{\operatorname{rank}(B)^2 (1 + \log \omega^{-1})}\right)$$

such that $|1 - \chi(x)| \leq \delta$ for all $\chi \in \operatorname{Spec}_\lambda(\mu_Y)$ and $x \in \operatorname{Bohr}(\Gamma \cup \Lambda, \rho') \subseteq B$.

References

[AS74] M. Ajtai and E. Szemerédi, *Sets of lattice points that form no squares*, Studia Sci. Math. Hungar. **9** (1974), 9–11.

[Beh46] F. A. Behrend, *On sets of integers which contain no three terms in arithmetical progression*, Proc. Nat. Acad. Sci. U.S.A. **32** (1946), 331–332.

[Bek24] Adrian Beker, *Improved bounds for skew corner-free sets*, arXiv preprint arXiv:2402.19169 (2024).

[BK12] Michael Bateman and Nets Hawk Katz, *New bounds on cap sets*, J. Amer. Math. Soc. **25** (2012), no. 2, 585–613.

[Blo16] T. F. Bloom, *A quantitative improvement for Roth's theorem on arithmetic progressions*, J. Lond. Math. Soc. (2) **93** (2016), no. 3, 643–663.

[Bou99] J. Bourgain, *On triples in arithmetic progression*, Geom. Funct. Anal. **9** (1999), no. 5, 968–984.

[Bou08] Jean Bourgain, *Roth's theorem on progressions revisited*, J. Anal. Math. **104** (2008), 155–192.

[BS20] Thomas F. Bloom and Olof Sisask, *Breaking the logarithmic barrier in Roth's theorem on arithmetic progressions*, arXiv preprint arXiv:2007.03528 (2020).

[BS23a] ———, *An improvement to the Kelley–Meka bounds on three-term arithmetic progressions*, arXiv preprint arXiv:2309.02353 (2023).

[BS23b] ———, *The Kelley–Meka bounds for sets free of three-term arithmetic progressions*, Essent. Number Theory **2** (2023), no. 1, 15–44.

[CFTZ22] Matthias Christandl, Omar Fawzi, Hoang Ta, and Jeroen Zuiddam, *Larger corner-free sets from combinatorial degenerations*, 13th Innovations in Theoretical Computer Science Conference, LIPIcs. Leibniz Int. Proc. Inform., vol. 215, Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2022, pp. Art. No. 48, 20.

[Cha02] Mei-Chu Chang, *A polynomial bound in Freiman's theorem*, Duke Math. J. **113** (2002), no. 3, 399–419.

[CLP17] Ernie Croot, Vsevolod F. Lev, and Péter Pál Pach, *Progression-free sets in $\mathbb{Z}_4^n$ are exponentially small*, Ann. of Math. (2) **185** (2017), no. 1, 331–337.

[CS10] Ernie Croot and Olof Sisask, *A probabilistic technique for finding almost-periods of convolutions*, Geom. Funct. Anal. **20** (2010), no. 6, 1367–1396.

[EG17] Jordan S. Ellenberg and Dion Gijswijt, *On large subsets of $\mathbb{F}_q^n$ with no three-term arithmetic progression*, Ann. of Math. (2) **185** (2017), no. 1, 339–343.

[Elk11] Michael Elkin, *An improved construction of progression-free sets*, Israel J. Math. **184** (2011), 93–128.

[EPS24] Christian Elsholtz, Laura Proske, and Lisa Sauermann, *New lower bounds for three-term progression free sets in $\mathbb{F}_p^n$*, arXiv preprint arXiv:2401.12802 (2024).

[ET36] Paul Erdös and Paul Turán, *On Some Sequences of Integers*, J. London Math. Soc. **11** (1936), no. 4, 261–264.

[FHR92] G. A. Freiman, H. Halberstam, and I. Z. Ruzsa, *Integer sum sets containing long arithmetic progressions*, J. London Math. Soc. (2) **46** (1992), no. 2, 193–201.

[FS11] Jacob Fox and Benny Sudakov, *Dependent random choice*, Random Structures Algorithms **38** (2011), no. 1-2, 68–99.

[Gow01] W. T. Gowers, *A new proof of Szemerédi's theorem*, Geom. Funct. Anal. **11** (2001), no. 3, 465–588.

[Gre05] Ben Green, *Finite field models in additive combinatorics*, Surveys in combinatorics 2005, London Math. Soc. Lecture Note Ser., vol. 327, Cambridge Univ. Press, Cambridge, 2005, pp. 1–27.

[Gre21] ———, *Lower bounds for corner-free sets*, New Zealand J. Math. **51** (2021), 1–2.

[GT09] Ben Green and Terence Tao, *New bounds for Szemerédi's theorem, I: progressions of length 4 in finite field geometries*, Proc. Lond. Math. Soc. (3) **98** (2009), no. 2, 365–392.

[GT12] ———, *New bounds for Szemerédi's theorem, Ia: Progressions of length 4 in finite field geometries revisited*, arXiv preprint arXiv:1205.1330 (2012).

[GT17] ———, *New bounds for Szemerédi's theorem, III: a polylogarithmic bound for $r_4(N)$*, Mathematika **63** (2017), no. 3, 944–1040.

[HB87] D. R. Heath-Brown, *Integer sets containing no arithmetic progressions*, J. London Math. Soc. (2) **35** (1987), no. 3, 385–394.

[Hun24] Zach Hunter, *New lower bounds for $r_3(n)$*, arXiv preprint arXiv:2401.16106 (2024).

[JLO24] Michael Jaber, Shachar Lovett, and Anthony Ostuni, *Strong bounds for skew corner-free sets*, arXiv preprint arXiv:2404.07380 (2024).

[KM23] Zander Kelley and Raghu Meka, *Strong bounds for 3-progressions*, 2023 IEEE 64th Annual Symposium on Foundations of Computer Science (FOCS), IEEE Computer Soc., Los Alamitos, CA, 2023, pp. 933–973.

[LM07] Michael T. Lacey and William McClain, *On an argument of Shkredov on two-dimensional corners*, Online J. Anal. Comb. (2007), no. 2, Art. 2, 21.

[LSS24a] James Leng, Ashwin Sah, and Mehtaab Sawhney, *Improved Bounds for Szemerédi's theorem*, arXiv preprint arXiv:2402.17995 (2024).

[LSS24b] ———, *Quasipolynomial bounds on the inverse theorem for the Gowers $U^{s+1}[N]$-norm*, arXiv preprint arXiv:2402.17994 (2024).

[Mes95] Roy Meshulam, *On subsets of finite abelian groups with no* 3*-term arithmetic progressions*, J. Combin. Theory Ser. A **71** (1995), no. 1, 168–172.
[Mil24] Luka Milićević, *Good bounds for sets lacking skew corners*, arXiv preprint arXiv:2404.07180 (2024).
[O'B11] Kevin O'Bryant, *Sets of integers that do not contain long arithmetic progressions*, Electron. J. Combin. **18** (2011), no. 1, Paper 59, 15.
[Pel23] Sarah Peluse, *Finite field models in arithmetic combinatorics—twenty years on*, arXiv preprint arXiv:2312.08100 (2023).
[Rot52] Klaus Roth, *Sur quelques ensembles d'entiers*, C. R. Acad. Sci. Paris **234** (1952), 388–390.
[Rot53] K. F. Roth, *On certain sets of integers*, J. London Math. Soc. **28** (1953), 104–109.
[RPBN$^+$24] Bernardino Romera-Paredes, Mohammadamin Barekatain, Alexander Novikov, Matej Balog, M. Pawan Kumar, Emilien Dupont, Francisco J. R. Ruiz, Jordan S. Ellenberg, Pengming Wang, Omar Fawzi, Pushmeet Kohli, and Alhussein Fawzi, *Mathematical discoveries from program search with large language models*, Nature **625** (2024), 468–475.
[San08] Tom Sanders, *Additive structures in sumsets*, Math. Proc. Cambridge Philos. Soc. **144** (2008), no. 2, 289–316.
[San11] ———, *On Roth's theorem on progressions*, Ann. of Math. (2) **174** (2011), no. 1, 619–636.
[San12a] ———, *On certain other sets of integers*, J. Anal. Math. **116** (2012), 53–82.
[San12b] ———, *On the Bogolyubov-Ruzsa lemma*, Anal. PDE **5** (2012), no. 3, 627–655.
[Sch21] Tomasz Schoen, *Improved bound in Roth's theorem on arithmetic progressions*, Adv. Math. **386** (2021), Paper No. 107801, 20.
[Shk05] I. D. Shkredov, *On a generalization of Szemerédi's theorem*, Dokl. Akad. Nauk **405** (2005), no. 3, 315–319.
[Sol03] József Solymosi, *Note on a generalization of Roth's theorem*, Discrete and computational geometry, Algorithms Combin., vol. 25, Springer, Berlin, 2003, pp. 825–827.
[SS42] R. Salem and D. C. Spencer, *On sets of integers which contain no three terms in arithmetical progression*, Proc. Nat. Acad. Sci. U.S.A. **28** (1942), 561–563.
[SS16] Tomasz Schoen and Olof Sisask, *Roth's theorem for four variables and additive structures in sums of sparse sets*, Forum Math. Sigma **4** (2016), Paper No. e5, 28.
[Sze75] E. Szemerédi, *On sets of integers containing no k elements in arithmetic progression*, Acta Arith. **27** (1975), 199–245.
[Sze90] ———, *Integer sets containing no arithmetic progressions*, Acta Math. Hungar. **56** (1990), no. 1-2, 155–158.
[Tao07] Terence Tao, *The dichotomy between structure and randomness, arithmetic progressions, and the primes*, International Congress of Mathematicians. Vol. I, Eur. Math. Soc., Zürich, 2007, pp. 581–608.
[TV06] Terence Tao and Van Vu, *Additive combinatorics*, Cambridge Studies in Advanced Mathematics, vol. 105, Cambridge University Press, Cambridge, 2006.
[Wol15] J. Wolf, *Finite field models in arithmetic combinatorics—ten years on*, Finite Fields Appl. **32** (2015), 233–274.
[Zha23] Yufei Zhao, *Graph theory and additive combinatorics—exploring structure and randomness*, Cambridge University Press, Cambridge, 2023.