A survery of Roth's theorem on progressions of length three

by

Yui Nishizawa

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Mathematics
in
Pure Mathematics

Waterloo, Ontario, Canada, 2011

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

For any finite set B and a subset $A \subset B$, we define the density of A in B to be the value $\alpha = |A|/|B|$. Roth's famous theorem, proven in 1953, states that there is a constant $C > 0$, such that if $A \subset \{1, ..., N\}$ for a positive integer N and A has density α in $\{1, ..., N\}$ with $\alpha > C/\log\log N$, then A contains a non-trivial arithmetic progression of length three (3AP). The proof of this relies on the following dichotomy: either 1) A looks like a random set and the number of 3APs in A is close to the probabilistic expected value, or 2) A is more structured and consequently, there is a progression P of about length $\alpha\sqrt{N}$ on which $A \cap P$ has $\alpha(1 + c\alpha)$ for some $c > 0$. If 1) occurs, then we are done. If 2) occurs, then we identify P with $\{1, ..., |P|\}$ and repeat the above argument, whereby the density increases at each iteration of the dichotomy. Due to the density increase in case 2), an argument of this type is called a density increment argument. The density increment is obtained by studying the Fourier transforms of the characterstic function of A and extracting a structure out of A. Improving the lower bound for α is still an active area of research and all improvements so far employ a density increment. Two of the most recent results are $\alpha > C(\log\log N/\log N)^{1/2}$ by Bourgain in 1999 and $\alpha > C(\log\log N)^5/\log N$ by Sanders in 2010. This thesis is a survey of progresses in Roth's theorem, with a focus on these last two results. Attention was given to unifying the language in which the results are discussed and simplifying the presentation.

Acknowledgements

First and foremost, I would like to thank my supervisor Yu-Ru Liu for everything that she has done for me in these past 2 years. I can only sum it up as “everything”, because there is so much that I possibly cannot list all in this small margin nor do I wish to undermine them by a systematic list. But I am especially indebted to her for her constant encouragement regarding both my work and life in general, and for being a kind, patient listener to my often confused ramblings. I learned a lot from her.

I would also like to thank my co-supervisor Wentang Kuo for helping me in the final stages of my thesis and for organizing the summer student number theory seminars which I gained a lot out of. The seminars greatly enriched my knowledge of the field I am working in.

I am thankful to my humorous officemates, Allen O’Hara and Nikita Nikolaev for always entertaining me and keeping my spirits high (although I never expressed it) despite my bad temper.

Finally, I wish to thank my family for nothing specific, but again just everything in general, because I cannot recall or list all the particular instances. Knowing of their existence alone was comforting to me.

Table of Contents

CHAPTER 1

Introduction

One of the recurring themes in contemporary number theory is a question of the form "what type of structure can randomness give rise to, and when?" Roth's theorem deals with a question of such type. Of course, to make this general question into a mathematical question, we need to decide what type of structure we are dealing with and what we mean by random. The structure we are concerned with is the following.

Definition 1.1. *Let k be a positive integer. A sequence of integers $x_1, \ldots, x_k$ is called an arithmetic progression of length k (kAP) if there exist integers a and d such that $x_i = a + (i-1)d$ for each $i = 1, \ldots, k$. A kAP is non-trivial if $d \neq 0$.*

We note that this definition can be extended to abelian groups, i.e. when $x_1, \ldots, x_k, a, d$ are elements of an abelian group G. In fact, we will be working in finite abelian groups most of the time throughout this thesis, and transfer the results obtained there back to the integers.

Now, on the other hand, the "when does this structure arise?" will be described using the following natural notion of density.

Definition 1.2. *Let B be a finite set, and let $A \subset B$ be any subset. Then the quantity $\alpha = |A|/|B|$ is called the density of A in B.*

The density of a set has no obvious relationship with APs in the set. In 1936, Erdös and Turán made the following conjecture. Here, $[N]$ is defined as the set $\{1, \ldots, N\}$.

Conjecture 1.3 (Erdös-Turán)**.** *Let k be any positive integer and let $\alpha > 0$. Then for any N sufficiently large, any $A \subset [N]$ with density at least α in $[N]$ contains a non-trivial kAP.*

A famous theorem of Roth in 1953 provided an answer to the non-trivial simplest case of this conjecture, namely when $k = 3$. Not only did his proof show that the conjecture is true for $k = 3$, but it also provided an explicit lower bound for α in terms of N. More precisely, Roth proved the following in [**14**].

Theorem 1.4 (Roth)**.** *There exists $C > 0$ such that if $A \subset [N]$ is of density α in $[N]$ where*

$$\alpha > C \frac{1}{\log \log N} \ ,$$

then A contains a non-trivial 3AP.

This lower bound for α, however, is not tight, and a lot of effort has been invested to improve it. For example, in 1987, Heath-Brown in [**13**] proved that it suffices to have

$$\alpha > C\frac{1}{(\log N)^{c'}},$$

for some $c' > 0$. In 1990, Szemerédi proved in [**22**] that we can take $c' = 1/20$. In 1999, Bourgain proved in [**3**] that we can have

$$\alpha > C\left(\frac{\log\log N}{\log N}\right)^{1/2}.$$

In 2008, Bourgain further improved it to

$$\alpha > C\frac{(\log\log N)^2}{(\log N)^{2/3}}$$

in the paper [**4**]. Recently, in 2010, Tom Sanders refined it in [**17**] to

$$\alpha > C\frac{(\log\log N)^5}{\log N}.$$

The techniques employed in all the results above are Fourier analytic and probabilistic, and the methods are not uncorrelated at all, but we can observe a progression of ideas. The main goal of this thesis is to lay out the ideas of the most recent results, namely the results by Bourgain and Sanders presented above, in a clear, simple, and coherent manner as possible. It is nowadays common knowledge that analysis can be used to prove things about a discrete structure like the integers, but this is remarkable nevertheless. We note that the conjecture has also been proven by a combinatorial approach such as in [**20**] by Szemerédi in 1969, but the bound obtained is not as good as those obtained analytically.

There is a converse version of Roth's theorem as well, by which we mean a lower bound for the largest α such that there is a subset of density α containing only trivial 3APs. The first of these results goes back to a result by Behrend in [**3**] from 1946.

Theorem 1.5 (Behrend)**.** *There exists $C > 0$ such that for any positive integer N, there is a subset $A \subset [N]$ of density α in $[N]$ where*

$$\alpha > C\frac{1}{\log^{1/4} N}\cdot\frac{1}{2^{2\sqrt{2}\sqrt{\log_2 N}}},$$

such that A has only trivial 3APs.

Unlike Roth's result, this result was not improved until 2010, and the improvement is also very small. Elkin proved in [**6**] that we can have

$$\alpha > \frac{C\log^{1/4} N}{2^{2\sqrt{2}\sqrt{\log_2 N}}}.$$

Green and Wolf shortened Elkin's proof in [**10**]. It seems to be commonly believed that this bound for a set free of non-trivial 3AP is almost tight. The problem of finding the precise boundary separating when a set is free of non-trivial 3AP and when it must contain one continues to draw great attention of mathematicians.

We will conclude the introduction by presenting a few other related results. Although Roth's theorem answers only the simplest $k = 3$ case of the Erdös-Turán conjecture, the general case has been proven by Szemerédi in 1975 in [**21**] by generalizing the combinatorial approach in [**20**]. Gowers further contributed to this by reproving the $k = 4$ case in [**11**] and generalizing this approach for any positive integer k in [**12**]; the importance of his result was that he extended the analytic approach inspired by Roth rather than using the combinatorial approach. This provided a better lower bound on the density.

Another significant result is a version of Roth's theorem restricted to the prime numbers. In 2005, Green proved the following in [**8**].

Theorem 1.6 (Green)**.** *Let $\alpha > 0$, and let P_N be the set of prime numbers in $[N]$. There exists $C > 0$ such that if $A \subset P_N$ has density α in P_N where*

$$\alpha > C \left(\frac{\log\log\log\log\log N}{\log\log\log\log N} \right)^{1/2},$$

then A contains a non-trivial 3AP.

We see that this lower bound on α is approaching 0 as $N \to \infty$, so this indeed provides an affirmative answer to the prime version of the conjecture. We note, however, that if we can find a lower bound $L(N)$ for the original version of Roth's theorem (that is, not the prime restriction, but in the natural numbers) such that $L(N)/(1/\log N) \to 0$ as $N \to \infty$, then this implies the prime restriction as well, because the primes have an asymptotic density of $1/\log N$ in $[N]$.

Notation: For any set A, we denote by 1_A its characteristic function and $|A|$ its cardinality. We denote by μ_A the function $1_A/|A|$ (or more specifically, a probability measure. See Chapter 2, Section 2.1 for more details). We define $2A$ to be the set $\{2a : a \in A\}$. For any real valued functions f and g with the same domain, say D, we write $f = O(g)$ to mean that there is a constant $C > 0$ such that for all $x \in D$, we have $|f(x)| \leq Cg(x)$. Throughout this entire paper, the implied constant C will be absolute. We write $f = \Omega(g)$ to mean that there is a constant $C > 0$ such that for all $x \in D$, we have $f(x) \geq C|g(x)|$. For any function f, we define $\tau_t(f)$ to be the function $\tau_t(f)(x) := f(x + t)$ wherever it makes sense. We define the function $e : \mathbb{R} \to \mathbb{C}$ to be the function $e(x) := e^{i2\pi x}$. We also define $e_y : \mathbb{R} \to \mathbb{C}$ to be the function $e_y(x) := e(yx) = e^{i2\pi yx}$. The domain of these two functions may change (to say $\mathbb{Z}_N$ for some positive integer N, but it should be clear from the context when this happens). For a positive integer N, we define $[N] := \{1, \ldots, N\}$.

CHAPTER 2

Roth's theorem

The first step forward in the Erdös-Turán Conjecture was presented by Roth in 1953 in [**14**].

Theorem 2.1. *There exists $C > 0$ such that if $A \subset [N]$ is of density α in $[N]$ where*

$$\alpha > C\frac{1}{\log\log N} ,$$

then A contains a non-trivial 3AP.

In this Chapter, we will provide a proof of Roth's result. It is not the original proof that appeared in his paper, but a streamlined version of it that evolved over the years such as that provided in [**18**]. The ideas employed in this proof will serve as groundwork for later chapters, where some of the ideas are re-employed or improved upon. Thus understanding the proof in this chapter is sure to help understanding the ideas in the later chapters.

Before moving on, we make one easy, but important observation about the definition of a 3AP. We defined a 3AP to be a sequence of the form $a, a+d, a+2d$ for some integers a and d, but there is another equivalent formulation of a 3AP which will be more convenient for us throughout this entire thesis. Setting $x = a$, $y = a + d$ and $z = a + 2d$, we see that a 3AP satisfies

$$x + z = 2y. \tag{2.1}$$

Conversely, any x, y, and z satisfying the above forms a 3AP.

2.1. Fourier analysis

In this section, we will review some of the main notions that will be used throughout this paper. Books such as [**19**] (Chapter 7) discusses Fourier analysis on abelian groups more in detail. Surprisingly, Fourier analysis is the main driving force in all the techniques to follow, so most of the following notions are related to Fourier analysis. We will primarily be working in a finite abelian group G, and we will denote the *dual group* of G by $\hat{G}$, i.e. the set of all homomorphisms $\gamma : G \to \mathbb{S}^1$ where $\mathbb{S}^1$ is the unit sphere in $\mathbb{C}$. These homomorphisms are also called *characters*. We will denote the set of all (continuous) functions $f : G \to \mathbb{C}$ by $C(G)$, and the set of all measures on G by $M(G)$. Recall that a measure is a non-negative function $\mu : 2^G \to \mathbb{R}$ with the following properties:

i) If $\{E_i\}_{i\in I}$ is a countable sequence of pairwise disjoint sets, then

$$\mu(\cup_{i\in I}E_i) = \sum_{i\in I}\mu(E_i).$$

ii) $\mu(\emptyset) = 0$.

μ is called a *probability measure* if, additionally, $\mu(G) = 1$. If $x \in G$, then we will abuse the notation and write $\mu(x)$ to mean $\mu(\{x\})$. If $A \subset G$, then μ_A is a probability measure.

Example If $G = \mathbb{Z}_N$, then $\hat{G}$ is the set of exponential functions $e_{n/N}$ for each $n \in \mathbb{Z}_N$. To see this, suppose $\gamma \in \hat{G}$ and let $a = \gamma(1)$. Then since 1 has order N in $\mathbb{Z}_N$, we have that $1 = \gamma(N) = a^N$; that is, a is an Nth root of unity, say $a = e^{n/N}$ for some $n \in \mathbb{Z}_N$. Then by the linearity of γ, we see that $\gamma(x) = e_{n/N}(x)$ for all $x \in \mathbb{Z}_N$. Conversely, if $\gamma = e_{n/N}$ for any $n \in \mathbb{Z}_N$, it is easy to verify that γ is a homomorphism. It follows that $\hat{G} = \left\{e_{n/N} : n \in \mathbb{Z}_N\right\}$ and $|\hat{G}| = N$.

Characters have the following general properties.

Proposition 2.2. [**1**, pg.133] *The dual group of a finite abelian group G has the properties below.*

i) (Orthogonality) For any $\gamma \in \hat{G}$, we have

$$\sum_{x\in G}\gamma(x) = \begin{cases} 0, & \text{if } \gamma \neq 1_G, \\ |G|, & \text{if } \gamma = 1_G. \end{cases}$$

ii) (Dual orthogonality) For any $x \in G$, we have

$$\sum_{\gamma\in \hat{G}}\gamma(x) = \begin{cases} 0, & \text{if } x \neq 0, \\ |G|, & \text{if } x = 0. \end{cases} .$$

iii) $\hat{G}$ is a group of order $|G|$ where the binary operation is multiplication, 1_G is the identity, and for any $\gamma \in \hat{G}$, we have $\gamma^{-1}(x) = \gamma(-x) = \bar{\gamma}(x)$.

With these characters, we define the discrete Fourier transform as follows.

Definition 2.3 (Fourier transform)**.** *Let $f \in C(G)$ and $\mu \in M(G)$. The Fourier transform of f with respect to μ is the function $\widehat{fd\mu} \in C(\hat{G})$ defined by*

$$\widehat{fd\mu}(\gamma) = \sum_{y\in G} f(y)\bar{\gamma}(y)\mu(x).$$

Let $h \in C(\hat{G})$. The inverse Fourier transform of h is the function $\check{h} \in C(G)$ defined by

$$\check{h}(x) = \sum_{\gamma\in\hat{G}} h(\gamma)\gamma(x).$$

As we will be working mainly with the measure μ_G, if no measure is specified, then it will be assumed that we are using this measure. That is,

$$\hat{f}(\gamma) = \frac{1}{|G|} \sum_{y \in G} f(y)\bar{\gamma}(y).$$

This convention will apply for all subsequent definitions as well. The generalized definition will be used only in Chapter 4.

The following are also important concepts in Fourier analysis.

Definition 2.4 (Convolution)**.** *Let $f, g \in C(G)$ and $\mu \in M(G)$. The convolution $f * gd\mu \in C(G)$ is the function defined by*

$$f * gd\mu(x) = \sum_{y \in G} f(y)g(x-y)\mu(y).$$

Let $h, k \in C(\hat{G})$ The convolution $h * k \in C(\hat{G})$ is the function defined by

$$h * k(\gamma) = \sum_{\nu \in \hat{G}} h(\nu)k(\gamma\bar{\nu}).$$

Definition 2.5 (L^ps norm and inner product)**.** *Let $f, g \in C(G)$, $\mu \in M(G)$ and let $p \in (0, \infty]$. For $p \neq \infty$, the $L^p(\mu)$ norm of f is the value*

$$||f||_{L^p(\mu)} = \left(\sum_{x \in G} |f(x)|^p \mu(x) \right)^{1/p}.$$

If $p = \infty$, the $L^\infty(\mu)$ norm is the value

$$||f||_{L^\infty} = ||f||_{L^\infty(\mu)} = \sup_{x \in G} |f(x)|.$$

The inner product of f and g is the value

$$\langle f, g \rangle_\mu = \sum_{x \in G} f(x)\bar{g}(x)\mu(x).$$

In particular,

$$||f||_{L^2(\mu)} = \sqrt{\langle f, f \rangle_\mu}.$$

Let $h, k \in C(\hat{G})$. For $p \neq \infty$, the $L^p(\hat{G})$ norm of h is the value

$$||h||_{L^p} = \left(\sum_{\gamma \in \hat{G}} |h(\gamma)|^p \right)^{1/p}.$$

For $p = \infty$, the $L^\infty(\hat{G})$ norm is the value

$$||h||_{L^\infty} = \sup_{\gamma \in \hat{G}} |h(\gamma)|.$$

The inner product of h and k is the value

$$\langle h, k\rangle = \sum_{\gamma\in\hat{G}} h(\gamma)\bar{k}(\gamma).$$

Again, we note that in most cases we are only concerned with the convolution and $L^p(\mu)$ norm when $\mu = \mu_G$, but we have defined them in the general way above for consistency.

Since G is a finite group, all functions on G have an $L^p(\mu)$ norm, but as is done conventionally, we will sometimes write $f \in L^p(\mu)$ to specify which norm we are using on f. We also remind the reader that the reason for the odd definition of L^∞ is due to the property

$$\lim_{p\to\infty} ||f||_{L^p(\mu)} = \lim_{p\to\infty} ||f||_{L^\infty} \left|\left| \frac{f}{||f||_{L^\infty}} \right|\right|_{L^p(\mu)} = ||f||_{L^\infty},$$

if f is not identically zero. The following proposition shows how the notions defined above are related to the Fourier transform.

Proposition 2.6. *Let $f, g \in C(G)$. Then*

i) (Parseval's identity) $||\hat{f}||_{L^2} = ||f||_{L^2}$.
ii) (Fourier inversion) $f = \check{\hat{f}}$.
iii) $\langle f, g\rangle = \langle \hat{f}, \hat{g}\rangle$.
iv) $\widehat{f * g} = \hat{f}\hat{g}$.

Proof: The proofs of all these are similar in that we begin with the Fourier transform, use the orthogonality property of Proposition 2.2, and simplify the result. We will prove iii) as an example. We see that by definition, we have

$$\langle \hat{f}, \hat{g}\rangle = \frac{1}{|G|^2} \sum_{\gamma\in\hat{G}} \sum_{x,y\in G} f(x)\bar{g}(y)\bar{\gamma}(x)\gamma(y) = \frac{1}{|G|^2} \sum_{x,y\in G} f(x)\bar{g}(y) \sum_{\gamma\in\hat{G}} \bar{\gamma}(x-y).$$

By ii) of Proposition 2.2, the inner sum is non-zero only when $x = y$. Hence,

$$\langle \hat{f}, \hat{g}\rangle = \frac{1}{|G|} \sum_{x\in G} f(x)\bar{g}(x) = \langle f, g\rangle,$$

as desired. □

Some basic properties of the convolution are the following. Properties i) and ii) hold also for the convolution on $C(\hat{G})$, but we will not be needing them. They follow directly from the definition.

Proposition 2.7. *Let $f, g, h : G \in C(G)$. Then*

i) The convolution operator $* : C(G) \times C(G) \to C(G)$ is commutative, associative, and linear.

ii) Let $\dot{h}(x) := \bar{h}(-x)$. *Then*

$$\langle f * h, g\rangle = \langle f, g * \dot{h}\rangle.$$

In particular, the operator $*h : C(G) \to C(G) : f \mapsto f * h$ *is self-adjoint if* h *is a real-valued even function.*

iii) If $A_1, A_2 \subset G$, *then*

$$1_{A_1} * 1_{A_2}(x) = \frac{|\{(a,b) \in A_1 \times A_2 : a + b = x\}|}{|G|} = \frac{|A_1 \cap (x - A_2)|}{|G|}.$$

Hölder's inequality is a standard tool in L^p spaces.

Proposition 2.8 (Hölder's inequality)**.** *Let* $1 \leq p \leq \infty$ *and* $\mu \in M(G)$. *Then for any* $f \in L^p(\mu)$ *and* $g \in L^q(\mu)$ *where* $q = p/(p-1)$, *we hvae*

$$||fg||_{L^1(\mu)} \leq ||f||_{L^p(\mu)}\, ||g||_{L^q(\mu)}\,.$$

More generally, let $p_1, \ldots, p_n \in (0, \infty]$, *and let* $r > 0$ *be the number such that*

$$\frac{1}{r} = \sum_{k=1}^{n} \frac{1}{p_k}.$$

Then for any $f_k \in L^{p_k}(\mu)$ *where* $k = 1, \ldots, n$, *we have*

$$\left|\left|\prod_{k=1}^{n} f_k\right|\right|_{L^r(\mu)} \leq \prod_{k=1}^{n} ||f||_{L^{p_k}(\mu)}\,.$$

Proof: The proof of Hölder's inequality can be found in many standard texts on real analysis with a L^p space section, such as [**15**, pg.119]. We will prove only the generalization part, and the proof is by induction. The claim clearly holds for $n = 1$, so suppose it is true for some $n > 1$. We have to consdier two cases, depending on whether $p_n = \infty$ or $p_n < \infty$. First, suppose $p_n = \infty$. Then

$$\left|\left|\prod_{k=1}^{n} f_k\right|\right|_{L^r(\mu)} \leq ||f_n||_{L^\infty} \left|\left|\prod_{k=1}^{n-1} f_k\right|\right|_{L^r(\mu)},$$

and we can just now apply the induction hypothesis. If $p < \infty$, we use the original Hölder's inequality with $p = p_n/r$ and $q = p_n/(p_n - r)$. Then

$$\left|\left|\prod_{k=1}^{n} f_k\right|\right|_{L^r(\mu)} = \left|\left|\prod_{k=1}^{n} f_k^r\right|\right|_{L^1(\mu)}^{1/r} \leq ||f_n^r||_{L^p(\mu)}^{1/r} \left|\left|\prod_{k=1}^{n-1} f_k^r\right|\right|_{L^q(\mu)}^{1/r} = ||f_n||_{L^{p_n}(\mu)} \left|\left|\prod_{k=1}^{n-1} f_k\right|\right|_{L^{rq}(\mu)}.$$

Notice that

$$\sum_{k=1}^{n-1} \frac{1}{p_k} = \frac{1}{r} - \frac{1}{p_n} = \frac{1}{rq}.$$

Hence, we can apply the induction hypothesis. This completes the proof. □

Hölder's inequality shows that $L^p(\mu)$ norms are convex in p as in the following corollary.

Corollary 2.9 (Interpolation)**.** *Let $p_1, \dots, p_n \in (0, \infty]$ and $\theta_1 \in [0,1]$ such that $\theta_1 + \cdots + \theta_n = 1$, and let r be the number such that*

$$\frac{1}{r} = \sum_{k=1}^{n} \frac{\theta_k}{p_k}.$$

Then for any $f \in L^r(\mu)$, we have

$$||f||_{L^r(\mu)} \leq \prod_{k=1}^{n} ||f||_{L^{p_k}(\mu)}^{\theta_k}.$$

In particular, if $n = 2$, then for any $\theta \in [0,1]$, we have

$$||f||_{L^r(\mu)} \leq ||f||_{L^{p_1}(\mu)}^{\theta} \, ||f||_{L^{p_2}(\mu)}^{1-\theta}.$$

Proof: Let $f_k = |f|^{\theta_k}$. Then by the preceding proposition, we have

$$||f||_{L^r(\mu)} = \left|\left| \prod_{k=1}^{n} f_k \right|\right|_{L^1 p(\mu)} \leq \prod_{k=1}^{n} ||f_k||_{L^{p_k}(\mu)} = \prod_{k=1}^{n} ||f||_{L^{p_k}(\mu)}^{\theta_k}.$$

□

The following Corollary shows how the different L^p norms are related. A proof of this can be found also in many real analysis texts, such as [**15**, pg.131].

Corollary 2.10 (Embedding)**.** *Let $0 < p < q$ and $\mu \in M(G)$. Then the operator $T : L^p(\mu) \to L^q(\mu)$ has operator norm $\mu(G)^{1/p-1/q}$. Hence, for any $f \in L^p(\mu)$,we have*

$$||f||_{L^p(\mu)} \leq \mu(G)^{1/p-1/q} ||f||_{L^q(\mu)}.$$

2.2. Proof of Roth's Theorem

Instead of discussing its underlying ideas, we will head in to the proof right away. There will be a discussion of the proof in the subsequent Section 2.3. Thus, the reader interested in the underlying ideas may wish to read that section first.

One simplification we use is that we work in $\mathbb{Z}_N$ instead of $\mathbb{Z}$.

Lemma 2.11. *Let $A \subset [N]$. Then all 3APs of A in $\mathbb{Z}_{3N}$ are also 3APs in $\mathbb{Z}$.*

Proof: Suppose $x, y, z \in A$ is a 3AP in $\mathbb{Z}_{3N}$ so without loss of generality, $x + y \equiv 2z \pmod{3N}$. That is, $3N | (2z - (x+y))$. Since $0 \leq x+y \leq 2N$, it must be that $x+y = 2z$. □

We will prove Theorem 2.1 in $\mathbb{Z}_N$. Then the same statement is true for $\mathbb{Z}$ with just a different constant C by the preceding lemma.

The key lemma in proving Roth's theorem is the following.

Lemma 2.12 (Density increment or 3AP)**.** *Let $G = \mathbb{Z}_N$ where $N > 64^2$ is odd. Let $A \subset G$ and set α to be its density in G i.e. $\alpha := \mu_G(A)$. Then at least one of the following occurs*

i) (Failure of size condition)

$$N \leq 4/3\alpha^2.$$

ii) A contains a non-trivial 3AP

iii) (Density increment) A has density $\alpha + \alpha^2/32$ on some progression $P \subset G$, where $|P| \geq \alpha^2\sqrt{N}/2$ and is of odd length.

Proof: Since G is of odd order, we have that $G = 2G$. Consider the difference

$$\epsilon := |\langle 1_A * 1_A, 1_{2A}\rangle - \alpha\langle 1_A * 1_A, 1_{2G}\rangle| \tag{2.2}$$

The left term $\langle 1_A * 1_A, 1_{2A}\rangle$ counts the number of 3APs in A due to (2.1) (with a weight $1/|G|^2$). By Proposition 2.7 iii), we see that

$$\langle 1_A * 1_A, 1_{2G}\rangle = \alpha^2.$$

Therefore,

$$\epsilon = |\langle 1_A * 1_A, 1_{2A}\rangle - \alpha^3|.$$

By Proposition 2.6 iii) and iv), we have

$$\langle 1_A * 1_A, 1_{2A}\rangle = \langle \hat{1}_A^2, \hat{1}_{2A}\rangle.$$

Now, as $\hat{1}_A(1_G) = \mu_G(A) = \alpha$, and since G is of odd order, we also have that $\hat{1}_{2A}(1_G) = \alpha$. Hence,

$$\langle 1_A * 1_A, 1_{2A}\rangle = \alpha^3 + \sum_{\gamma \neq 1_G} \hat{1}_A(\gamma)^2\hat{1}_{2A}(\bar{\gamma})$$

It follows that

$$\epsilon = \left|\sum_{\gamma \neq 1_G} \hat{1}_A(\gamma)^2\hat{1}_{2A}(\bar{\gamma})\right|$$

This can be interpreted as the error between the actual and expected number of 3APs. We will now consider two cases.

Case 1: $\epsilon < \alpha^3/4$

By the triangle inequality, we see from (2.2) that

$$\langle 1_A * 1_A, 1_{2A}\rangle \geq \frac{3\alpha^3}{4}.$$

If A contains only trivial 3APs, then $\langle 1_A * 1_A, 1_{2A}\rangle = |A|/|G|^2 = \alpha/|G|$. Hence, if $3\alpha^3/4 < \alpha/|G|$, or equivalently, if

$$N = |G| > \frac{4}{3\alpha^2},$$

then we have a non-trivial 3AP. That is, if N is sufficiently large relative to the density, then we are guaranteed a non-trivial 3AP. Conditions as such will be referred to as the *size condition* in the future. This condition will yield the condition on the density in the statement of Theorem 2.1.

Case 2: $\epsilon \geq \alpha^3/4$
By Parseval's identity, we have

$$\frac{\alpha^3}{4} \leq \left| \sum_{\gamma \neq 1_G} \hat{1}_A(\gamma)^2 \hat{1}_{2A}(\bar{\gamma}) \right| \leq \sup_{\gamma \neq 1_G} |\hat{1}_{2A}(\gamma)| \sum_{\gamma \in \hat{G}} |\hat{1}_A(\gamma)|^2 = \sup_{\gamma \neq 1_G} |\hat{1}_{2A}(\gamma)| \alpha.$$

For any $\gamma \in \hat{G}$, the function $\gamma_2(x) := \gamma(2x)$ is also a automorphism. It follows that $\hat{1}_{2A}(\gamma) = \hat{1}_A(\gamma_2)$. Furthermore, since N is odd, if $\gamma = e_{n/N}$ is not the identity function, then neither is γ_2. Therefore, we can replace the $\hat{1}_{2A}$ in the sup with $\hat{1}_A$, giving

$$\frac{\alpha^2}{4} \leq |\hat{1}_A(\gamma)| \text{ where } \gamma \neq 1_G.$$

By Proposition 2.2 i), we see that,

$$\frac{\alpha^2}{4} \leq |\hat{1}_A(\gamma)| = \left| \frac{1}{G} \sum_{x \in G} (1_A(x) - \alpha)\gamma(x) \right|. \tag{2.3}$$

Heuristically, the Fourier Transform being large suggests that there is a certain type of structure in A related to the characters γ. It will turn out that there is a progression P on which A has an increased density $\alpha + c\alpha^2$. We extract this structure as follows. Recall that $\gamma = e_{n/N}$ for some $n \not\equiv 0 (\text{mod} N)$. Let $Q < N$ be a positive integer, to be determined later. By Dirichlet's Approxomation Theorem we can find a non-negative integer r and a positive integer $q \leq Q$ such that

$$\frac{n}{N} = \frac{r}{q} + \theta,$$

where $|\theta| \leq 1/qQ$. We now consider the progression $P = \{0, q, 2q, \ldots, (L-1)q\}$ for some L to be determined later. In particular, if $0 \leq l < L$ then by the mean value theorem, we have

$$|\gamma(lq) - 1| = |e(lq\theta) - 1| \leq 2\pi lq|\theta| \leq \frac{8L}{Q}. \tag{2.4}$$

We would like to use this to 'annhilate' γ from (2.3), by transferring from a sum over G to a sum over translates of P. Let $f = 1_A - \alpha$. We have, by (2.3),

$$\frac{\alpha^2}{4} \leq |\langle f\bar{\gamma}, 1_G \rangle|. \tag{2.5}$$

Now, by Proposition 2.7 ii), we have

$$\langle f\bar{\gamma} * 1_P, 1_G \rangle = \langle f\bar{\gamma}, 1_G * 1_{-P} \rangle = \mu_G(P)\langle f\bar{\gamma}, 1_G \rangle, \tag{2.6}$$

where we recall that $\mu_G(P) = |P|/|G|$. Therefore,

$$\frac{\alpha^2}{4}\mu_G(P) \leq \langle |f\bar{\gamma} * 1_P|, 1_G \rangle. \tag{2.7}$$

We can then 'annihilate' γ, because γ is almost constant on translates of P due to (2.4) as follows:

$$\begin{aligned}
\langle |f\bar{\gamma} * 1_P|, 1_G\rangle &= \frac{1}{|G|^2}\sum_{x\in G}\left|\sum_{y\in P-x} f(y)\bar{\gamma}(y)\right| \\
&\le \frac{1}{|G|^2}\sum_{x\in G}\left|\sum_{y\in P-x} f(y)\bar{\gamma}(-x)\right| + \frac{1}{|G|^2}\sum_{x\in G}\sum_{y\in P-x} f(y)\left|\bar{\gamma}(y)-\bar{\gamma}(-x)\right| \\
&\le \frac{1}{|G|^2}\sum_{x\in G}\left|\sum_{y\in P-x} f(y)\right| + \frac{1}{|G|^2}\sum_{x\in G}\sum_{y\in P-x}\frac{8L}{Q} \\
&= \langle |f * 1_P|, 1_G\rangle + \frac{8L}{Q}\mu_G(P).
\end{aligned}$$

Combining this with (2.7), we have

$$\left(\frac{\alpha^2}{4} - \frac{8L}{Q}\right)\mu_G(P) \le \langle |f * 1_P|, 1_G\rangle.$$

If we now set $Q = 64\sqrt{N}$ and $L = \lfloor \alpha^2\sqrt{N}\rfloor$ (recall that we need $Q < N$), the above becomes

$$\frac{\alpha^2}{8}\mu_G(P) \le \langle |f * 1_P|, 1_G\rangle.$$

Now, notice that by Proposition 2.7 ii),

$$\langle f * 1_P, 1_G\rangle = \langle f * 1_G, 1_{-P}\rangle = \langle 0, 1_{-P}\rangle = 0.$$

Hence, if we set B to be all $x \in G$ such that $f * 1_P(x) \ge 0$, then

$$2\langle f * 1_P, 1_B\rangle = \langle |f * 1_P|, 1_G\rangle + \langle f * 1_P, 1_G\rangle \ge \frac{\alpha^2}{8}\mu_G(P).$$

In particular, we have

$$\frac{\alpha^2}{16}\mu_G(P) \le \langle f * 1_P, 1_B\rangle.$$

Hence, there is some $x \in G$ such that

$$\frac{\alpha^2}{16}\mu_G(P) \le (1_A - \alpha) * 1_P(x).$$

In other words,

$$(\alpha + \alpha^2/16)\mu_G(P) \le 1_A * 1_P(x),$$

so A has a higher density on $x - P$, or what is the same, $x - A$ on P. This increase in density is called a *density increment*. We require P to be of odd length, so we may need to add another point to P. Nevertheless, we will still have a density increment of $(\alpha + \alpha^2/32)$. □

Now we can complete the proof of Roth's result.

Proof of Theorem 2.1:
The proof is by iterating Lemma 2.12. We will suppose that i), the failure of the size condition, does not occur. If ii) occurs then we are done. In the case when iii) occurs, we can repeat the argument by identifying $P = \{0, q, 2q, \ldots\}$ with $G_1 = \mathbb{Z}_{|P|}$ and $A \cap P$ with $A_1 \subset G_1$ in the natural way, i.e. q with $1 \in G_1$, $2q$ with $2 \in G_1$, and etc. Note that this identification preserves 3APs. We now apply Lemma 2.12 on G_1 and A_1, and continue in this manner. Let L_i and α_i be the length of the progression and the density of (the set equivalent to) A in the progression at the ith iteration.

Let us first see how the density increases. This can be approximated quite efficiently by a sort of dyadic decomposition. After $i \geq 32/\alpha$ steps, we have

$$\alpha_i \geq \alpha + \frac{\alpha^2}{32} \cdot \frac{32}{\alpha} = 2\alpha.$$

Hence, for higher i, the density increases at least by 2α. After another additional $16/\alpha$ steps, we have

$$\alpha_i \geq 2\alpha + \frac{(2\alpha)^2}{32} \cdot \frac{16}{\alpha} = 4\alpha.$$

In general, after

$$\frac{32}{\alpha} \sum_{j=0}^{k} 2^{-j}$$

steps, the density is at least $2^{k+1}\alpha$. Since the density cannot exceed 1, ii) must occur at some point in the iteration. This gives a non-trivial 3AP in A. Note that we can have at most $64/\alpha$ iterations.

All the above was however assuming that the size condition never failed during the iterations. Thus, we must guarantee that it is indeed met at each step of the iteration. First, let us see how L_i decreases. We have

$$L_2 \geq \frac{\alpha_1^2}{2} \sqrt{L_1} \geq \frac{\alpha^{2+1}}{2^{1+1/2}} N^{1/4}.$$

Continuing in this manner, we have

$$L_i \geq \frac{\alpha^{2+1+\cdots+1/2^{i-2}}}{2^{2+1+\cdots+1/2^{i-2}}} N^{1/2^i} \geq \frac{\alpha^4}{16} N^{1/2^{(64/\alpha)}}.$$

The size condition was that we need $L_i \geq 4/3\alpha_i^2$ for each i in order for the iterations to work. Since L_i is minimum at $i = \lfloor 64/\alpha \rfloor$, which is the largest possible i whereas, α_i is minimum at $i = 0$, it suffices to have

$$\frac{\alpha^4}{16} N^{1/2^{(64/\alpha)}} \geq \frac{4}{3\alpha^2}.$$

This is equivalent to

$$N \geq \left(\frac{64}{3\alpha}\right)^{2^{64/\alpha^6}}.$$

Taking logarithms twice gives

$$\log \log N \geq \frac{64}{\alpha} \log 2 + \log \log \frac{64}{3\alpha^6}.$$

Therefore, it suffices to have

$$\log \log N \geq \frac{C}{\alpha},$$

where $C > 0$ is some absolute constant. This proves the claim for the version of the Theorem in $\mathbb{Z}_N$. By Lemma 2.11, the same result is true in $\mathbb{Z}$, but just with a different constant. □

2.3. Discussion of the proof

The idea is as follows. We begin with the quantity

$$\epsilon = |\langle 1_A * 1_A, 1_{2A}\rangle - \alpha\langle 1_A * 1_A, 1_{2G}\rangle|.$$

The left term is the number of 3APs in A (weighted by $1/|G|^2$), and the right side is the expected value α^3. We then consider two cases.

1) ϵ is small (Pseudorandom case): Heurstically, this can be interpreted as A behaving like a typical random subset of density α. In the future, we shall call this the *pseudorandom case*. In this case, we use the triangle inequality to bound $\langle 1_A * 1_A, 1_{2A}\rangle$ from below by about α^3. On the other hand, if A contains only trivial 3APs, then $\langle 1_A * 1_A, 1_{2A}\rangle = \alpha/|G|$. Therefore, given that α is sufficiently large relative to $|G|$, we can conclude that A contains a non-trivial 3AP in this case. This last condition will be referred to as the *size condition*. This is the condition which ultimately produces the condition on the density in the statement of Theorem 2.1.
2) ϵ is large : In the second case, we have that $|\hat{1}_A(\gamma)|$ is large for some $\gamma \neq 1_G$. What can we make of this? Recall the definition that

$$\hat{1}_A(\gamma) = \frac{1}{|G|}\sum_{x\in A} \bar{\gamma}(x).$$

On the other hand, the orthogonality property of characters says that $\sum_{x\in G} \bar{\gamma}(x) = 0$. Hence, the above quantity being large implies that A has a certain bias. The main difficulty is extracting this bias. There are three things that we will use.
   a) *Annihilating set of* γ: In order to extract data about the structure of A, we first find a set P, such that $|\gamma(x) - 1|$ is small on P, and hence $\gamma(x)$ is almost constant on all translates of P.
   b) *Average transfer from G to P:* In order to use P, we move from $\hat{1}_A(\gamma) = \langle 1_A\gamma, 1_G\rangle$ to $\langle 1_A\gamma * 1_P, 1_G\rangle$, as done in (2.6).

c) *Annihilating γ:* We move from $\langle 1_A \gamma * 1_P, 1_G \rangle$ to $\langle 1_A * 1_P, 1_G \rangle$ by a), which completes the annihihation of γ. We find that A has a *density increment* i.e. A has density greater than α on some translate of P.

If case 1) occurrs, then we are done, and if case 2) occurs, we can iterate the same argument from the beginning, taking G to be $\mathbb{Z}_{|P|}$. Due to the density increment, only finitely many iterations are possible until the density of A exceeds 1. Hence, the iteration must terminate at some step by the occurrence of 1).

The overall flow of proofs in the subsequent chapters improving Roth's result remains the same (in particular, this dichotomy of either getting a density increment or a non-trivial 3AP). In the next chapter, we will optimize the points a) and b) of case 2). Instead of arithmetic progressions, we will be using a set known as a Bohr set, which by definition achieves a), the annihilation of γ. However, this change will also call for a technique to achieve the average transfer in b) using a Bohr set.

CHAPTER 3

Bourgain's improvement

In 1999, Jean Bourgain improved Theorem 2.1 to the following.

Theorem 3.1. *There exists $C > 0$ such that if $A \subset [N]$ is of density α in $[N]$ where*

$$\alpha > C \left(\frac{\log \log N}{\log N} \right)^{1/2},$$

then A contains a non-trivial 3AP.

In fact, what we will prove is the following generalization of the above result.

Theorem 3.2. *Let G be a group of odd order. Then there exists an absolute $C > 0$ such that if $A \subset G$ is of density α in G where*

$$\alpha > C \left(\frac{\log \log |G|}{\log |G|} \right)^{1/2},$$

then A contains a non-trivial 3AP.

By Lemma 2.11, we can easily obtain Theorem 3.1 from Theorem 3.2, .

In Chapter 2, Section 2.3 we discuss the ideas involved in the proof of Theorem 2.1, and how it can be improved. There are two important concepts which are used to extract the structure of A from its Fourier transform: the annihilation set, and the average transfer. Bourgain's main achievement in his 1999 paper is the success of using Bohr sets as the annihilation set. A Bohr set, as we shall see, is by its definition suited for annihilation. The difficulty is to understand what such a set looks like, and show how other parts of the method, in particular the average transfer, can also be achieved with Bohr sets. We will first generalize the average transfer technique.

3.1. Average transfer

As discussed in Chapter 2, Section 2.3, the average transfer refers to the idea of moving from a sum $\langle f, 1_S \rangle$ to a sum $\langle f * 1_{-T}, 1_S \rangle$ (the negative sign is just there for a technical reason) for some subsets $S, T \subset G$ and function $f \in C(G)$. In the proof of Lemma 2.12 this was easy to achieve since we had $S = G$, the whole group. In this case, we have

$$\mu_G(T) \langle f, 1_S \rangle = \langle f * 1_{-T}, 1_S \rangle.$$

However, it is not so simple if S is not the entire group. Looking at this closely, we have

$$\langle f * 1_{-T}, 1_S\rangle = \langle f, 1_S * 1_T\rangle = \frac{1}{|G|}\sum_{x\in G} f(x)(1_S * 1_T)(x).$$

If we want to relate this to $\mu_G(T)\langle f, 1_S\rangle$, what we want then is approximately

$$1_S * 1_T(x) = |S \cap (x - T)|/|G| \approx \begin{cases} \mu_G(T) & \text{if } x \in S \\ 0 & \text{if } x \notin S \end{cases}.$$

Intuitively, this can be interpreted as saying that the neighborhood of each point on S looks like T. This motivates the following definition used by Green in [**9**].

Definition 3.3. *Let $S, T \subset G$. We say that S is η locally-like $T \subset G$ if*

$$(3.1)\qquad ||1_S * 1_T - \mu_G(T)1_S||_{L^1} = \frac{1}{|G|}\sum_{x\in G}|1_S * 1_T(x) - \mu_G(T)1_S(x)| \le \eta\mu_G(S)\mu_G(T),$$

We note that

$$||1_S * 1_T - \mu_G(T)1_S||_{L^1} \le \frac{1}{|G|}\sum_{x\in G}(1_S * 1_T(x) + \mu_G(T)1_S(x)) = 2\mu_G(S)\mu_G(T),$$

so local-likeness is meaningful only when $\eta < 2$.

Example 1 Let $T \subset G$. Then G is 0-locally-like T.

Example 2 Let $P \subset G$ be an arithmetic progression, say $P = \{a, a+d, a+2d, \ldots\}$ for some $a, d \in G$. Let $P' = \{a, a+d, a+2d, \ldots\}$ as well, but let it be shorter than P. Let $\alpha = |P|/|P'|$.

Claim: P is 2α locally-like $-P'$
Proof of Claim: We need to look at $(x+P')\cap P$ for each x. If $x \in P$, then $|(x+P')\cap P| = |P'|$ except at the last $|P'|$ points of P. We see that in the last case, the intersection is at most $|P'|$. For $x \notin P$, there are also at most $|P'|$ many ways such that $(x+P')\cap P \neq \emptyset$ (for each $y \in P'$, there is at most one $x \notin P$ such that $x + y$ is the first point on $x + P'$ intersecting P) and the intersection again has at most size $|P'|$. Therefore,

$$\frac{1}{|G|}\sum_{x\in G}|1_P * 1_{P'}(x) - \mu_G(P')1_P(x)| \le 2\mu_G(P')^2 = 2\alpha\mu_G(P)\mu_G(P').$$

□Claim

Let us now show how an average of $f \in C(G)$ over S transfers to an average over T by the concept of local-likeness, where S is η locally-like T. This can be in fact achieved quite easily. Suppose that we have

$$|\langle f, 1_S\rangle| \ge \kappa\mu_G(S),$$

for some $\kappa > 0$. We consider the difference

$$\begin{aligned}|\langle f * 1_{-T}, 1_S\rangle - \mu_G(T)\langle f, 1_S\rangle| &= |\langle f, 1_S * 1_T - \mu_G(T)1_S\rangle| \\ &\leq \frac{1}{|G|}\sum_{x\in G}|f(x)||1_S * 1_T(x) - \mu_G(T)1_S(x)| \\ &\leq \eta\,||f||_\infty\,\mu_G(S)\mu_G(T),\end{aligned}$$

where the last inequality is by local-likeness. Hence, we have

$$\langle f * 1_{-T}, 1_S\rangle = \mu_G(T)\langle f, 1_S\rangle + O(\eta\,||f||_\infty\,\mu_G(S)\mu_G(T)),$$

where the implied constant can be taken as 1. We also have the explicit lower bound

$$\begin{aligned}|\langle f * 1_{-T}, 1_S\rangle| &= |\mu_G(T)\langle f, 1_S\rangle + \langle f * 1_{-T}, 1_S\rangle - \mu_G(T)\langle f, 1_S\rangle| \\ &\geq |\mu_G(T)\langle f, 1_S\rangle| - |\langle f * 1_{-T}, 1_S\rangle - \mu_G(T)\langle f, 1_S\rangle| \\ &\geq (\kappa - \eta\,||f||_\infty)\mu_G(S)\mu_G(T).\end{aligned}$$

It follows that

$$\max_{x\in S}\{|f * 1_T(x)|\}\,\mu_G(S) \geq |\langle f * 1_{-T}, 1_S\rangle| \geq (\kappa - \eta\,||f||_\infty)\mu_G(S)\mu_G(T).$$

We summarize this in the following lemma.

Lemma 3.4 (Average transfer from S to T)**.** *Let $f \in C(G)$ and let $S, T \subset G$. Suppose that S is η locally-like T for some $\eta > 0$. Also suppose that on S, we have*

$$|\langle f, 1_S\rangle| \geq \kappa\mu_G(S).$$

for some $\kappa > 0$. Then we have the following.

i) (Average transfer):

$$\langle f * 1_{-T}, 1_S\rangle = \mu_G(T)\langle f, 1_S\rangle + O(\eta\,||f||_\infty\,\mu_G(S)\mu_G(T)),$$

where the implied constant is at most 1.

ii) (Lower bound):

$$\langle f * 1_{-T}, 1_S\rangle \geq (\kappa - \eta\,||f||_\infty)\mu_G(S)\mu_G(T).$$

iii) (Maximum of transfer): There exists $x \in S$ such that

$$|f * 1_{-T}(x)| \geq (\kappa - \eta\,||f||_\infty)\mu_G(T).$$

Example (Counting 3APs) We can use Lemma 3.4 to count the number of 3APs of a certain type. Suppose S is symmetric and S is η locally-like $2T$. Then with $f = 1_S$, using Lemma 3.4 ii) we have

$$\langle 1_S * 1_S, 1_{2T}\rangle = \langle 1_S * 1_{-2T}, 1_{-S}\rangle \geq (1-\eta)\mu_G(S)\mu_G(T).$$

Now $|G|^2\langle 1_S * 1_S, 1_{2T}\rangle$ is precisely the number of 3APs $(x, y, z) \in S \times T \times S$ with $x+z = 2y$, so the above says that the number of such 3APs is at least $(1-\eta)|S||T|$, which is close to the largest possible value of $|S||T|$.

3.1.1. Density transfer. We will now show an application of Lemma 3.4, which will be used in later proofs. If S is locally-like T, and if we know the density of A on S, then it is not unreasonable to hope to get some knowledge of the density of A on T.

Lemma 3.5 (Density transfer)**.** *Let $S, T \subset G$. Suppose that S is η locally-like T. Also suppose that a set A has density α on S. Let $\alpha(x)$ be the density of A on $x+T$ for each $x \in S$ i.e.*

$$\alpha(x) := \frac{|A \cap (x+T)|}{|x+T|} = \mu_G(T)^{-1}(1_A * 1_{-T})(x).$$

Then

$$\left|\sum_{x\in S}(\alpha(x) - \alpha)\right| \leq \eta|S|.$$

Moreover, there is a $x \in S$ such that $\alpha(x) \geq \alpha - \eta$.

Proof: We simply apply of Lemma 3.4 i) with $f(n) = 1_A(n)$ and $\kappa = \alpha$. For we would then have $\langle 1_A, 1_S\rangle = \alpha\mu_G(S)$, so

$$\eta\mu_G(S)\mu_G(T) \geq |\langle 1_A * 1_{-T}, 1_S\rangle - \mu_G(T)\langle 1_A, 1_S\rangle| = \frac{\mu_G(T)}{|G|}\left|\sum_{x\in S}(\alpha(x) - \alpha)\right|.$$

Cancelling the appropriate factors gives the desired expression. The lower bound on $\alpha(x)$ follows from applying iii) of Lemma 3.4. $\square$

Local-likeness alone is insufficient to show that there is a density increment. We will need some specific knowledge about the structure of S and T that are involved. A slight improvement of the previous lemma is the following.

Lemma 3.6 (Density transfer 2)**.** *Let $S, T_1, T_2 \subset G$. Suppose that S is η locally-like T_1 and T_2, and suppose a set A has density α on S. Let $\alpha_1(x)$ and $\alpha_2(x)$ be the density of A on $T_1 + x$ and $T_2 + x$ respectively. Then there exists $x \in S$ such that one of the following holds.*

i) $\alpha_1(x) \geq \alpha + \eta$, *or*
ii) $\alpha_2(x) \geq \alpha + \eta$, *or*
iii) $\alpha_1(x) \geq \alpha - 4\eta$ *and* $\alpha_2(x) \geq \alpha - 4\eta$.

If either i) or ii) occur, then we obtain a density increment of A; otherwise, we have T_1 and T_2 on which the density of $A - x$ is not too small simulatenously. In application, when the latter case happens, we will work with A on T_1 and T_2 instead of S. which give us more freedom. This lemma can be easily generalized to more than just two sets T_1 and T_2.

Proof: If i) or ii) hold, then there is nothing to show, so let us suppose otherwise. For a contradiction, we also assume that iii) does not hold. This implies that either there are at least $|S|/2$ values of x where $\alpha_1(x) < \alpha - 4\eta$, or $|S|/2$ values of x where $\alpha_2(x) < \alpha - 4\eta$. Without loss of generality, we assume it is the former. Then, since i) does not hold for all x, we have

$$\sum_{x\in S,\ \alpha_1(x)\geq\alpha} (\alpha_1(x) - \alpha) < \eta|S|.$$

We also have

$$\left|\sum_{x\in S,\ \alpha_1(x)<\alpha} (\alpha_1(x) - \alpha)\right| > 4\eta\frac{|S|}{2} = 2\eta|S|.$$

Hence

$$\left|\sum_{x\in S,\ \alpha_1(x)<\alpha} (\alpha_1(x) - \alpha) + \sum_{x\in S,\ \alpha_1(x)\geq\alpha} (\alpha_1(x) - \alpha)\right| > 2\eta|S| - \eta|S| = \eta|S|,$$

which contradicts Lemma 3.5. □

3.1.2. Annihilating a character by average transfer. Recall from Chapter 2, Section 2.3 that the average transfer was used to annihilate the character from a Fourier transform in the proof of Lemma 2.12. We can generalize the method used in that proof as follows.

Lemma 3.7 (Annihilation by average transfer)**.** *Let $S, T \subset G$. Suppose that S is η locally-like T. Let $\gamma \in \hat{G}$ and $f \in C(G)$. Suppose that for all $x \in T$, we have*

$$|1 - \gamma(x)| \leq \epsilon.$$

If

$$|\langle f\bar{\gamma}, 1_S\rangle| \geq \kappa\mu_G(S),$$

then

$$\langle |f * 1_{-T}|, 1_S\rangle \geq (\kappa - \eta - \epsilon\,||f||_\infty)\mu_G(S)\mu_G(T).$$

That is, assuming that T annihilates γ, we can achieve an annihilation of γ from the Fourier transform (if f is supported on S, then $\langle f\bar{\gamma}, 1_S\rangle = \hat{f}(\gamma)$). In the next section on Bohr sets, we will study more about sets that satisfy the condition above on T.

Proof: By Lemma 3.4, we have

$$(\kappa - \eta\,||f||_\infty)\,\mu_G(S)\mu_G(T) \leq |\langle f\bar{\gamma} * 1_{-T}, 1_S\rangle| \leq \langle |f\bar{\gamma} * 1_{-T}|, 1_S\rangle. \tag{3.2}$$

On the otherhand, since $|1-\gamma(x)| \leq \epsilon$ for all $x \in T$, we have

$$\begin{aligned}
\langle |f\bar{\gamma} * 1_{-T}|, 1_S\rangle &= \frac{1}{|G|^2}\sum_{x\in S}\left|\sum_{y\in x+T} f(y)\bar{\gamma}(y)\right| \\
&\leq \frac{1}{|G|^2}\sum_{x\in S}\left|\sum_{y\in x+T} f(y)\bar{\gamma}(x)\right| + \frac{1}{|G|^2}\sum_{x\in S}\left|\sum_{y\in x+T} f(y)(\bar{\gamma}(x)-\bar{\gamma}(y))\right| \\
&\leq \frac{1}{|G|^2}\sum_{x\in S}\left|\sum_{y\in x+T} f(y)\right| + \frac{1}{|G|^2}\sum_{x\in S}\sum_{y\in x+T} ||f||_\infty \epsilon \\
&= \langle |f * 1_T|, 1_S\rangle + \epsilon\, ||f||_\infty\, \mu_G(S)\mu_G(T).
\end{aligned}$$

Combined with (3.2), we have

$$(\kappa - \eta\, ||f||_\infty - \epsilon\, ||f||_\infty)\, \mu_G(S)\mu_G(T) \leq \langle |f * 1_{-T}|, 1_S\rangle.$$

□

3.2. Annihilating set: Bohr set

In this section, we will discuss some results about Bohr sets. Before defining a Bohr set, it will be useful to have the following terminology.

Definition 3.8. *Let $B \subset G$ and let $\gamma \in \hat{G}$. We say that B ϵ-annihilates γ if for all $x \in B$, we have $|1-\gamma(x)| \leq \epsilon$. If $\Gamma \subset \hat{G}$, then we say that B ϵ-annihilates Γ if B ϵ-annihilates γ for each $\gamma \in \Gamma$.*

A Bohr set is by definition a set which ϵ-annihilates a set of characters.

Definition 3.9. *Let $\Gamma \subset \hat{G}$ be non-empty and $0 \leq \epsilon \leq 2$. We define the Bohr set Λ with frequency set Γ and width ϵ as the set*

$$\Lambda = \{x \in G : |1-\gamma(x)| \leq \epsilon,\ \forall \gamma \in \Gamma\}. \tag{3.3}$$

$|\Gamma|$ is called the dimension of Λ. For any $\rho > 0$ we define Λ_ρ to be the Bohr set with frequency set Γ and width $\rho\epsilon$.

Below are some easy consequences of the definition.

Lemma 3.10. *Let Λ be a Bohr set with frequency set Γ and width ϵ. Then*

i) Λ is a symmetric set.
ii) For any $\rho > 0$, if $x \in \Lambda$ and $y \in \Lambda_\rho$, then $x + y \in \Lambda_{1+\rho}$.
iii) If Λ' is another Bohr set with frequency set Γ' and width ϵ, then $\Lambda \cap \Lambda'$ is the Bohr set with frequency set $\Gamma \cup \Gamma'$ and width ϵ.

Proof: i) and iii) are obvious. For ii), if $\gamma \in \Gamma$, then

$|1-\gamma(x+y)| = |\gamma(-x)-\gamma(y)| \leq |\gamma(-x)-1|+|1-\gamma(y)| = |1-\gamma(x)|+|1-\gamma(y)| \leq (1+\rho)\epsilon.$

□

Example 1 G itself is a Bohr set with frequency set $\emptyset$ and width of any $\epsilon > 0$.

Example 2 Let $G = \mathbb{Z}_N$ for some integer N. If $k|N$, then multiples of k form a Bohr set. Indeed, let $\gamma = e_{(N/k)/N} = e_{1/k}$, and consider the Bohr set Λ with frequency set $\{\gamma\}$ and width 0. Then $x \in \Lambda$ if and only if $k|x$.

Example 3 Let $G = \mathbb{Z}_N$. The following is a generalization of the previous example. Let $k_1, \ldots, k_l$ be pairwise coprime and $k_i|N$ for each i. Set $M = k_1 \cdots k_l$; note that $M|N$. Let γ_i be the character $\gamma_i = e_{(N/k_i)/N} = e_{1/k_i}$ for each $i = 1, \ldots, l$. Let us consider the Bohr set Λ with frequency set $\{\gamma_1, \ldots, \gamma_l\}$ and width $1/2 > \epsilon > 0$. Let us write $-k/2 < \{x\}_k \leq k/2$ to mean the mod k representation of x. We can now give a precise description of Λ. Observe that for $x \in \mathbb{Z}_{k_i}$, the value $|1 - \gamma_i(x)|$ is increasing as x runs from 0 to $k_i/2$. Let x_i be the largest such x such that $|1 - \gamma_i(x)| \leq \epsilon$, and set $r_i = |1 - \gamma(x_i)|$. Now let $a = (a_1, \ldots, a_l)$ be such that $|\{a_i\}_{k_i}| \leq r_i$ for each $i = 1, \ldots, l$. There are $R = \prod_{i=1}^{l}(2r_i + 1)$ possible such choices of vector a. Now by the Chinese remainder theorem, there exists exactly one $m_a \in \mathbb{Z}_M$ such that

$$m \equiv m_a \pmod M \iff \forall i \; m \equiv a_i \pmod{k_i}.$$

Such m_a lies in Λ. Conversely, if $m \in \mathbb{Z}_M$ is in Λ, then it must satisfy the right condition for some vector a. Thus the vectors a completely characterize all $m \in \mathbb{Z}_M$ which are in Λ, and there are R many of them. N contains N/M copies of M, so Λ has period M and $|\Lambda| = NR/M$.

The following is a general estimate of the size of a Bohr set

Proposition 3.11 (Cardinality of a Bohr set)**.** *Let Λ be a Bohr set of dimension d and width ϵ. Then*

$$\left(\frac{\epsilon}{3\pi}\right)^d M \leq |\Lambda| < 4^d |\Lambda_{1/2}|.$$

Proof: Let $\gamma_1, \ldots, \gamma_d$ be the characters in its frequency set.

Lower bound: Divide the unit sphere $\mathbb{S}^1 \subset \mathbb{C}$ into $\lceil 2\pi/\epsilon \rceil$ intervals of equal arc length $\Delta = 2\pi/\lceil 2\pi/\epsilon \rceil$, so $\mathbb{S}^d$ has $\lceil 2\pi/\epsilon \rceil^d$ divisions. Note that $2\pi = \Delta\lceil 2\pi/\epsilon \rceil \geq \Delta(2\pi/\epsilon)$ so $\epsilon \geq \Delta$. Also,

$$\lceil 2\pi/\epsilon \rceil^d \leq \left(\frac{2\pi}{\epsilon} + 1\right)^d = \left(\frac{2\pi}{\epsilon}\right)^d (1 + \frac{\epsilon}{2\pi})^d < \left(\frac{2\pi}{\epsilon}\right)^d \left(\frac{3}{2}\right)^d,$$

so there are at most $(3\pi/\epsilon)^d$ many divisions. Now for each $x \in G$, consider the point $P_x = (\gamma_1(x), \ldots, \gamma_d(x)) \in \mathbb{S}^d$. Then by the pigeonhole principle, one of the divisions must contain at least $(\epsilon/3\pi)^d|G|$ many such P_x's. Let Q be the set of such points. Then for each $x, y \in Q$ and any γ_i for $i = 1, \ldots, d$, we have

$$|1 - \gamma_i(x - y)| = |\gamma_i(x) - \gamma_i(y)| \leq \Delta \leq \epsilon,$$

(recalling that the distance between two points on an arc is less than the arc length between them). Therefore, the set $Q - Q = \{x - y : x, y \in Q\}$ is contained in Λ. The claim for the lower bound follows.

Upper bound: Consider the arc R with midpoint 1 on $\mathbb{S}$ with arc length 2ϵ, and partition this arc into 4 arcs of length $\epsilon/2$ each. Then for any γ_i and $x \in \Lambda$, $\gamma_i(x)$ must lie somewhere on this arc since $|1 - \gamma_i(x)| \leq \epsilon$ and hence in one of these divisions. Considering all the γ_i's in the frequency set, there are 4^d possible divisions of R^d. Let Q be the set of all $x \in \Lambda$ lying in one of these divisions. Then for any $x, y \in Q$, we have

$$|1 - \gamma_i(x - y)| = |\gamma_i(x) - \gamma_i(y)| \leq \frac{\epsilon}{2}.$$

Therefore, $Q - Q \subset \Lambda_{1/2}$. There are 4^d many such Q's, so the claim for the upper bound follows. □

3.2.1. Bohr sets and average transfer. We would like to perform an average transfer from a large Bohr set to translates of a small Bohr set. In the previous section, we saw that this can be achieved by local-likeness. Therefore, what we would like to know now is, given a Bohr set Λ and Λ_ρ for some $0 < \rho < 1$, how much the former is locally-like the latter.

In order to obtain a general method to calculate the local-likeness parameter, we introduce one intermediate concept.

Definition 3.12. *Let Λ be a Bohr set of dimension $d > 0$. This set is said to be regular if for any $|\rho| \leq 1/100d$, we have*

$$1 - 100d|\rho| \leq \frac{|\Lambda_{1+\rho}|}{|\Lambda|} \leq 1 + 100d|\rho|.$$

Informally, a regular Bohr set is one where small perturbations in its width does not drastically change the structure of the set. The factor 100 is not critical, and can be replaced by other large numbers. In fact, according to an analyst that the author is acquainted with, the use of 100 is a standard implicit joke amongst harmonic analysts, representing the stance "I don't care, it's just some large number".

Example 1 Consider the Bohr set Λ with frequency set $\{1_G\}$ and width 1. Then Λ is just the entire group G, and for any $\rho > 0$, $\Lambda_{1+\rho}$ is still G. Hence Λ is regular.

Example 2 Here is an example of a non-regular Bohr set. Let $G = \mathbb{Z}_N$ and let k be an integer such that $k|N$. Set $l = N/k$ and $\gamma = e_{l/N} = e_{1/k}$. We consider the Bohr set with frequency set $\{\gamma\}$ and width $\epsilon = |1 - \gamma(1)| - \eta$ where $\eta = \rho(|1 - \gamma(1)|)/(\rho + 1)$ and $\rho = 1/100$. Since $\eta > 0$, it is easy to see that this Bohr set consists exactly all the multiples of k in $\mathbb{Z}_N$. Hence $|\Lambda| = l$. On the other hand, $\Lambda_{1+\rho}$ also contains all the integers $\pm 1 \pmod{k}$, since if x is such an integer, we have

$$|1-\gamma(x)| = |1-\gamma(1)| = \epsilon + \rho\epsilon = (1+\rho)\epsilon.$$

Therefore, $|\Lambda_{1+\rho}| = 3l = 3|\Lambda|$. Thus, Λ is not regular.

The following proposition calculates the local-likeness parameter for a regular Bohr set.

Proposition 3.13 (Local-likeness of Bohr sets)**.** *Let Λ be a Bohr set of dimension d. If Λ is regular, then for any $T \subset \Lambda_\rho$ with $0 < \rho \leq 1/100d$, Λ is $200d\rho$ locally-like T. Hence, for any $\eta < 1$, if $0 < \rho < \eta/200d$, then Λ is η locally-like T.*

Proof: We need to calculate

$$\frac{1}{|G|}\sum_n |1_\Lambda * 1_T(n) - \mu_G(T)1_\Lambda(n)|.$$

Recall that

$$1_\Lambda * 1_T(n) = |\{(m,t) \in \Lambda \times T : m+t = n\}|.$$

Let Γ be the frequency set of Λ.

(1) If $n \notin \Lambda_{1+\rho}$, then $1_\Lambda * 1_T(n) = 0$: We will prove the contrapositive. Suppose $1_\Lambda * 1_T(n) \neq 0$, so there exists $m \in \Lambda$ and $t \in T$ such that $m+t = n$. This implies that $n \in \Lambda_{1+\gamma}$ by Lemma 3.10 ii).

(2) If $n \in \Lambda_{1-\rho}$, then $1_\Lambda * 1_T(n) = \mu_G(T)$: Suppose $m \in T - n$, so $m = t - n$ for some $t \in T$. This implies that $m \in \Lambda$ by Lemma 3.10 ii).

In particular, in the above two cases, we have

$$1_\Lambda * 1_T(n) - \mu_G(T)1_\Lambda(n) = 0.$$

Furthermore, since Λ is regular, we have

$$\Lambda_{1+\rho} \leq (1+100d\rho)|\Lambda|$$

and

$$\Lambda_{1-\rho} \geq (1-100d\rho)|\Lambda|.$$

Therefore

$$\begin{aligned}
\frac{1}{|G|}\sum_n |1_\Lambda * 1_T(n) - \mu_G(T)\Lambda(n)| &= \frac{1}{|G|}\sum_{n \in \Lambda_{1+\rho}\setminus\Lambda_{1-\rho}} |1_\Lambda * 1_T(n) - \mu_G(T)\Lambda(n)| \\
&\leq \frac{1}{|G|}\mu_G(T)|\Lambda_{1+\rho} \setminus \Lambda_{1-\rho}| \\
&\leq \frac{1}{|G|}\mu_G(T)|\Lambda|(1+100d\rho - (1-100d\rho)) \\
&= 200d\rho\mu_G(T)\mu_G(\Lambda).
\end{aligned}$$

□

Equipped with a way of calculating the local-likeness parameter of Bohr sets, it remains to show how to obtain a regular Bohr set. Intuitively, if a Bohr set is irregular, it means that the width ϵ is near the borderline case, where small perturbations dramatically change the size of the set. Hence, if we modify the width slightly to move away from this value of ϵ, then we can reasonably expect to obtain a regular set.

Proposition 3.14 (Obtaining a regular Bohr set)**.** *If Λ is not regular, then there exists $\alpha \in [1/2, 1)$ such that Λ_α is regular.*

Proof: We assume otherwise and seek a contradiction. The idea is to compute the ratio of $|\Lambda|$ and $|\Lambda_{1/2}|$ by a sort of telescoping product of $|\Lambda_{\alpha(1+\rho)}|$ and $|\Lambda_{\alpha(1-\rho)}|$ for each $\alpha \in [1/2, 1]$ and some $\rho > 0$. If the latter ratios are large (which can be forced since Λ_α is irregular), then the former ratio violates Proposition 3.11. We will now make this precise. Let $f(x) = |\Lambda_x|$. , By the definition of regularity, for each $\alpha \in [1/2, 1)$, there exists at least one $0 < \rho \leq 1/100d$ such that one of the following holds. Either

$$\frac{f((1+\rho)\alpha)}{f(\alpha)} > 1 + 100d\rho$$

or

$$\frac{f((1-\rho)\alpha)}{f(\alpha)} < 1 - 100d\rho.$$

In the latter case, since f is increasing, we obtain that

$$1 + 100d\rho < \frac{1}{1 - 100d\rho} < \frac{f(\alpha)}{f((1-\rho)\alpha)} \leq \frac{f((1+\rho)\alpha)}{f((1-\rho)\alpha)}.$$

In any case, since f is increasing, we have

$$e^{50d\rho} < 1 + 100d\rho < \frac{f((1+\rho)\alpha)}{f((1-\rho)\alpha)}.$$

Let $t_\alpha = \rho$ for each $\alpha \in [1/2, 1)$.

Claim: If the intervals $I_1, \ldots, I_k$ cover $[0, 1]$, then there is a disjoint union (except maybe at endpoints) of a subcollection of these intervals such that the measure of this union is at least $1/2$.

Suppose for now that the claim is true, and consider the cover of

$$I = [1/2 + 1/100d, 1 - 1/100d],$$

by the set of intervals

$$I_\alpha = [(1 - t_\alpha)\alpha, (1 + t_\alpha)\alpha],$$

for each $\alpha \in I$. Since I is compact, there is a finite subcover generated by $\alpha_1, \ldots, \alpha_k \in I$. Then by the claim above, there is a disjoint subcollection of these intervals, say $\alpha_1, \ldots, \alpha_l$

without loss of generality in increasing order, such that the union has measure at least $1/2(1/2 - 2/100d) > 1/5$. The exact measure of this union is

$$\sum_{i=1}^{k} 2t_{\alpha_i}\alpha_i.$$

Hence

$$\sum_{i=1}^{k} t_{\alpha_i} \geq \sum_{i=1}^{k} t_{\alpha_i}\alpha_i > \frac{1}{10}.$$

It follows that

$$\prod_{i=1}^{k} \frac{f((1+t_{\alpha_i})\alpha_i)}{f((1-t_{\alpha_i})\alpha_i)} > e^{50d\sum t_{\alpha_i}} > e^{5d}.$$

On the other hand, since the I_αs are disjoint, we know that

$$(1+t_{\alpha_i})\alpha_i \leq (1-t_{\alpha_{i+1}})\alpha_{i+1}$$

for any $i = 1, \ldots, k-1$. Thus,

$$\prod \frac{f((1+t_{\alpha_i})\alpha_i)}{f((1-t_{\alpha_i})\alpha_i)} \leq \frac{f((1+t_{\alpha_k})\alpha_k)}{f((1-t_{\alpha_1})\alpha_1)}.$$

Furthermore, by the choice of I, we have that $1/2 < (1+t_{\alpha_i})\alpha_i < 1$ for each i, so the above is bounded above by $f(1)/f(1/2)$. Now, by Proposition 3.11, we know that $f(1) \leq 4^{d+1} f(1/2)$. Thus, combining above estimates gives

$$e^{5d} < \frac{f(1)}{f(1/2)} < 4^d,$$

but this inequality does not hold for any $d \geq 1$. We have obtained a contradiction. It now remains to prove the Claim.

Proof of Claim: Without loss of generality, suppose that $I_1, \ldots, I_k$ is a minimal cover of $[0,1]$. We first observe that each $x \in [0,1]$ is contained in at most two intervals. To see this, again without loss of generality, we suppose that $x \in I_1 \cap I_2$, and that $x \in I_3$. Let us write $I_j = [a_j, b_j]$ for each $j = 1, 2, 3$. We may also suppose that $a_1 < a_2$, and consequently that $b_1 < b_2$; for if $b_2 \leq b_1$, then $I_2 \subset I_1$, which contradicts the minimality of the cover. Now since $x \in I_3$ as well, I_3 must satisfy one of the following:

(1) $a_3 \leq a_1$ and $b_3 \leq b_2$, so $I_1 \subset I_2 \cup I_3$, or
(2) $a_3 \leq a_1$ and $b_2 \leq b_3$, so $I_1 \cup I_2 \subset I_3$, or
(3) $a_1 \leq a_3$ and $b_3 \leq b_2$, so $I_3 \subset I_1 \cup I_2$, or
(4) $a_1 \leq a_3$ and $b_2 \leq b_3$, so $I_2 \subset I_1 \cup I_3$.

All contradict the minimality of the cover. Hence, each x is contained in at most two intervals. Now, assuming that a_j's are in increasing order, consider the collections

$$\{I_1, I_3, \ldots\}, \ \{I_2, I_4, \ldots\}.$$

The intervals in each collection is pairwise-disjoint, and one of these collections must have a totalmeasure of at least 1/2. □

3.2.2. A Bohr set annihilates its own spectrum. Let Λ be a Bohr set with frequency set Γ and width ϵ. Then by definition, Λ annihilates Γ. But can it annihilate more characters? If so, which ones?

We first notice that 1_Λ has a relatively large Fourier transform on Γ, for if $\gamma \in \Gamma$, then

$$|\hat{1}_\Lambda(\gamma)| = \frac{1}{|G|}\left|\sum_{x\in\Lambda}\bar{\gamma}(x)\right| \geq \frac{1}{|G|}\sum_{x\in\Lambda}1 - \frac{1}{|G|}\sum_{x\in\Lambda}|1-\gamma(x)| \geq (1-\epsilon)\mu_G(\Lambda).$$

We first consider the following definition.

Definition 3.15. *Let $f \in C(G)$, $\mu \in M(G)$, and $\kappa > 0$. The κ-spectrum of f with respect to μ is the set*

$$\mathrm{Spec}_\kappa(f,\mu) = \left\{\gamma \in \hat{G} : |f\hat{d}\mu(\gamma)| \geq \kappa\,||f||_{L^1(\mu)}\right\}.$$

Again, as with the definitions in Chapter 2 Section 2.1, we are usually only concerned with spectrum case when $\mu = \mu_G$. The generalized definition comes into play, however, in Chapter 2.

What we have shown then is that

$$\Gamma \subset \mathrm{Spec}_{1-\epsilon}(1_\Lambda).$$

Naturally, we could then ask the converse; that is, does Λ annihilate $\mathrm{Spec}_{1-\epsilon}(1_\Lambda)$?

Lemma 3.16 (Annihilating the spectrum)**.** *Let Λ be a regular Bohr set of frequency set Γ and width ϵ. Then*

$$\Gamma \subset \mathrm{Spec}_{1-\epsilon}(1_\Lambda),$$

and, for any $\kappa > 0$ and $\rho < 1/100|\Gamma|$,

$$\mathrm{Spec}_\kappa(1_\Lambda) \subset \left\{\gamma \in \hat{G} : |1-\gamma(x)| < 200d\rho\kappa^{-1}\ \forall x \in \Lambda_\rho\right\}.$$

In other words, Λ_ρ $200d\rho\kappa^{-1}$-annihilates $\mathrm{Spec}_\kappa(1_\Lambda)$.

So in a sense, a Bohr set annihilates its own spectrum. This will be the key in the proof of Bourgain's result.

Proof: We have already proved the first inclusion. Let $\gamma \in \mathrm{Spec}_\kappa 1_\Lambda$ and $x \in \Lambda_\rho$. Then

$$|1-\gamma(x)|\kappa\mu_G(\Lambda) \leq |\hat{1}_\Lambda(\gamma) - \hat{1}_\Lambda(\gamma)\gamma(x)| = \frac{1}{|G|}\left|\sum_{y\in\Lambda}\bar{\gamma}(y) - \sum_{y\in\Lambda}\bar{\gamma}(y-x)\right|.$$

Note that if $y \in \Lambda$, then $y - x \in \Lambda_{1+\rho}$. If $y - x \in \Lambda$, then $\bar{\gamma}(y-x)$ in the second sum gets cancelled by some $\bar{\gamma}(y')$ for $y' \in \Lambda$ in the first sum. Therefore, at most $|\Lambda_{1+\rho} \setminus \Lambda|$ terms do not get cancelled in the second sum. Since both sums have the same number of terms, this means that the first sum also has the same number of terms that do not get cancelled. Thus, by the regularity of Λ, we have

$$|1-\gamma(x)|\kappa\mu_G(\Lambda) \leq \frac{1}{|G|}2|\Lambda_{1+\rho} \setminus \Lambda| \leq 200d\rho\mu_G(\Lambda).$$

The lemma follows. □

3.3. Proof of Bourgain's Theorem

With these tools, we now present a proof of Theorem 3.2. The basic idea is for the most part the same as that in Chapter 2, Section 2.2, except that now we iterate on Bohr sets rather than progressions. This adds a little complication, since a Bohr set is not necessarily a group. Hence, for instance, counting the number of 3APs in a Bohr set is not easy, nevermind even counting the number of 3APs in a subset A of it. Nevertheless, the main idea is that, as in Chapter 2, Section 2.2, either the size condition fails, A contains a non-trivial 3AP or A has a density increment on some smaller Bohr set, as seen in the following lemma.

Lemma 3.17. *Let G be a group of odd order. Let $\Lambda \subset G$ be a regular d dimensional Bohr set with width ϵ, and let $A \subset \Lambda$ have density α in Λ. The one of the following occurs.*

i) (Size condition failure)

$$|G| \leq (2\alpha^{-1}d\epsilon^{-1})^{25d}.$$

ii) A contains a non-trivial 3AP.

iii) (Density increment) There is a $x \in \Lambda$ such that $x - A$ has density at least $\alpha + 2^{-10}\alpha^2$ on a Bohr set Λ' of dimension at most $d+1$ and width at least $\rho\epsilon$, where $\rho = 2^{-57}\alpha^6 d^{-3}$.

In comparison to Lemma 2.12, the improvement is in the cardinality of the set on which a density increment of A is found. This makes sense, since Lemma 2.12 used a progression for its annihilating set, whereas the present lemma uses a Bohr set which is optimized for annihilation. The reason we require G to be odd is due to the following property.

Lemma 3.18. *Let G be a group of odd order. Then for any $A \subset G$, we have $|2A| = |A|$.*

Proof: Let $a, b \in A$. If $a + a = b + b$, then $b - a$ has order 2. But since G is of odd order, it must be that $b = a$. □

We now prove Lemma 3.17.

Proof of Lemma 3.17: Let $\Lambda' = \Lambda_{\rho'}$ and $\Lambda'' = \Lambda'_{\rho''}$ for some $\rho' < \eta'/200d$ and $\rho'' < \eta''/400d$ where $\eta', \eta'' > 0$. These values will be specified more later. By Proposition 3.14, we can assume that all these Bohr sets are regular. Also, by Proposition 3.13, we know that Λ is η' locally-like Λ' and hence Λ is also η' locally-like Λ'' since $\Lambda'' \subset \Lambda'$. If case i) or case ii) of Lemma 3.6 occur, then we have a density increment of $\alpha + \eta'$ on either Λ' or Λ'', and we are done, given that η' is chosen adequately (later), so let us suppose that case iii) occurs.

Then there is a x such that $A - x$ has density at least $\alpha - 4\eta'$ on both Λ' and Λ''. We will let $A' = (A-x) \cap \Lambda'$ and $A'' = (A-x) \cap \Lambda''$, and denote by α' and α'' the density of A' in Λ' and A'' in Λ'' respectively. We will require that $\eta' < \alpha/8$, so α' and α'' are at least $\alpha/2$.

As in the proof of Theorem 2.1, our starting point is the quantity

$$\epsilon := |\langle 1_{A'} * 1_{A'}, 1_{2A''} \rangle - \langle 1_{A'} * \alpha' 1_{\Lambda'}, \alpha'' 1_{2\Lambda''} \rangle|,$$

where the left term counts the 3APs $(x, y, z) \in A' \times A' \times A''$ (with weight $1/|G|^2$), and the right term is the expected value. Note that $2A''$ has density α'' in $2\Lambda''$ due to Lemma 3.18. Now, $2\Lambda'' \subset \Lambda'_{2\rho''}$, so Λ' is η'' locally-like $2\Lambda''$ by Proposition 3.13 (recall that $\rho'' < \eta''/400d$). Applying Lemma 3.4, we have

$$\langle 1_{A'} * \alpha' 1_{\Lambda'}, \alpha'' 1_{2\Lambda''} \rangle = \alpha'^2 \alpha'' \mu_G(\Lambda') \mu_G(\Lambda'') + O(\eta'' \alpha' \alpha'' \mu_G(\Lambda') \mu_G(\Lambda'')),$$

where the implied constant can be taken as 1, so if $\eta'' < \alpha'/4$, then

$$\langle 1_{A'} * \alpha' 1_{\Lambda'}, \alpha'' 1_{2\Lambda''} \rangle \geq \frac{3}{4} \alpha'^2 \alpha'' \mu_G(\Lambda') \mu_G(\Lambda'').$$

Now, we split into two cases.

Case 1: (Pseudorandom case) $\epsilon < \alpha'^2 \alpha'' \mu_G(\Lambda') \mu_G(\Lambda'')/2$
Then by the triangular inequality, we obtain that

$$\langle 1_{A'} * 1_{A'}, 1_{2A''} \rangle \geq \frac{1}{4} \alpha'^2 \alpha'' \mu_G(\Lambda') \mu_G(\Lambda'').$$

Now, if A only contains trivial 3APs, then

$$\langle 1_{A'} * 1_{A'}, 1_{2A''} \rangle = \frac{|A''|}{|G|^2} = \alpha'' \frac{\mu_G(\Lambda'')}{|G|}.$$

Hence if the above lower bound is greater than $\alpha'' \mu_G(\Lambda'')/|G|$, then we are guaranteed a non-trivial 3AP. This is satisfied if

$$|\Lambda'| \geq \frac{4}{\alpha'^2}. \tag{3.4}$$

Since $\alpha' \geq \alpha/2$, we can replace the right side by $16/\alpha^2$. As for $|\Lambda'|$, Proposition 3.11 implies that

$$|\Lambda'| \geq \left(\frac{\rho' \epsilon}{3\pi} \right)^d |G|.$$

Thus, (3.4) is satisfied if

$$\left(\frac{\rho' \epsilon}{3\pi} \right)^d |G| \geq \frac{16}{\alpha^2}. \tag{3.5}$$

We will come back to this after once we have ρ' determined.

Case 2: $\epsilon \geq \alpha'^2\alpha''\mu_G(\Lambda')\mu_G(\Lambda'')/2$
We will seek a density increment of A in this case. First, observe that

$$\begin{aligned} \alpha'^2\alpha''\mu_G(\Lambda')\mu_G(\Lambda'')/2 &\leq |\langle 1_{A'} * 1_{A'}, 1_{2A''}\rangle - \langle 1_{A'} * \alpha' 1_{\Lambda'}, \alpha'' 1_{2\Lambda''}\rangle| \\ &\leq |\langle 1_{A'} * 1_{A'}, (1_{2A''} - \alpha'' 1_{2\Lambda''})\rangle| + |\langle 1_{A'} * (1_{A'} - \alpha' 1_{\Lambda'}), \alpha'' 1_{2\Lambda''}\rangle| \end{aligned}$$

Hence we must have at least one of

$$\frac{1}{4}\alpha'^2\alpha''\mu_G(\Lambda')\mu_G(\Lambda'') \leq |\langle 1_{A'} * 1_{A'}, (1_{2A''} - \alpha'' 1_{2\Lambda''})\rangle| \tag{3.6}$$

or

$$\frac{1}{4}\alpha'^2\alpha''\mu_G(\Lambda')\mu_G(\Lambda'') \leq |\langle 1_{A'} * (1_{A'} - \alpha' 1_{\Lambda'}), \alpha'' 1_{2\Lambda''}\rangle|. \tag{3.7}$$

We shall consider two subcases

3.3.1. Density increment from $\hat{1}_{A'} - \alpha'\hat{1}_{\Lambda'}(\gamma)$. This section will deal with the case (3.6). It is a direct analogue of the method in Chapter 2 Section 2.2. By Proposition 2.6 iii), the inner product of (3.6) becomes

$$\frac{1}{4}\alpha'^2\alpha''\mu_G(\Lambda')\mu_G(\Lambda'') \leq \sum_{\gamma\in\hat{G}} \hat{1}_{A'}(\gamma)^2\overline{(\hat{1}_{2A''} - \alpha''\hat{1}_{2\Lambda''})(\gamma)}.$$

By Parseval's identity, this reduces to

$$\frac{1}{4}\alpha'^2\alpha''\mu_G(\Lambda')\mu_G(\Lambda'') \leq \left|\left|(\hat{1}_{2A''} - \alpha''\hat{1}_{2\Lambda''})\right|\right|_{L^\infty}\mu_G(A').$$

Therefore, there is a $\gamma \in \hat{G}$ such that

$$\frac{1}{4}\alpha'\alpha''\mu_G(\Lambda'') \leq |(\hat{1}_{2A''} - \alpha''\hat{1}_{2\Lambda''})(\gamma)|.$$

For any $\gamma \in \Gamma$, the function $\gamma'(x) := \gamma(2x)$ is also a homomorphism. Thus, for any set $S \subset G$, we then have $\hat{1}_{2S}(\gamma) = \hat{1}_S(\gamma')$. This means that we can, without loss of generality, replace $\hat{1}_{2A''}$ and $\hat{1}_{2\Lambda''}$ with $\hat{1}_{A''}$ and $\hat{1}_{\Lambda''}$. Furthermore, recall that α' and α'' are at least $\alpha/2$ each. We have

$$\frac{1}{16}\alpha^2\mu_G(\Lambda'') \leq |(\hat{1}_{A''} - \alpha''\hat{1}_{\Lambda''})(\gamma)| = |\langle(1_{A''} - \alpha'')\bar{\gamma}, 1_{\Lambda''}\rangle| \tag{3.8}$$

This is the same form of expression as (2.3), except that now we are working with Bohr sets. We can apply the same type of argument. Let $\Gamma' = \Gamma \cup \{\gamma\}$ where Γ is the frequency set of Λ. and let Λ''' be the Bohr set with frequency set Γ' and width $\rho'''\rho''\epsilon$, where $\rho''' \leq \eta'''/200d$ for some $\eta''' > 0$ to be determined later. By Proposition 3.14, we can assume that Λ''' is regular, and by Proposition 3.13, Λ'' is η''' locally-like Λ'''. Hence, by setting $f := 1_{A''} - \alpha''$, we see from Lemma 3.7 that

$$\left(\frac{1}{16}\alpha^2 - \eta''' - \rho'''\right)\mu_G(\Lambda'')\mu_G(\Lambda''') \leq \langle|f * 1_{\Lambda'''}|, 1_{\Lambda''}\rangle.$$

We now apply the same trick as in the proof of Lemma 2.12 to remove the absolute value on $f * 1_{\Lambda'''}$. Since

$$f * 1_{\Lambda'''}(x) = \frac{|A'' \cap (x - \Lambda''')| - \alpha''|\Lambda'''|}{|G|},$$

Lemma 3.5 says that

$$|\langle f * 1_{\Lambda'''}, 1_{\Lambda''} \rangle| \leq \eta''' \mu_G(\Lambda'')\mu_G(\Lambda''').$$

As before, let B be the set of all $x \in \Lambda''$ such that $f * 1_{\Lambda'''}(x) \geq 0$. Then

$$2\langle f * 1_{\Lambda'''}, 1_B \rangle = \langle |f * 1_{\Lambda'''}|, 1_{\Lambda''} \rangle + \langle f * 1_{\Lambda'''}, 1_{\Lambda''} \rangle \geq \left(\frac{1}{16}\alpha^2 - 2\eta''' - \rho''' \right) \mu_G(\Lambda'')\mu_G(\Lambda''').$$

Finally, we will set $\eta''' = \alpha^2/128$. Then

$$\frac{1}{32}\alpha^2 \mu_G(\Lambda'')\mu_G(\Lambda''') \leq \langle f * 1_{\Lambda'''}, 1_B \rangle.$$

This implies that for some $x \in B$,

$$\frac{1}{32}\alpha^2 \mu_G(\Lambda''') \leq f * 1_{\Lambda'''}(x) = 1_{A''} * 1_{\Lambda'''}(x) - \alpha'' s \mu_G(\Lambda''').$$

Since $\alpha'' > \alpha - 4\eta'$, if $\eta' \leq 2^{-8}\alpha^2$, then we have the density increment

$$\left(\alpha + \frac{1}{64}\alpha^2 \right) \mu_G(\Lambda''') \leq f * 1_{\Lambda'''}(x) = 1_{A''} * 1_{\Lambda'''}(x).$$

3.3.2. Density increment from high average of $\hat{1}_{A'} - \alpha'\hat{1}_{\Lambda'}(\gamma)$ on a spectrum. We now deal with the case of (3.7). We will first perform some reductions to isolate Fourier transforms of the $1_{A'} - \alpha' 1_{\Lambda'}$ term.

Lemma 3.19. *Let $B \subset G$, and let $A \subset B$ be of density α in B. If $f, g \in C(G)$ and*

$$\nu \mu_G(A)^{1/2} \, ||g||_{L^1} \, ||f||_{L^2} \leq |\langle f * g, 1_A \rangle - \langle f * g, \alpha 1_B \rangle|, \tag{3.9}$$

for some $\nu > 0$, then for any $\kappa > 0$, we have

$$(\nu^2 - \kappa^2)\mu_G(A) \leq ||(\hat{1}_A - \alpha \hat{1}_B) 1_{\mathrm{Spec}_\kappa(g)}||_{L^2}^2. \tag{3.10}$$

This lemma picks out Fourier transforms $\hat{1}_{A'} - \alpha'\hat{1}_{\Lambda'}(\gamma)$ as we did in the previous case, except that now instead of just one transform, we consider transforms over a spectrum. The proof is essentially a series of basic manipulation of inequalities to try to isolate the Fourier transform.

Proof: By the Cauchy-Schwarz inequality, followed by Parseval's identity, we have

$$\begin{aligned}
\langle 1_A, f * g\rangle - \langle \alpha 1_B, f * g\rangle &= \langle (\hat{1}_A - \alpha\hat{1}_B), \hat{f}\hat{g}\rangle \\
&= \sum_{\gamma\in\hat{G}} (1_A - \alpha 1_B)\hat{\ }(\gamma)\overline{\hat{f}(\gamma)\hat{g}(\gamma)} \\
&\leq \left(\sum_{\gamma\in\hat{G}} |(1_A - \alpha 1_B)\hat{\ }(\gamma)|^2|\hat{g}(\gamma)|^2\right)^{1/2} \left(\sum_{\gamma\in\hat{G}} |\hat{f}(\gamma)|^2\right)^{1/2} \\
&= \left(\sum_{\gamma\in\hat{G}} |(1_A - \alpha 1_B)\hat{\ }(\gamma)|^2|\hat{g}(\gamma)|^2\right)^{1/2} ||f||_{L^2}.
\end{aligned}$$

Hence

$$\nu^2\mu_G(A)\,||g||_{L^1}^2 \leq \sum_{\gamma\in\hat{G}} |(1_A - \alpha 1_B)\hat{\ }(\gamma)|^2|\hat{g}(\gamma)|^2.$$

The contribution of $\gamma \notin \mathrm{Spec}_\kappa(g)$ on the sum is

$$\begin{aligned}
\sum_{y\notin \mathrm{Spec}_\kappa(g)} |(1_A - \alpha 1_B)\hat{\ }(\gamma)|^2|\hat{g}(\gamma)|^2 &\leq \kappa^2\,||g||_{L^1}^2 \sum_{\gamma\in\hat{G}} |(1_A - \alpha 1_B)\hat{\ }(\gamma)|^2 \\
&= \kappa^2\,||g||_{L^1}^2\,\frac{1}{N}\sum_{x\in G} |(1_A - \alpha 1_B)(\gamma)|^2 \\
&= \kappa^2\,||g||_{L^1}^2\,\alpha(1-\alpha)\mu_G(B) \\
&\leq \kappa^2\,||g||_{L^1}^2\,\mu_G(A).
\end{aligned}$$

Thus,

$$||g||_{L^1}^2\,\mu_G(A)(\nu^2 - \kappa^2) \leq \sum_{\gamma\in \mathrm{Spec}_\kappa(g)} |(1_A - \alpha 1_B)\hat{\ }(\gamma)|^2|\hat{g}(\gamma)|^2.$$

Now, $|\hat{g}(\gamma)| \leq ||g||_{L^1}$ for any $\gamma \in \hat{G}$, so cancelling these gives the desired expression. $\square$

We now apply Lemma 3.19. Recall that the expression we have is

$$\frac{1}{4}\alpha'^2\mu_G(\Lambda')\mu_G(\Lambda'') \leq |\langle 1_{A'} * (1_{A'} - \alpha' 1_{\Lambda'}), 1_{2\Lambda''}\rangle|.$$

Applying Lemma 3.19 with $f = 1_{A'}$, $g = 1_{2\Lambda''}$, $\nu = \alpha'/4$ and $\kappa = \alpha'/16$, we obtain

$$\frac{1}{32}\alpha'^3\mu_G(\Lambda') \leq \sum_{\gamma\in \mathrm{Spec}_\kappa(1_{2\Lambda''})} |(\hat{1}_{A'} - \alpha'\hat{1}_{\Lambda'})(\gamma)|^2 \tag{3.11}$$

Thus, as in Section 3.3.1, we have isolated the $\hat{1}_{A'} - \alpha'\hat{1}_{\Lambda'}$ term, but now we are examining more than just one value of γ. Therefore, we wish to find a Bohr set which annihilates the entire spectrum.

The key is an argument of the type in Lemma 3.16. Let $\rho'''' < 1/100d$ to be chosen later, and set $\Lambda'''' = \Lambda''_{\rho''''}$. We may suppose that Λ'''' is regular. Then by the same argument as in the proof of Lemma 3.16, we have

$$\mathrm{Spec}_\kappa 1_{2\Lambda''} \subset \left\{ \gamma \in \hat{G} : |1-\gamma(x)| \leq 3200 d\rho''''\alpha^{-1} \ \forall x \in 2\Lambda'''' \right\}.$$

Suppose $\rho'''' \leq 2^{-13}\alpha'/d$ so that $|1-\gamma(x)| \leq 1/2$ in the above set. Thus, $2\Lambda''''$ 1/2-annihilates the spectrum. We can now use an argument similar to Lemma 3.7 where we transfer from a sum over $2\Lambda''$ to $2\Lambda''''$ for each term in the sum of (3.11), but we will take a more elegant approach here, which completes the transfer in one step by Parseval's identity. Both methods, however, are essentially doing the same thing, and hence will produce the same density increment.

Lemma 3.20 (Density increment from large average on annihilated characters)**.** *Let $B \subset G$. Let $A \subset B$ be of density α in B. Let $D \subset G$ be such that B is η locally-like both D and $-D$. If $\Theta \subset \hat{G}$ is a set of characters which are 1/2-annihilated by D and*

$$\kappa\mu_G(A) \leq ||(\hat{1}_A - \alpha\hat{1}_B)1_\Theta||_{L^2}^2, \tag{3.12}$$

then there exists $x \in G$ such that $x - A$ has density at least $\alpha + \kappa/4 - 6\eta$ on D.

Proof: For any $\theta \in \Theta$, as in Lemma 3.16, we have,

$$|\hat{1}_D(\theta)| = \frac{1}{|G|}\left|\sum_{x\in D}\bar{\theta}(x)\right| = \frac{1}{|G|}\left|\sum_{x\in D}1 + \sum_{x\in D}\bar{\theta}(x) - 1\right| \geq \frac{\mu_G(D)}{2}$$

since D 1/2-annihilates Θ. In particular,

$$\frac{\mu_G(D)}{2}1_\Theta \leq \hat{1}_D.$$

Thus, by Parseval's identity, we have

$$\kappa\frac{\mu_G(A)\mu_G(D)^2}{4} \leq ||(\hat{1}_A - \alpha\hat{1}_B)\hat{1}_D||_{L^2}^2 = ||((1_A - \alpha 1_B) * 1_D)\hat{\ }||_{L^2}^2 = ||(1_A - \alpha 1_B) * 1_D||_{L^2}^2. \tag{3.13}$$

Now, by applying Lemma 3.4 twice, we have

$$\begin{aligned}
\langle 1_A * 1_D, 1_B * 1_D\rangle &= \langle 1_A * 1_D * 1_{-D}, 1_B\rangle \\
&= \mu_G(D)\langle 1_A * 1_D, 1_B\rangle + O_1(\eta\mu_G(D)^2\mu_G(B)) \\
&= \mu_G(D)(\mu_G(D)\langle 1_A, 1_B\rangle + O_1(\eta\mu_G(D)\mu_G(B)) + O_1(\eta\mu_G(D)^2\mu_G(B)) \\
&= \alpha\mu_G(D)^2\mu_G(B) + O_1(2\eta\mu_G(D)^2\mu_G(B)).
\end{aligned}$$

where we have used O_1 to indicate that the implied constant is 1. Similarly, we can show that

$$\langle 1_B * 1_D, 1_B * 1_D\rangle = \mu_G(D)^2\mu_G(B) + O_1(2\eta\mu_G(D)^2\mu_G(B)).$$

Expanding the L^2 norm in (3.13) and inserting the above values for the inner product gives

$$\frac{\mu_G(A)\mu_G(D)^2}{4}\kappa \le ||1_A * 1_D||_{L^2}^2 - \alpha^2\mu_G(D)^2\mu_G(B) + O_1(6\alpha\eta\mu_G(D)^2\mu_G(B))$$

Thus,

$$\begin{aligned} \mu_G(A)\mu_G(D)^2(\alpha + \kappa/4 - 6\eta) &\le ||1_A * 1_D||_{L^2}^2 \\ &\le ||1_A * 1_D||_{L^\infty} ||1_A * 1_D||_{L^1} \\ &= ||1_A * 1_D||_{L^\infty} \mu_G(A)\mu_G(D). \end{aligned}$$

This proves the claim. □

We will apply this lemma with $D = 2\Lambda''''$. Recall that $\pm 2\Lambda'''' \subset 2\Lambda'' \subset \Lambda_{2\rho''}$, so Λ' is η'' locally-like $\pm 2\Lambda''''$ by Proposition 3.13. Since $2\Lambda''''$ 1/2-annihilates $\mathrm{Spec}_\kappa(1_{2\Lambda''})$, the lemma applied on (3.11) says that there is a $x \in G$ such that

$$\left(\alpha' + \frac{\alpha'^2}{128} - 6\eta''\right)\mu_G(\Lambda'''') \le (1_{A'} * 1_{2\Lambda''''})(x).$$

Suppose $\eta'' \le 2^{-12}\alpha^2$. Also, recall that $\alpha' \ge \alpha - 4\eta'$, so supposing that $\eta' \le 2^{-10}\alpha^2$, we obtain

$$\left(\alpha + \frac{\alpha^2}{256}\right)\mu_G(\Lambda'''') \le (1_{A'} * 1_{2\Lambda''''})(x),$$

which is the desired density increment. We note that $2\Lambda''''$ is not a Bohr set, but since Λ'''' is, we can just use Λ'''' and $A'''' := \{y \in \Lambda'''' : 2y \in (x - A') \cap 2\Lambda''''\}$. Also note that $|2\Lambda''''| = |\Lambda''''|$ and $|A''''| = |(x - A) \cap 2\Lambda''''|$ by Lemma 3.18.

3.3.3. Completing the proof of Lemma 3.18. It remains to choose the values of η' and η'', and hence $c_1 := \rho$, $c_2 = \rho''\rho'$, $c_3 = \rho'''\rho''\rho'$, and $c_4 = \rho''''\rho''\rho'$ that will allow all the argument to work. We saw that it suffices to take $\eta' \le 2^{-10}\alpha^2$. Hence we can choose

$$c_1 \in [2^{-19}\alpha^2 d^{-1}, 2^{-18}\alpha^2 d^{-1}],$$

such that $\Lambda' = \Lambda_{c_1}$ is regular. It also sufficed to take $\eta'' \le 2^{-12}\alpha^2$ and $\rho'' \le 2^{-9}\eta'' d^{-1}$, so we can choose

$$c_2 \in [2^{-41}\alpha^4 d^{-2}, 2^{-40}\alpha^4 d^{-2}],$$

such that $\Lambda'' = \Lambda_{c_2}$ is regular. Since $\eta''' = 2^{-6}\alpha^2$ and it sufficed to take $\rho''' \le \eta''' 2^{-9} d^{-1}$, we can choose

$$c_3 \in [2^{-57}\alpha^6 d^{-3}, 2^{-56}\alpha^6 d^{-3}],$$

such that $\Lambda''' = \Lambda_{c_3}$ is regular. Lastly, since it sufficed to take $\rho'''' \le 2^{-15}\alpha^2 d^{-1}$ (again since $\alpha/2 \le \alpha'$), we can take

$$c_4 \in [2^{-57}\alpha^6 d^{-3}, 2^{-56}\alpha^6 d^{-3}],$$

so that $\Lambda'''' = \Lambda_{c_4}$ is regular. The smallest density increment that we found possible was $\alpha + 2^{-10}\alpha^2$, which occurred right at the beginning on either Λ' or Λ''. On the other hand, the smallest Bohr set on which we found the density increment is either Λ''' or Λ''''. Also, returning to the size condition (3.5), we have

$$|G| \geq \frac{16}{\alpha^2}\left(\frac{\rho'\epsilon}{3\pi}\right)^{-d}.$$

This indeed is satisfied if

$$|G| \geq (2^{-1}\alpha d^{-1}\epsilon)^{-25d}.$$

This completes the proof of Lemma 3.17. □

3.3.4. Completing the proof of Theorem 3.2. Theorem 3.2 follows from Lemma 3.17 in the same way as Theorem 2.1 did from Lemma 2.12 in Chapter 2, Section 2.2.

Proof of Theorem 3.2: We will let Λ_0 be the Bohr set with frequency set $\{\gamma\}$ for some $\gamma \in \hat{G}$ and width 2 so that $\Lambda_0 = G$. As in the proof of Theorem 2.1, we will suppose that the failure of the size condition i) does not occur, and this will ultimately yield the condition on the density. If ii) occurs, then we are done, so let us suppose iii) occurs. We can then reapply Lemma 3.17 with Λ' and $x - A$. We itertate Lemma 3.17 in this manner. Let Λ_i be the Bohr set at the ith iteration, let d_i be the dimension of Λ_i, ϵ_i the width of Λ_i, and α_i the density of (a translate of) A on Λ_i. The density increment is always at least $\alpha_i + c_1\alpha_i^2$ for some fixed c_1 at the ith iteration, which is of the same form as Lemma 2.12. So by the same dyadic decomposition as in the proof of Theorem 2.1 in Chapter 2 Section 2.2, we can only iterate Lemma 3.17 at most $i_0 = c_2/\alpha$ times for some absolute constant $c_2 > 0$. This means that ii) must occur somewhere during the iteration, granted that the size condition i) is always met.

Thus, we only need to ensure that the size condition is satsified at each step i.e.

$$|G| \geq (2\alpha_i^{-1}d_i\epsilon_i^{-1})^{25d_i},$$

for each $i \leq i_0$. Now, α_i is smallest at $i = 0$, d_i is largest at $i = i_0$ and d_{i_0} is at most i_0, and ϵ_i is smallest at $i = i_0$. Therefore, it suffices to satisfy

$$|G| \geq (2\alpha^{-1}i_0\epsilon_{i_0}^{-1})^{20i_0}, \tag{3.14}$$

or a little more crudely, for some absolute constant $c_3 > 0$,

$$|G| \geq (c_3\alpha^{-2}\epsilon_{i_0}^{-1})^{c_3\alpha^{-1}}. \tag{3.15}$$

We require a lower bound of ϵ_{i_0}. By Lemma 3.17, we have

$$\epsilon_{i+1} \geq (2^{-57}\alpha_i^6 d_i^{-3})\epsilon_i \geq (2\alpha_i^{-1}i)^{-60}\epsilon_i.$$

This suggests that we can have the explicit lower bound

$$\epsilon_i \geq (2\alpha^{-1}i)^{-60i}\epsilon. \tag{3.16}$$

These estimates are very rough, but it will suffice since what will be decisive is i in the exponent (hence the number of iterations) rather than the terms below it. The above explicit bound can be easily proved by induction. Indeed, by Lemma 3.17, we have

$$\epsilon_{i+1} \geq (2^{-57}\alpha_i^6 i^{-3})\epsilon_i \geq (2^{-57}\alpha^6 i^{-3})(2\alpha^{-1}i)^{-60i}\epsilon \geq (2\alpha^{-1}(i+1))^{-60(i+1)}\epsilon.$$

Hence

$$\epsilon_{i_0} \geq (2c_2\alpha^{-2})^{-60c_2\alpha^{-1}}.$$

This means that there is some absolute $c_4 > 0$ such that (3.15) is satisfied if

$$|G| \geq (c_4\delta)^{c_4\delta},$$

where we have set $\delta := \alpha^{-2}$. Taking logarithms, we obtain

$$\log|G| \geq c_4\delta\log(c_4\delta).$$

This implies that we can take $\delta = \log|G|/c_4\log\log|G|$. The claim now follows. □

CHAPTER 4

Sanders's Improvement

In 2010, Sanders improved Bourgain's result to the following in [**17**].

Theorem 4.1 (Sanders)**.** *There exists $C > 0$ such that if $A \subset [N]$ is of density α in $[N]$ with*

$$\alpha > C\frac{(\log\log N)^5}{\log N},$$

then A contains a non-trivial 3AP.

Again, as in the previous chapter, this theorem follows from the more general form below.

Theorem 4.2 (Sanders)**.** *Let G be a group of odd order. Then there exists an absolute $C > 0$ such that if $A \subset G$ is of density α in G with*

$$\alpha > C\frac{(\log\log |G|)^5}{\log |G|},$$

then A contains a non-trivial 3AP.

Theorem 4.1 is a straightforward consequence of Theorem 4.2 by Lemma 2.11. Also, as in Theorem 3.2, the reason we require G to be of odd order is due to Lemma 3.18.

In Chapter 3, Section 3.3.2, we see a new way of obtaining a density increment. It is as follows. Suppose Λ and Λ' are Bohr sets where Λ is locally-like Λ', and let $A \subset \Lambda$ be of density α in Λ. We consider the case when

$$\epsilon = |\langle f * g, 1_A\rangle - \langle f * g, \alpha 1_\Lambda\rangle| \tag{4.1}$$

is large, with $g = 1_{2\Lambda'}$ and $f = 1_A$. By Lemma 3.19, we can convert this into an expression of the form

$$||(\hat{1}_A - \alpha\hat{1}_\Lambda)1_{\mathrm{Spec}_\eta(g)}||_{L^2}^2,$$

and obtain a density increment from here by Lemma 3.20. Historically, this method of using the L^2 norm goes back to Heath-Brown in [**13**] and Szemerédi in [**22**]. However, if we review the argument in the proof, we see that it works for any $f \in C(G)$. Sanders exploits this idea and approaches the problem in a different way. The starting point of both Roth's and Bourgain's methods is to compute the number of 3APs in A directly, as discussed in Section 2.2.3. Sanders, however begins with (4.1) with f not necessarily related to A directly. In fact, Sanders generalizes this method so that we can take g to be the characteristic function of any subset of Λ'. With this tool, he computes the number of 3APs in A in smaller incremental steps rather than in one sweep as in the previous methods.

4.1. Density increment from high energy on a spectrum

The focus of this section is to generalize the method used in Chapter 3, Section 3.3.2 to obtain a density increment. All the results in this section are due to Sanders in [**17**]. Let $B, C \subset G$, and let $A \subset B$ be of density α in A. Let $f \in C(G)$. Our starting point is an expression of the form

$$\nu \mu_G(A)^{1/2} \mu_G(C) \,||f||_{L^2} \leq |\langle f * 1_C, 1_A \rangle - \langle f * 1_C, \alpha 1_B \rangle| \tag{4.2}$$

for some $\nu > 0$. By Lemma 3.19, this immediately implies that for any $\kappa > 0$ we have

$$(\nu^2 - \kappa^2)\mu_G(A) \leq ||(\hat{1}_A - \alpha \hat{1}_B) 1_{\mathrm{Spec}_\kappa(1_C)}||_{L^2}^2. \tag{4.3}$$

We would like to obtain a density increment from this by using Lemma 3.20. The only obstacle here is that we do not know a suitable set that will annihilate the spectrum $\mathrm{Spec}_\kappa(1_C)$. When C is a Bohr set, this is simple because by Lemma 3.16, a Bohr set can annihilate its own spectrum. In this section, we will consider the case when C is a subset of a Bohr set, or more precisely, we will prove the following lemma.

Lemma 4.3 (Sanders)**.** *Let Λ be a regular Bohr set with dimension d and width ϵ. Let $T \subset \Lambda$ be of density τ in Λ. Then for any $\kappa > 0$, there exists a Bohr set $\Lambda' \subset \Lambda$ with dimension at most $d + O(\kappa^{-2} \log 2\tau^{-1})$ and width $\rho\epsilon$ where $\rho = \Omega(\kappa^2/d^2 \log 2\tau^{-1})$ such that Λ' 1/2-annihilates* $\mathrm{Spec}_\kappa(1_T)$.

From this lemma, the following corollary follows immediately.

Corollary 4.4. *Let Λ be a regular Bohr set with dimension d. Let $\Lambda' \subset \Lambda_{\rho'}$ for some $\rho' \leq 1/100d$ be a regular Bohr set of dimension d' and width ϵ'. Let $A \subset \Lambda$ and $T \subset \Lambda'$ have density α in Λ and τ in Λ' respectively. If, for some $f \in C(G)$ and $\nu > 0$, we have*

$$\nu \mu_G(A)^{1/2} \mu_G(T) ||f||_{L^2} \leq |\langle f * 1_T, 1_A \rangle - \langle f * 1_T, \alpha 1_\Lambda \rangle|,$$

then for any $\kappa > 0$, there exists a regular Bohr set Λ''' of dimension at most

$$d' + O(\kappa^{-2} \log 2\tau^{-1})$$

and width at least $\rho''\epsilon'$ where

$$\rho'' = \Omega\left(\frac{\kappa^2}{d'^2 \log 2\tau^{-1}}\right),$$

such that for some $x \in G$, $x - A$ has density at least $\alpha + (\nu^2 - \kappa^2)/2 - 800d\rho'$ on Λ'''.

Proof: We begin at (4.3) with $B = \Lambda$ and $C = T$. Then by Lemma 4.3, there is a Bohr set $\Lambda'' \subset \Lambda'$ of dimension at most $d' + O(\kappa^{-2} \log 2\tau^{-1})$ and width $\rho''\epsilon$ where $\rho'' = \Omega(\kappa^2/d'^2 \log 2\tau^{-1})$ such that Λ'' 1/2 annihilates $\mathrm{Spec}_\kappa(1_T)$. By Proposition 3.13, Λ is $200d\rho'$ locally-like Λ'' since $\Lambda'' \subset \Lambda'$. The claim follows by Lemma 3.20. □

4.1.1. Annihilating a spectrum. The main task in this section is to prove Lemma 4.3. We first observe that if we are just interested in 1/2 annihilating $\mathrm{Spec}_\kappa(1_T)$ with a Bohr set, this is pretty easy to do. For example, we can simply let Λ' be the Bohr set with frequency set $\mathrm{Spec}_\kappa(1_T)$ and width 1/2. The problem with this is that we do not know how large $\mathrm{Spec}_\kappa(1_T)$ is, which gives us little control over this Bohr set. Instead, we will try to 'decompose' the spectrum. It will be useful to have the following concept.

Definition 4.5. *Let $\Delta \subset \hat{G}$.* SpanΔ *is the set of all linear combinations $\delta_1 \cdots \delta_k$ where $k \leq |\Delta|$, and $\delta_i \in \Delta \cup \bar{\Delta}$ for $i = 1, \dots, k$. Here $\bar{\Delta}$ is the set of all conjugates of elements in Δ.*

By Lemma 3.16, we know how Λ and $\mathrm{Spec}_c(1_\Lambda)$ are related for any $c > 0$. The motivating idea is that since $T \subset \Lambda$, $\mathrm{Spec}_\kappa(1_T)$ cannot be too far from $\mathrm{Spec}_c(1_\Lambda)$. Now, since $\mathrm{Spec}_c(1_{\Lambda_\rho})$ increases as $\rho \to 0$, we could decrease ρ until it contains $\mathrm{Spec}_\kappa(1_T)$, but this is not very efficient. What we will do, on the other hand, is to try to find $\Theta \subset \hat{G}$ and ρ such that the product set $\mathrm{Spec}_c(1_{\Lambda_\rho}) \cdot \mathrm{Span}\Theta$ contains $\mathrm{Spec}_\kappa(1_T)$. That is, we will use SpanΘ to widen our range. Then we can set Λ' to be a Bohr set with frequency set Θ combined with that of Λ.

Now, in seeking for a good Θ, suppose $\gamma \in \mathrm{Spec}_\kappa(1_T)$ and $\theta \in \hat{G}$. If $|\hat{1}_{\Lambda_\rho}(\theta)|$ is small, then $|\hat{1}_T(\gamma\theta)|$ is probably not much smaller than $|\hat{1}_T(\gamma)|$; hence, it is likely that $\gamma\theta$ is in $\mathrm{Spec}_\kappa(1_T)$ as well. Thus, we want to choose Θ so that $\hat{1}_{\Lambda_\rho}(\theta)$ is small for all $\theta \in \mathrm{Span}\Theta$. The following definition builds on this idea to find such Θ.

Definition 4.6. *Let $\Delta \subset \hat{G}$ and $\eta > 0$. Let $\mu \in M(G)$. We let $\mathbb{D}$ denote the unit disk in $\mathbb{C}$. Then Δ is said to be η dissociated with respect to μ if for any function $w : \Delta \to \mathbb{D}$ and*

$$p_{w,\Delta}(x) := \prod_{\delta \in \Delta} (1 + \Re(w(\delta)\delta(x))),$$

we have

$$\sum_{x \in G} p_{w,\Delta}(x)\mu(x) \leq e^\eta.$$

Δ is said to have (η, μ)-entropy k if k is the cardinality of the largest $\Theta \subset \Delta$ such that Θ is η dissociated with respect to μ.

One of the main properties of $p_{w,\Delta}$ in the definition of dissociation is the following. This property will ultimately allow us to extract the span that we discussed above.

Lemma 4.7. *Let $\Delta \subset \hat{G}$ and let $w : \Delta \to \mathbb{D}$. Then $\hat{p}_{w,\Delta}(\chi) \neq 0$ only if $\chi \in$* SpanΔ.

Proof: We recall that

$$p_{w,\Delta} = \prod_{\delta \in \Delta} (1 + \Re(w(\delta)\delta)) = \sum_{l \leq |\Delta|} \sum_{\substack{\Xi \subset \Delta \\ |\Xi| = l}} \prod_{\xi \in \Xi} \Re(w(\xi)\xi).$$

Suppose Ξ consists of $\xi_1, \ldots, \xi_l$, and let $z_i = w(\xi_i)\xi_i$ for each $i = 1, \ldots, l$. Then

$$\prod_{\xi \in \Xi} \Re(w(\xi)\xi) = \frac{1}{2^l} \sum_{\substack{z'_i \in \{z_i, \bar{z}_i\} \\ i=1,\ldots,l}} \prod_{i=1}^{l} z'_i.$$

Thus,

$$p_{w,\Delta} = \sum_{l \leq |\Delta|} \sum_{\substack{\Xi \subset \Delta \\ |\Xi| = l}} \frac{1}{2^l} \sum_{\substack{z'_i \in \{z_i, \bar{z}_i\} \\ i=1,\ldots,l}} \prod_{i=1}^{l} z'_i.$$

Therefore,

$$\hat{p}_{w,\Delta}(\chi) = \frac{1}{|G|} \sum_{l \leq |\Delta|} \sum_{\substack{\Xi \subset \Delta \\ |\Xi| = l}} \sum_{\substack{z'_i \in \{z_i, \bar{z}_i\} \\ i=1,\ldots,l}} \frac{1}{2^l} \sum_{x \in G} \left(\chi \prod_{i=1}^{l} z'_i \right)(x).$$

The inner sum is always zero by the orthogonality property of Proposition 2.2 provided that $\chi \notin \mathrm{Span}\Delta$. $\square$

Recall that our task is to determine a Bohr set Λ' that 1/2-annihilates $\mathrm{Spec}_\kappa(1_T)$ where $T \subset \Lambda$ with density τ. We have two problems to solve: 1) determine the entropy of $\mathrm{Spec}_\kappa(1_T)$ (with respect to some measure μ), and 2) given the entropy, find Λ'. We begin with the latter in the next section, as this will show how the definition is relevant for annihilation.

4.1.2. A Bohr set annihilating a set of characters. Given a set of characters Δ and its entropy with respect to some measure, we will show in this section how to find a Bohr set which 1/2-annihilates Δ.

Lemma 4.8. *Let Λ be a regular Bohr set of dimension d. Let $\Delta \subset \hat{G}$ be a set of characters with (η, μ) entropy k where $\mu := 1_\Lambda/|\Lambda|$. Then there is a $\Theta \subset \Delta$ of cardinality at most k such that for any $\nu > 0$, $\rho \ll 1/(\log(\eta^{-1}6^k)d)$ and $\rho' > 0$, we have*

$$|1 - \gamma(x)| \ll k\nu + \rho' d,$$

for all $x \in \Lambda_{\rho\rho'} \cap \Lambda'$ and $\gamma \in \Delta$, where Λ' is the Bohr set with frequency set Θ and width ν.

We can make the intersection $\Lambda_{\rho\rho'} \cap \Lambda'$ into a Bohr set by combining the frequency sets of $\Lambda_{\rho\rho'}$ and Λ' into one. This leads to the following corollary.

Corollary 4.9. *Let Λ be a regular Bohr set of dimension d and width ϵ. Let $\Delta \subset \hat{G}$ be a set of characters with (η, μ) entropy k, where $\mu := 1_\Lambda/|\Lambda|$. Then there exists a regular Bohr set $\Lambda' \subset \Lambda$ with dimension at most $d + k$ and width $\epsilon' = \Omega(\epsilon/(\log(\eta^{-1}6^k)d^2))$ such that for all $\gamma \in \Delta$ and $x \in \Lambda'$, we have*

$$|1 - \gamma(x)| \leq \frac{1}{2}.$$

We will now prove Lemma 4.8.

Proof: Let L be a positive integer to be chosen later. Let $\rho > 0$ be a real number such that $\rho \ll 1/Ld$, $\Lambda_{1+L\rho}$ is regular, and

$$|\Lambda_{1+L\rho}| \leq (1+\eta/3)|\Lambda|.$$

We set

$$f = \frac{\mu_G(\Lambda_\rho)^{-L}}{|\Lambda_{1+L\rho}|}(1_{\Lambda_{1+L\rho}} * 1_{\Lambda_\rho} * \cdots * 1_{\Lambda_\rho}),$$

where the convolution contains L copies of 1_{Λ_ρ}. Then notice that for any $x \in B$, we have that $x - (\Lambda_\rho + \cdots + \Lambda_\rho) \subset \Lambda_{1+L\rho}$ (where the sum has again L copies), so

$$f(x) = \frac{1}{|\Lambda_{1+L\rho}|}.$$

Therefore,

$$\mu \leq \frac{|\Lambda_{1+L\rho}|}{|\Lambda|} f \leq \left(1 + \frac{\eta}{3}\right) f.$$

In particular, suppose that $\Theta \subset \hat{G}$ is $\eta/2$ dissociated with respect to f. Then for any $w : \Theta \to \mathbb{D}$, we have

$$\sum_{x \in G} p_{w,\Theta}(x) f(x) \leq e^{\eta/2},$$

which implies that

$$\sum_{x \in G} p_{w,\Theta}(x)\mu(x) \leq \left(1 + \frac{\eta}{3}\right) e^{\eta/2} \leq e^{\eta}.$$

Hence Θ is η dissociated with respect to μ. Therefore, Δ has $(\eta/2, f)$ entropy at most k. But we would like to be a bit more precise as follows. Let $\eta_0 := 0$ and $\Theta_0 = \emptyset$. For $i > 0$, let $\eta_i := i\eta/(2(k+1))$. If there exists $\gamma \in \Delta \setminus \Delta_{i-1}$ such that $\Delta_{i-1} \cup \{\gamma\}$ is η_i dissociated with respect to f, then set $\Delta_i := \Delta_{i-1} \cup \{\gamma\}$. This must terminate at some $i < k+1$, say i_0. Let $\Theta = \Delta_{i_0}$. With this Θ, we claim the following.

Claim: If $L = \lceil \log 2 \cdot 6^k (k+1)\eta^{-1} \rceil$, then for any $\gamma \in \Delta$, there exists $\chi \in \mathrm{Span}\Theta \cup \{1_G\}$ such that

$$|\hat{1}_{\Lambda_\rho}(\gamma\chi)| \geq \frac{1}{2}.$$

Proof of Claim: If $\gamma \in \Theta$, then we can just take $\chi = \bar{\gamma}$ and we are done, so let us assume otherwise. Then by our choice of Θ, there exists some $c \in \mathbb{D}$ and $w : \Theta \to \mathbb{D}$ such that

$$e^{\eta_{i_0+1}} < \sum_{x \in G} p_{w,\Theta}(x)(1 + \Re(c\gamma(x))) f(x).$$

On the other hand, since Θ is η_{i_0} dissociated, we have

$$\sum_{x \in G} p_{w,\Theta}(x)(1 + \Re c\gamma(x)) f(x) \leq e^{\eta_{i_0}} + \sum_{x \in G} p_{w,\Theta}(x)\Re(c\gamma(x)) f(x).$$

Hence, by combining these two inequalities, we get

$$\sum_{x\in G} p_{w,\Theta}(x)\Re(c\gamma(x))f(x) > e^{\eta_{i_0+1}} - e^{\eta_{i_0}} > \eta_{i_0+1} - \eta_{i_0} = \frac{\eta}{2(k+1)}.$$

Now, taking the Fourier transform, we obtain

$$\begin{aligned}\frac{\eta}{2(k+1)} &\leq \sum_{\chi\in\hat{G}} |\hat{p}_{w,\Theta}(\chi)||\hat{f}(\bar{\gamma}\bar{\chi})||G| \\ &= \mu_G(\Lambda_{1+L\rho})^{-1}\mu_G(\Lambda_\rho)^{-L}\sum_{\chi\in\hat{G}} |\hat{p}_{w,\Theta}(\chi)||\hat{1}_{\Lambda_{1+L\rho}}(\bar{\gamma}\bar{\chi})||\hat{1}_{\Lambda_\rho}(\bar{\gamma}\bar{\chi})|^L \\ &\leq \mu_G(\Lambda_\rho)^{-L}\sum_{\chi\in\hat{G}} |\hat{p}_{w,\Theta}(\chi)||\hat{1}_{\Lambda_\rho}(\bar{\gamma}\bar{\chi})|^L\end{aligned} \tag{4.4}$$

By Lemma 4.7, we see that $\hat{p}_{w,\Theta}(\chi) \neq 0$ only if $\chi \in \mathrm{Span}\Theta$. Now, this span has at most 3^k elements and $|\hat{p}_{w,\Theta}| \leq 2^k$, so we have

$$\frac{\eta}{2(k+1)} \leq \mu_G(\Lambda_\rho)^{-L}\sup_{\chi\in\mathrm{Span}\Theta}|\hat{1}_{\Lambda_\rho}(\bar{\gamma}\bar{\chi})|^L \sum_{\chi\in\mathrm{Span}\Theta}|\hat{p}_{w,\Theta}(\chi)| \leq \mu_G(\Lambda_\rho)^{-L}\sup_{\chi\in\mathrm{Span}\Theta}|\hat{1}_{\Lambda_\rho}(\bar{\gamma}\bar{\chi})|^L 6^k.$$

This gives

$$\left(\frac{\eta}{2\cdot(k+1)6^k}\right)^{1/L}\mu_G(\Lambda_\rho) \leq |\hat{1}_{\Lambda_\rho}(\bar{\gamma}\bar{\chi})| = |\hat{1}_{\Lambda_\rho}(\gamma\chi)| \tag{4.5}$$

for some $\chi \in \mathrm{Span}\Theta$. Thus, if

$$L = \lceil \log 2\cdot 6^k(k+1)\eta^{-1}\rceil,$$

the claim follows. □Claim

We now just need to verify that this Θ does the job. Let Λ' be the Bohr set with frequency set Θ and width $\nu > 0$. Let $\rho' > 0$. Let $\gamma \in \Delta$ and $x \in \Lambda_{\rho\rho'} \cap \Lambda'$. Then by what was proven above, there exists $\chi \in \mathrm{Span}\Theta \cup \{1_G\}$ such that

$$\frac{1}{2}\mu_G(\Lambda_\rho) \leq |\hat{1}_{\Lambda_\rho}(\gamma\chi)|.$$

Therefore, by Lemma 3.7, we have

$$|1-\gamma\chi(x)| \ll \rho' d.$$

Since $\chi = \chi_1\cdots\chi_l$ for some $l \leq k$ where each χ_i is an element or the conjugate of an element of Θ, we see that (with $\chi_0 := 1_G$ for notation's sake),

$$|1-\chi(x)| \leq \sum_{i=1}^{l} |\chi_1\cdots\chi_{i-1}(x) - \chi_i\cdots\chi_i(x)| = \sum_{i=1}^{l}|1-\chi_i(x)| \leq 2k\nu.$$

Therefore,

$$|1-\gamma(x)| \leq |1-\gamma\chi(x)| + |\gamma\chi(x)-\gamma(x)| \ll \rho' d + k\nu.$$

□

Remark It is worth commenting on the convolution technique here. If we use $f = 1_\Lambda$ andd carry out the same calculation, we would obtain

$$\left(\frac{\eta}{4 \cdot (k+1)6^k}\right) \mu_G(\Lambda_\rho) \leq |\hat{1}_{\Lambda_\rho}(\gamma\chi)|$$

in place of (4.5). Then in order to get the annihilation result of Corollary 4.9, we need ϵ' to be about $\epsilon\eta/(k6^k)d^2$, which is quite worse than that which was proven. The saving in our proof comes from the fact that convolving Λ_ρ L many times has a linear cost (we had to take $\rho < 1/Ld$), but a gain of a power of L as in (4.4).

4.1.3. Entropy of a spectrum. The remaining problem is to calculate the entropy of $\mathrm{Spec}_\kappa(1_T)$. The approach is to observe that for any $\Delta \subset \mathrm{Spec}_\kappa(1_T)$, by the definition of the spectrum we have

$$\kappa^2 ||1_T||_{L^1}^2 |\Delta| \leq \sum_{\delta \in \Delta} |\hat{1}_T(\delta)|^2.$$

Now, suppose Δ is 1-dissociated with respect to $\mu = 1_\Lambda/|\Lambda|$ (with Corollary 4.9 in mind, as we would like to result obtained in this subsection in order to invoke Corollary 4.9) and consider the map $V : L^k(\mu) \to L^2(\Delta) : f \mapsto \hat{fd\mu}|_\Delta$. We have

$$\kappa^2 ||1_T||_{L^1}^2 |\Delta| \leq ||V||^2 ||1_T||_{L^k}^2,$$

where $||V||$ is the operator norm of V, so

$$|\Delta| \leq \kappa^{-2} ||V||^2 ||1_T||_{L^k}^2 ||1_T||_{L^1}^{-2}.$$

This gives an upper bound on $|\Delta|$, and hence on the $(1, \mu)$ entropy of $\mathrm{Spec}_\kappa(1_T)$. The problem is to determine $||V||$. Since Δ is 1-dissociated with respect to μ, each $\hat{d\mu}(\delta)$ for $\delta \in \Delta$ is not too large. Thus, we would expect that neither is $||Vf||_2$. We will be needing the following lemma to help us compute $||V||$.

Lemma 4.10. *Suppose that Δ is K-dissociated with respect to $\mu \in M(G)$. Then for any $g \in C(\Delta)$, we have*

$$\sum_{x \in G} |\exp(\check{g}(x))| \mu(x) \leq \exp(K + ||g||_2^2/2).$$

Proof: The proof relies on the convexity of exp. First, observe that

$$|\exp(\check{g})| = \exp(\Re\check{g}) = \prod_{\lambda \in \Delta} \exp(\Re g(\delta)\delta).$$

Now, since $\exp(x)$ is convex in x, for any $t \in \mathbb{R}$ and $y \in [-1, 1]$, we have

$$e^{ty} \leq e^{-t} + (y+1)\frac{e^t - e^{-t}}{2} = \cosh(t) + y\sinh(t).$$

Therefore, with $t = |g(\delta)|$, we have

$$|\exp(\check{g})| \leq \prod_{\delta \in \Delta} \left(\cosh |g(\delta)| + \frac{\Re g(\delta)\delta}{|g(\delta)|} \sinh |g(\delta)| \right).$$

Hence,

$$\sum_{x \in G} |\exp(\check{g}(x))| \mu(x) \leq \prod_{\delta \in \Delta} \cosh |g(\delta)| \sum_{x \in G} \prod_{\delta \in \Delta} \left(1 + \Re \frac{g(\delta) \sinh |g(\delta)|}{|g(\delta)| \cosh |g(\delta)|} \delta(x) \right).$$

By setting

$$w(\delta) := \frac{g(\delta) \sinh |g(\delta)|}{|g(\delta)| \cosh |g(\delta)|},$$

we have

$$\sum_{x \in G} |\exp(\check{g}(x))| \mu(x) \leq \prod_{\delta \in \Delta} \cosh |g(\delta)| \sum_{x \in G} p_{w,\Theta} \mu(x) \leq e^K \prod_{\delta \in \Delta} \cosh |g(\delta)|.$$

Since $\cosh(t) \leq e^{t^2/2}$, the result follows. □

With this, we can now compute the entropy of the spectrum. The lemma below is more general than we need, but this brings forth the idea better.

Lemma 4.11. *Let $f \in L^2(\mu)$ and let $L_f = ||f||_{L^2(\mu)} ||f||_{L^1(\mu)}^{-1}$. Then $\mathrm{Spec}_\kappa(f, \mu)$ has $(1, \mu)$ entropy $O(\kappa^{-2} \log 2L_f)$.*

Proof: Let $\Delta \subset \mathrm{Spec}_\kappa(f, \mu)$. As noted earlier, we first observe that by the definition of the spectrum, we have

$$\kappa^2 \, ||f||_{L^1(\mu)}^2 \, |\Delta| \leq \sum_{\delta \in \Delta} |\hat{fd\mu}(\delta)|^2.$$

Now suppose Δ is also 1-dissociated with respect to μ. For some $k > 0$, consider the map $V : L^k(\mu) \to L^2(\Delta)$, defined by

$$V(f) = \hat{fd\mu}|_\Delta.$$

We then have

$$\kappa^2 \, ||f||_{L^1(\mu)}^2 \, |\Delta| \leq ||V||^2 ||f||_{L^k(\mu)}^2,$$

where $||V||$ is the operator norm of V. This gives us an upper bound for Δ in terms of f and κ. Now, if we set $k = 2l/(2l-1)$ for some positive integer l and apply Corollary 2.9 with $p_1 = 1$, $p_2 = 2$, $\theta_1 = 1 - 1/l$ and $\theta_2 = 1/l$, we obtain

$$\kappa^2 \, ||f||_{L^1(\mu)}^2 \, |\Delta| \leq ||V||^2 ||f||_{L^1(\mu)}^{2(1-1/l)} ||f||_{L^2(\mu)}^{2/l} = ||V||^2 ||f||_{L^1(\mu)}^2 L_f^{2/l}.$$

Setting $l = \lceil \log L_f \rceil$, we see that

$$|\Delta| \ll \kappa^{-2} ||V||^2. \tag{4.6}$$

As observed at the beginning of this subsection, this is an upper bound on the cardinality of any Δ which is 1-dissociated with respect to μ, and hence bounds the entropy of $\mathrm{Spec}_\kappa(f,\mu)$. It remains to determine $||V||$. First, we observe that for any $g \in C(\Delta)$ and $f \in C(G)$,

$$\begin{aligned}
\langle g, Vf\rangle &= \sum_{\delta\in\Delta} g(\delta)\overline{\left(\sum_{x\in G} f(x)\mu(x)\bar{\delta}(x)\right)} \\
&= \sum_{x\in G}\left(\sum_{\delta\in\Delta} g(\delta)\delta(x)\right)\bar{f}(x)\mu(x) \\
&= \langle \check{g}, f\rangle_\mu .
\end{aligned}$$

Now let $V^* : L^2(\Delta) \to L^{k'}(\mu)$ be defined by $V^*(g) = \check{g}$, where $k' = k/(k-1)$. Then by the extremal property, we see that

$$\begin{aligned}
||V|| &= \sup_{||f||_{L^k(\mu)}\leq 1} ||Vf||_{L^2} \\
&= \sup_{||f||_{L^k(\mu)}\leq 1}\ \sup_{||g||_{L^2}\leq 1} \langle Vf, g\rangle \\
&= \sup_{||f||_{L^k(\mu)}\leq 1}\ \sup_{||g||_{L^2}\leq 1} \langle f, V^*g\rangle \\
&= \sup_{||g||_{L^2}\leq 1} ||V^*g||_{L^{k'}(\mu)} \\
&= ||V^*||.
\end{aligned}$$

We will now determine $||V^*||$. We will only need the case when k' is a positive integer. Indeed, by our choice of k, we have $k' = 2l$. Let $g \in L^2(\Delta)$ and suppose that, to begin with a simple case, we have $||g||_2 = c\sqrt{k'}$, where $c > 0$ is some absolute constant to be determined soon. First we determine the values of x such that

$$e^x \geq \frac{|x|^{k'}}{k'!}.$$

This is equivalent to

$$x \geq k' \log|x| - \log k'! = k'\log|x| - \sum_{i=1}^{k'} \log i.$$

It is well-known that

$$\sum_{n=1}^{N} \log n = N\log N - N + O(\log N),$$

for any positive integer N (for instance, see [**1**, pg.76]). Therefore, it suffices to have

$$x \geq k'(\log|x| + 1 - \log k') + O(\log k').$$

From here, it is easy to see that it suffices to have

$$x \geq -c'k',$$

for some $c' > 0$. Let $c = c'$. In particular, since

$$|\check{g}| \leq ||g||_1 \leq ||g||_2,$$

we have that $|\check{g}| \geq -ck'$. Hence, by Corollary 2.10, we have

$$|\Re\check{g}|^{k'} \leq k'! \exp \Re\check{g}.$$

Therefore, by Lemma 4.10, we have

$$||\Re\check{g}||_{L^{k'}(\mu)} \leq (k'!)^{1/k'} \exp(1/k' + ||g||_{L^2}^2 /k') = O((k'!)^{1/k'}) = O(k') = O(\sqrt{k'}||g||_{L^2})c.$$

Furthermore, since $\Im\check{g} = \Re(-i\check{g})$, we have $||\check{g}||_{L^{k'}(\mu)} = O(\sqrt{k'}||g||_{L^2})$. Now, for any general $g \in L^2(\Delta)$, let $a = a(g)$ be the positive real such that $||ag||_2 = c\sqrt{k}$. Then

$$a||\check{g}||_{L^{k'}(\mu)} = ||a\check{g}||_{L^{k'}(\mu)} = O(\sqrt{k}\,||ag||_2) = aO(\sqrt{k'}\,||g||_{L^2}).$$

We have thus shown that $||V|| = ||V^*|| = O(\sqrt{k'})$. Recalling that $k' = 2l$, the result follows from (4.6). □

4.1.4. Proof of Lemma 4.3. The proof of the lemma is now straight forward. Note that $\mathrm{Spec}_\kappa(1_T) = \mathrm{Spec}_\kappa(1_T, \mu)$ since $T \subset \Lambda$. Now, by Lemma 4.11, $\mathrm{Spec}_\kappa(1_T, \mu)$ has $(1, \mu)$ entropy $k = O(\kappa^{-2} \log 2\tau^{-1})$. The lemma follows by applying Corollary 4.9. □

4.2. Proof of Sanders's theorem

With this general technique of obtaining a density increment, we can now discuss Sanders's proof of Theorem 4.1 in [**17**]. The method goes as follows: suppose that $A \subset \Lambda$ has density α. By using the techniques from Section 4.1, we will try to construct a very dense set $L \subset \Lambda$ (density $\Omega(1)$ in Λ) and a set $S \subset \Lambda$ such that for every x, we have

$$1_L * 1_{2S}(x) \ll 1_A * 1_{-2A}(x).$$

The construction of L is incremental, adding more elements at each step. Once this is complete, notice that

$$\langle 1_L * 1_{2S}, 1_{-A}\rangle \ll \langle 1_A * 1_{-2A}, 1_{-A}\rangle.$$

The right hand side is precisely the number of 3APs in A (with weight $1/|G|^2$). We will compute a lower bound of the left hand side. If all of this goes well, then we end up with a lower bound of the number of 3APs in A. It will turn out that we can either complete this construction, or we will find that A has density at least $\alpha + c\alpha$ on some smaller Bohr set. In the methods of Chapters 2 and 3, we estimated the number of 3APs in A in one step, but the present method breaks it down into smaller, more careful steps, which allows a larger density increment.

4.2.1. A sumset transforms. The focus of this subsection is to construct the sets L and S. The construction of such L and S proceeds step by step. The following is a heuristic provided by Sanders on the construction. Suppose $B, B' \subset G$, and also suppose that $A, L \subset B$ and $A', S \subset B'$ with density α, λ, α', and σ respectively in their respective sets. Suppose that $L + S \subset A + A'$. By averaging, we can expect to find many x such that

$$|S \cap (x - 2A)| \geq \frac{\sigma\alpha}{2}|B'|. \tag{4.7}$$

On the other hand, we do not expect to find too many x such that

$$|L \cap (A - x)| \geq \frac{\alpha}{2}|B|, \tag{4.8}$$

since this lower bound here does not take λ into account. If there is a x satisfying (4.7) but not (4.8), we shall expand L and slightly shrink S by setting

$$L' := L \cup (A - x)$$

and

$$S' := S \cap (x + A').$$

Then

$$|L'| = |L| + |A - x| - |L \cap (A - x)| \geq (\lambda + \alpha/2)|\Lambda]$$

and

$$|S'| \geq \frac{\sigma\alpha}{2}|\Lambda|.$$

In particular, L' is larger than L. We preserve the property $L' + S' \subset A + A'$ since

$$(L' + S') \subset (L + S') \cup ((A - x) + S') \subset (L + S) \cup ((A - x) + (x + A')) \subset (A + A').$$

By repeatedly applying this procedure, we can expand L while also controlling the size of S. This transforms $A + A'$ into $L + S$; hence the section name. This technique is motivated by Katz and Koester in [**7**].

We will need a relativized version of this construction. The following lemma does precisely this.

Lemma 4.12 (Incremental construction of L and S)**.** *Let G be a group of odd order. Let Λ be a regular Bohr set of dimension d, $\Lambda' \subset \Lambda_{\rho'}$ a regular Bohr set of dimension d' and width ϵ' and $\Lambda'' \subset \Lambda_{\rho''}$. Let $A \subset \Lambda$ and $A' \subset \Lambda'$ have densities α in Λ and α' in Λ' respectively. Furthermore, suppose that there is a $L \subset \Lambda$ and $S \subset \Lambda''$ of density λ in Λ and σ in Λ'' respectively. If $\lambda < 1/8$, $\rho' \ll \alpha/d$ and $\rho'' \ll' \alpha'/d'$, then one of the following occurs.*

i) (Density increment) There is a regular Bohr set Λ''' of dimension at most $d' + m$ and width $\rho'''\epsilon'$, where

$$\rho''' = \Omega\left(\frac{1}{d'^2 m}\right) \text{ and } m = O(\alpha^{-1} \log 2\alpha'^{-1})$$

such that for some $x \in G$, $x - A$ has density $\alpha(1 + \Omega(1))$ on Λ'''.

ii) (Expand L) There are sets $L' \subset \Lambda$ and $S' \subset \Lambda''$ with density at least $\lambda + \alpha/4$ in Λ and $\alpha'\sigma/2$ in Λ'' respectively, such that for all $x \in G$, we have

$$1_{L'} * 1_{2S'}(x) \le 1_L * 1_{2S}(x) + 1_A * 1_{2A'}(x). \tag{4.9}$$

.

By repeated application of this lemma until the density of L exceed $1/8$, we obtain our final sets L and S as in the proposition below.

Proposition 4.13. *Let G be a group of odd order. Let Λ be a regular Bohr set of dimension d, $\Lambda' \subset \Lambda_{\rho'}$ a regular Bohr set of dimension d' and width ϵ' and $\Lambda'' \subset \Lambda_{\rho''}$. Let $A \subset \Lambda$ and $A' \subset \Lambda'$ have densities α in Λ and α' in Λ' respectively. If $\rho' \ll \alpha/d$ and $\rho'' \ll \alpha'/d'$, then one of the following occurs.*

i) (Density increment) There is a regular Bohr set Λ''' of dimension at most $d' + m$ and width $\rho'''\epsilon'$, where

$$\rho''' = \Omega\left(\frac{1}{d'^2 m}\right) \text{ and } m = O(\alpha^{-1}\log 2\alpha'^{-1})$$

such that for some $x \in G$, $x - A$ has density $\alpha(1 + \Omega(1))$ on Λ'''.

ii) (Existence of L and S) There are sets $L \subset \Lambda$ and $S \subset \Lambda''$ with density at least $1/8$ in Λ and $(\alpha'/2)^{O(\alpha^{-1})}$ in Λ'' respectively, such that for all $x \in G$, we have

$$1_L * 1_{2S}(x) \ll \alpha^{-1} 1_A * 1_{2A'}(x). \tag{4.10}$$

Proof of Proposition 4.13: To start the iteration, we will set $L = \emptyset$ and $S \subset \Lambda''$ to be any subset of density between $\alpha'/2$ and α' in Λ''. Applying Lemma 4.12, if i) occurs, then we are done, so let us suppose otherwise. Then there exist $L_1 \subset \Lambda$ and $S_1 \subset \Lambda''$ of density at least $\alpha/4$ and $(\alpha'/2)^2$ respectively such that

$$1_{L_1} * 1_{2S_1}(x) \le 1_A * 1_{2A'}(x).$$

We now repeat this argument with L_1 and S_1. Note that in this iteration, the only objects that change are L and S; everything else remains the same. Now iterating in this way for $I = O(\alpha^{-1})$ many times, either i) occurs in the process and we are done, or we obtain $L_I \subset \Lambda$ and $S_I \subset \Lambda''$ of density at least $1/8$ in Λ and $(\alpha'/2)^{O(\alpha^{-1})}$ in Λ'' respectively such that

$$\begin{aligned} 1_{L_I} * 1_{2S_I}(x) &\le 1_{L_{I-1}} * 1_{2S_{I-1}}(x) + 1_A * 1_{2A'}(x) \\ &\le 1_{L_{I-2}} * 1_{2S_{I-2}}(x) + 1_A * 1_{2A'}(x) + 1_A * 1_{2A'}(x) \le \dots \\ &\le O(\alpha^{-1}) 1_A * 1_{2A'}(x). \end{aligned}$$

□

Now it remains to prove the Lemma.

Proof of Lemma 4.12: Let

$$\mathcal{L} = \left\{ x \in \Lambda' : 1_{-L} * 1_A(x) \geq \frac{\alpha}{2}\mu_G(\Lambda) \right\}$$

and

$$\mathcal{S} = \left\{ x \in \Lambda' : 1_{2S} * 1_{-2A'}(x) \geq \frac{\alpha'\sigma}{2}\mu_G(\Lambda'') \right\}.$$

Note that by Lemma 3.18, $2A'$ has density α' in $2\Lambda'$. By averaging, we can show that $\mathcal{S}$ is quite large. Indeed, if $\rho'' \leq 1/100d'$, then Λ' is $200d'\rho''$ locally-like $-S$ by Proposition 3.11, so $2\Lambda'$ is $200d'\rho''$ locally-like $-2S$. By Lemma 3.4, we have

$$\langle 1_{2S} * 1_{-2A'}(x), 1_{2\Lambda'} \rangle = \alpha'\sigma\mu_G(\Lambda')\mu_G(\Lambda'') + O(\sigma d'\rho''\mu_G(\Lambda')\mu_G(\Lambda'')).$$

Hence, if $\rho'' \ll \alpha'/d'$ is small enough, then we have the lower bound

$$\langle 1_{2S} * 1_{-2A'}(x), 1_{2\Lambda'} \rangle > \frac{3\alpha'\sigma}{4}\mu_G(\Lambda')\mu_G(\Lambda'').$$

On the other hand, by the definition of $\mathcal{S}$, we have an upper bound

$$\begin{aligned} \langle 1_{2S} * 1_{-2A'}(x), 1_{2\Lambda'} \rangle &\leq \mu_G(\mathcal{S})\mu_G(S) + (\mu_G(\Lambda') - \mu_G(\mathcal{S}))\frac{\alpha'\sigma}{2}\mu_G(\Lambda'') \\ &< \sigma\mu_G(\mathcal{S})\mu_G(\Lambda'') + \frac{\alpha'\sigma}{2}\mu_G(\Lambda')\mu_G(\Lambda''). \end{aligned}$$

By combining these two bounds, we see that

$$\frac{|A'|}{4} \leq |\mathcal{S}|.$$

Thus, if it happens that $\mathcal{L} < \frac{|A'|}{4}$, then we can find some $x \in \mathcal{S} \setminus \mathcal{L}$. We now treat two separate cases.

Case 1: $\mathcal{L} \leq \frac{|A'|}{8}$
In this case, let $x \in \mathcal{S} \setminus \mathcal{L}$ and set

$$L' := L \cup ((A - x)) \cap \Lambda)$$

and

$$S' := \{ y \in S : 2y \in 2S \cap (2A' + x) \}.$$

By Lemma 3.18, $|S'| = |2S \cap (2A' + x)|$. Thus, since $x \in \mathcal{S}$, the density of S' in Λ' is at least $\alpha'\sigma/2$. As for L', we have

$$|L'| = |L| + |(A - x) \cap \Lambda| - |L \cap ((A - x) \cap \Lambda)|.$$

Now, since $x \notin \mathcal{L}$, we have that

$$|L \cap ((A - x) \cap \Lambda)| \leq |L \cap (A - x)| < \alpha|\Lambda|/2.$$

On the other hand, notice that $(A - x) \subset \Lambda_{1+\rho'}$ since $x \in \Lambda'$. By the regularity of Λ, if $\rho' < 1/100d$, then

$$||\Lambda_{1+\rho'}| - |\Lambda|| \leq 100d\rho'|\Lambda|.$$

Hence, if $\rho' \ll \alpha/d$ and the implied constant is sufficiently small, then the above difference is at most $\alpha|\Lambda|/4$. This implies that

$$|(A-x)\cap\Lambda| > |(A-x)| - \frac{\alpha}{4}|\Lambda| = \frac{3\alpha'}{4}|\Lambda|.$$

From this, we see that

$$|L'| \geq \left(\lambda + \frac{\alpha}{4}\right)|\Lambda|$$

as required in the lemma. Finally, for any $y \in G$, we have

$$\begin{aligned}
|G|\cdot 1_{L'} * 1_{2S'}(y) &= |L' \cap (y-2S')| \\
&\leq |L \cup (A-x)) \cap (y-2S')| \\
&\leq |L \cap (y-2S')| + |(A-x)\cap(y-2S')| \\
&\leq |L \cap (y-2S)| \cup |(A-x)\cap(y-(x+2A'))| \\
&\leq |L \cap (y-2S)| \cup |A \cap (y-2A')| \\
&= |G|(1_L * 1_{2S}(y) + 1_A * 1_{2A'}(y)).
\end{aligned}$$

This completes the present case.

Case 2: $\mathcal{L} > \frac{|A'|}{8}$

In this case, we first observe that by the definition of $\mathcal{L}$, we have

$$\langle 1_{-L} * 1_A, 1_{\mathcal{L}}\rangle \geq \frac{\alpha}{2}\mu_G(\Lambda)\mu_G(\mathcal{L}).$$

Now, since $\mathcal{L} \subset \Lambda'$, Λ is $200d\rho'$ locally-like $\mathcal{L}$ if $\rho' \leq 1/100d$ by Proposition 3.11. Hence, by Lemma 3.4, we have

$$\begin{aligned}
\alpha\langle 1_{-L} * 1_\Lambda, 1_{\mathcal{L}}\rangle &= \alpha\langle 1_\Lambda, 1_{\mathcal{L}} * 1_L\rangle) \\
&= \alpha\lambda\mu_G(\Lambda)\mu_G(\mathcal{L}) + O(d\rho'\lambda\mu_G(\Lambda)\mu_G(\mathcal{L})).
\end{aligned}$$

If $\rho' \ll \alpha/d$ for a sufficiently small implied constant, then the last error term is at most $\alpha\mu_G(\Lambda)\mu_G(\mathcal{L})/4$. Hence, we see that

$$\langle 1_{-L} * 1_A, 1_{\mathcal{L}}\rangle - \alpha\langle 1_{-L} * 1_\Lambda, 1_{\mathcal{L}}\rangle \geq \alpha\mu_G(\Lambda)\mu_G(\mathcal{L})\left(\frac{1}{4} - \lambda\right).$$

If $\lambda \leq 1/8$, then

$$\langle 1_{-L} * 1_A, 1_{\mathcal{L}}\rangle - \alpha\langle 1_{-L} * 1_\Lambda, 1_{\mathcal{L}}\rangle \geq \frac{\alpha}{8}\mu_G(\Lambda)\mu_G(\mathcal{L}).$$

Now we apply Corollary 4.4 with $f = 1_L$, $T = \mathcal{L}$ (so $\tau \geq \alpha'/8$), $\nu = \sqrt{\alpha}/8$, and $\kappa = \sqrt{\alpha}/16$. This gives the desired density increment. $\square$

4.2.2. The Croot-Sisask Lemma. As discussed at the beginning of this section, our task now is to obtain a lower bound for $\langle 1_L * 1_{2S}, 1_{-A}\rangle$.

A natural first approach, in relation to our methods from the previous chapters, would be to directly compare it with the expected value. That is, we compare $\langle 1_L * 1_{2S}, 1_{-A}\rangle$ with $\langle 1_L * 1_{2S}, \alpha 1_\Lambda\rangle$. Here, we are using the same notation from Propsotion 4.13. By Lemma 3.4, we have

$$\langle 1_L * 1_{2S}, \alpha 1_\Lambda\rangle = \alpha\mu_G(L)\mu_G(S) + O(\alpha d\rho'\mu_G(\Lambda)\mu_G(S)).$$

Since L has density $\Omega(1)$ by Proposition 4.13, we see that if $\rho' \ll 1/d$ is sufficiently small, then the right hand side is at least $\alpha\mu_G(L)\mu_G(S)/2$. Hence, if

$$|\langle 1_L * 1_{2S}, 1_{-A}\rangle - \langle 1_L * 1_{2S}, \alpha 1_\Lambda\rangle| \leq \frac{\alpha}{4}\mu_G(L)\mu_G(S), \tag{4.11}$$

we obtain that

$$\langle 1_L * 1_{2S}, 1_{-A}\rangle \geq \frac{\alpha}{4}\mu_G(L)\mu_G(S),$$

and we have the desired lower bound. On the other hand, if the inequality of (4.11) is reversed, then by applying Corollary 4.4 with $f = 1_L$, $T = 2S$, $\nu = \sqrt{\alpha\lambda}/4$, $\kappa = \sqrt{\alpha\lambda}/8$ and some $\rho' \ll 1/100d$ sufficiently small, we obtain a density increment of $\alpha(1+\Omega(\lambda))$ on some Bohr neighborhood Λ''' with dimension at most

$$k + O(\alpha^{-1}\log 2\sigma^{-1})$$

and width $\rho'''\epsilon'$, where

$$\rho''' = \Omega\left(\frac{\alpha}{d'^2\log\sigma^{-1}}\right).$$

Recalling that $\sigma \geq (\alpha'/2)^{O(\alpha)}$, we now see that the α above gains a square. It turns out that this is not good enough.

The problem comes from using $T = 2S$; since S is already fixed, this gives us little freedom to work with the spectrum. In order to gain an additional degree of freedom, we will convolve $1_L * 1_S$ with g, where g itself is a convolution of characterstic functions of some set T, and perform the above approach with $\langle 1_L * 1_{2S} * g, 1_{-A}\rangle$ instead. This use of convolutions is similar to the technique employed in Lemma 4.8. However, we need to relate this new $\langle 1_L * 1_{2S} * g, 1_{-A}\rangle$ to our original $\langle 1_L * 1_{2S}, 1_{-A}\rangle$. This is achieved by the following lemma of Croot and Sisask in [**5**]. A proof of it is given in Sanders's paper [**17**] as well.

Lemma 4.14 (Croot-Sisask)**.** *Let $f \in C(G)$ and let $A, B \subset G$. Let $p > 1$ and $\epsilon > 0$. If $|A+B| \leq K|B|$ for some $K > 0$, then there exist $a \in A$ and $T \subset A$ such that T has density at least $(2K)^{-O(\epsilon^{-2}p)}$ in A and for all $t \in T - a$, we have*

$$||\tau_t(f * 1_A) - f * 1_A||_{L^p} \leq \epsilon\, ||f||_{L^p}\,\mu_G(B).$$

Recall that $\tau_t(f) \in C(G)$ is defined as $\tau_t(f)(x) := f(t+x)$. Using this we can compute a lower bound for $\langle 1_L * 1_{2S}, 1_{-A}\rangle$.

Proposition 4.15. *Let Λ be a regular Bohr set of dimension d, and $\Lambda' \subset \Lambda'_\rho$ a regular Bohr set of dimension d' and width ϵ'. Let $L, A \subset \Lambda$ and $S \subset \Lambda'$ have density λ in Λ, α in Λ, and σ in Λ' respectively. If $\rho' \ll \alpha\lambda/d$, then one of the following occurs.*

i) (Lower bound)

$$\langle 1_L * 1_{2S}, 1_A \rangle \geq \frac{\lambda\alpha\sigma}{2}\mu_G(\Lambda)\mu_G(\Lambda').$$

ii) (Density increment) There is a regular Bohr set Λ'' of dimension at most $2d' + m$ and width $\rho''\epsilon'$, where

$$\rho'' = \Omega\left(\frac{1}{d'^4 m^2}\right)$$

and

$$m = O(\lambda^{-2}(\log 2\lambda^{-1}\alpha^{-1})^2(\log 2\alpha^{-1})(\log 2\sigma^{-1})),$$

such that for some $x \in G$, $x - A$ has density at least $\alpha(1 + \Omega(\lambda))$ on Λ''.

Proof: As discussed earlier in this subsection, we shall convolve $1_L * 1_{2S}$ with another function g. Let $l > 1$, p and ϵ be positive parameters to be chosen later. We now apply Lemma 4.14 with $f = 1_L$ and $B = 2S$. Choose some ρ'' such that $\Lambda'' := \Lambda'_{\rho''/2l}$ is regular. By Proposition 3.14, we know that we can take $\rho' = \Omega(1/d')$. By the regularity of Λ', we have

$$|2\Lambda'' + B| \leq |\Lambda'' + \Lambda'| \leq |\Lambda_{1+\rho''/2l}| \leq 2|\Lambda'| = 2\sigma^{-1}|B|.$$

Hence, by applying Lemma 4.14 with $A = 2\Lambda''$, we obtain $T \subset \Lambda''$ and $a \in 2\Lambda''$ such that T has density τ at least $(2\sigma^{-1})^{O(\epsilon^{-2}p)}$ in Λ'', and for all $t \in 2T - a$,

$$||\tau_t(1_L * 1_{2S}) - 1_L * 1_{2S}||_{L^p} \leq \epsilon\mu_G(L)^{1/p}\mu_G(S).$$

Note that for any $f \in C(G)$ and $t \in G$, we have

$$||\tau_t(f)||_{L^p} = ||f||_{L^p},$$

and hence

$$||\tau_{-t}(f) - f||_{L^p} = ||\tau_{-t}(f - \tau_t(f)||_{L^p} = ||\tau_t(f) - f||_{L^p}.$$

Now, choose any $t_1, \ldots, t_{2l} \in 2T - a$, and set

$$\begin{aligned} r_0 &= 0, \\ r_1 &= t_1, \\ r_2 &= r_1 + t_2, \\ &\vdots \\ r_l &= r_{l-1} + t_l \\ r_{l+1} &= r_l - t_{l+1}, \\ &\vdots \\ r_{2l} &= r_{2l-1} - t_{2l} \end{aligned}$$

so in particular,

$$r := r_{2l} = t_1 + \cdots + t_l - t_{l+1} - \cdots - t_{2l} \in (2T - 2T) + \cdots + (2T - 2T),$$

where the sum has l copies of $(2T - 2T)$. Set $h := 1_L * 1_{2S}$. Note that each t_i is from $2T - a$, but the a's cancel out in the above term. Then

$$\begin{aligned} ||\tau_r(h) - h||_{L^p} &\leq \sum_{i=0}^{2l-1} \left|\left|\tau_{r_{i+1}}(h) - \tau_{r_i}(h)\right|\right|_{L^p} \\ &= \sum_{i=0}^{2l-1} \left|\left|\tau_{t_{i+1}}(h) - h\right|\right|_{L^p} \\ &\leq 2l\epsilon\mu_G(L)^{1/p}\mu_G(S) \end{aligned}$$

Furthermore, if $g = 1_{2T} * 1_{2T} * \cdots * 1_{-2T} * 1_{-2T}$ where there are l copies of 1_{2T} and 1_{-2T} each, then

$$\begin{aligned} & g * 1_L * 1_{2S}(x) - \mu_G(T)^{2l} 1_L * 1_{2S}(x) \\ = \; & \frac{1}{|G|^{2l}} \sum_{t_1,\ldots,t_l \in 2T} \sum_{t'_1,\ldots,t'_l \in 2T} \tau_{t'_1+\cdots+t'_l-t_1-\cdots-t_l}(h)(x) - h(x). \end{aligned}$$

Note again that in the above some, we took t_i, t'_i from $2T$ and not $2T - a$ because the a cancels anyway. Hence, by the triangular inequality, we have

$$||g * 1_L * 1_{2S} - \mu_G(T)^{2l} 1_L * 1_{2S}||_{L^p} \leq 2l\epsilon\mu_G(L)^{1/p}\mu_G(S)\mu_G(T)^{2l}.$$

Finally, by Hölder's inequality, we have

$$|\langle g * 1_L * 1_{2S}, 1_A\rangle - \mu_G(T)^{2l}\langle 1_L * 1_{2S}, 1_A\rangle| \leq 2l\epsilon\mu_G(L)^{1/p}\mu_G(S)\mu_G(T)^{2l}\mu_G(A)^{(p-1)/p}.$$

If we take $p := 2 + \log\alpha^{-1}$, then $\alpha^{(p-1)/p} \leq 2\alpha$. Also, by setting $\epsilon := \lambda/32l$, we obtain

$$|\langle g * 1_L * 1_{2S}, 1_A\rangle - \mu_G(T)^{2l}\langle 1_L * 1_{2S}, 1_A\rangle| \leq \frac{\lambda\alpha\sigma}{8}\mu_G(\Lambda)\mu_G(\Lambda')\mu_G(T)^{2l}. \tag{4.12}$$

Thus, the above relates $\langle g * 1_L * 1_{2S}, 1_A\rangle$ and $\langle 1_L * 1_{2S}, 1_A\rangle$. We now calculate the expected value of $\langle g * 1_L * 1_{2S}, 1_A\rangle$. By Proposition 3.11, we know that Λ is $400d\rho'\rho''/2l$ locally-like $2T$ (since $2T \subset 2\Lambda_{\rho'\rho''} \subset \Lambda_{2\rho'\rho'}$) and $400d\rho'$ locally-like $2S$ (since $2S \subset 2\Lambda_{\rho'} \subset \Lambda_{2\rho'}$). Thus, the expected value of this is, by $2l + 1$ repeated applications of Lemma 3.4,

$$\begin{aligned} \langle g * 1_L * 1_{2S}, \alpha 1_\Lambda\rangle &= \mu_G(T)^{2l}\langle 1_{2S} * 1_L, \alpha 1_\Lambda\rangle + O(\alpha\rho'\rho'' d\mu_G(\Lambda)\mu_G(T)^{2l}) \\ &= \alpha\mu_G(S)\mu_G(L)\mu_G(T)^{2l} + O(\alpha\rho' d\mu_G(S)\mu_G(\Lambda)\mu_G(T)^{2l}) \\ &= \alpha\lambda\sigma\mu_G(\Lambda)\mu_G(\Lambda')\mu_G(T)^{2l} + +O(\alpha\sigma\rho' d\mu_G(\Lambda)\mu_G(\Lambda')\mu_G(T)^{2l}) \end{aligned}$$

Hence, if $\rho' \ll \lambda/d$ for some sufficiently small implied constant, we have

$$\langle g * 1_L * 1_{2S}, \alpha 1_\Lambda\rangle \geq \frac{3\alpha\lambda\sigma}{4}\mu_G(\Lambda)\mu_G(\Lambda')\mu_G(T)^{2l}. \tag{4.13}$$

Now, if

$$|\langle g * 1_L * 1_{2S}, \alpha 1_\Lambda\rangle - \langle g * 1_L * 1_{2S}, 1_A\rangle| \leq \frac{\alpha\lambda\sigma}{8}\mu_G(\Lambda)\mu_G(\Lambda')\mu_G(T)^{2l}, \tag{4.14}$$

then by combining this with (4.12) and (4.13), we have

$$\langle 1_L * 1_{2S}, 1_A\rangle \geq \frac{\alpha\lambda\sigma}{2}\mu_G(\Lambda)\mu_G(\Lambda')$$

as desired. On the other hand, suppose that the inequality of (4.14) is reversed. Then

$$\begin{aligned}\frac{\alpha\lambda\sigma}{8}\mu_G(\Lambda)\mu_G(\Lambda')\mu_G(T)^{2l} &\leq |\langle (g * 1_L * 1_{2S})\hat{}, (\alpha 1_\Lambda - 1_A)\hat{}\,\rangle| \\ &\leq \langle |\hat{g}\hat{1}_L|, |(\alpha 1_\Lambda - 1_A)\hat{}\,|\rangle \mu_G(S)\end{aligned}$$

Hence, we have

$$\frac{\alpha\lambda}{8}\mu_G(\Lambda)\mu_G(T)^{2l} \leq \langle |\hat{g}\hat{1}_L|, |(\alpha 1_\Lambda - 1_A)\hat{}\,|\rangle.$$

By applying the same argument as in the proof of Lemma 3.19 with $f = 1_L$, $\nu = \sqrt{\alpha\lambda}/8$ and $\kappa = \sqrt{\alpha\lambda}/16$, we obtain

$$\frac{\lambda\alpha}{128}\mu_G(A) \leq \left|\left|(\hat{1}_A - \alpha\hat{1}_\Lambda)1_{\mathrm{Spec}_\kappa(g)}\right|\right|_{L^2}^2.$$

Now since $\hat{g} = \hat{1}_{2T}^{2l}$, it is easy to see that

$$\mathrm{Spec}_\kappa(g) = \mathrm{Spec}_{\kappa^{1/2l}}(1_{2T}).$$

Thus, we have

$$\frac{\lambda\alpha}{128}\mu_G(A) \leq ||(\hat{1}_A - \alpha\hat{1}_\Lambda)1_{\mathrm{Spec}_{\kappa^{1/2l}}(1_{2T})}||_{L^2}^2.$$

By setting $l = \lceil \log 2\alpha^{-1}\lambda^{-1}\rceil$, we have that $\kappa^{1/2l} = \Omega(1)$. Now, we would like to apply Lemma 4.3, but $2\Lambda''$ in which $2T$ lies is not a Bohr set. However, we see that

$$2T \subset 2\Lambda'' = 2\Lambda_{\rho''/2l} \subset \Lambda_{\rho''/l}.$$

Let τ' be the density of $2T$ in $\Lambda_{\rho''/l}$. Then by Proposition 3.11, we see that

$$|\Lambda_{\rho''/l}| \leq 4^{d'}|\Lambda''|.$$

Hence,

$$\tau' = \frac{|2T|}{|\Lambda_{\rho''/l}|} \geq \frac{\tau}{4^{d'}}.$$

Applying Lemma 4.3 on $2T$ and $\Lambda_{\rho''/l}$, we obtain a Bohr set Λ''' of dimension at most

$$d' + O(\log 2\tau'^{-1} = 2d' + O(\log 2\tau^{-1}) = 2d' + O(\epsilon^{-2}p\log 2\sigma^{-1}) = 2d' + O(m)$$

and width at least $\rho'''\rho''\epsilon'/2l$, where

$$\rho''' = \Omega\left(\frac{1}{d'^2 3\log 2\tau^{-1}}\right) = \Omega\left(\frac{1}{d'^3 m}\right)$$

and

$$m = O(\lambda^{-2}(\log 2\alpha^{-1}\lambda^{-1})^2(\log 2\alpha^{-1})(\log 2\sigma^{-1}))$$

such that Λ''' 1/2-annihilates $\mathrm{Spec}_{\kappa^{1/2l}}(1_T)$. We will write the width as $\rho\epsilon'$ where, a little crudely, we have

$$\rho = \Omega\left(\frac{1}{d'^4 m^2}\right).$$

Finally, if $\rho' \ll \alpha\lambda/d$, then by Lemma 3.20, we see that a translate of A has density $\alpha(1+\Omega(\lambda))$ on Λ'''. □

4.2.3. Completing the proof of the Theorem 4.2. First we summarize our findings.

Lemma 4.16. *Let G be a group of odd order. Let Λ be a regular Bohr set of dimension d, $\Lambda' \subset \Lambda_{\rho'}$ a regular Bohr set of dimension d' and width ϵ' and $\Lambda'' \subset \Lambda'_{\rho''}$ a regular Bohr set. Let $A \subset \Lambda$ and $A' \subset \Lambda'$ be of density α and α' in their respective sets. If $\rho' \ll \alpha/d$ and $\rho'' \ll \alpha'/d'$, then one of the following occurs.*

i) (Large inner product)

$$\langle 1_A * 1_{2A'}, 1_{-A}\rangle \geq \left(\frac{\alpha'}{2}\right)^{O(\alpha^{-1})}\mu_G(\Lambda)\mu_G(\Lambda'').$$

ii) (Density increment) There is a Bohr set Λ''' of dimension at most $2d'+m$ and width at least $\rho'''\epsilon'$, where

$$\rho''' = \Omega\left(\frac{1}{d'^4 m^2}\right)$$

and

$$m = O(\alpha^{-1}(\log 2\alpha^{-1})^3(\log 2\alpha'^{-1})),$$

such that for some $x \in G$, $x - A$ has density $\alpha(1+\Omega(1))$ on Λ'''.

Proof: This is simply an application of Proposition 4.13 and Proposition 4.15. By Proposition 4.13, we either obtain a density increment of A on some Bohr set which is larger than the one stated in our present lemma, or otherwise we find sets $L \subset \Lambda$ of density at least 1/8 and $S \subset \Lambda''$ of density at least $(\alpha'/2)^{O(\alpha)}$ such that

$$1_L * 1_{2S}(y) \ll \alpha^{-1} 1_A * 1_{2A'}(y)$$

for all $y \in G$. By Proposition 4.15, we either find a density increment of A again on some Bohr set which is larger than the one stated in the present lemma, or we obtain a lower bound

$$\langle 1_L * 1_{2S}, 1_{-A}\rangle \geq \frac{\alpha\sigma}{8}\mu_G(\Lambda)\mu_G(\Lambda'') \geq \left(\frac{\alpha'}{2}\right)^{O(\alpha^{-1})}\mu_G(\Lambda)\mu_G(\Lambda'').$$

The claim follows. □

The rest of the proof now just flows essentially in the same way as we've seen in the same way Theorem 2.1 followed from Lemma 2.12.

Proof of Theorem 4.2: We iterate Lemma 4.16. We let $B^{(0)}$ be a Bohr set with frequency set $\{1_G\}$ and width $\epsilon_0 = 1$, so $B^{(0)} = G$. Let α_0 be the density of A in G. We will let $B^{(i)}$ be the Bohr set at the ith iteration, d_i its dimension and ϵ_i its width, and α_i the largest density on $B^{(i)}$ achieved by a translate of A. We begin with $i = 0$ and perform the following iteratively. Let $\Lambda = B^{(i)}_\rho$, $\Lambda' = \Lambda_{\rho'}$ and $\Lambda'' = \Lambda'_{\rho''}$ be regular Bohr sets with $\rho = \Omega(\alpha_i/200d_i)$ such that $B^{(i)}$ is $\alpha_i/8$ locally-like Λ, Λ' and Λ'' by Proposition 3.11. ρ' and ρ'' will be specified soon.

Now, by Lemma 3.6, if a translate of A has density $\alpha_i(1 + 1/8)$ on either Λ or Λ', then we just set $B^{(i+1)}$ to be the Bohr set on which this density increment is achieved and we are done with this step. So let us suppose otherwise. Then there exists x such that $A - x$ has density at least $\alpha/2$ on both Λ and Λ'. Let $\tilde{A} = (A - x) \cap \Lambda$ and $A' = (A - x) \cap \Lambda'$. We now apply Lemma 4.16. Here, we will set $\rho' = \Omega(\alpha_i/d_i)$ and $\rho'' = \Omega(\alpha_i/d_i)$ satisfy the condition imposed on them in this lemma. First, let us consider the large inner product case, which gives

$$\langle 1_{\tilde{A}} * 1_{-2A'}, 1_{-\tilde{A}} \rangle \geq \left(\frac{\alpha_i}{2}\right)^{O(\alpha_i^{-1})} \mu_G(\Lambda)\mu_G(\Lambda'').$$

Now, by the cardinality estimate of Proposition 3.11, we can get a more explicit bound

$$\langle 1_{\tilde{A}} * 1_{-2A'}, 1_{-\tilde{A}} \rangle \geq \left(\frac{\alpha_i}{2}\right)^{O(\alpha_i^{-1})} O(\rho^2\rho'\rho''\epsilon_i^2)^{d_i}.$$

We can reduce this to

$$\langle 1_{\tilde{A}} * 1_{-2A'}, 1_{-\tilde{A}} \rangle \geq \left(\frac{\alpha_i\epsilon_i}{2d_i}\right)^{O(\alpha_i^{-1}+d_i)} \geq \left(\frac{\alpha\epsilon_i}{2d_i}\right)^{O(\alpha^{-1}+d_i)}. \tag{4.15}$$

The last inequality is by the fact that α_i is increasing due to the density increment. On the other hand, if the density increment case occurs, then there is a Bohr set Λ''' of rank $d + m$ and width $\rho'''\epsilon_i$, where

$$\rho''' = \Omega\left(\frac{\rho\rho'}{d_i^4 m^2}\right)$$

and

$$m = O(\alpha_i^{-1}(\log 2\alpha_i^{-1})^4) = O(\alpha_i^{-1}(\log 2\alpha^{-1})^4)$$

such that a translate of A has density $\alpha + c\alpha$ for some constant $c < 1/8$ on Λ'''. In this case, we set $B^{(i+1)} = \Lambda'''$ so it has dimension

$$d_{i+1} = 2d_i + O(\alpha_i^{-1}(\log 2\alpha^{-1})^4) \tag{4.16}$$

and width

$$\epsilon_{i+1} = \rho'''\epsilon_i = \Omega\left(\frac{\alpha}{d_i \log 2\alpha^{-1}}\right)^8 \epsilon_i.$$

Also, we have

$$\alpha_{i+1} \geq \alpha_i(1+c) \geq \alpha(1+c)^i.$$

This shows that the iteration must stop in at most $O(\log \alpha^{-1})$ many steps, for otherwise the density will exceed 1. Thus, i) of Lemma 4.3 must occur at some point.

This completes the setup. By summing the geometric progression of (4.16), we see that we have the uniform bound

$$d_i \leq O(\alpha^{-1}(\log 2\alpha^{-1})^4)$$

for all i, and hence

$$\epsilon_{i+1} = \Omega\left(\frac{\alpha}{\log 2\alpha^{-1}}\right)^{O(\log \alpha^{-1})}$$

for all i. Combining this with (4.15), we obtain

$$\langle 1_A * 1_{-2A}, 1_{-A}\rangle \geq \left(\frac{\alpha}{2\log 2\alpha^{-1}}\right)^{O(\alpha^{-1}(\log 2\alpha^{-1})^4)}.$$

If A contains only trivial 3APs, then $\langle 1_A * 1_{-2A}, 1_{-A}\rangle = \alpha/|G|$. Hence it suffices to ensure that, if D is the implied constant in the above term, then

$$\left(\frac{\alpha}{2\log 2\alpha^{-1}}\right)^{D\alpha^{-1}(\log 2\alpha^{-1})^4} > \frac{\alpha}{|G|}.$$

Taking logarithms and rearraging the terms gives

$$D\alpha^{-1}(\log 2\alpha^{-1})^4(\log 2\log 2\alpha^{-1} - \log \alpha) + \log \alpha < \log |G|.$$

Hence it suffices to have

$$2D\alpha^{-1}(\log 2\alpha^{-1})^5 < \log |G|.$$

The theorem follows. □

Bibliography

[1] T. Apostol. *Introduction to Analytic Number Theory.* Undergraduate Texts in Mathematics, Springer, USA, 1976.

[2] F. Behrend. *On sets of integers which contain no three terms in arithmetic progression*, Proc. Nat. Acad. Sci., 32:331-332, 1946.

[3] J. Bourgain. *On triples in arithmetic progression.* Geom. Func. Anal., 9(5):968-984, 1999.

[4] J. Bourgain. *Roth's theorem on progressions revisited.* Journal D'analyse Mathematique, 104:155-192, 2008.

[5] E. S. Croot and O. Sisask. *A probabilistic technique for finding almost-periods of convolutions.* Geom. Funct. Anal., 20(6):13671396, 2010.

[6] M. Elkin. *An improved construction of progression-free sets.* Symposium on Discrete Algorithms, pg. 886-905, 2010.

[7] N. H. Katz and P. Koester. *On additive energy and doubling.* SIAM J. Discrete Math., 24(4):1684-1693, 2010.

[8] B. J. Green. *Roth's theorem in the primes.* Ann. of Math.(2), 161(3):1609-1936, 2005.

[9] B. J. Green. *On triples in arithmetic progression, an exposition.* Cambridge University, 1999. http://www.dpmms.cam.ac.uk/ bjg23/notes.html

[10] B. J. Green. and J. Wolf. *A note on Elkin's improvement of Behrend's construction.* Additive number theory: Festschrift in honor of the sixieth birthday of Melvyn B. Nathanson, pg. 141-144, Springer-Verlag, 2010.

[11] T. Gowers. *A new proof of Szemerédi's theorem for arithmetic progressions of length four.* Geom. Funct. Anal., 8:529-551, 1998.

[12] T. Gowers. *A new proof of Szemerédi's theorem.* Geom. Funct. Anal., 11:465-588, 2001.

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

[14] K. Roth. *On certain sets of integers.* J. London Math. Soc., 28:104-109, 1953.

[15] H. L. Royden. *Real Analysis (3rd ed)*, Prentice Hall, New Jersey, 1988.

[16] T. Sanders. *On certain sets of integers.* to appear in J. Anal. Math., arXiv:1007.5444v2

[17] T. Sanders. *On Roth's Theorem on progressions.* to appear in Ann. of Math., arXiv:1011.0104v3

[18] K. Soundararajan. *Additive Combinatorics.* Course notes, Stanford University, 2007. http://math.stanford.edu/˜ksound/

[19] E. M. Stein and R. Shakarchi. *Fourier Analysis, An Introduction.* Princeton University Press, New Jersey, 2003.

[20] E. Szemerédi. *On sets of integers containing no four elements in arithmetic progression.* Acta Math. Hung. 20:89-104, 1990.

[21] E. Szemerédi. *On sets of integers containing no k elements in arithmetic progression.* Acta Arith. 27:299-345, 1975.

[22] E. Szemerédi. *Integer sets containing no arithmetic progressiona.* Acta. Math. Hung. 56(1-2):155-158, 1990.

[23] T. Tao and H. V. Vu. *Additive Combinatorics*, volume 105 of Cambridge studies in advanced mathematics, Cambridge University Press, Cambridge, 2006.