ROTH'S THEOREM ON ARITHMETIC PROGRESSIONS

ADAM LOTT

ABSTRACT. The goal of this paper is to present a self-contained exposition of Roth's celebrated theorem on arithmetic progressions. We also present two different stronger versions of Roth's theorem for two different notions of optimal sets.

CONTENTS

1. INTRODUCTION

1.1. **History of the problem.** A central question in additive number theory is to ask what conditions need to be imposed on a subset of the integers to guarantee that it contains an arithmetic progression. A natural first step is to guess that if a set is big enough, it will contain an arithmetic progression. In order to talk about the size of subsets of the integers in a meaningful way, we define the notion of upper density for a subset of the integers.

Definition 1.1 (Upper density)**.** *The upper density of a set $A \subset \mathbb{Z}$ is defined as*

$$\limsup_{N\to\infty} \frac{\#(A \cap [1, N])}{N}. \tag{1.1}$$

Date: April 29, 2017.

Remark 1.2. *Lower density can be defined in the same way, with* $\liminf$ *replacing* $\limsup$. *If the upper and lower densities of a set are equal, the set is said to have an asymptotic density.*

One of the first results along these lines was proven by van der Waerden in 1927 [vdW], who showed that for any positive integers r and k, there exists an $N(r,k)$ such that for any partition of $\{1,2,\ldots,N\}$ into r distinct color classes, there exists a monochromatic arithmetic progression of length k. In 1936, Erdős and Turán [ET] made the stronger conjecture that any set of integers with positive upper density contains arbitrarily long arithmetic progressions. This conjecture was resolved by Roth in 1953 [Ro1] for progressions of length three, and that result is the focus of this paper. In 1969, Szemerédi [Sz1] extended the result to progressions of length four. Roth used analytic methods, while Szemerédi used an intricate combinatorial argument. In 1972, Roth [Ro2] extended his analytic method to work for four-term progressions, and in 1975 Szemerédi [Sz2] extended his combinatorial method to resolve the conjecture for arbitrarily long progressions. The affirmative answer to Erdős and Turán's original conjecture is now known as Szemerédi's Theorem. Many different proofs of Szemerédi's Theorem have been produced, including an ergodic theoretical proof due to Furstenberg, Katznelson, and Ornstein [FKO]. In the late 1990s, Timothy Gowers developed new analytic machinery to work for four-term progressions [Go1] and arbitrarily long progressions [Go2], work for which he was awarded a Fields Medal. The theorem has also been generalized to cases where the ambient set is some other subset of the integers. For example, Ben Green [Gr1] showed that any set with positive upper *relative* density within the prime numbers contains a three-term progression, and Green and Tao later extended this to arbitrarily long progressions [GT].

More recently, attention has been given to trying to reduce the density threshold. Roth's original work actually showed that

$$\frac{\#(A\cap[1,N])}{N} \;\geq\; \frac{c}{\log\log N} \tag{1.2}$$

for some constant c is sufficient to guarantee the existence of a three-term progression in A. This bound has been slowly lowered over the years by Heath-Brown [He], Bourgain [Bo], and Sanders [Sa]. The current record is due to Bloom [Bl], who has shown that

$$\frac{\#(A\cap[1,N])}{N} \;\geq\; \frac{c(\log\log N)^4}{\log N} \tag{1.3}$$

suffices. The largest constructed example of a set containing no three-term progressions is due to Behrend [Be] and satisfies

$$\frac{\#(A\cap[1,N])}{N} \;\geq\; \exp(-c\sqrt{\log N}), \tag{1.4}$$

so the gap between the best known upper and lower bounds is still quite large. It is another conjecture of Erdős [Er] that any set A such that

$$\sum_{n\in A}\frac{1}{n} \;=\; \infty \tag{1.5}$$

contains arbitrarily long arithmetic progressions. This conjecture is still wide open and Erdős himself offered a \$3000 reward for a solution.

For the remainder of the paper, we focus our attention on Roth's original theorem.

Theorem 1.3 (Roth)**.** *Let A be a subset of $\mathbb{Z}$ with positive upper density. Then A contains a three term arithmetic progression.*

The theorem is often phrased in the following equivalent form, which is easier to work with.

Theorem 1.4 (Roth, finitary form)**.** *For every $\delta > 0$, there exists an $N_0(\delta)$ such that for every $N \geq N_0$ and every $A \subseteq \{1, 2, \ldots, N\}$ with $\#A \geq \delta N$, A contains a three term arithmetic progression.*

1.2. **Notation.** We denote by $\mathbb{Z}_N$ the additive cyclic group $\mathbb{Z}/N\mathbb{Z}$. Throughout the paper, we identify sets with their characteristic functions in the sense that for any set S, we define the function $S(x) := 1$ if $x \in S$ and $S(x) := 0$ if $x \notin S$. We also write $\#S$ for the cardinality of a finite set S. We write $[N]$ to denote the set $\{1, 2, \ldots, N\} \subset \mathbb{Z}$. Also, from now on, if the range of summation is not specified, it is understood to be all of $\mathbb{Z}_N$ (N will always be clear from context). We denote by χ the additive character $\chi(t) := \exp(2\pi i t/N)$.

2. Preliminaries

We will make heavy use of the following well-known identity.

Proposition 2.1. *We have*

$$\sum_{m \in \mathbb{Z}_N} \chi(mu) \;=\; \begin{cases} 0 & u \neq 0 \\ N & u = 0. \end{cases} \tag{2.1}$$

Proof. Let

$$S \;=\; \chi(0u) + \chi(1u) + \ldots + \chi((N-1)u) \;=\; 1 + \chi(u) + \ldots + \chi((N-1)u). \tag{2.2}$$

Then

$$\chi(u)S \;=\; \chi(u) + \ldots + \chi((N-1)u) + 1 \;=\; S, \tag{2.3}$$

which implies that either $S = 0$ or $\chi(u) = 1$, which happens if and only if $u = 0$. When $u = 0$ the sum is obviously equal to N, so the proof is complete. □

Our main tool throughout will be the Fourier transform, which is defined as follows.

Definition 2.2 (Fourier Transform)**.** *Given a function $f : \mathbb{Z}_N \to \mathbb{C}$, its Fourier transform $\widehat{f}$ is defined by*

$$\widehat{f}(m) \;:=\; N^{-1} \sum_{x \in \mathbb{Z}_N} \chi(-xm) f(x). \tag{2.4}$$

The Fourier transform has many useful properties. The following theorem summarizes a few of them which will be most useful for proving Roth's Theorem.

Theorem 2.3. *Let $f : \mathbb{Z}_N \to \mathbb{C}$. The Fourier transform $\widehat{f}$ possesses the following properties.*

(a) Inversion formula. For any $x \in \mathbb{Z}_N$,

$$f(x) \;=\; \sum_{m \in \mathbb{Z}_N} \chi(xm) \widehat{f}(m). \tag{2.5}$$

(b) *Pointwise estimate. For any $m \in \mathbb{Z}_N$, we have*

$$|\widehat{f}(m)| \;\le\; N^{-1} \sum_{x \in \mathbb{Z}_N} |f(x)|. \tag{2.6}$$

(c) *Plancherel's identity.*

$$\sum_{m \in \mathbb{Z}_N} |\widehat{f}(m)|^2 \;=\; N^{-1} \sum_{x \in \mathbb{Z}_N} |f(x)|^2. \tag{2.7}$$

(d) *Convolution identity. For f, $g : \mathbb{Z}_N \to \mathbb{C}$, define the convolution*

$$(f * g)(x) \;:=\; \sum_{y \in \mathbb{Z}_N} f(y) g(x - y). \tag{2.8}$$

Then for any m,

$$(\widehat{f * g})(m) \;=\; N \widehat{f}(m) \widehat{g}(m). \tag{2.9}$$

Proof.

(a) We have, by Proposition 2.1,

$$\begin{aligned}
\sum_m \chi(xm) \widehat{f}(m) \;&=\; N^{-1} \sum_{m,y} \chi(xm) \chi(-ym) f(y) &(2.10)\\
&=\; N^{-1} \sum_y f(y) \sum_m \chi(m(x - y)) &(2.11)\\
&=\; f(x). &(2.12)
\end{aligned}$$

(b) By the triangle inequality,

$$|\widehat{f}(m)| \;\le\; N^{-1} \sum_x |\chi(-xm)| |f(x)| \;=\; N^{-1} \sum_x |f(x)|. \tag{2.13}$$

(c) We have

$$\begin{aligned}
\sum_m |\widehat{f}(m)|^2 \;&=\; N^{-2} \sum_{m,x,y} \chi(-xm) f(x) \overline{\chi(-ym) f(y)} &(2.14)\\
&=\; N^{-2} \sum_{x,y} f(x) \overline{f(y)} \sum_m \chi(m(y - x)) &(2.15)\\
&=\; N^{-1} \sum_x |f(x)|^2. &(2.16)
\end{aligned}$$

(d) We have

$$\begin{aligned}
(\widehat{f * g})(m) \;&=\; N^{-1} \sum_x \chi(-xm) (f * g)(x) &(2.17)\\
&=\; N^{-1} \sum_{x,y} \chi(-xm + ym - ym) f(y) g(x - y) &(2.18)\\
&=\; N^{-1} \sum_{x,y} \chi(-ym) f(y) \chi(-(x - y)m) g(x - y) &(2.19)\\
&=\; N^{-1} \sum_y \chi(-ym) f(y) \sum_z \chi(-zm) g(z) &(2.20)
\end{aligned}$$

$$= N\widehat{f}(m)\widehat{g}(m). \quad \square \tag{2.21}$$

$\square$

3. Proof of Roth's Theorem

The general strategy of the proof is as follows. If the Fourier coefficients $|\widehat{A}(m)|$ are small for all $m \neq 0$, then we can find many arithmetic progressions in A by a simple counting argument. On the other hand, if there is an $m \neq 0$ such that $|\widehat{A}(m)|$ is large, then we can find a long arithmetic progression P such that the density of $A \cap P$ in P is greater than the density of A in $[N]$. Then, since arithmetic progressions are preserved under affine transformations, we can identify P with $[\#P]$ and repeat the process with $A \cap P$ as a subset of P. Since the density increases each time, eventually the density of A in an arithmetic progression becomes greater than 1, which is impossible. So we must eventually reach a point where all of the Fourier coefficients are small, implying the existence of arithmetic progressions inside A. The philosophy behind this approach is that the Fourier transform of the indicator function of a set encodes arithmetic information about that set. Specifically, if all of the Fourier coefficients are small, then the set is approximately random and thus likely to contain an arithmetic progression. And if a Fourier coefficient is large, it indicates that the set has more structure along certain arithmetic progressions. The proof presented here follows the proof given in [IMS] with the exception of the density increment construction, which is adapted from [Ly].

Let $A \subseteq [N]$ and let $\delta = (\#A)/N$. For technical reasons later, we will want 2 to be an invertible element in $\mathbb{Z}_N$, so we assume that N is odd. This does not cause any problems because if N is even, then the density of A in $[N+1]$ is only negligibly different from the density of A in $[N]$. We will consider A as a subset of $\mathbb{Z}_N$. The advantage of this is that we now have all of the tools of Fourier analysis at our disposal. The downside is that arithmetic progressions in $\mathbb{Z}_N$ are not necessarily progressions in $\mathbb{Z}$ (they might "wrap around"), but this will be taken care of. We will sometimes refer to a three-term arithmetic progression as a 3AP, and if "in $\mathbb{Z}_N$" is not specified, then we mean a 3AP in $\mathbb{Z}$. Note that $\{x, y, z\}$ is a 3AP in $\mathbb{Z}_N$ if $x + z \equiv 2y \pmod{n}$. To guarantee that we only count 3APs in $\mathbb{Z}$, we define $B := A \cap [N/3, 2N/3]$ and note that if $x, y \in B$, then $\{x, y, z\}$ is a 3AP in $\mathbb{Z}$. Denoting the number of 3APs contained in A by Q, we can count

$$Q \geq \#\{(x, y, z) \in B \times B \times A : x + z \equiv 2y \pmod{n}\} \tag{3.1}$$

$$= \sum_{x,y,z,\ x+z\equiv 2y} B(x)B(y)A(z) \tag{3.2}$$

$$= N^{-1} \sum_{x,y,z} \sum_{m} B(x)B(y)A(z)\chi(-m(x + z - 2y)) \tag{3.3}$$

$$= N^{-1} \sum_{m} \sum_{x} B(x)\chi(-mx) \sum_{y} B(y)\chi(2my) \sum_{z} A(z)\chi(-mz) \tag{3.4}$$

$$= N^2 \sum_{m} \widehat{B}(m)\widehat{B}(-2m)\widehat{A}(m) \tag{3.5}$$

$$= N^{-1}(\#B)^2(\#A) + N^2 \sum_{m\neq 0} \widehat{B}(m)\widehat{B}(-2m)\widehat{A}(m) \tag{3.6}$$

$$= N^{-1}(\#B)^2(\#A) + E. \tag{3.7}$$

At this point, we may assume that $\#B \geq \#A/5$. If this is not the case, then either $A \cap [0, N/3]$ or $A \cap [2N/3, N-1]$ must have size at least $2(\#A)/5$ and hence has relative density at least $6\delta/5 > \delta$ in its ambient progression. Thus the density increment argument kicks in (i.e., we may replace A by $A \cap [0, N/3]$ and $[N]$ by $[N/3]$ and repeat the same argument).

3.1. **Small Fourier coefficients.** When all of the nonzero Fourier coefficients $|\widehat{A}(m)|$ are small, we can use (3.7) to establish the existence of a three-term progression directly.

Theorem 3.1. *If $|\widehat{A}(m)| < \delta^2/10$ for all $m \neq 0$, then A contains a three term arithmetic progression.*

Proof. We prove this by showing that the error term E in (3.7) is small. By the Cauchy-Schwartz inequality and Plancherel's identity, we have

$$|E| \leq N^2 \max_m |\widehat{A}(m)| \left| \sum_{m\neq 0} \widehat{B}(m)\widehat{B}(-2m) \right| \tag{3.8}$$

$$\leq N^2 \max_m |\widehat{A}(m)| \left| \sum_{m} \widehat{B}(m)\widehat{B}(-2m) \right| \tag{3.9}$$

$$\leq N^2 \max_{m\neq 0} |\widehat{A}(m)| \left(\sum_m |\widehat{B}(m)|^2 \right)^{1/2} \left(\sum_m |\widehat{B}(-2m)|^2 \right)^{1/2} \tag{3.10}$$

$$= N \max_{m\neq 0} |\widehat{A}(m)| \left(\sum_x |B(x)|^2 \right)^{1/2} \left(\sum_x |B(-2x)|^2 \right)^{1/2} \tag{3.11}$$

$$\leq \frac{N\delta^2}{10}(\#B) \leq \frac{1}{2}N^{-1}(\#B)^2(\#A) \tag{3.12}$$

because of our assumption that $\#B \geq (\#A)/5$. This together with (3.7) implies that

$$Q \geq \frac{1}{2}N^{-1}(\#B)^2(\#A) \geq \frac{\delta^3}{50}N^2. \tag{3.13}$$

Since the count Q does not exclude the trivial progression $x = y = z$, we have to subtract those off, but there are only $\#A = \delta N$ of those, so the number of nontrivial 3APs contained in A is at least

$$\frac{\delta^3}{50}N^2 - \delta N, \tag{3.14}$$

which is positive if N is sufficiently large.

□

3.2. **Large Fourier coefficients.** If A has a large Fourier coefficient, then the estimate given by (3.7) is not useful, so we must use the additional arithmetic information that the large Fourier coefficient encodes. In this section, it will be convenient to work with the so-called balanced function $f(x) := A(x) - \delta$. We quickly describe the important properties of f.

Proposition 3.2. *The balanced function f possesses the following properties.*

(a) $\sum_x f(x) = 0$.
(b) $\widehat{f}(m) = \widehat{A}(m)$ *for all* $m \neq 0$ *and* $\widehat{f}(0) = 0$.

Proof. We have

$$\sum_x f(x) \;=\; \sum_x A(x) - \delta \;=\; \#A - \delta N \;=\; 0$$

and

$$\widehat{f}(m) \;=\; N^{-1}\sum_x \chi(-mx)f(x) \;=\; N^{-1}\sum_x \chi(-mx)A(x) - \delta N^{-1}\sum_x \chi(-mx) \;=\; \widehat{A}(m)$$

if $m \neq 0$. Also,

$$\widehat{f}(0) \;=\; N^{-1}\sum_x f(x) \;=\; 0.$$

□

In this section we will prove the following theorem, which is the statement that if A has a large Fourier coefficient, then there is a long arithmetic progression P on which A has increased density.

Theorem 3.3. *Suppose that* $|\widehat{A}(r)| > \epsilon > 0$ *for some* $r \neq 0$. *Then there exists an arithmetic progression P satisfying* $\#P \geq (\epsilon/64)\sqrt{N}$ *and* $\#(A \cap P) \geq (\delta + \epsilon/8)(\#P)$.

The proof consists of three basic steps.

Step 1: Construct a long $\mathbb{Z}_N$-arithmetic progression P such that $\widehat{P}(r)$ is also large.
Step 2: Show that this implies that A has a large intersection with some translate of P, call it P'.
Step 3: Lift the $\mathbb{Z}_N$-progression P' to a $\mathbb{Z}$-progression P'' without losing too much length or too much intersection with A.

Proof of Theorem 3.3.
Step 1. Suppose that $|\widehat{A}(r)| > \epsilon$ for some $r \neq 0$. Consider the set of points

$$\{(0,0), (1,r), (2,2r), \ldots, (N-1, (N-1)r)\} \subseteq [N-1]^2.$$

By subdividing $[N-1]^2$ into a $\lfloor\sqrt{N-1}\rfloor \times \lfloor\sqrt{N-1}\rfloor$ grid, there are fewer than N total boxes so the pigeonhole principle implies that two points are in the same box, i.e., there exist p, q such that both $p - q \leq \sqrt{N}$ and $r(p-q) \leq \sqrt{N}$ when considered modulo N (see Figure 1).

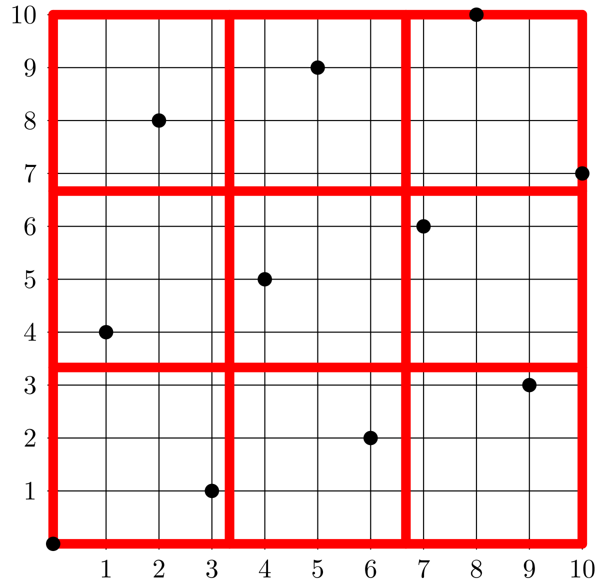

FIGURE 1. An example of the above construction with $N = 11$, $r = 4$, $\lfloor\sqrt{N-1}\rfloor = 3$. Notice that there are two points in the bottom left and top right boxes.

Let $d = p - q$. Let P be the progression $\{\ldots, -d, 0, d, \ldots\}$ of length $\lfloor\sqrt{N}/8\rfloor$ where the endpoints of P are chosen so that it is as symmetric around 0 as possible. We want $|\widehat{P}(r)|$ to be large, so consider

$$\begin{aligned}
\left|N\widehat{P}(r) - \#P\right| &= \left|\sum_x (P(x)\chi(-rx) - P(x))\right| && (3.15)\\
&\leq \sum_{x\in P} |\chi(-rx) - 1| && (3.16)\\
&\leq \sum_{|\ell|\leq(\#P)/2} |\chi(rd\ell) - 1| && (3.17)\\
&\leq \#P \cdot \max_{|\ell|\leq(\#P)/2} |\chi(rd\ell) - 1|. && (3.18)
\end{aligned}$$

By construction, we have that $rd\ell \leq \sqrt{N} \cdot (\#P)/2 \leq N/16$. Thus, $\chi(rd\ell)$ is an Nth root of unity which is no more than $1/16$th of the way around the unit circle. Hence for any ℓ, $|\chi(rd\ell) - 1| \leq 2\pi/16 < 1/2$. Thus we conclude that

$$\left|N \cdot \widehat{P}(r) - \#P\right| < \frac{1}{2}, \tag{3.19}$$

so that $|\widehat{P}(r)| > (1/2)N^{-1}(\#P)$.

Step 2. We now want to show that this implies that A has a large intersection with some translate of P. Define $G(x) := (f * P)(x) = \sum_y f(y)P(x-y)$ where f is the balanced function of A. Note that G has mean value zero because

$$\sum_x G(x) \;=\; \sum_{x,y} f(y)P(x-y) \;=\; \sum_y f(y)\sum_x P(x-y) \;=\; \sum_y (\#P)f(y) \;=\; 0. \tag{3.20}$$

By the pointwise estimate and the convolution identity of Theorem 2.3, we have

$$N^{-1}\sum_x |G(x)| \;\geq\; |\widehat{G}(r)| \;=\; N|\widehat{f}(r)||\widehat{P}(r)| \;\geq\; \frac{1}{2}\epsilon(\#P). \tag{3.21}$$

Since G has mean value zero, we can write the above as

$$\sum_x |G(x)| + G(x) \;\geq\; \frac{1}{2}N\epsilon(\#P), \tag{3.22}$$

so that $|G(x)| + G(x) \geq \frac{1}{2}\epsilon(\#P)$ for some x, which implies that $G(x) \geq \frac{1}{4}\epsilon(\#P)$. Expanding out definitions, we get

$$\begin{aligned} \frac{1}{4}\epsilon(\#P) \;&\leq\; \sum_y f(y)P(x-y) && (3.23)\\ &=\; \sum_y A(y)P(x-y) - \delta\sum_y P(x-y) && (3.24)\\ &=\; \#(A\cap(x-P)) - \delta(\#(x-P)) && (3.25) \end{aligned}$$

so that

$$\#(A\cap(x-P)) \;\geq\; (\delta + \frac{1}{4}\epsilon)(\#(x-P)). \tag{3.26}$$

Step 3. Let $P' = x - P$. The only remaining obstacle is that although P' is an arithmetic progression in $\mathbb{Z}_N$, it may not be an arithmetic progression in $\mathbb{Z}$. We need to find an arithmetic progression in $\mathbb{Z}$ contained in P' that is still relatively long and still has a relatively large intersection with A. We accomplish this in the following way. Note that since $\#P' \leq \sqrt{N}$ and the common difference of P' is also at most $\sqrt{N}$, the last term of P' does not "wrap around" the first term of P' modulo N. Hence P' can be written as $P' = P_1 \cup P_2$ where both P_1 and P_2 are arithmetic progressions in $\mathbb{Z}$. Without loss of generality, suppose that $\#P_1 \leq \#P_2$. If $\#P_1 \leq (1/8)\epsilon(\#P')$, then

$$\begin{aligned} \#(A\cap P_2) \;&\geq\; \#(A\cap P') - \#P_1 && (3.27)\\ &\geq\; (\delta + \frac{1}{4}\epsilon)(\#P') - \frac{1}{8}\epsilon(\#P') && (3.28)\\ &=\; (\delta + \frac{1}{8}\epsilon)(\#P'). && (3.29) \end{aligned}$$

If, on the other hand, $\#P_1 \geq (1/8)\epsilon(\#P')$, then both P_1 and P_2 have length at least $(1/8)\epsilon(\#P')$ and since $P' = P_1 \cup P_2$, A must have density of at least $\delta + (1/4)\epsilon$ on one of them. Thus we have established the existence of an arithmetic progression P'' in $\mathbb{Z}$ such that $\#P'' \geq (1/8)\epsilon(\#P')$ and $\#(A\cap P'') \geq (\delta + (1/8)\epsilon)(\#P'')$. Since $\#P' = \lfloor\sqrt{N}/8\rfloor$, this completes the proof. $\square$

3.3. **Completing the proof via density increment argument.** Suppose A is a subset of $[N]$ containing no 3APs. By Theorem 3.1, this implies that for some $r \neq 0$, $|\widehat{A}(r)| \geq \delta^2/10$. Then by Theorem 3.3 with $\epsilon = \delta^2/10$, there exists an arithmetic progression P_1 such that $\#P_1 \geq (\delta^2/640)\sqrt{N}$ and $\#(A \cap P_1) \geq (\delta + \delta^2/80)(\#P_1)$. Let $A_1 = A \cap P_1$. Since arithmetic progressions are preserved under affine transformations, we can identify P_1 with $[\#P_1]$ and A_1 with a subset of $[\#P_1]$ of density $\delta_1 \geq \delta + \delta^2/80$. Since we assumed A contains no 3APs, obviously neither does A_1, so we may repeat the argument and obtain another progression P_2 and $A_2 := (A_1 \cap P_2)$ where the density of A_2 in P_2 is $\delta_2 \geq \delta_1 + \delta_1^2/80$. Repeating this process, we get a sequence of progressions P_k and subsets A_k with relative densities δ_k satisfying

$$\#P_k \;\geq\; \frac{\delta_{k-1}^2}{640}\sqrt{\#P_{k-1}}, \qquad \delta_k \;\geq\; \delta_{k-1} + \frac{\delta_{k-1}^2}{80}. \tag{3.30}$$

Notice that $\delta_k \geq \delta + k(\delta^2/80)$. Thus after $k = 80/\delta$ iterations, we have $\delta_k \geq 2\delta$. Now for $k > 80/\delta$, we have $\delta_k \geq 2\delta + k((2\delta)^2/80)$, so that after $k = 80/(2\delta)$ more iterations, the density has increased from 2δ to 4δ. In general, the density will have increased to $2^a\delta$ after no more than $(80/\delta)(1+1/2+1/4+\ldots+1/2^a) \leq 160/\delta$ iterations. Picking a to be sufficiently large, we see that the density has increased past 1 after a finite number of iterations, which is a contradiction. This completes the proof of Roth's Theorem.

Remark 3.4. *This method also yields an upper bound for the constant $N_0(\delta)$ mentioned in Theorem 1.4. We have shown that at most $k = 160/\delta$ iterations are needed to arrive at a contradiction, so N_0 just needs to be large enough so that $\#P_k > 0$ after $k = 160/\delta$ iterations (see equation* (3.30)*). In* [Ly]*, it is shown that*

$$N_0(\delta) \;=\; \exp(\exp(C\delta^{-1}))$$

suffices for some constant $C > 0$.

4. Roth's Theorem for Salem sets

4.1. **Statement and proof of density threshold.** In the proof of Roth's Theorem, it becomes clear that when all of the nonzero Fourier coefficients $\widehat{A}(m)$ are small in absolute value, it is very easy to establish the existence of three term arithmetic progressions contained in A. This suggests the following question. If the Fourier coefficients of A are "optimally small", can we get away with a smaller density threshold to guarantee the existence of three term arithmetic progressions contained in A? To answer this question, we first need to understand how small is "optimally small". Plancherel's identity tells us

$$\sum_m |\widehat{A}(m)|^2 \;=\; N^{-1}\sum_x |A(x)|^2 \;=\; N^{-1}(\#A), \tag{4.1}$$

which tells us that, even in theory, the smallest order of magnitude for $|\widehat{A}(m)|$ that we can hope for is $N^{-1}\sqrt{\#A}$. This observation leads to the following definition.

Definition 4.1 (Salem set)**.** *Let $\{A_N\}_{N=1}^{\infty}$ be a family of sets with $A_N \subseteq \mathbb{Z}_N$. The family is said to be Salem if there exists a universal constant C such that for all N and all nonzero $m \in \mathbb{Z}_N$,*

$$|\widehat{A_N}(m)| \;\leq\; CN^{-1}\sqrt{\#A_N}. \tag{4.2}$$

For Salem sets, it is possible to guarantee the existence of three term arithmetic progressions when the density is much smaller. Since the Salem condition only makes sense within the framework of $\mathbb{Z}_N$, for this section and the next, we will only consider arithmetic progressions in $\mathbb{Z}_N$. The argument given here is presented in [IMS].

Theorem 4.2. *Let $\{A_N\}_{N=1}^{\infty}$ be a Salem family. If there exists a universal constant c such that $\#A_N \geq cN^{2/3}$, then for all sufficiently large N, A_N contains a three term arithmetic progression.*

Proof. Let $Q_N(t) := \{x, y, z \in A_N : y - x = z - y = t\}$ be the number of three term arithmetic progressions in A_N with common difference t. The number of three term arithmetic progressions contained in A_N is then equal to

$$\begin{aligned}
\sum_t Q_N(t) &= \sum_{x,t} A_N(x-t)A_N(x)A_N(x+t) && (4.3)\\
&= \sum_{x,t} A_N(x) \sum_m \chi((x-t)m)\widehat{A_N}(m) \sum_\ell \chi((x+t)\ell)\widehat{A_N}(\ell) && (4.4)\\
&= \sum_{x,m,\ell} A_N(x)\widehat{A_N}(m)\widehat{A_N}(\ell)\chi(x(m+\ell)) \sum_t \chi(t(\ell-m)) && (4.5)\\
&= N\sum_{x,m} A_N(x)\chi(2mx)\widehat{A_N}(m)\widehat{A_N}(m) && (4.6)\\
&= N^2 \sum_m \widehat{A_N}(m)\widehat{A_N}(m)\widehat{A_N}(-2m) && (4.7)\\
&= N^{-1}(\#A_N)^3 + N^2 \sum_{m\neq 0} \widehat{A_N}(m)\widehat{A_N}(m)\widehat{A_N}(-2m) && (4.8)\\
&= N^{-1}(\#A_N)^3 + E && (4.9)
\end{aligned}$$

To bound the error term E, we can use the Salem condition to get

$$|E| \leq N^2 \sum_{m\neq 0} |\widehat{A_N}(m)|^2 |\widehat{A_N}(-2m)| \leq C^3(\#A_N)^{3/2}. \tag{4.10}$$

If c is sufficiently large with respect to C and $\#A_N \geq cN^{2/3}$, then (4.10) implies that $|E| \leq (1/2)N^{-1}(\#A_N)^3$. Hence the number of nontrivial three term arithmetic progressions in $\mathbb{Z}_N$ is at least

$$\begin{aligned}
\frac{1}{2}N^{-1}(\#A_N)^3 - (\#A_N) &\geq (\#A_N)((1/2)N^{-1}(\#A_N)^2 - 1) && (4.11)\\
&\geq cN^{2/3}((c^2/2)N^{1/3} - 1) && (4.12)
\end{aligned}$$

which is positive if N is sufficiently large. □

4.2. **Example of Salem set.** At first glance, it is not obvious that Salem sets actually exist. In this section, we provide an example of a Salem family of subsets of $\mathbb{Z}_p$ as p ranges over the primes. First we need a lemma.

Lemma 4.3. *For any prime p and any $a \neq 0$, let $G(a) = \sum_{x\in\mathbb{Z}_p} \chi(ax^2)$. Then we have*

$$|G(a)| = \sqrt{p}. \tag{4.13}$$

Proof. We have

$$|G(a)|^2 \;=\; \sum_{x,y\in\mathbb{Z}_p} \chi(ax^2)\overline{\chi(ay^2)} \;=\; \sum_{x,y\in\mathbb{Z}_p} \chi(a(x^2-y^2)). \tag{4.14}$$

The change of variables $t = x - y$, $u = x + y$ is bijective and $tu = x^2 - y^2$, so we have

$$|G(a)|^2 \;=\; \sum_{t,u\in\mathbb{Z}_p} \chi(atu) \tag{4.15}$$

$$= \sum_{r\in\mathbb{Z}_p} \sum_{\substack{t,u \\ tu=r}} \chi(ar) \tag{4.16}$$

$$= \sum_{r\in\mathbb{Z}_p} \chi(ar)m(r) \tag{4.17}$$

where $m(r) := \#\{(t,u) \in \mathbb{Z}_p^2 : tu = r\}$. We know $tu = 0$ if and only if $t = 0$ or $u = 0$, so $m(0) = 2p - 1$. If $r \neq 0$, then t can be any nonzero element and u is determined, so $m(r) = p - 1$. Hence

$$|G(a)|^2 \;=\; m(0) + \sum_{r\neq 0} \chi(ar)m(r) \tag{4.18}$$

$$= 2p - 1 + (p-1)\sum_{r\neq 0} \chi(ar) \tag{4.19}$$

$$= 2p - 1 - (p-1) + (p-1)\sum_{r} \chi(ar) \tag{4.20}$$

$$= p. \tag{4.21}$$

□

Theorem 4.4. *Define $E_p := \{t^2 : t \in \mathbb{Z}_p\}$. Then $\{E_p\}$ is a Salem family.*

Proof. For any $m \neq 0$, we have

$$\widehat{E_p}(m) \;=\; p^{-1}\sum_{x} \chi(-mx)E_p(x) \;=\; p^{-1}\sum_{x\in E_p} \chi(-mx). \tag{4.22}$$

Since exactly half of the nonzero residue classes in $\mathbb{Z}_p$ are squares and every nonzero square has exactly two square roots, this becomes

$$\widehat{E_p}(m) \;=\; p^{-1}\chi(0) + \frac{1}{2}p^{-1}\sum_{t\neq 0} \chi(-mt^2) \tag{4.23}$$

$$= p^{-1}\chi(0) + \frac{1}{2}p^{-1}\sum_{t} \chi(-mt^2) - \frac{1}{2}p^{-1}\chi(0) \tag{4.24}$$

$$= \frac{1}{2}p^{-1}\left(\chi(0) + \sum_{t} \chi(-mt^2)\right). \tag{4.25}$$

By Lemma 4.3, we have

$$\left|\widehat{E_p}(m)\right| \;\le\; \frac{1}{2}p^{-1}(1+\sqrt{p}) \;\le\; p^{-1/2}. \tag{4.26}$$

Since $\#E_p = (p+1)/2$, we have

$$p^{-1}\sqrt{\#E_p} \;\geq\; \frac{1}{2}p^{-1/2} \;\geq\; \frac{1}{2}\left|\widehat{E_p}(m)\right| \tag{4.27}$$

for every $m \neq 0$, so $\{E_p\}$ is a Salem family. □

5. Roth's Theorem for U^2-optimal sets

5.1. **Preliminaries.** In the previous section, we defined a notion of what it means for a set to be "optimally" random and showed that sets satisfying that condition require a much smaller density threshold in order to guarantee the existence of a three-term progression. In this section, we will introduce a weaker notion of randomness and show that sets with this condition also get away with a smaller density threshold to guarantee the existence of a three-term progression. The relevant concepts were first introduced by Gowers in [Go2] in his proof of Szemerédi's Theorem.

Definition 5.1 (L^p norm)**.** *For any function $f : \mathbb{Z}_N \to \mathbb{C}$ and any $p > 0$, the L^p norm of f is defined as*

$$||f||_p \;:=\; \left(\sum_x |f(x)|^p\right)^{1/p}. \tag{5.1}$$

Definition 5.2 (Gowers uniformity norm)**.** *For any function $f : \mathbb{Z}_N \to \mathbb{C}$, the Gowers U^2 norm is defined by the formula*

$$||f||_{U^2} \;:=\; \left(N^{-3}\sum_{x,h_1,h_2} f(x)\overline{f(x+h_1)f(x+h_2)}f(x+h_1+h_2)\right)^{1/4}. \tag{5.2}$$

Remark 5.3. *The U^2 norm is the norm that is best suited for handling three-term progressions. In fact, Gowers introduced U^k norms for all $k \geq 2$ which are more suitable for handling longer progressions. However, the difficulty of the argument increases dramatically for $k \geq 3$.*

Remark 5.4. *The U^2 norm satisfies all the necessary properties of a norm and even satisfies a generalized Cauchy-Schwartz inequality, but we never need to use these facts, so we will not prove them.*

The U^2 norm has an interpretation in terms of the Fourier transform, as the following proposition shows. This is motivation for why it is useful for studying three-term progressions. The reason that the cases of longer progressions are so much harder is that for $k > 2$ the U^k norm no longer has a nice Fourier analytic interpretation.

Proposition 5.5. *For any function $f : \mathbb{Z}_N \to \mathbb{C}$, we have*

$$||f||_{U^2} \;=\; ||\widehat{f}||_4. \tag{5.3}$$

Proof. We have

$$||f||_{U^2}^4 \;=\; N^{-3}\sum_{x,h_1,h_2} f(x)\overline{f(x+h_1)f(x+h_2)}f(x+h_1+h_2) \tag{5.4}$$

$$= N^{-3} \sum_{\substack{a,b,c,d \\ a+d=b+c}} f(a)f(d)\overline{f(b)f(c)} \tag{5.5}$$

$$= N^{-4} \sum_{a,b,c,d} f(a)f(d)\overline{f(b)f(c)} \sum_m \chi(m(b+c-(a+d))) \tag{5.6}$$

$$= N^{-4} \sum_m \sum_{a,b,c,d} f(a)\chi(-ma)f(d)\chi(-md)\overline{f(b)\chi(-mb)f(c)\chi(-mc)} \tag{5.7}$$

$$= \sum_m |\widehat{f}(m)|^4 \;=\; ||\widehat{f}||_4^4. \tag{5.8}$$

□

In order to prove anything about optimal sets, it is first necessary to determine how small the U^2 norm of a characteristic function can be. Since the U^2 norm is related to the Fourier coefficients, we can observe that for any set A,

$$||A||_{U^2}^4 \;=\; \sum_m |\widehat{A}(m)|^4 \;=\; (\#A)^4 N^{-4} + \sum_{m \neq 0} |\widehat{A}(m)|^4 \tag{5.9}$$

so that

$$||A||_{U^2}^4 - (\#A)^4 N^{-4} \;\geq\; \left(\sup_{m \neq 0} |\widehat{A}(m)| \right)^2 \sum_{m \neq 0} |\widehat{A}(m)|^2 \tag{5.10}$$

$$\geq\; \left(\sup_{m \neq 0} |\widehat{A}(m)| \right)^2 \sum_m |\widehat{A}(m)|^2 \tag{5.11}$$

$$=\; \left(\sup_{m \neq 0} |\widehat{A}(m)| \right)^2 N^{-1} \sum_x |A(x)|^2 \tag{5.12}$$

$$=\; \left(\sup_{m \neq 0} |\widehat{A}(m)| \right)^2 N^{-1}(\#A). \tag{5.13}$$

Recall from the Salem condition (which was deduced directly from Plancherel's identity) that $\sup_{m \neq 0} |\widehat{A}(m)|$ can never be smaller than $(\#A)^{1/2}N^{-1}$, so we are left with

$$||A||_{U^2}^4 - (\#A)^4 N^{-4} \;\geq\; (\#A)^2 N^{-3}. \tag{5.14}$$

At this point, it will be useful to recall the definition of the balanced function

$$f(x) := A(x) - (\#A)N^{-1}$$

introduced in Section 3.2. The following lemma relates the norms of the characteristic function A and its balanced function f.

Lemma 5.6. *Let A be a subset of $\mathbb{Z}_N$, let $\delta = \#A/N$, and let f be the balanced function of A. Then*

$$||f||_{U^2}^4 \;=\; ||A||_{U^2}^4 - \delta^4. \tag{5.15}$$

Proof. We have

$$||A||_{U^2}^4 \;=\; N^{-3} \sum_{x,h_1,h_2} A(x)A(x+h_1)A(x+h_2)A(x+h_3) \tag{5.16}$$

$$= N^{-3} \sum_{x,h_1,h_2} (\delta + f(x))(\delta + f(x+h_1))(\delta + f(x+h_2))(\delta + f(x+h_1+h_2)) \tag{5.17}$$

$$= N^{-3} \left[\sum_{x,h_1,h_2} \delta^4 + f(x)f(x+h_1)f(x+h_2)f(x+h_1+h_2) + (\text{other terms}) \right] \tag{5.18}$$

where each of the terms in (other terms) is a product of the form

$$g_1(x)g_2(x+h_1)g_3(x+h_2)g_4(x+h_1+h_2) \tag{5.19}$$

where at most three of the g_i are equal to f and the others are equal to δ. Thus, by changing variables, since f has mean value zero (recall Proposition 3.2), each of the terms in (other terms) vanishes. So we have

$$||A||_{U^2}^4 = \delta^4 + N^{-3} \sum_{x,h_1,h_2} f(x)f(x+h_1)f(x+h_2)f(x+h_1+h_2) \tag{5.20}$$

$$= \delta^4 + ||f||_{U^2}^4. \tag{5.21}$$

□

Equation (5.14) and Lemma 5.6 motivate the following definition.

Definition 5.7 (U^2-optimal set)**.** *Let $\{A_N\}_{N=1}^{\infty}$ be a family of subsets of $\mathbb{Z}_N$ and let f_N be the balanced function of A_N. We say the family is U^2-optimal if there exists a universal constant C such that*

$$||f_N||_{U^2}^4 = ||A_N||_{U^2}^4 - (\#A)^4 N^{-4} \leq C(\#A)^2 N^{-3}. \tag{5.22}$$

We also prove here one more lemma, adapted from [Gr2], which will be necessary to prove the density threshold for U^2-optimal sets.

Lemma 5.8. *Let $f : \mathbb{Z}_N \to [-1, 1]$ be any function. Then*

$$\left| \sum_{y,d} f(y)f(y+d)f(y+2d) \right| \leq N^2 ||f||_{U^2}. \tag{5.23}$$

Proof. Let Q denote the quantity $\left|\sum_{y,d} f(y)f(y+d)f(y+2d)\right|$. We have, by the change of variables $x = 2(y+d)$ and repeated application of the Cauchy-Schwartz inequality,

$$Q = \left| \sum_{x,y} f(y)f(x/2)f(x-y) \right| \tag{5.24}$$

$$\leq \left(\sum_x |f(x/2)|^2 \right)^{1/2} \left(\sum_x \left| \sum_y f(y)f(x-y) \right|^2 \right)^{1/2} \tag{5.25}$$

$$\leq N^{1/2} \left(\sum_{x,y,z} f(y)f(x-y)f(z)f(x-z) \right)^{1/2} \tag{5.26}$$

$$\leq\ N^{1/2}\left[\sum_y f(y)\left(\sum_z |f(z)|^2\right)^{1/2}\left(\sum_z\left|\sum_x f(x-y)f(x-z)\right|^2\right)^{1/2}\right]^{1/2} \tag{5.27}$$

$$\leq N^{1/2}\left[N^{1/2}\sum_y f(y)\left(\sum_{x,z,t} f(x-y)f(x-z)f(t-y)f(t-z)\right)^{1/2}\right]^{1/2} \tag{5.28}$$

$$\leq N^{3/4}\left[\left(\sum_y |f(y)|^2\right)^{1/2}\left(\sum_y\left|\sum_{x,z,t} f(x-y)f(x-z)f(t-y)f(t-z)\right|\right)^{1/2}\right]^{1/2} \tag{5.29}$$

$$\leq\ N^{3/4}N^{1/4}\left(\sum_{x,y,z,t} f(x-y)f(x-z)f(t-y)f(t-z)\right)^{1/4} \tag{5.30}$$

$$=\ N\left(N\sum_{a+b=c+d} f(a)f(b)f(c)f(d)\right)^{1/4} \tag{5.31}$$

$$=\ N^2\left(N^{-3}\sum_{a+b=c+d} f(a)f(b)f(c)f(d)\right)^{1/4} \tag{5.32}$$

$$=\ N^2\left(N^{-3}\sum_{x,h_1,h_2} f(x)f(x+h_1)f(x+h_2)f(x+h_1+h_2)\right)^{1/4} \tag{5.33}$$

$$=\ N^2||f||_{U^2}. \tag{5.34}$$

□

5.2. **Statement and proof of density threshold.** We now have the proper notions in place to state the main theorem of this section.

Theorem 5.9. *Let $\{A_N\}_{N=1}^{\infty}$ be a U^2-optimal family of sets. There exists a constant c such that if $\#A_N \geq cN^{9/10}$, then for all sufficiently large N, A_N contains a three-term arithmetic progression.*

Proof. Let $Q_N := \sum_{x,d} A_N(x)A_N(x+d)A_N(x+2d)$ denote the number of three-term progressions contained in A_N. Also let $\delta = \#A/N$ and let f_N be the balanced function of A_N. We have

$$Q_N\ =\ \sum_{x,d}(\delta+f_N(x))(\delta+f_N(x+d))(\delta+f_N(x+2d)) \tag{5.35}$$

$$=\ N^2\delta^3+\sum_{x,d} f_N(x)f_N(x+d)f_N(x+2d)+\text{(other terms)} \tag{5.36}$$

where, as in Lemma 5.6, each of the terms in (other terms) is a product

$$g_1(x)g_2(x+d)g_3(x+2d) \tag{5.37}$$

where at most two of the g_i are equal to f_N and the others are equal to δ. So again, every term in (other term) vanishes when summed over x and d. So we have by Lemma 5.6, Lemma 5.8, and the U^2-optimal hypothesis that

$$\left|Q_N - N^2\delta^3\right| = \left|\sum_{x,d} f_N(x)f_N(x+d)f_N(x+2d)\right| \tag{5.38}$$

$$\leq N^2||f_N||_{U^2} \tag{5.39}$$

$$\leq CN^2((\#A_N)^2N^{-3})^{1/4} \tag{5.40}$$

This shows that Q_N is guaranteed to be at least $(1/2)N^2\delta^3 = (1/2)(\#A_N)^3N^{-1}$ provided that $CN^2(\#A_N)^{1/2}N^{-3/4} < (1/2)N^2\delta^3$. This happens if

$$\#A_N > (2C)^{2/5}N^{9/10}, \tag{5.41}$$

so for any A_N satisfying (5.41), the number of nontrivial three-term progressions it contains is at least

$$\frac{1}{2}(\#A_N)^3N^{-1} - (\#A_N) = (\#A_N)\left(\frac{1}{2}(\#A_N)^2N^{-1} - 1\right), \tag{5.42}$$

which is positive for sufficiently large N. □

REFERENCES

[Be] F. Behrend. *On the sets of integers which contain no three in arithmetic progression.* Proc. Nat. Acad. Sci. **23** (1946), 331 – 332.

[Bl] T. F. Bloom. *A quantitative improvement for Roth's theorem on arithmetic progressions.* J. London Math. Soc. **93** (2016), 643 – 663.

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

[Er] P. Erdős. *Problems in number theory and combinatorics.* Proc. 6th Manitoba Conference on Numerical Math. (1976), 35 – 58.

[ET] P. Erdős and P. Turán. *On some sequences of integers.* J. London Math. Soc. **11** (1936), 261 – 264.

[FKO] H. Furstenberg, Y. Katznelson, and D. Ornstein. *The ergodic theoretical proof of Szemerédi's Theorem.* Bull. Amer. Math. Soc. **7**, no. 3 (1982), 527 – 552.

[Go1] W. T. Gowers. *A new proof of Szemerédi's Theorem for arithmetic progressions of length four.* Geom. Funct. Anal. **8** (1998), 529 – 551.

[Go2] W. T. Gowers. *A new proof of Szemerédi's Theorem.* Geom. Funct. Anal. **11** (2001), 465 – 588.

[Gr1] B. Green. *Roth's Theorem in the primes.* Ann. Math. **161** (2005), 1609 – 1636.

[Gr2] B. Green. *Montreal Lecture Notes on Quadratic Fourier Analysis*. Lecture notes. https://arxiv.org/pdf/math/0604089v2.pdf

[GT] B. Green and T. Tao. *The primes contain arbitrarily long arithmetic progressions.* Ann. Math. **167** (2008), 481 – 547.

[He] D. R. Heath-Brown. *Integer sets containing no arithmetic progressions.* J. London Math. Soc. **35** (1987), 383 – 394.

[IMS] A. Iosevich, A. Mayeli, and S. Senger. *Fourier bases: an elementary viewpoint on a variety of applications*. In preparation.

[Ly] N. Lyall. *Roth's Theorem, the Fourier analytic approach*. Lecture notes. http://alpha.math.uga.edu/%7Elyall/REU/Roth.pdf

[Ro1] K. Roth. *On certain sets of integers*. J. London Math Soc. **28** (1953), 104 – 109.

[Ro2] K. Roth. *Irregularities of sequences relative to arithmetic progressions. IV.* Period. Math. Hungar. **2** (1972), 301 – 326.

[Sa] T. Sanders. *On Roth's theorem on progressions.* Ann. Math. **174** (2011), 619 – 636.

[Sz1] E. Szemerédi. *On sets of integers containing no* 4 *elements in arithmetic progression.* Acta Math. Acad. Sci. Hungar. **20** (1969), 89 – 104.

[Sz2] E. Szemerédi. *On sets of integers containing no* k *elements in arithmetic progression.* Acta Arith. **27** (1975), 199 – 245.

[vdW] B. L. van der Waerden. *Beweis einer Baudetschen Vermutung.* Nieuw Arch. Wisk. **15** (1927), 212 – 216.