Graph Theory and Additive Combinatorics

Lecturer: Prof. Yufei Zhao

6 Roth's theorem

In Chapter 3.3, we proved Roth's theorem using Szemerédi regularity lemma via the triangle removal lemma. In this chapter, we will be instead be studying Roth's original proof of Roth's Theorem using Fourier analysis. First, let us recall the statement of Roth's Theorem. Let $r_3([N])$ denote the maximum size of a 3-AP-free subset of $[N]$. Then Roth's theorem states that $r_3([N]) = o(N)$.

One of the drawbacks of using Szemerédi regularity which shows an upper bound that is something like $\frac{N}{\log^* N}$. Roth's Fourier analytic proof would instead give us an upper bound of something like $\frac{N}{\log \log N}$, which is a much more reasonable bound.

Sanders (2011)
Bloom (2016)

Remark 6.1. The current best upper bound known is $r_3([N]) \leq N(\log N)^{1-o(1)}$ and the best lower bound known is $r_3(N) \geq Ne^{-O(\sqrt{\log N})}$ due to the Behrend construction. There is some evidence that seem to suggest that the lower bound is closer to truth, but closing the gap is still an open problem.

6.1 Roth's theorem in finite fields

We will begin by examining a finite field analogue to Roth's Theorem. Finite field models are a good sandbox for testing methods before applying to general integer cases; in particular, it is a good starting point because a lot of technicalities go away.

Let $r_3(\mathbb{F}_3^n)$ denote the maximum size of 3 AP-free subset of $\mathbb{F}_3^n$. Note that given x, y, z in $\mathbb{F}_3^n$, the following are equivalent:

- x, y, z for a 3 term arithmetic progression
- $x - 2y + z = 0$
- $x + y + z = 0$
- x, y, z form a line

- for all i, the ith coordinate of x, y, z are all distinct or all equal.

This is relevant to the game of SET, which can be thought of as finding 3 APs in $\mathbb{F}_3^4$.

We will state and prove a version of Roth's theorem in the finite field model. The proof is in the same spirit as the general Roth's theorem, but is slightly easier.

Meshulam (1995)

Theorem 6.2.

$$r_3(\mathbb{F}_3^n) = O\left(\frac{3^n}{n}\right)$$

The proof using triangle removal lemma copies verbatim so we can get $r_3(\mathbb{F}_3^n) = o(3^n)$ but the above theorem gives a better dependence.

We comment briefly on the history of this problem. In 2004, Edel found that $r_3(\mathbb{F}_3^n) \geq 2.21^n$. It was open for a long time whether $r_3(\mathbb{F}_3^n) = (3 - o(1))^n$. Recently, a surprising breakthrough showed that $r_3(\mathbb{F}_3^n) \leq 2.76^n$.

Edel (2004)

Croot, Lev, Pach (2016)
Ellenberg and Gijswijt (2016)

We had an energy increment argument during the proof of Szemerédi Regularity lemma. The strategy for Roth's theorem is a variant of energy increment. Instead, we will consider density increment. Given $A \subset \mathbb{F}_3^n$, we employ the follow strategy.

1. If A is pseudorandom (which we will see is equivalent to it being Fourier uniform, which roughly translates to all its Fourier coefficients are small) then there is a counting lemma which will show that A has lots of 3-AP.

2. If A is not pseudorandom, then we will show that A has a large Fourier coefficient. Then we can find a codimension 1 affine subspace (i.e. hyperplane) where density of A will increase. Now we consider A restricted to this hyperplane, and repeat the previous steps.

3. Each time we repeat, we obtain a density increment. Since density is bounded above by 1, this gives us a bounded number of steps.

Next, we recall some Fourier analytic ideas that will be important in our proof. In $\mathbb{F}_3^n$, we consider the Fourier characters $\gamma_r : \mathbb{F}_3^n \to \mathbb{C}$, indexed by $r \in \mathbb{F}_3^n$, which are defined to be $\gamma_r(x) = \omega^{r \cdot x}$ where $\omega = e^{2\pi i/3}$ and $r \cdot x = r_1 x_1 + \cdots + r_n x_n$. We define a ***Fourier transform***. For $f : \mathbb{F}_3^n \to \mathbb{C}$, the Fourier transform is given by $\widehat{f} : \mathbb{F}_3^n \to \mathbb{C}$ where

$$\widehat{f}(r) = \mathop{\mathbb{E}}_{x \in \mathbb{F}_3^n} f(x) \omega^{-r \cdot x} = \langle f, \gamma_r \rangle.$$

Effectively, the fourier transform is the inner product of f and the Fourier characters.

Remark 6.3. We use the following convention on normalization: in a finite group, for a physical space we will use average measure but in frequency we will always use sum measure.

We note some key properties of the Fourier transform.

Proposition 6.4. • $\widehat{f}(0) = \mathbb{E} f$

- *(Plancheral/Parseval)* $\mathbb{E}_{x \in \mathbb{F}_3^n} f(x)\overline{g(x)} = \sum_{r \in \mathbb{F}_3^n} \widehat{f}(r)\overline{\widehat{g}(r)}$.
- *(Inversion)* $f(x) = \sum_{r \in \mathbb{F}_3^n} \widehat{f}(r)\omega^{r \cdot x}$
- *(Convolution) Define* $(f * g)(x) = \mathbb{E}_y f(y)g(x-y)$. *Then we claim that* $\widehat{f * g}(x) = \widehat{f}(x)\widehat{g}(x)$.

To prove these properties notice that Fourier characters form an orthonormal basis. Indeed, we can check

$$\langle \gamma_r, \gamma_s \rangle = \mathbb{E}_x \gamma_r(x)\overline{\gamma_s(x)} = \mathbb{E}_x \omega^{-(r-s)\cdot x} = \begin{cases} 1 \text{ if } r = s, \\ 0 \text{ otherwise.} \end{cases}$$

If we think of Fourier transform as a unitary change of basis, inversion and Parseval's follows immediately. To see the formula for convolution, note that

$$\mathbb{E}_x (f * g)\omega^{r \cdot x} = \mathbb{E}_{x,y} f(y)g(x-y)\omega^{-r(y+(x-y))} = \mathbb{E}_r f(x)\omega^{-r\cdot x} \mathbb{E}_s g(x)\omega^{-s \cdot x}.$$

The following key identity relates Fourier transform with 3-APs.

Proposition 6.5. *If* $f, g, h : \mathbb{F}_3^n \to \mathbb{C}$, *then*

$$\mathbb{E}_{x,y} f(x)g(x+y)h(x+2y) = \sum_r \widehat{f}(r)\widehat{g}(-2r)\widehat{h}(r).$$

We will give two proofs of this proposition, with the second being more conceptual.

First proof. We expand the LHS using the formula for Fourier inversion.

$$\begin{aligned} LHS &= \mathbb{E}_{x,y} \left(\sum_{r_1} \widehat{f}(r_1)\omega^{r_1 \cdot x} \right) \left(\sum_{r_2} \widehat{g}(r_2)\omega^{r_2 \cdot (x+y)} \right) \left(\sum_{r_3} \widehat{h}(r_3)\omega^{r_3 \cdot (x+2y)} \right) \\ &= \sum_{r_1, r_2, r_3} \widehat{f}(r_1)\widehat{g}(r_2)\widehat{h}(r_3) \mathbb{E}_x \omega^{x \cdot (r_1+r_2+r_3)} \mathbb{E}_y \omega^{y \cdot (r_2 + 2r_3)} \\ &= \sum_r \widehat{f}(r)\widehat{g}(-2r)\widehat{h}(r) \end{aligned}$$

The last equality follows because

$$\mathbb{E}_x \omega^{x \cdot (r_1+r_2+r_3)} = \begin{cases} 1 & \text{if } r_1 + r_2 + r_3 = 0, \\ 0 & \text{otherwise} \end{cases}$$

and

$$\mathop{\mathbb{E}}_{y} \omega^{y\cdot(r_2+2r_3)} = \begin{cases} 1 & \text{if } r_2 + 2r_3 = 0, \\ 0 & \text{otherwise.} \end{cases} \qquad \square$$

Second proof. In this proof, we think of the LHS as a convolution.

$$\begin{aligned} \mathop{\mathbb{E}}_{x,y,z:x+y+z=0} f(x)g(y)h(z) &= (f * g * h)(0) \\ &= \sum_r \widehat{f * g * h}(r) \\ &= \sum_r \widehat{f}(r)\widehat{g}(r)\widehat{h}(r) \square \end{aligned}$$

In particular, note that if we take $f, g, h = 1_A$ where $A \subset \mathbb{F}_3^n$, then

$$3^{-2n}\#\{(x,y,z) \in A^3 : x + y + z = 0\} = \sum_r \widehat{1}_A(r)^3. \tag{6.1}$$

Remark 6.6. If $A = -A$ then this gives the same formula that counts closed walks of length 3 in Cayley graphs. In particular, $\{\widehat{1}_A(r) = r\}$ correspond eigenvalues of Cayley(G, A).

Lemma 6.7 (Counting Lemma). *If $A \subset \mathbb{F}_3^n$ with $|A| = \alpha 3^n$, let $\Lambda_3(A) = \mathbb{E}_{x,y} 1_A(x)1_A(x+y)1_A(x+2y)$. Then,*

$$\left|\Lambda_3(A) - \alpha^3\right| \le \alpha \max_{r\neq 0} \left|\widehat{1_A}(r)\right|.$$

Proof. By Proposition 6.5,

$$\Lambda_3(A) = \sum_r \widehat{1_A}(r)^3 = \alpha^3 + \sum_{r\neq 0} \widehat{1_A}(r)^3.$$

Therefore,

$$\begin{aligned} \left|\Lambda_3(A) - \alpha^3\right| &\le \sum_{r\neq 0} \left|\widehat{1_A}(r)\right|^3 \\ &\le \max_{r\neq 0} \left|\widehat{1_A}(r)\right| \cdot \sum_r \left|\widehat{1_A}(r)\right|^2 \\ &= \max_{r\neq 0} \left|\widehat{1_A}(r)\right| \cdot \mathbb{E}1_A^2 && \text{(Parseval)} \\ &= \alpha \max_{r\neq 0} \left|\widehat{1_A}(r)\right|. \end{aligned}$$

$\square$

Proof of Theorem 6.2. Let $N = 3^n$, the number of elements in $\mathbb{F}_3^n$.

Step 1. If the set is 3-AP free, then there is a large Fourier coefficient.

Lemma 6.8. *If A is 3-AP-free and $N \ge 2\alpha^{-2}$, then there is $r \neq 0$ such that $\left|\widehat{1_A}(r)\right| \ge \alpha^2/2$.*

Proof. By counting lemma and the fact that $\Lambda_3(A) = \dfrac{|A|}{N^2} = \dfrac{\alpha}{N}$,

$$\alpha \max_{r \neq 0} \left| \widehat{1_A}(r) \right| \geq \alpha^3 - \frac{\alpha}{N} \geq \frac{\alpha^3}{2}.$$

□

Step 2. Large Fourier coefficient implies density increment on a hyperplane.

Lemma 6.9. *If* $\left|\widehat{1_A}(r)\right| \geq \delta$ *for some* $r \neq 0$*, then* A *has density at least* $\alpha + \frac{\delta}{2}$ *when restricted to some hyperplane.*

Proof. We have

$$\begin{aligned} \widehat{1_A}(r) &= \mathop{\mathbb{E}}_{x \in \mathbb{F}_3^n} 1_A(x) w^{-r \cdot x} \\ &= \frac{1}{3}(\alpha_0 + \alpha_1 w + \alpha_2 w^2) \end{aligned}$$

where $\alpha_0, \alpha_1, \alpha_2$ are densities of A on the cosets of $r^{\perp}$. Notice that $\alpha = \frac{\alpha_0 + \alpha_1 + \alpha_2}{3}$. By triangle inequality,

$$\begin{aligned} 3\delta &\leq \left| \alpha_0 + \alpha_1 w + \alpha_2 w^2 \right| \\ &= \left| (\alpha_0 - \alpha) + (\alpha_1 - \alpha) w + (\alpha_2 - \alpha) w^2 \right| \\ &\leq \sum_{j=0}^{2} |\alpha_j - \alpha| \\ &\leq \sum_{j=0}^{2} (|\alpha_j - \alpha| + (\alpha_j - \alpha)). \end{aligned}$$

(This final step is a trick that will be useful in the next section.) Note that every term in the last summation is non-negative. Consequently, there exists j such that $\delta \leq |\alpha_j - \alpha| + (\alpha_j - \alpha)$. Then, $\alpha_j \geq \alpha + \dfrac{\delta}{2}$. □

Step 3 : Iterate density increment.

So far, we have that if A is 3-AP-free and $N \geq 2\alpha^{-2}$, then A has density at least $\alpha + \alpha^2/4$ on some hyperplane. Let our initial density be $\alpha_0 = \alpha$. At the i-th step, we restrict A to some hyperplane, so that the restriction of A inside the smaller space has density

$$\alpha_i \geq \alpha_{i-1} + \alpha_{i-1}^2/4.$$

Let $N_i = 3^{n-i}$. We can continue at step i as long as $N_i \geq 2\alpha_i^{-2}$.

We note that the first index i_1 such that $\alpha_{i_1} \geq 2\alpha_0$ satisfies $i_1 \leq \frac{4}{\alpha} + 1$. This is because $\alpha_{i+1} \geq \alpha + i\frac{\alpha^2}{4}$. Similar calculations shows that if i_ℓ is the first index such that $\alpha_{i_\ell} \geq 2^\ell \alpha_0$ then

$$i_\ell \leq \frac{4}{\alpha} + m\frac{2}{\alpha} + \cdots + \frac{4}{2^{\ell-1}\alpha} + \ell \leq \frac{8}{\alpha} + \log_2 \frac{1}{\alpha}.$$

Suppose the process terminates after m steps with density α_m. Then we find that the size of the subspace in the last step is given by $3^{n-m} < 2\alpha_m^{-2} \leq 2\alpha^{-2}$. So

$$n \leq \frac{8}{\alpha} + \log_3\left(\frac{2}{\alpha^2}\right) = O\left(\frac{1}{\alpha}\right)$$

Thus $\frac{|A|}{N} = \alpha = O(1/n)$. Equivalently, $|A| = \alpha N = O\left(\frac{3^n}{n}\right)$ as desired. □

Remark 6.10. This proof is much more difficult in integers, because there is no subspace to pass down to.

A natural question is whether this technique can be generalized to bound 4-AP counts. In the regularity-based proof of Roth's theorem, we saw that the graph removal lemma was not sufficient, and we actually needed hypergraph regularity and a hypergraph removal lemma to govern 4-AP counts. Similarly, while the counting lemma developed here shows that Fourier coefficients control 3-AP counts, they do not in fact control 4-AP counts. For example, consider the set $A = \{x \in \mathbb{F}_5^n : x \cdot x = 0\}$. One can show that the nonzero Fourier coefficients corresponding to A are all small. However, one can also show that A has the wrong number of 4-APs, thus implying that Fourier coefficients cannot control 4-AP counts. The field of higher-order Fourier analysis, namely quadratic Fourier analysis, was developed by Gowers specifically to extend this proof of Roth's Theorem to prove Szemeredi's Theorem for larger APs. An example of quadratic Fourier analysis is given by the following theorem.

Gowers (1998)

Theorem 6.11 (Inverse theorem for quadratic Fourier analysis)**.** *For all $\delta > 0$, there exists a constant $c(\delta) > 0$ such that if $A \subset \mathbb{F}_5^n$ has density α, and $|\Lambda_4(A) - \alpha^4| > \delta$, then there exists a non-zero quadratic polynomial $f(x_1, \ldots, x_n)$ over $\mathbb{F}_5$ satisfying*

$$|\mathbb{E}_{x \in \mathbb{F}_5^n} 1_A(x) \omega^{f(x)}| \geq c(\delta).$$

6.2 *Roth's proof of Roth's theorem in the integers*

In Section 6.1 we saw the proof of Roth's theorem in the finite field setting, specifically for the set $\mathbb{F}_3^n$. We will now extend this analysis to prove the following bound, which will imply Roth's theorem in the integers:

Theorem 6.12.

Roth (1953)

$$r_3([N]) = O\left(\frac{N}{\log\log N}\right)$$

.

The subsequent proof of this bound is the original one given by Roth himself. Recall that the proof of Roth's theorem in finite fields had the following 3 steps:

1. Show that a 3-AP-free set admits a large Fourier coefficient.
2. Deduce that there must exist a subspace with a density increment.
3. Iterate the density increment to upper bound the size of a 3-AP free set.

The proof of Roth's theorem on the integers will follow the same 3 steps. However, the execution will be quite different. The main difference lies in step 2, where there is no obvious notion of a subspace of $[N]$.

Previously we defined Fourier analysis in terms of the group $\mathbb{F}_3^n$. There is a general theory of Fourier analysis on Abelian groups which relates a group G to its set of characters $\widehat{G}$, also referred to as its dual group. For now, however, we work with the group $\mathbb{Z}$.

The dual group of $\mathbb{Z}$ is $\widehat{\mathbb{Z}} = \mathbb{R}/\mathbb{Z}$. The Fourier Transform of a function $f : \mathbb{Z} \to \mathbb{C}$ is given by the function $\widehat{f} : \mathbb{R}/\mathbb{Z} \to \mathbb{C}$ satisfying

$$\widehat{f}(\theta) = \sum_{x \in \mathbb{Z}} f(x) e(-x\theta),$$

where $e(t) = e^{2\pi i t}$. This is commonly referred to as the Fourier series of f.

As they were in $\mathbb{F}_3^n$, the following identities are also true in $\mathbb{Z}$. Their proofs are the same.

- $\widehat{f}(0) = \sum_{x \in \mathbb{Z}} f(x)$
- (Plancherel/Parseval) $\sum_{x \in \mathbb{Z}} f(x)\overline{g(x)} = \int_0^1 \widehat{f}(\theta)\overline{\widehat{g}(\theta)} d\theta$
- (Inversion) $f(x) = \int_0^1 \widehat{f}(\theta) e(x\theta) d\theta$
- Define $\Lambda(f, g, h) = \sum_{x,y \in \mathbb{Z}} f(x)g(x+y)h(x+2y)$. Then

$$\Lambda(f, g, h) = \int_0^1 \widehat{f}(\theta)\widehat{g}(-2\theta)\widehat{h}(\theta) d\theta.$$

In the finite field setting, we defined a counting lemma, which showed that if two functions had similar Fourier transforms, then they had a similar number of 3-APs. We can define an analogue to the counting lemma in $\mathbb{Z}$ as well.

Theorem 6.13 (Counting Lemma). *Let $f, g : \mathbb{Z} \to \mathbb{C}$ such that $\sum_{n \in \mathbb{Z}} |f(n)|^2, \sum_{n \in \mathbb{Z}} |g(n)|^2 \leq M$. Define $\Lambda_3(f) = \Lambda(f, f, f)$. Then*

$$|\Lambda_3(f) - \Lambda_3(g)| \leq 3M \left\| \widehat{f - g} \right\|_\infty .$$

Proof. We can rewrite

$$\Lambda_3(f) - \Lambda_3(g) = \Lambda(f-g, f, f) + \Lambda(g, f-g, f) + \Lambda(g, g, f-g).$$

We want to show that each of these terms is small when $f-g$ has small Fourier coefficients. We know that

$$\begin{aligned}
|\Lambda(f-g, f, f)| &= \left| \int_0^1 \widehat{(f-g)}(\theta)\widehat{f}(-2\theta)\widehat{f}(\theta)d\theta \right| \\
&\leq \left\| \widehat{f-g} \right\|_\infty \left| \int_0^1 \widehat{f}(-2\theta)\widehat{f}(\theta)d\theta \right| \quad \text{(triangle inequality)} \\
&\leq \left\| \widehat{f-g} \right\|_\infty \left(\int_0^1 |\widehat{f}(-2\theta)|^2 d\theta \right)^{1/2} \left(\int_0^1 |\widehat{f}(\theta)|^2 d\theta \right)^{1/2} \\
&\qquad \text{(Cauchy-Schwarz)} \\
&\leq \left\| \widehat{f-g} \right\|_\infty \left(\sum_{x\in\mathbb{Z}} |f(x)|^2 \right) \quad \text{(Plancherel)} \\
&\leq M \left\| \widehat{f-g} \right\|_\infty .
\end{aligned}$$

Bounding the other two terms is identical. □

We can now proceed with proving Roth's Theorem.

Proof of Theorem 6.12. We follow the same 3 steps as in the finite field setting.

Step 1: 3-AP free sets induce a large Fourier coefficient

Lemma 6.14. *Let $A \subset [N]$ be a 3-AP free set, $|A| = \alpha N$, $N \geq 5/\alpha^2$. Then there exists $\theta \in \mathbb{R}$ satisfying*

$$\left| \sum_{n=1}^{N} (1_A - \alpha)(n) e(\theta n) \right| \geq \frac{\alpha^2}{10} N$$

Proof. Since A has no 3-AP, the quantity $1_A(x)1_A(x+y)1_A(x+2y)$ is nonzero only for trivial APs, i.e. when $y=0$. Thus $\Lambda_3(1_A) = |A| = \alpha N$. Now consider $\Lambda_3(1_{[N]})$. This counts the number of 3-APs in $[N]$. We can form a 3-AP by choosing the first and third elements from $[N]$, assuming they are the same parity. Therefore $\Lambda_3(1_{[N]}) \geq N^2/2$. Now, we apply the counting lemma to $f = 1_A, g = \alpha 1_{[N]}$

Remark 6.15. The spirit of this whole proof is the theme of structure versus pseudorandomness, an idea we also saw in our discussion graph regularity. If A is "pseudorandom", then we wish to show that A has small Fourier coefficients. But that would indicate that f and g have similar Fourier coefficients, implying that A has many 3-AP counts, which is a contradiction. Thus A cannot be pseudorandom, it must have some structure.

Applying Theorem 6.13 yields (where we use the notation $f^\wedge = \widehat{f}$)

$$\frac{\alpha^3 N^2}{2} - \alpha N \leq 3\alpha N \left\| (1_A - \alpha 1_{[N]})^\wedge \right\|_\infty$$

and thus

$$\begin{aligned} \left\| (1_A - \alpha 1_{[N]})^\wedge \right\|_\infty &\geq \frac{\frac{1}{2}\alpha^3 N^2 - \alpha N}{3\alpha N} \\ &= \frac{1}{6}\alpha^2 N - \frac{1}{3} \\ &\geq \frac{1}{10}\alpha^2 N, \end{aligned}$$

where in the last inequality we used the fact that $N \geq 5/\alpha^2$. Therefore there exists some θ with

$$\left| \sum_{n=1}^{N} (1_A - \alpha)(n) e(\theta n) \right| = (1_A - \alpha 1_{[N]})^\wedge(\theta) \geq \frac{1}{10}\alpha^2 N,$$

as desired. □

Step 2: A large Fourier coefficient produces a density increment.

In the finite field setting our Fourier coefficients corresponded to hyperplanes. We were then able to show that there was a coset of a hyperplane with large density. Now, however, θ is a real number. There is no concept of a hyperplane in $[N]$, so how can we chop up $[N]$ in order to use the density increment?

On each coset of the hyperplane each character was exactly constant. This motivates us to partition $[N]$ into sub-progressions such that the character $x \mapsto e(x\theta)$ is roughly constant on each subprogression.

As a simple example, assume that θ is a rational a/b for some fairly small b. Then $x \mapsto e(x\theta)$ is constant on arithmetic progressions with common difference b. Thus we could partition $[N]$ into arithmetic progressions with common difference b.

Before formalizing this idea, we require the following classical lemma from Dirichlet.

Lemma 6.16. *Let $\theta \in \mathbb{R}$ and $0 < \delta < 1$. Then there exists a positive integer $d \leq 1/\delta$ such that $\|d\theta\|_{\mathbb{R}/\mathbb{Z}} \leq \delta$ (here, $\|\cdot\|_{\mathbb{R}/\mathbb{Z}}$ is defined as the distance to the nearest integer).*

Proof. Pigeonhole principle. Let $m = \left\lfloor \frac{1}{\delta} \right\rfloor$. Consider the $m + 1$ numbers $0, \theta, \cdots, m\theta$. By the pigeonhole principle, there exist i, j such that the fractional parts of $i\theta$ and $j\theta$ differ by at most δ. Setting $d = |i - j|$ gives us $\|d\theta\|_{\mathbb{R}/\mathbb{Z}} \leq \delta$, as desired. □

The next lemma formalizes our previous intuition for partitioning $[N]$ into subprogressions such that the map $x \mapsto e(x\theta)$ is roughly constant on each progression.

Lemma 6.17. *Let $0 < \eta < 1$ and $\theta \in \mathbb{R}$. Suppose $N > C\eta^{-6}$ (for some universal constant C). Then one can partition $[N]$ into sub-APs P_i, each with length $N^{1/3} \leq |P_i| \leq 2N^{1/3}$, such that $\sup_{x,y\in P_i} |e(x\theta) - e(y\theta)| < \eta$ for all i.*

Proof. By Lemma 6.16, there exists an integer $d \leq \frac{4\pi N^{1/3}}{\eta}$ such that $\|d\theta\|_{\mathbb{R}/\mathbb{N}} \leq \frac{\eta}{4\pi N^{1/3}}$. Since $N > C\eta^{-6}$, for $C = (4\pi)^6$ we get that $d < \sqrt{N}$. Therefore we can partition $[N]$ into APs with common difference d, each with lengths between $N^{1/3}$ and $2N^{1/3}$. Then inside each sub-AP P, we have that

$$\sup_{x,y\in P} |e(x\theta) - e(y\theta)| \leq |P||e(d\theta) - 1| \leq 2N^{1/3} \cdot 2\pi \,\|d\theta\|_{\mathbb{R}/\mathbb{Z}} \leq \eta,$$

where we get the inequality $|e(d\theta) - 1| \leq 2\pi \,\|d\theta\|_{\mathbb{R}/\mathbb{Z}}$ from the fact that the length of a chord is at most the length of the corresponding arc. □

We can now apply this lemma to obtain a density increment.

Lemma 6.18. *Let $A \subset [N]$ be 3-AP-free, with $|A| = \alpha N$ and $N > C\alpha^{-12}$. Then there exists a sub-AP $P \subset [N]$ with $|P| \geq N^{1/3}$ and $|A \cap P| \geq (\alpha + \alpha^2/40)|P|$.*

Proof. By Lemma 6.14, there exists θ satisfying $|\sum_{x=1}^{N} (1_A - \alpha)(x)e(x\theta)| \geq \alpha^2 N/10$. Next, apply Lemma 6.17 with $\eta = \alpha^2/20$ to obtain a partition $P_1, \ldots, P_k$ of $[N]$ satisfying $N^{1/3} \leq |P_i| \leq 2N^{1/3}$. We then get that

$$\frac{\alpha^2}{10}N \leq \left|\sum_{x=1}^{N} (1_A - \alpha)(x)e(x\theta)\right| \leq \sum_{i=1}^{k} \left|\sum_{x\in P_i} (1_A - \alpha)(x)e(x\theta)\right|.$$

For $x, y \in P_i$, $|e(x\theta) - e(y\theta)| \leq \alpha^2/20$. Therefore we have that

$$\left|\sum_{x\in P_i} (1_A - \alpha)(x)e(x\theta)\right| \leq \left|\sum_{x\in P_i} (1_A - \alpha)(x)\right| + \frac{\alpha^2}{20}|P_i|.$$

Altogether,

$$\begin{aligned}\frac{\alpha^2}{10}N &\leq \sum_{i=1}^{k} \left(\left|\sum_{x\in P_i} (1_A - \alpha)(x)\right| + \frac{\alpha^2}{20}|P_i|\right) \\ &= \sum_{i=1}^{k} \left|\sum_{x\in P_i} (1_A - \alpha)(x)\right| + \frac{\alpha^2}{20}N\end{aligned}$$

Thus

$$\frac{\alpha^2}{20}N \leq \sum_{i=1}^{k} \left| \sum_{x \in P_i} (1_A - \alpha)(x) \right|$$

and hence

$$\frac{\alpha^2}{20} \sum_{i=1}^{k} |P_i| \leq \sum_{i=1}^{k} \big| |A \cap P_i| - \alpha |P_i| \big|.$$

We want to show that there exists some P_i such that A has a density increment when restricted to P_i. Naively bounding the RHS of the previous sum does not guarantee a density increment, so we use the following trick

$$\begin{aligned}\frac{\alpha^2}{20} \sum_{i=1}^{k} |P_i| &\leq \sum_{i=1}^{k} \big| |A \cap P_i| - \alpha |P_i| \big| \\ &= \sum_{i=1}^{k} \left(\big| |A \cap P_i| - \alpha |P_i| \big| + |A \cap P_i| - \alpha |P_i| \right).\end{aligned}$$

Thus there exists an i such that

$$\frac{\alpha^2}{20}|P_i| \leq \big| |A \cap P_i| - \alpha |P_i| \big| + |A \cap P_i| - \alpha |P_i|.$$

Since the quantity $|x| + x$ is always strictly greater than 0, this i must satisfy $|A \cap P_i| - \alpha |P_i| \geq 0$, and thus we have

$$\frac{\alpha^2}{20}|P_i| \leq 2(|A \cap P_i| - \alpha |P_i|),$$

which yields

$$|A \cap P_i| \geq (\alpha + \frac{\alpha^2}{40})|P_i|.$$

Thus we have found a subprogression with a density increment, as desired. □

Step 3: Iterate the density increment.

Step 3 is very similar to the finite field case. Let our initial density be $\alpha_0 = \alpha$, and the density after each iteration be α_i. We have that $\alpha_{i+1} \geq \alpha_i + \alpha_i^2/40$, and that $\alpha_i \leq 1$. We double α (i.e. reach T such that $\alpha_T \geq 2\alpha_0$) after at most $40/\alpha + 1$ steps. We double α again (i.e. go from $2\alpha_0$ to $4\alpha_0$) after at most $20/\alpha + 1$ steps. In general, the kth doubling requires at most $\frac{40}{2^{k-1}\alpha}$ steps. There are at most $\log_2(1/\alpha) + 1$ doublings, as α must remain less than 1. Therefore the total number of iterations must be $O(1/\alpha)$.

Lemma 6.18 shows that we can pass to a sub-AP and increment the density whenever $N_i > C\alpha^{-12}$. Therefore if the process terminates at step i, we must have $N_i \leq C\alpha_i^{-12} \leq C\alpha^{-12}$. Each iteration reduces the size of our set by at most a cube root, so

$$N \leq N_i^{3^i} \leq (C\alpha^{-12})^{3^{O(1/\alpha)}} = e^{e^{O(1/\alpha)}}.$$

Therefore $\alpha = O(1/\log\log N)$ and $|A| = \alpha N = O(N/\log\log N)$, as desired. □

Remark 6.19. This is the same proof in spirit as last time. A theme in additive combinatorics is that the finite field model is a nice playground for most techniques.

Let us compare this proof strategy in both $\mathbb{F}_3^n$ and $[N]$. We saw that $r_3(\mathbb{F}_3^n) = O(N/\log N)$. However, the bound for $[N]$ is $O(N/\log\log N)$, which is weaker by a log factor. Where does this stem from? Well, in the density increment step for $\mathbb{F}_3^n$, we were able to pass down to a subset which had size a constant factor of the original one. However, in $[N]$, each iteration gives us a subprogression which has size equal to the cube root of the previous subspace. This poses a natural question—is it possible to pass down to subsprogressions of $[N]$ which look more like subspaces? It turns out that this is indeed possible.

For a subset $S \subset \mathbb{F}_3^n$, we can write its *orthogonal complement* as

$$U_S = \{x \in \mathbb{F}_3^n : x \cdot s = 0 \text{ for all } s \in S\}.$$

In $[N]$, the analogous concept is known as a ***Bohr set***, an idea developed by Bourgain to transfer the proof in Section 6.1 to $\mathbb{Z}$. This requires us to work in $\mathbb{Z}/N\mathbb{Z}$. For some subset $S \subset \mathbb{Z}/N\mathbb{Z}$, we can define its Bohr set as

Bourgain, 1999

$$\operatorname{Bohr}(S, \epsilon) = \left\{x \in \mathbb{Z}/N\mathbb{Z} : \left\|\frac{sx}{N}\right\| \le \epsilon \text{ for all } s \in S\right\}.$$

This provides a more natural analogy to subspaces, and is the basis for modern improvements on bounds to Roth's Theorem. We will study Bohr sets in relation to Freiman's Theorem in Chapter 7.

6.3 *The polynomial method proof of Roth's theorem in the finite field model*

Currently, the best known bound for Roth's Theorem in $\mathbb{F}_3^n$ is the following:

Theorem 6.20. $r_3(\mathbb{F}_3^n) = O(2.76^n)$.

Ellenberg and Gijswijt (2017)

This bound improves upon the $O(3^n/n^{1+\epsilon})$ bound (for some $\epsilon > 0$) proved earlier by Bateman and Katz. Bateman and Katz used Fourier-analytic methods to prove their bound, and until very recently, it was open whether the upper bound could be improved to a power-saving one (one of the form $O(c^n)$ for $c < 3$), closer to the lower bound given by Edel of 2.21^n.

Bateman and Katz (2012)

Edel (2004)

Croot–Lev–Pach gave a similar bound for 3-APs over $(\mathbb{Z}/4\mathbb{Z})^n$, proving that the maximum size of a set in $(\mathbb{Z}/4\mathbb{Z})^n$ with no 4-APs is

$O(3.61^n)$. They used a variant of the polynomial method, and their proof was made easier by the fact that there are elements of order 2. Ellenberg and Gijswijt used the Croot–Lev–Pach method, as it is often referred to in the literature, to prove the bound for $\mathbb{F}_3^n$.

Croot, Lev, and Pach (2017)

We will use a formulation that appears on Tao's blog.

Tao (2016)

Let $A \subseteq \mathbb{F}_3^n$ be 3-AP-free (this is sometimes known as a *cap set* in the literature). Then we have the identity

$$\delta_0(x+y+z) = \sum_{a \in A} \delta_a(x)\delta_a(y)\delta_a(z) \tag{6.2}$$

for $x, y, z \in A$, where δ_a is the Dirac delta function, defined as follows:

$$\delta_a(x) := \begin{cases} 1 & \text{if } x = a, \\ 0 & \text{if } x \neq a. \end{cases}$$

Note that (6.2) holds because $x+y+z=0$ if and only if $z-y = y-x$ in $\mathbb{F}_3^n$, meaning that x, y, z form an arithmetic progression, which is only possible if $x = y = z = a$ for some $a \in \mathbb{F}_3^n$.

We will show that the left-hand side of (6.2) is "low-rank" and the right-hand side is "high-rank" in a sense we explain below.

Recall from linear algebra the classical notion of rank: given a function $F\colon A \times A \to \mathbb{F}$, for a field $\mathbb{F}$, we say F is rank 1 if it is nonzero and can be written in the form $F(x,y) = f(x)g(y)$ for some functions $f, g\colon A \to \mathbb{F}$. In general, we define $\operatorname{rank} F$ to be the minimum number of rank 1 functions required to write F as a linear combination of rank 1 functions. We can view F as a matrix.

How should we define the rank of a function $F\colon A \times A \times A \to \mathbb{F}$? We might try to extend the above notion by defining such a function F to be rank 1 if $F(x,y,z) = f(x)g(y)h(z)$, known as *tensor rank*, but this is not quite what we want. Instead, we say that F has ***slice-rank* 1** if it is nonzero and it can be written in one of the forms $f(x)g(y,z)$, $f(y)g(x,z)$, or $f(z)g(x,y)$. In general, we say the ***slice-rank*** of F is the minimum number of slice-rank 1 functions required to write F as a linear combination. For higher powers of A, we generalize this definition accordingly.

What is the rank of a diagonal function? Recall from linear algebra that the rank of a diagonal matrix is the number of nonzero entries. A similar result holds true for the slice-rank.

Lemma 6.21. *If $F\colon A \times A \times A \to \mathbb{F}$ equals*

$$F(x,y,z) = \sum_{a \in A} c_a \delta_a(x)\delta_a(y)\delta_a(z),$$

then

$$\operatorname{slice-rank} F = |\{a \in A : c_a \neq 0\}|.$$

Here the coefficients c_a correspond to diagonal entries.

Proof. It is clear that slice-rank $F \leq |\{a \in A : c_a \neq 0\}|$, as we can write F as a sum of slice-rank 1 functions by

$$F(x,y,z) = \sum_{\substack{a \in A \\ c_a \neq 0}} c_a \delta_a(x)(\delta_a(y)\delta_a(z)).$$

For the other direction, assume that all diagonal entries are nonzero; if $c_a = 0$ for some a, then we can remove a from A without increasing the slice-rank. Now suppose slice-rank $F < |A|$. So we can write

$$\begin{aligned} F(x,y,z) = {} & f_1(x)g_1(y,z) + \cdots + f_\ell(x)g_\ell(y,z) \\ & + f_{\ell+1}(y)g_{\ell+1}(x,z) + \cdots + f_m(y)g_m(x,z) \\ & + f_{m+1}(z)g_{m+1}(x,y) + \cdots + f_{|A|-1}(z)g_{|A|-1}(x,y). \end{aligned}$$

Claim 6.22. There exists $h \colon A \to \mathbb{F}_3$ with $|\operatorname{supp} h| > m$ such that

$$\sum_{z \in A} h(z) f_i(z) = 0 \tag{6.3}$$

for all $i = m+1, \ldots, |A| - 1$.

Here $\operatorname{supp} h$ is the set $\{z \in A : h(z) \neq 0\}$.

Proof. In the vector space of functions $A \to \mathbb{F}_3$, the set of h satisfying (6.3) for all $i = m+1, \ldots, |A| - 1$ is a subspace of dimension greater than m. Furthermore, we claim that every subspace of dimension $m + 1$ has a vector whose support has size at least $m + 1$. For a subspace X of dimension $m + 1$, suppose we write $m + 1$ vectors forming a basis of X in an $|A| \times (m+1)$ matrix Y. Then, this matrix has rank $m + 1$, so there must be some non-vanishing minor of order $m + 1$; that is, we can delete some rows of Y to get an $(m+1) \times (m+1)$ matrix with nonzero determinant. If the column of this matrix are the vectors v_1 through v_{m+1}, then these vectors generate all of $\mathbb{F}_3^{m+1}$. In particular, some linear combination of $v_1, v_2, \ldots, v_{m+1}$ is equal to the vector of all ones, which has support $m + 1$. So, taking that linear combination of the original vectors (the columns of Y) gives a vector of support at least $m + 1$. □

Pick the h from the claim. We find

$$\sum_{z \in A} F(x,y,z)h(z) = \sum_{a \in A} \sum_{z \in A} c_a \delta_a(x)\delta_a(y)\delta_a(z)h(z) = \sum_{a \in A} c_a h(a)\delta_a(x)\delta_a(y),$$

but also

$$\begin{aligned} \sum_{z \in A} F(x,y,z)h(z) = {} & f_1(x)\widetilde{g_1}(y) + \cdots + f_\ell(x)\widetilde{g_\ell(y)} \\ & + f_{\ell+1}(y)\widetilde{g_{\ell+1}}(x) + \cdots + f_m(y)\widetilde{g_m}(x), \end{aligned}$$

where $\widetilde{g_i}(y) = \sum_{z\in A} g_i(y,z)h(z)$ for $1 \le i \le \ell$, and
$\widetilde{g_i}(x) = \sum_{z\in A} g_i(x,z)h(z)$ for $\ell + 1 \le i \le m$. Thus

$$\sum_{a\in A} c_a h(a)\delta_a(x)\delta_a(y) = f_1(x)\widetilde{g_1}(y) + \cdots + f_\ell(x)g_\ell(y)$$
$$+ f_{\ell+1}(y)\widetilde{g_{\ell+1}}(x) + \cdots + f_m(y)\widetilde{g_m}(x).$$

Note the left-hand side has more than m diagonal entries (namely the a where $h(a) \neq 0$), but the left-hand side has rank at most m, which is a contradiction as we have reduced to the 2-dimensional case. □

Using induction, we can easily generalize (from 3 variables) to any finite number of variables, the proof of which we omit.

We have thus proved that the slice-rank of the right hand side of (6.2) is $|A|$, and is therefore "high-rank." We now show that the left hand side has "low-rank."

Lemma 6.23. *Define* $F\colon A \times A \times A \to \mathbb{F}_3$ *as follows:*

$$F(x + y + z) := \delta_0(x + y + z).$$

Then slice-rank $F \le 3M$*, where*

$$M := \sum_{\substack{a,b,c\ge 0\\ a+b+c=n\\ b+2c\le 2n/3}} \frac{n!}{a!b!c!}.$$

Proof. In $\mathbb{F}_3$, one has $\delta_0(x) = 1 - x^2$. Applying this coordinate-wise,

$$\delta_0(x + y + z) = \prod_{i=1}^{n}(1 - (x_i + y_i + z_i)^2), \tag{6.4}$$

where the x_i are the coordinates of $x \in \mathbb{F}_3^n$, and so on. If we expand the right-hand side, we obtain a polynomial in $3n$ variables with degree $2n$. We find a sum of monomials, each of the form

$$x_1^{i_1} \cdots x_n^{i_n} y_1^{j_1} \cdots y_n^{j_n} z_1^{k_1} \cdots z_n^{k_n},$$

where $i_1, i_2, \ldots, i_n, j_1, \ldots, j_n, k_1, \ldots, k_n \in \{0, 1, 2\}$. Group these monomials. For each term, by the pigeonhole principle, at least one of $i_1 + \cdots + i_n, j_1 + \cdots + j_n, k_1 + \cdots + k_n$ is at most $2n/3$.

We can write (6.4) as a sum of monomials, which we write explicitly as

$$\prod_{i=1}^{n}(1 - (x_i + y_i + z_i)^2) = \sum_{\substack{i_1,i_2,\ldots,i_n\\ j_1,j_2,\ldots,j_n\\ k_1,k_2,\ldots,k_n}} c_{i_1,\ldots,i_n,j_1,\ldots,j_n,k_1,\ldots,k_n} x_1^{i_1} \cdots x_n^{i_n} y_1^{j_1} \cdots y_n^{j_n} z_1^{k_1} \cdots z_n^{k_n} \tag{6.5}$$

where $c_{i_1,\ldots,i_n,j_1,\ldots,j_n,k_1,\ldots,k_n}$ is a coefficient in $\mathbb{F}_3$. Then, we can group terms to write (6.5) as a sum of slice-rank 1 functions in the following way:

$$\begin{aligned}\prod_{i=1}^{n}(1-(x_i+y_i+z_i)^2) = &\sum_{i_1+\cdots+i_n\le\frac{2n}{3}} x_1^{i_1}\cdots x_n^{i_n} f_{i_1,\ldots,i_n}(y,z)\\ &+\sum_{j_1+\cdots+j_n\le\frac{2n}{3}} y_1^{j_1}\cdots y_n^{j_n} g_{j_1,\ldots,j_n}(x,z)\\ &+\sum_{k_1+\cdots+k_n\le\frac{2n}{3}} z_1^{k_1}\cdots z_n^{k_n} h_{k_1,\ldots,k_n}(x,y),\end{aligned}$$

where

$$f_{i_1,\ldots,i_n}(y,z) = \sum_{\substack{j_1,j_2,\ldots,j_n\\ k_1,k_2,\ldots,k_n}} c_{i_1,\ldots,i_n,j_1,\ldots,j_n,k_1,\ldots,k_n} y_1^{j_1}\cdots y_n^{j_n} z_1^{k_1}\cdots z_n^{k_n},$$

and $g_{j_1,\ldots,j_n}(x,z)$ and $h_{k_1,\ldots,k_n}(x,y)$ are similar except missing some terms to avoid overcounting.

So, each monomial with degree at most $2n/3$ contributes to the slice-rank 3 times, and the number of such monomials is at most M. Thus the slice-rank is at most $3M$. □

We would like to estimate M. If we let $0 \le x \le 1$, we see that $Mx^{2n/3} \le (1+x+x^2)^n$ if we expand the right-hand side. Explicitly,

$$Mx^{2n/3} \le \sum_{\substack{a,b,c\ge 0\\ a+b+c=n\\ b+2c\le 2n/3}} x^{b+2c}\frac{n!}{a!b!c!} \le (1+x+x^2)^n.$$

So

$$M \le \inf_{0<x<1}\frac{(1+x+x^2)^n}{x^{2n/3}} \le (2.76)^n,$$

where we plug in $x = 0.6$.

Alternatively, we could Stirling's formula, which would give the same bound.

When this proof came out, people were shocked; this was basically a four-page paper, and demonstrated the power of algebraic methods. However, these methods seem more fragile compared to the Fourier-analytic methods we used last time. It is an open problem to extend this technique to prove a power-saving upper-bound for the size of a 4-AP-free subset of $\mathbb{F}_5^n$ (in the above arguments, we can replace $\mathbb{F}_3$ with any other finite field, so the choice of field does not really matter). It is also open to extend the polynomial method to corner-free sets in $\mathbb{F}_2^n \times \mathbb{F}_2^n$, where corners are sets of the form $\{(x,y),(x+d,y),(x,y+d)\}$, or to the integers.

6.4 Roth's theorem with popular differences

After giving a new method for 3-APs in $\mathbb{F}_3^n$ that gave a much better bound than Fourier analysis, we will now give a different proof that gives a much worse bound, but has strong consequences.

This theorem involves a "popular common difference."

Theorem 6.24. *For all $\epsilon > 0$, there exists $n_0 = n_0(\epsilon)$ such that for $n \geq n_0$ and every $A \subseteq \mathbb{F}_3^n$ with $|A| = \alpha 3^n$, there exists $y \neq 0$ such that*

Green (2005)

$$|\{x : x, x+y, x+2y \in A\}| \geq (\alpha^3 - \epsilon)3^n.$$

Here y is the popular common difference; this theorem obtains a lower bound on the number of 3-APs with common difference y in A. Note that $\alpha^3 3^n$ is roughly the expected number of 3-APs with common difference y if A is a random subset of $\mathbb{F}_3^n$ with size $\alpha 3^n$. The theorem states we can find some y such that the number of 3-APs with common difference y is close to what we expect in a random set, and suggests that it is not true that the number of 3-APs is at least what we would expect in a random set.

Green showed that the theorem is true with $n_0 = \text{tow}((1/\epsilon)^{O(1)})$. This bound was improved by Fox–Pham to $n_0 = \text{tow}(O(\log \frac{1}{\epsilon}))$, using the regularity method. They showed that this bound is tight; this is an instance in which the regularity method gives the right bounds, which is interesting. This is the bound we will show.

Fox and Pham (2019+)

Lemma 6.25 (Bounded increments)**.** *Let $\alpha, \epsilon > 0$. If $\alpha_0, \alpha_1, \ldots \in [0,1]$ such that $\alpha_0 \geq \alpha$, then there exists $k \leq \lceil \log_2 \frac{1}{\epsilon} \rceil$ such that $2\alpha_k - \alpha_{k+1} \geq \alpha^3 - \epsilon$.*

Proof. Otherwise, $\alpha_1 \geq 2\alpha_0 - \alpha^3 + \epsilon \geq \alpha^3 + \epsilon$. Similarly $\alpha_2 \geq 2\alpha_1 - \alpha^3 + \epsilon \geq \alpha^3 + 2\epsilon$. If we continue this process, we find $\alpha_k \geq \alpha^3 + 2^{k-1}\epsilon$ for all $1 \leq k \leq \lceil \log_2 \frac{1}{\epsilon} \rceil$+1. Thus $\alpha_k > 1$ if $k = \lceil \log_2 \frac{1}{\epsilon} \rceil + 1$, which is a contradiction. □

Let $f \colon \mathbb{F}_3^n \to \mathbb{C}$, and let $U \leq \mathbb{F}_3^n$; this notation means that U is a subspace of $\mathbb{F}_3^n$. Let $f_U(x)$ be the average of $f(x)$ on the U-coset that x is in.

The lemma below is related to an arithmetic analog of the regularity lemma.

Lemma 6.26. *For all $\epsilon > 0$, there exists $m = \text{tow}(O(\log \frac{1}{\epsilon}))$ such that for all $f \colon \mathbb{F}_3^n \to [0,1]$, there exist subspaces $W \leq U \leq \mathbb{F}_3^n$ with* codim $W \leq m$ *such that*

$$\|\widehat{f - f_W}\|_\infty \leq \frac{\epsilon}{|U^\perp|}$$

and

$$2\|f_U\|_3^3 - \|f_W\|_3^3 \geq (\mathbb{E}f)^3 - \epsilon.$$

Proof. Let $\epsilon_0 := 1$ and $\epsilon_{k+1} := \epsilon 3^{-1/\epsilon_k^2}$ for integers $k \geq 0$. Using the recursion, we find that the recursion says $\epsilon_{k+1}^{-2} = \epsilon^{-2} 3^{2/\epsilon_k^2}$, so that

$$\epsilon_{k+1}^{-2} \leq 2^{2^{\epsilon_k^{-2}}}$$

for sufficiently large k. Let

$$R_k := \{r \in \mathbb{F}_3^n : |\hat{f}(r)| \geq \epsilon_k\}.$$

Then $|R_k| \leq \epsilon_k^{-2}$, since by Parseval's identity, $\sum_r |\hat{f}(r)|^2 = \mathbb{E}[f^2] \leq 1$. Now define $U_k := R_k^\perp$ and $\alpha_k := \|f_{U_k}\|_3^3$. Note $\alpha_k \geq (\mathbb{E}f)^3$ by convexity. So by the previous lemma, there exists $k = O(\log \frac{1}{\epsilon})$ such that $2\alpha_k - \alpha_{k+1} \geq (\mathbb{E}f)^3 - \epsilon$. For this choice of k, let $m := \epsilon_{k+1}^{-2}$. With some computation we find $m = \operatorname{tow}(O(\log \frac{1}{\epsilon}))$.

It is not too hard to check that

$$\widehat{f_W}(r) = \begin{cases} \hat{f}(r) & \text{if } r \in W^\perp, \\ 0 & \text{if } r \notin W^\perp. \end{cases}$$

So $\|\widehat{f - f_{U_{k+1}}}\|_\infty \leq \max_{r \notin R_{k+1}} |\hat{f}(r)| \leq \epsilon_{k+1} \leq 3^{-|R_k|}\epsilon \leq \epsilon/|U_k^\perp|$. So if we take $W = U_{k+1}$ and $U = U_k$, we are done, as $\operatorname{codim} U_{k+1} \leq |R_{k+1}| \leq m$. □

With a regularity lemma comes a counting lemma, which is left as an exercise (it is fairly easy to prove). Define

$$\Lambda_3(f; U) = \mathbb{E}_{x \in \mathbb{F}_3^n, y \in U} f(x) f(x+y) f(x+2y).$$

Lemma 6.27 (Counting lemma)**.** *Let $f, g \colon \mathbb{F}_3^n \to [0,1]$ and $U \leq \mathbb{F}_3^n$. Then*

$$|\Lambda_3(f;U) - \Lambda_3(g;U)| \leq 3|U^\perp| \cdot \|\widehat{f-g}\|_\infty.$$

Lemma 6.28. *Let $f \colon \mathbb{F}_3^n \to [0,1]$, with subspaces $W \leq U \leq \mathbb{F}_3^n$. Then*

$$\Lambda_3(f_W; U) \geq 2\|f_U\|_3^3 - \|f_W\|_3^3.$$

Proof. We use Schur's inequality: $a^3 + b^3 + c^3 + 3abc \geq a^2(b+c) + b^2(a+c) + c^2(a+b)$ for $a, b, c \geq 0$. We find

$$\begin{aligned} \Lambda(f_W; U) &= \mathbb{E}_{\substack{x,y,z \\ \text{form a 3-AP in} \\ \text{the same } U\text{-coset}}} f_W(x) f_W(y) f_W(z) \\ &\geq 2\mathbb{E}_{x,y \text{ in same } U\text{-coset}} f_W(x)^2 f_W(y) - \mathbb{E} f_W^3 \\ &\geq 2\mathbb{E} f_W^2 f_U - \mathbb{E} f_W^3 \\ &\geq 2\mathbb{E} f_U^3 - \mathbb{E} f_W^3, \end{aligned}$$

where the first inequality follows from Schur's inequality and the last follows from convexity. □

Theorem 6.29. *For all $\epsilon > 0$, there exists $m = \operatorname{tow}(O(\log \frac{1}{\epsilon}))$ such that if $f \colon \mathbb{F}_3^n \to [0,1]$, then there exists $U \leq \mathbb{F}_3^n$ with codimension at most m such that*

$$\Lambda_3(f;U) \geq (\mathbb{E}f)^3 - \epsilon.$$

Note if n is large enough, then $|U|$ is large enough, so there exists a nonzero "common difference" y.

Proof. Choose U, W as in the regularity lemma. Then

$$\Lambda_3(f;U) \geq \Lambda_3(f_W;U) - 3\epsilon \geq 2\|f_U\|_3^3 - \|f_W\|_3^3 - 3\epsilon \geq (\mathbb{E}f)^3 - 4\epsilon.$$

□

The corresponding statement for popular differences is true in $\mathbb{Z}$ as well.

Theorem 6.30. *For all $\epsilon > 0$, there exists $N_0 = N_0(\epsilon)$ such that if $N > N_0$ and $A \subseteq [N]$ with $|A| = \alpha N$, then there exists $y > 0$ such that*

Green (2005)

$$|\{x : x, x+y, x+2y \in A\}| \geq (\alpha^3 - \epsilon)N.$$

A similar statement also holds for 4-APs in $\mathbb{Z}$:

Theorem 6.31. *For all $\epsilon > 0$, there exists $N_0 = N_0(\epsilon)$ such that if $N > N_0$ and $A \subseteq [N]$ with $|A| = \alpha N$, then there exists $y > 0$ such that*

Green and Tao (2010)

$$|\{x : x, x+y, x+2y, x+3y \in A\}| \geq (\alpha^4 - \epsilon)N.$$

Remark 6.32. Surprisingly, the corresponding statement for 5-APs (or longer) in $\mathbb{Z}$ is false.

Bergelson, Host, and Kra (2005) with appendix by Ruzsa

18.217 Graph Theory and Additive Combinatorics
Fall 2019