18.099 PROJECT: 4-TERM ARITHMETIC PROGRESSIONS

MARK SELLKE AND VICTOR WANG

1. Introduction

In this expository paper, we discuss Gowers’ Fourier-analytic proof of Szemerédi’s theorem for arithmetic progressions of length four (4-APs), following [1]. Recall that Szemerédi’s theorem, in its usual form, states that for any positive integer \( k \), any subset \( A \subseteq \mathbb{Z} \) of positive upper density

\[
0 < \delta(A) := \limsup_{N \to \infty} \frac{|\{1, 2, \ldots, N\} \cap A|}{N}
\]

contains infinitely many non-constant arithmetic progressions \((a, a + d, \ldots, a + (k - 1)d)\) of length \( k \). It is easy and elementary to reduce this theorem to showing that for \( N > n(\delta) \), if \( A \subseteq \mathbb{Z}_N \) has size \(|A| > \delta N\), then \( A \) contains a proper arithmetic progression of length \( k \), that is, a progression \((a, a + d, \ldots, a + (k - 1)d)\) consisting of distinct numbers. Szemerédi’s theorem is a natural generalization of the famous and much easier Van der Waerden’s theorem, which states that any \( r \)-coloring of \( \mathbb{Z} \) contains arbitrarily long monochromatic arithmetic progressions.

The original proof of this theorem by Szemerédi himself was a difficult tour de force of intricate combinatorics, which introduced and made use the important regularity lemma. However, this dependence on the regularity lemma meant that the proof gave tower-type bounds on \( n \) in terms of \( \delta \).

The (already non-trivial) case \( k = 3 \) had a proof using discrete Fourier analysis on the group \( \mathbb{Z}_N \). This proof goes as follows. One first shows directly that if all nonzero Fourier coefficients of \( 1_A \) are small (i.e. \( A \) is linearly uniform), then \( A \) has many 3-APs. One also shows that if some Fourier coefficient is large, some large subprogression has significantly higher density than \( \delta \), allowing for an inductive argument. This proof was, with much effort, extended to a full proof of Szemerédi’s theorem by Gowers in 2001 ([2]), resulting in much improved bounds of doubly exponential type. This proof is quite difficult and runs over 100 pages. However, the \( k = 4 \) case is already quite nontrivial and illustrates the core ideas of the generalization with fewer technical details.

Section 2 generalizes the fact that dense, linearly uniform sets contain the expected number of 3-APs (used in the standard proof of Roth’s theorem) to the fact that dense, so-called quadratically uniform sets (see Proposition-Definition 2.5) contain the expected number of 4-APs.

The main difficulty is to obtain a density increment on a sizable subprogression in the quadratically non-uniform case. Section 3 begins the search for an inverse theorem for quadratically non-uniform sets \( A \) (of density \( \delta > 0 \) and with balanced function \( f = 1_A - \delta \)), in particular identifying many large Fourier coefficients of the form \([x \mapsto f(x)f(x-k)](\phi(k))\)

for some frequency function \( \phi \). In fact, such a function \( \phi \) must have many additive quadruples.

\textit{Date:} May 13, 2016.
(Definition 3.2). Section 4 derives a quantitative strengthening of the Balog–Szemerédi theorem by Gowers, which in particular allows one to find a substantial linear restriction of $\phi$. Section 5 leverages the linear restriction to deduce quadratic bias: correlation with local quadratic phase functions. Note that the analogous property for linearly non-uniform, namely correlation with linear phase functions, is essentially tautologically true by definition of the Fourier transform. Suitable equidistribution results from Section 6 then easily lead to a density increment, as explained in Section 7.

2. **Quadratic Uniformity**

In this section, we define linear and quadratic uniformity, and state some easy equivalences, following Gowers’ original paper [1]. Ultimately we show that dense, quadratically uniform sets contain the expected number of 4-APs (while it appears\(^1\) to be open whether dense, linearly uniform sets contain at least the expected number of 4-APs). In later sections we discuss the structure theory of quadratically non-uniform sets (which is significantly more involved than that for linear non-uniformity in the proof of Roth’s theorem).

We first establish some basic (sometimes nonstandard) notation. Let $D = \{ |z| \leq 1 \} \subset \mathbb{C}$ denote the complex unit disk. For a positive integer $N$ let $\mathbb{Z}_N$ denote the cyclic group of integers modulo $N$, and $\omega = \exp(2\pi i / N)$ a primitive $N$th root of unity. For a function $f : \mathbb{Z}_N \to \mathbb{C}$ define the discrete Fourier transform by

$$\tilde{f}(r) := \sum_{x \in \mathbb{Z}_N} f(x) \omega^{-r \cdot x}.$$

Define the (non-standard!) convolution $f * g$ for functions $f, g : \mathbb{Z}_N \to \mathbb{C}$ by

$$f * g(s) = \sum_{t - u = s} f(t) g(u).$$

We then have

$$\sum_{r \in \mathbb{Z}_N} |\tilde{f}(r)|^2 = N \sum_{s \in \mathbb{Z}_N} |f(s)|^2$$

and

$$(f * g)\overline{\omega}(r) = \tilde{f}(r) \overline{g(r)}.$$

We will study subsets $A \subseteq \mathbb{Z}_N$ via their Fourier transforms, and for this it is useful to introduce the notion of the balanced function of a set $A$ of size $|A| = \delta N$, given by

$$f(s) = 1_A(s) - \delta.$$

The idea here is that we want to study the uniformity properties of a set $A \subseteq \mathbb{Z}_N$ via its Fourier transform, where large Fourier coefficients correspond to non-uniformity. Since any $A$ satisfies $\mathbb{I}_A(0) = \frac{|A|}{N}$, indicator functions $1_A$ of dense sets always have at least one large Fourier coefficient, and working instead with the balanced function eliminates this rather silly problem.

\(^1\)at least as of 2001; see [2, Conjecture 4.1]
We now turn to the definitions of linear and quadratic uniformity.

**Proposition-Definition 2.1** (Linear uniformity, [I Lemma 1]). Let \( f : \mathbb{Z}_N \to D \). Then the following are equivalent up to power-dependence of \( c_i \) constants.

1. \( \sum_r |\hat{f}(r)|^4 \leq c_1 N^4. \)
2. \( \max_r |\hat{f}(r)| \leq c_2 N. \)
3. \( \sum_k \sum_s \sum_r f(s) \overline{f(s-k)}^2 \leq c_3 N^3. \)
4. \( \sum_k \sum_s \sum_r f(s) \overline{g(s-k)} \leq c_3 N^2 \|g\|_2^2 \) for every function \( g : \mathbb{Z}_N \to \mathbb{C}. \)

If \( f \) satisfies (i) with \( c_1 = \alpha \), then we say \( f \) is (linearly) \( \alpha \)-uniform.

**Remark 2.2.** Due to the use of Parseval it is important here that the exponent 4 in (i) is greater than 2, but otherwise the choice in the definition is not too important.

**Remark 2.3.** Note that (i) is equivalent to (iii) with power dependence \( c_1 = c_3 \) (by Parseval). By expanding (iii) we see that \( f \) is linearly \( \alpha \)-uniform if and only if \( \alpha \geq \|f\|_{U^2(\mathbb{Z}_N)}^2 \), where \( \|f\|_{U^2(\mathbb{Z}_N)} \) is the Gowers uniformity norm of order 2, as defined in [I Definition 11.2].

The following definition will be convenient throughout the paper (especially the Fourier-analytic parts).

**Definition 2.4.** For a function \( f : \mathbb{Z}_N \to \mathbb{C} \) and \( k \in \mathbb{Z}_N \), define \( \Delta(f;k)(x) := f(x)\overline{f(x-k)}. \)

**Proposition-Definition 2.5** (Quadratic uniformity, [I Lemma 2]). Let \( f : \mathbb{Z}_N \to D \). Then the following are equivalent up to power-dependence of \( c_i \) constants.

1. \( \sum_{u,v} \sum_s f(s)\overline{f(s-u)}\overline{f(s-v)}f(s-u-v) \leq c_1 N^4 \) (analogous to (iii) from 2.1).
2. \( \sum_{k,r} |\Delta(f;k)^\sim(r)|^4 \leq c_2 N^5 \) (analogous to (i) from 2.1).
3. \( |\Delta(f;k)^\sim| \geq c_3 N \) for at most \( c_3^2 N \) pairs \( (k,r) \).
4. For all but \( c_4 N \) values of \( k \) the function \( \Delta(f;k) \) is \( c_4 \)-uniform.

If \( f \) satisfies (i) with \( c_1 = \alpha \), then we say \( f \) is quadratically \( \alpha \)-uniform.

**Remark 2.6.** By expanding (i) we see that \( f \) is quadratically \( \alpha \)-uniform if and only if \( \alpha \geq \|f\|_{U^3(\mathbb{Z}_N)}^2 \), where \( \|f\|_{U^3(\mathbb{Z}_N)} \) is the Gowers uniformity norm of order 3, as defined in [I Definition 11.2].

The following two results are special cases of the generalized von Neumann theorem [4 Lemma 11.4] (proven by induction with the van der Corput lemma). Recall that the first is used in the standard Fourier proof of Roth’s theorem.

**Proposition 2.7** (\( U^2 \)-norm upper bound). If \( f_1, f_2, f_3 : \mathbb{Z}_N \to D \) (where \( N \) is relatively prime to \( 2! = 2 \)), and \( f_3 \) is \( \alpha \)-uniform, then \( |E_{a,d}f_1(a)f_2(a+d)f_3(a+2d)| \leq \alpha^{1/4}. \)

This is usually proven systematically from a Fourier perspective (in view of the interpretation of \( (a, a+d, a+2d) \) as solutions to \( x - 2y + z = 0 \)), so we instead give the proof from [4 Lemma 11.4].

**Proof.** The main point is that \( f_1, f_2 \) can be inductively removed from the picture by triangle inequality and Cauchy–Schwarz.² Indeed,

\[
|E_{a,d}f_1(a)f_2(a+d)f_3(a+2d)| \leq E_a|E_df_2(a+d)f_3(a+2d)|
\leq (E_a|E_df_2(a+d)f_3(a+2d)|^2)^{1/2}.
\]

²Of course, the removal must be done suitably “independently” from the other functions: the naive bound \( E_{a,d}|f_1(a)||f_2(a+d)||f_3(a+2d)| \) would certainly be too weak.
But we can simply expand
\[
E_a |E_d f_2(a + d) f_3(a + 2d)|^2 = E_{a,d,d'} f_2(a + d) f_2(a + d + d') f_3(a + 2d) f_3(a + 2d + 2d')
\]
\[
= E_d E_{a,d} f_2(a + d) f_2(a + d + d') f_3(a + 2d) f_3(a + 2d + 2d')
\]
\[
= E_d E_{a,d} f_2(a + d) f_2(a + d + d') f_3(a + d) f_3(a + d + d')
\]
(where we have switched the order of expectation and translated \(a\) by \(d\)).

Now inductively proceeding\(^3\) on the inner \(E_{a,d}\) (for fixed \(d'\)), we obtain
\[
|E_{a,d} f_2(a + d) f_3(a + d + d') f_3(a + d + d')| \leq E_a |E_d f_3(a + d) f_3(a + d + d')| \leq (E_a |E_d f_3(a + d) f_3(a + d + d')|^2)^{1/2}.
\]

But \(E_d f_3(a + d) f_3(a + d + d') = E_x f_3(x) f_3(x + d')\) is independent of \(a\), so finally we have
\[
|E_{a,d} f_1(a) f_2(a + d) f_3(a + 2d)| \leq (E_a |E_d f_2(a + d) f_3(a + 2d)|^2)^{1/2}
\]
\[
\leq (E_d (|E_x f_3(x) f_3(x + d')|^2)^{1/2})\]
\[
\leq (E_d |E_x f_3(x) f_3(x + d')|^2)^{1/4} \leq \alpha^{1/4},
\]
by one last Cauchy–Schwarz, followed by definition\(^2\) (iii) of linear \(\alpha\)-uniformity. \(\square\)

By the recursive nature of the quadratic \((U^3,.)\)-norm, the previous proposition similarly implies the following key estimate for this section. Again, for further generalization see \[4\] Lemma 11.4.

**Lemma 2.8** \((U^3,.)\)-norm upper bound. If \(f_1, f_2, f_3, f_4 : \mathbb{Z}_N \to D\) (where \(N\) is relatively prime to \(3! = 6\)), and \(f_4\) is quadratically \(\alpha\)-uniform, then \(|E_{a,d} f_1(a) f_2(a + d) f_3(a + 2d) f_4(a + 3d)| \leq \alpha^{1/8}\).

**Proof sketch.** We continue the induction from the \(U^2,.)\)-norm upper bound. First we get rid of \(f_1\) by triangle inequality and Cauchy–Schwarz:
\[
|E_{a,d} f_1(a) f_2(a + d) f_3(a + 2d) f_4(a + 3d)| \leq E_a |E_d f_2(a + d) f_3(a + 2d) f_4(a + 3d)| \leq (E_a |E_d f_2(a + d) f_3(a + 2d) f_4(a + 3d)|^2)^{1/2}.
\]

But we can simply expand
\[
E_a |E_d f_2(a + d) \ldots f_4(a + 3d)|^2 = E_{a,d,d'} f_2(a + d) f_2(a + d + d') \ldots f_4(a + 3d) f_4(a + 3d + 3d')
\]
\[
= E_d E_{a,d} f_2(a + d) f_2(a + d + d') \ldots f_4(a + 3d) f_4(a + 3d + 3d')
\]
\[
= E_d E_{a,d} f_2(a + d) f_2(a + d + d') \ldots f_4(a + 2d + 3d')
\]
(where we have switched the order of expectation and translated \(a\) by \(d\)).

Now the point is that \(f_4(x) f_4(x + 3d')\) is on average (over all \(d'\)) linearly uniform, since \(\gcd(N, 3) = 1\) and the definition\(^2\) of quadratic \(\alpha\)-uniformity has a suitably recursive nature.\(^4\) Thus we may apply the previous proposition to each \(E_{a,d} f_2(a + d) f_2(a + d + d') \ldots f_4(a + 2d + 3d')\), and finish by Cauchy–Schwarz. \(\square\)

---

3. i.e. using the \(U^1,.)\)-norm upper bound

4. This is slightly easier to see in the presentation of \[3\] Lemma 11.4, where the definition of Gowers \(U^{k-1,.)}\)-norm is explicitly used there.
The previous $4 - 2 = 2$ results (together) immediately yield the main theorem of this section.

**Theorem 2.9.** If $A, B, C, D$ are subsets of $\mathbb{Z}_N$ of density $\alpha, \beta, \gamma, \delta$ (where $N$ is relatively prime to $3! = 6$), respectively, with $D$ quadratically $\eta$-uniform and $C$ linearly $\eta^{1/2}$-uniform, then $|E_{a,d}A(a)B(a + d)C(a + 2d)D(a + 3d) - \alpha\beta\gamma\delta| \leq 2\eta^{1/8}$.

**Proof.** Split $D = (D - \delta) + \delta$ and use the lemma on the $D - \delta$ term; then split $C = (C - \gamma) + \gamma$ and use the proposition on the $C - \gamma$ term. We get an upper bound of

$$\eta^{1/8} + \delta[(\eta^{1/2})^{1/4} + \gamma|E_{a,d}A(a)B(a + d) - \alpha\beta|].$$

But $E_{a,d}A(a)B(a + d) = E_{a,b}A(a)B(b) = (E_a A(a))(E_b B(b)) = \alpha\beta$. So

$$|E_{a,d}A(a)B(a + d)C(a + 2d)D(a + 3d) - \alpha\beta\gamma\delta| \leq \eta^{1/8} + \delta\eta^{1/8} + 0 \leq 2\eta^{1/8},$$

as desired. \qed

In order to apply this theorem to the problem of finding 4-APs, it will be convenient to observe that quadratic uniformity is stronger than linear uniformity (up to power dependence of the tolerance).

**Proposition 2.10.** If $f$ is quadratically $\alpha$-uniform, then it is also (linearly) $\alpha^{1/2}$-uniform.

**Proof.** Write $\sum_r |\tilde{f}(r)|^4 = N\sum_{x,k,l} f(x)\overline{f}(x - k)f(x - l)f(x - k - l)$ by $4 = 2 + 2$ and Parseval. Then by Cauchy–Schwarz,

$$\sum_r |\tilde{f}(r)|^4 \leq N(N^2 \sum_{k,l} f(x)\overline{f}(x - k)f(x - l)f(x - k - l))^2 \cdot \frac{1}{2} \leq N(N^2\alpha N^4)^{1/2} = \alpha^{1/2}N^4,$$

as desired. \qed

**Example 2.11** (Linear uniformity does not imply quadratic uniformity, nor expected number of 4-APs). Let $A$ be the set $\{x \in \mathbb{Z}_N : |x^2| \leq N/10\}$. It turns out [2] beginning of Section 4 that $A$ is uniform but not quadratically uniform. Furthermore, $A$ does not contain the expected number of arithmetic progressions; in fact, it contains more than expected (due to the linear finite-difference relation $(a + 3d)^2 - 3(a + 2d)^2 + 3(a + d)^2 - a^2 = 0$).

**Remark 2.12.** Gowers gives two examples, including the preceding one, to motivate his full argument (for all $k \geq 3$, not just $k = 3, 4$) in [2] Section 4, “Two Motivating Examples”.

In any case, we may use the theorem to finally prove that quadratic uniformity implies existence of 4-APs.

**Corollary 2.13** ([11 Corollary 8]). Let $A_0 \subseteq [N]$ of density $\delta$. If $A_0$ is quadratically $\eta$-uniform with $\eta \leq \delta^{112}/2^{208}$ and $N > 200/\delta^3$, then $A_0$ contains a 4-AP.

**Proof.** In the theorem, take $A, B$ to be $A_0 \cap [2N/5, 3N/5]$ (to avoid wrap-around issues), and take $C, D$ to be $A_0$ (which is both quadratically $\eta$-uniform and, by the previous proposition, linearly $\eta^{1/2}$-uniform). Since $A_0$ is $\eta^{1/2}$-uniform, the upper bound on $\eta$ forces $\alpha, \beta \geq \delta/10$.

---

5Strictly speaking, it suffices to have a suitable inverse theorem for both linear and quadratic non-uniformity (and the linear inverse theorem already exists from the proof of Roth’s theorem!), but it is conceptually clearer to make the observation.
(we delay the proof to the end). By the theorem, there are at least $(\delta/10)^2 \delta^2 N^2 - 12\eta^{1/8} N^2$ tuples $(a, a + d, a + 2d, a + 3d)$ in $A \times B \times C \times D$, but at most $\delta N$ with $d = 0$. So we are done if

$$\delta^4/100 - 2\eta^{1/8} > \delta/N;$$

by our assumptions on $\eta, N$ this boils down to

$$\delta^4/100 > 2\delta^{112/8}/2^{208/8} + \delta^4/200,$$

or $\delta^4/200 > 2\delta^{14}/2^{26}$, or $2^{26} > 2 \cdot 200^{10}$, which is obvious (as $\delta \leq 1$).

Proof of the promised $\alpha, \beta \geq \delta/10$. Let $f := 1_{A_0} - \delta$ be the balanced function, so $\sum_{x \in [2N/5,3N/5]} f(x) = (\alpha - \delta) N$. But by Fourier inversion,

$$\sum_{x \in [2N/5,3N/5]} f(x) = N^{-1} \sum_r \tilde{f}(r) \sum_{x \in [2N/5,3N/5]} \omega^{r \cdot x},$$

which by the triangle inequality and geometric series is bounded in magnitude by

$$N^{-1} \sum_r |\tilde{f}(r)| \min(N, 1 + 1/|1 - \omega^r|) \leq (N^{-1} \sum_r |\tilde{f}(r)|^4)^{1/4} (N^{-1} \sum_r \min(N, 1 + 1/|1 - \omega^r|)^{4/3})^{3/4}. $$

(The exponent 4 is not terribly important; essentially anything greater than 2 would work.) Note that for $r \neq 0$ in $(-N/2, N/2)$, we have $2/|1 - \omega^r| = 1/\sin(\pi |r|/N) \leq (2/\pi |r|/N)^{-1} = N/|2r|$, and $N^{-1} \sum_r \min(N, N/|2r|)^{4/3} \leq N^{1/3} (1 + 2 \cdot (1/2)^{4/3}(1 + 3)) < 16N^{1/3}$. But by definition of $\eta^{1/2}$-uniformity, $\sum_r |\tilde{f}(r)|^4 \leq \eta^{1/2} N^4$, so we deduce

$$|\alpha - \delta|N \leq (N^{-1} \sum_r |\tilde{f}(r)|^4)^{1/4} (16N^{1/3})^{3/4} \leq \eta^{1/8} N^{3/4} \cdot 8N^{1/4}. $$

Thus

$$|\alpha - \delta| \leq 8\eta^{1/8} \leq 8(\delta^{14}/2^{26}) < 9\delta/10,$$

so indeed $\alpha \geq \delta/10$, as desired.

\[ \square \]

3. ADDITIVE QUADRUPLES

We now begin the inverse structure-chasing (in the quadratically non-uniform case), again following Gowers [1]. Whereas linear non-uniformity immediately provides a density increment in the proof of Roth’s theorem, quadratic non-uniformity requires a bit more care.

**Proposition 3.1** ([1 Section 7, beginning of proof of Theorem 20]). If $f : \mathbb{Z}_N \to D$ is not quadratically $\alpha$-uniform, then there exists a set $B \subseteq \mathbb{Z}_N$ of cardinality at least $\alpha N/2$, and $\phi : B \to \mathbb{Z}_N$ such that $|\Delta(f; k)^{(\phi(k))}| \geq (\alpha/2)^{1/2} N$ for all $k \in B$. In particular, $\sum_{k \in B} |\Delta(f; k)^{(\phi(k))}|^2 \geq (\alpha/2)^2 N^3$.

**Proof.** If $f$ is not quadratically $\alpha$-uniform, then $\sum_{k \neq r} |\Delta(f; k)^{(r)}|^4 > \alpha N^5$ (by definition of uniformity, followed by expansion and Parseval). In particular, by pruning, there are more than $\alpha N/2$ values of $k$ such that $\sum_r |\Delta(f; k)^{(r)}|^4 > (\alpha/2)N^4$ (cf. implication from (iv) to (ii) in [2.5]). The set of such $k$ will form our set $B$. To define $\phi$, note that for each such $k$,

---

\[ ^6 \text{Cf. [2] Lemma 5.2}. \text{ Again, this is not strictly necessary for our purposes, because we could directly obtain a density increment otherwise, but it is still conceptually cleaner to prove it.} \]
the fact that $4 - 2 = 2$ yields, for $g = \Delta(f; k) : \mathbb{Z}_N \to D$, a standard Fourier estimate (the implication from (ii) to (i) in 2.1)

$$\sum_r |\tilde{g}(r)|^4 \leq \max_r |\tilde{g}(r)|^{4 - 2} \cdot \sum_r |\tilde{g}(r)|^2 = \max_r |\tilde{g}(r)|^2 \cdot N \sum_x |g(x)|^2 \leq N^2 \max_r |\tilde{g}(r)|^2,$$

so $\phi(k) := \arg \max_r |\Delta(f; k)^{(r)}(r)|$ will do. \hfill \square

It turns out that the previous proposition yields enough inverse structure, beginning as follows. The point is that the function $\phi$ constructed is far from arbitrary. It must have many additive quadruples $(a, b, c, d) \in B^4$, defined as follows.

**Definition 3.2** (Additive quadruple). A quadruple $(a, b, c, d) \in B^4$ is an additive quadruple (of the function $\phi : B \to \mathbb{Z}_N$) if $a + b = c + d$ and $\phi(a) + \phi(b) = \phi(c) + \phi(d)$.

**Proposition 3.3** ([2 Section 3, Proposition 9]). Let $\alpha > 0$, let $f : \mathbb{Z}_N \to D$, let $B \subseteq \mathbb{Z}_N$ and let $\phi : B \to \mathbb{Z}_N$ be a function such that

$$\sum_{k \in B} |\Delta(f; k)^{(r)}(\phi(k))|^2 \geq \alpha N^3.$$

Then $\phi$ has at least $\alpha^4 N^3$ additive quadruples.

**Proof.** We will essentially show that the functions $B(k)\omega^{-\phi(k)u} : \mathbb{Z}_N \to D$ (functions of $k$, for various fixed $u$) are, on average, not linearly uniform. This will then easily give the desired conclusion.

We start by expanding the LHS (with $\Delta(f; k)^{(r)}(\phi(k)) = \sum_x f(x)\overline{f(x-k)}\omega^{-\phi(k)x}$) to get

$$\sum_k \sum_{x,y} f(x)\overline{f(x-k)}f(y)\overline{f(y-k)}B(k)\omega^{-\phi(k)(x-y)} \geq \alpha N^3.$$

Switching the order of summation (to get $k$ on the inside) and using the trivial estimate $|f(x)f(y)| \leq 1$ for all $x, y$, we obtain

$$\sum_{x,y} \sum_k |\overline{f(x-k)}f(y-k)B(k)\omega^{-\phi(k)(x-y)}| \geq \alpha N^3.$$

After Cauchy and rewriting in terms of $u = x - y$, we have

$$\sum_{u,x} \sum_k |B(k)\omega^{-\phi(k)u}\Delta(f; u)(x-k)|^2 \geq (\alpha N^3)^2/N^2 = \alpha^2 N^4,$$

i.e. $\sum \gamma(u) \geq \alpha^2 N$, where $\gamma(u) := N^{-2} \sum_x |\sum_k B(k)\omega^{-\phi(k)u}\Delta(f; u)(x-k)|^2$.

Since $|\Delta(f; u)(x)| \leq 1$ everywhere, one of the easy equivalences for linear uniformity (implication from (i) to (iv) of 2.1 proved using Parseval followed by Cauchy–Schwarz) implies that the function $B(k)\omega^{-\phi(k)u}$ (of $k$) is not $\gamma(u)^2$-uniform:

$$\sum_r |B(k)\omega^{-\phi(k)u}\omega^{-r-k}|^4 \geq \gamma(u)^2 N^4$$

for each $u$. Summing over all $u$, and expanding with conjugates, we find that

$$\sum_{u,r} \sum_{a,b,c,d \in B} \omega^{-u(\phi(a)+\phi(b)-\phi(c)-\phi(d))}\omega^{-r(a+b-c-d)} \geq \alpha^4 N^5.$$

But the LHS is $N^2$ times the number of additive quadruples for $\phi$, so indeed $\phi$ has at least $\alpha^4 N^3$ additive quadruples, as desired. \hfill \square
4. Linearization

In this section, we show how to transform our information on additive quadruples $\phi$ into the statement that $\phi$ is exactly affine on a sizable set. We first recall a famous theorem of Freiman and state a variant.

**Definition 4.1.** A *generalized arithmetic progression* (of dimension $d$) in an abelian group $G$ is a set $P = P(a, v_1, \ldots, v_d, c_1, \ldots, c_d) = \{a + \sum_{k=1}^{d} b_k v_k : 0 \leq b_k < c_k\}$, for $a, v_1, \ldots, v_d \in G$ and $c_1, \ldots, c_d \in \mathbb{Z}^+$. $P$ is *proper* when all possible sums are distinct, i.e. $|P| = \prod_{k=1}^{d} c_k$.

**Theorem 4.2** ([3]; originally due to Freiman). Let $C$ be a constant. There exist constants $d = d(C)$ and $K = K(C)$ such that, whenever $A \subseteq \mathbb{Z}$ satisfies $|A| = m$ and $|A - A| \leq CM$, there exists a generalized arithmetic progression $Q$ of dimension at most $d$ such that $|Q| \leq Km$ and $A \subseteq Q$.

A dense subset $A$ of a generalized arithmetic progression automatically has $\frac{|A - A|}{|A|}$ small, so Freiman’s theorem essentially classifies such sets $A$. We actually use a version of Freiman’s theorem due to Ruzsa, which guarantees a *proper* generalized progression, but settles for a large intersection instead of containment.

**Theorem 4.3** ([3] see proof of Theorem 1.1). If $A \subseteq \mathbb{Z}^D$ satisfies $|A - A| \leq C|A|$, then there is a proper generalized progression $Q$ of dimension $d = O(C^{O(1)})$ and size $\Omega(C^{-O(C)}|A|)$ such that $|A \cap Q| = \Omega_{C}(|Q|)$ (which implies $|Q| = O_{C}(|A|)$).

We will apply Ruzsa’s theorem to the graph $\Gamma$ of $\phi$. However, the existence of many additive quadruples in no way guarantees that $\Gamma$ satisfies the hypotheses of the above theorems; for example, $\Gamma$ could be a disordered set together with a progression. However, it does turn out that the preponderance of additive quadruples ensures that $\Gamma$ contains a large subset $A$ with $\frac{|A - A|}{|A|}$ bounded. We now prove this, the Balog-Szemerédi-Gowers theorem, starting with a lemma.

**Lemma 4.4.** Let $X$ be a set of size $m$, let $\delta > 0$, and let $A_1, \ldots, A_n$ be subsets of $X$ such that

$$\sum_{x,y=1}^{n} |A_x \cap A_y| \geq \delta^2 mn^2.$$ 

Then there is a subset $K \subseteq [n]$ with $|K| \geq 2^{-1/2} \delta^5 n$ such that for at least 90% of pairs $(x, y) \in K^2$, the inequality $|A_x \cap A_y| \geq \frac{\delta mn}{2}$ holds. In particular, the result holds if $|A_x| \geq \delta m$ for all $x$.

**Proof.** The idea is to restrict to the subsets $A_x$ which contain a small, random set; these subsets $A_x$ will be biased towards having large pairwise intersection. For each $j \leq m$ let $B_j = \{i \leq n : j \in A_i\}$ be the indices of the subsets containing $j$ and let $E_j = B_j \times B_j$. Pick (with replacement) uniformly random, independent integers $1 \leq j_1, j_2, j_3, j_4, j_5 \leq m$ and let $S = E_{j_1} \cap \cdots \cap E_{j_5}$ be the set of pairs $(x, y)$ such that $j_k \in A_x \cap A_y$ for each $k \leq 5$. Note that $S = K \times K$ for $K := B_{j_1} \cap \cdots \cap B_{j_5}$.

First we show $S$ is large on average. Indeed, each pair $(x, y)$ has probability $p_{xy} = \frac{|A_x \cap A_y|}{m}$ to be in $S$, so linearity of expectation followed by Jensen’s inequality shows
\[ \mathbb{E}[|S|] = \sum_{1 \leq x, y \leq n} \left( \frac{|A_x \cap A_y|}{m} \right)^5 \geq n^{-8} \left( \sum_{1 \leq x, y \leq n} \frac{|A_x \cap A_y|}{m} \right)^5 \geq \delta^{10} n^2. \]

Now we show that there are on average few pairs \((x, y) \in S\) with \(|A_x \cap A_y|\) small. Let \(T \subseteq S\) be defined by \(T = \{(x, y) \in S : |A_x \cap A_y| < \frac{\delta m}{2}\}\), or equivalently \(T = \{(x, y) \in S : p_{xy} < \frac{\delta}{2}\}\). The same linearity of expectation argument shows

\[ \mathbb{E}[|T|] \leq n^2 \left( \frac{\delta^2}{2} \right)^5 = \frac{\delta^{10} n^2}{32}. \]

Therefore \(\mathbb{E}[|S| - 16|T|] \geq \frac{\delta^{10} n^2}{2}\). So with positive probability we have \(|S| - 16|T| \geq \frac{\delta^{10} n^2}{2}\). This means \(|K| = \sqrt{|S|} \geq 2^{-\frac{1}{2}} \delta^5 n\) and \(|T| \leq \frac{|S|}{16} \leq \frac{|S|}{10}\), which are exactly the desired properties.

\[ \square \]

**Theorem 4.5** (Balog-Szemerédi-Gowers). Let \(A \subseteq \mathbb{Z}^D\) satisfy \(|A| = m\) and \(|A \ast A|^2 \geq c_0 m^3\). There are constants \(c = c(c_0), C = C(c_0)\) and a subset \(A'' \subseteq A\) of cardinality \(|A''| \geq cm\) such that \(|A'' - A''| \leq Cm\).

**Remark 4.6.** The Balog-Szemerédi-Gowers theorem was originally proved by Balog and Szemerédi using the regularity lemma. Gower’s proof (given below) avoids the regularity lemma, resulting in hugely improved bounds.

**Proof.** The idea is as follows. We will find a large subset \(A'' \subseteq A\) such that for any \(a_1'' \in A''\), the difference \(d = a_1'' - a_2''\) has \(\Omega(|A|^8)\) representations as an alternating sum \(d = a_1 - a_2 + a_3 - a_4 + a_5 - a_6 + a_7 - a_8\) for \(a_1, \ldots, a_8 \in A\). The pigeonhole principle will then show that \(|A'' - A''| = O(|A|)\). To carry out this plan, we will define \(A''\) as a “popular” subset of \(A\), which is itself a “popular” subset of \(A\), and use these subsets to rewrite \(a_1'' - a_2''\) as \(a_1'' - a_2'' = (a_1'' - a_1) + (a_2 - a_2''\) for \(a_1, a_2 \in A\) and \(a_1', a_2' \in A'\). For \(\Theta(1)\)-fraction of these choices, each of the \(4\) atomic differences will be a popular difference, i.e. a number \(z\) which has many representations \(z = a_1 - a_2\) for \(a_1, a_2 \in A\). This will show that there are many representations \(a_1'' - a_2'' = a_1 - a_2 + a_3 - a_4 + a_5 - a_6 + a_7 - a_8\).

First define the non-negative function \(f : \mathbb{Z}^D \rightarrow \mathbb{Z}\) by \(f(x) = A \ast A(x)\). The given information translates to \(|f|_\infty \leq m\), \(|f|^2 \geq c_0 m^3\), and \(|f|_1 = m^2\). This easily implies that \(f(x) \geq \frac{c_0 m^2}{2}\) for at least \(\frac{c_0 m^2}{2}\) values of \(x\). Call \(x\) a popular difference if \(f(x) \geq \frac{c_0 m^2}{2}\).

Now define a graph structure \(G\) on \(A\), with \(a_1, a_2\) connected when \(a_1 - a_2\) (equivalently \(a_2 - a_1\)) is a popular difference. Since there are at least \(\frac{c_0 m^2}{2}\) distinct popular differences, there are at least \(\frac{c_0 m^2}{8}\) edges (where we lose a factor of \(2\) since each edge is counted in both directions). Therefore the average degree of \(G\) is at least \(\frac{c_0 m^2}{4}\), and so there are at least \(\frac{c_0 m^2}{8}\) vertices of degree at least \(\frac{c_0 m^2}{8}\).

Now let \(\delta = \frac{a^2}{8}\) and for \(n \geq \delta m\) let \(a_1, \ldots a_n\) be vertices of degree at least \(\delta m\) in \(G\) with neighborhoods \(A_1, \ldots, A_n\). Applying the lemma to the sets \(A_i\) gives a set \(A' \subseteq \{a_1, \ldots, a_n\}\) with \(|A'| \geq 2^{-\frac{1}{2}} \delta^5 n\) such that at least \(90\%\) of intersections \(A_i \cap A_j\) for \(i, j \in A'\) have size at least \(\delta^2 m\). Setting \(\alpha = 2^{-\frac{1}{2}} \delta^6\) we have \(|A'| \geq \alpha m\).

We now construct \(A''\). Define the graph \(H\) on the vertices of \(A'\), with \(a_i\) and \(a_j\) adjacent if \(|A_i \cap A_j| \geq \frac{\delta^2 m}{2}\). Since the average \(H\)-degree is at least \(\frac{9}{10} |A'|\), at least \(\frac{2}{3} |A'|\) vertices have
Corollary 4.7. Let \( |A| = \frac{1}{4}|A'| \). Define \( A'' \) to be the set of such vertices. We immediately see that any \( a''_1, a''_2 \in A'' \) have at least \( \frac{3}{5}|A'| \) common \( H \)-neighbors.

Now we use these subsets to write \( a''_1 - a''_2 \) in many ways as outlined above. For fixed \( a''_1, a''_2 \in A'' \) with neighborhoods \( A''_1, A''_2 \), we have that there are at least \( \frac{2}{5}|A'| \) common \( H \)-neighbors \( a' \in A' \) of \( a''_1, a''_2 \). For each such \( a' \) (with neighborhood \( A' \)), by definition of \( H \) we have \( |A_1'' \cap A'|, |A_2'' \cap A'| \geq \frac{2}{5}m \). Therefore there are \( \Omega_3(1) \) choices each for numbers \( a_1 \in A_1'' \cap A' \) and \( a_2 \in A_2'' \cap A' \). For all such numbers, by definition of the graph \( G \) we have that \( a''_1 - a_1, a''_2 - a_2, a''_1 - a_2, a''_2 - a_1 \) are all common differences.

Now, given \( a''_1, a''_2 \), the differences \( a''_1 - a_1, a''_2 - a_2, a''_1 - a_2, a''_2 - a_1 \) clearly determine \( a_1, a', a_2 \) uniquely. Therefore, we have \( \Omega(\delta^O(1)m^7) \) representations of \( a''_1 - a''_2 = a_1 - a_2 + a_3 - a_4 + a_5 - a_6 + a_7 - a_8 \) for \( a_1, \ldots, a_8 \in A \). Since there are only \( m^8 \) such octuples, this implies that \( |A'' - A''| = O(m^{5-O(1)}) \) as desired.

\[ \square \]

Corollary 4.7. Let \( A \subseteq \mathbb{Z}^D \) satisfy \( |A| = m \) and \( |A \ast A|^2_{L^2} \geq c_0m^3 \). For constants \( C = C(c_0), d = d(c_0), c = c(c_0) \) there is a generalized arithmetic progression \( Q \) of cardinality at most \( Cm \) and dimension at most \( d \) such that \( |A \cap Q| \geq cm \).

Proof. Just use Balog-Szemerédi-Gowers to find the set \( A'' \) with small doubling, and then apply Freiman’s theorem.

Using Ruzsa’s theorem instead of Freiman’s theorem gives the same result, except that \( Q \) can be taken to be a proper progression.

Corollary 4.8. Let \( A \subseteq \mathbb{Z}^D \) satisfy \( |A| = m \) and \( |A \ast A|^2_{L^2} \geq c_0m^3 \). For constants \( C = C(c_0), d = d(c_0), c = c(c_0) \) there is a proper generalized arithmetic progression \( Q \) of cardinality at most \( Cm \) and dimension at most \( d \) such that \( |A \cap Q| \geq cm \).

For induction purposes, we really just want almost-linear behavior on a 1-dimensional progression, and this can easily be extracted from the preceding.

Corollary 4.9. Let \( B \subseteq \mathbb{Z}_N \) with \( |B| = \beta N \), and let \( \phi : B \rightarrow \mathbb{Z}_N \) be a function with at least \( c_0N^3 \) additive quadruples. Then there exist constants \( \gamma(\beta, c_0), \eta(\beta, c_0) \) and a mod-\( N \) arithmetic progression \( P \subseteq \mathbb{Z}_N \) of cardinality at least \( N^\gamma \) and a linear function \( \psi : P \rightarrow \mathbb{Z}_N \) such that \( \phi(s) \) is defined and \( \phi(s) = \psi(s) \) for at least \( \eta|P| \) values of \( s \in P \).

Proof. Let \( \Gamma \subseteq \mathbb{Z}^2 \) be the graph of \( \phi \). By the above corollary we can find a proper progression \( Q = P_1 + \cdots + P_d \) with \( |Q| = O_{\beta, c_0}(|B|) \) and \( |
\Gamma \cap Q| = \Omega_{\beta, c_0}(|B|) \). Then some \( P_i \) has size \( \Omega(N^{\frac{1}{4}+\epsilon}) \) and averaging over the translates of \( P_i \) shows that some translate of \( P_i \) intersects \( \Gamma \) in \( \Omega_{\beta, c_0}(|P_i|) \) places. Since \( \Gamma \) is a graph, \( P_i \) can easily be shortened to pass the vertical line test, and then we obtain from it a linear function \( \psi : P \rightarrow \mathbb{Z}_N \) with the desired properties.

\[ \square \]

5. Quadratic Bias

In this section, the result from the previous section is leveraged to show correlation with local quadratic phase functions. In more detail:
Proposition 5.1. Let $A \subseteq \mathbb{Z}_N$ have size $|A| = \delta N$ and let $f = 1_A - \delta$, where $N$ is odd. Let $P = \{x + d, x + 2d, \ldots, x + Td\}$. Suppose there exist $\lambda, \mu$ with

$$\sum_{k \in P} |\Delta(f; k)^{\ast}(\lambda k + \mu)|^2 \geq \beta N^2 T.$$

Then there exist quadratic polynomials $\psi_0, \ldots, \phi_{N-1}$ such that

$$\sum_x | \sum_{z \in P + x} f(z) e^{-i \psi_x(z)} | \geq \frac{\beta NT}{\sqrt{2}}.$$

Proof. Start, as we did when finding additive quadruples, by expanding the assumed inequality in frequency variables $x,y$, and substituting $u = x - y$, to obtain

$$\sum_{k \in P} \sum_{x,u} f(x) \overline{f(x - k)} f(x - u) \overline{f(x - k - u)} \omega^{-(\lambda k + \mu)u} \geq \beta N^2 T.$$

For the sake of symmetry in the phase exponent, note that each $u \in \mathbb{Z}_N$ can be written in exactly $T$ ways as $v + l$ with $v \in \mathbb{Z}_N$ and $l \in P$, so we have the equivalent inequality

$$\sum_{k,l \in P} \sum_{x,v} f(x) \overline{f(x - k)} f(x - v - l) f(x - v - k - l) \omega^{-(\lambda k + \mu)(l + v)} \geq \beta N^2 T^2.$$

In particular, using the trivial bound $|f(x)| \leq 1$ for all $x$, we find $E_{x,v} \gamma(x,v) \geq \beta$, where

$$\sum_{k \in P} \frac{f(x - k) f(x - v - l) f(x - v - k - l) \omega^{-(\lambda k + \mu)(l + v)}}{\sqrt{2}} =: \gamma(x,v) T^2.$$

For fixed $x,v$, let $f_1(k), f_2(l), f_3(k + l)$ be functions on $P, P, P + P$, respectively, defined as $f(x - k), f(x - v - l), f(x - v - k - l)$ for $k, l \in P$. The phase exponent expands as $-(\lambda kl + \lambda v k + \mu l + \mu v)$; let $(a, b, -2c) = (\lambda v, \mu, \lambda)$, respectively, so that

$$| \sum_{k \in P} f_1(k) f_2(l) f_3(k + l) \omega^{-(ak + bl + 2cl)} | = \gamma(x,v) T^2.$$

But $2cl = c[(k + l)^2 - k^2 - l^2]$, so letting $h_1(k), h_2(l), h_3(k + l)$ be (functions on $P, P, P + P$, respectively, defined as) $f_1(k) \omega^{-(ak + cl)}$, $f_2(l) \omega^{-(bl + cl)}$, $f_3(k + l) \omega^{(k + l)^2}$, we deduce the Fourier-analytic estimate

$$\sum_{r} \sum_{k,l \in P} \sum_{m \in P + P} h_1(k) h_2(l) h_3(m) \omega^{-r(k + l - m)} = N \cdot \gamma(x,v) T^2.$$

The LHS is just the product of $g_1(r) = \sum_{k \in P} h_1(k) \omega^{-rk}$, $g_2(r) = \sum_{l \in P} h_2(l) \omega^{-rl}$, and $g_3(r) = \sum_{m \in P + P} h_3(m) \omega^{rm}$ (all essentially Fourier coefficients at the frequency $\pm r$), so by the standard use of triangle inequality and Cauchy–Schwarz, we obtain

$$\gamma(x,v) T^2 N = \left| \sum_{r} g_1(r) g_2(r) g_3(r) \right| \leq \| g_1 \|_\infty \| g_2 \|_2 \| g_3 \|_2.$$

Yet $g_2, g_3$ are (essentially) Fourier coefficients of functions $\mathbb{Z}_N \to D$ supported on $P, P + P$, respectively, so by Parseval their square-norms are bounded by $\sqrt{NT}, \sqrt{N(2T)}$. So $|g_1(r)| \geq \gamma(x,v) T / \sqrt{2}$ for some frequency $r$, whence there exists a quadratic polynomial $\psi_{x,v}$ such that $\sum_{k \in P} f(x - k) \omega^{-\psi_{x,v}(k)} \geq \gamma(x,v) T / \sqrt{2}$ (as $g_1(r) = \sum_{k \in P} h_1(k) \omega^{-rk} = \sum_{k \in P} f(x - k) \omega^{-(ak + cl)} \omega^{-rk}$).
Now for fixed $x$, take $v_x$ with $\gamma(x,v_x)$ maximal. Then
\[ \sum_x \sum_{k \in P} f(x - k) \omega^{-\psi(x,v_x(k))} \beta N : T/\sqrt{2}, \]
which is the desired conclusion up to translation and re-indexing. \qed

6. Weyl’s inequality (technical ingredient)

Theorem 6.1. Let $N$ be sufficiently large and let $a \in \mathbb{Z}_N$. For any $t \leq N$ there exists $p \leq t$ such that $|p^2a| \leq Ct^{-\frac{1}{2}}N$, where $C$ is an absolute constant.

Outline. Ignoring the explicit term $t^{-\frac{1}{2}}$, this is essentially asserting that for any irrational $\alpha \in (0,1)$ the sequence $(\alpha, 2^2\alpha, 3^2\alpha \ldots)$ is equidistributed modulo 1, meaning that for each interval $(a,b)$ the density of numbers inside that interval is $(b-a)$, or equivalently that each non-zero Fourier coefficient of the probability measure $\mu = \frac{1}{N} \sum_{k=1}^{N} \delta_{k^2\alpha}$ converges to 0. A well-known result of Van der Corput states that if, for each fixed $m \in \mathbb{Z}$, the sequence $(a_{n+m} - a_n)$ is equidistributed, then so is the sequence $(a_n)$. So the equidistribution of $(k^2\alpha)$ reduces to the equidistribution of the sequence $(ka')$ which is clear.

Obtaining the explicit estimate given is just a matter of chasing through the dependencies in the proof of Van der Corput’s theorem. For details, we refer the reader to section 5 of [1].

Lemma 6.2. Let $\phi : \mathbb{Z}_N \rightarrow \mathbb{Z}_N$ be affine (so $\phi(x) = ax + b$) and let $r, s \leq N$. For some $m \leq (\frac{2rN}{s})^{\frac{1}{2}}$, the set $\{0, 1, \ldots, r-1\}$ can be partitioned into arithmetic progressions $P_1, \ldots, P_m$ such that the diameter of $\phi(P_j)$ is at most $s$ for each $j$. Moreover the sizes of the $P_j$ differ pairwise by at most 1.

Proof. Though there are lots of variables, the idea is simply to use the usual rational approximation theorem of Dirichlet to find $u$ with $|\phi(u) - \phi(0)|$ small, and break up $\{0, 1, \ldots, r-1\}$ along multiples of $u$. Let $t \geq \sqrt{\frac{2rN}{s}} \geq \sqrt{r}$ be an integer. By pigeonhole there is some $u \leq t$ with $|\phi(u) - \phi(0)| \leq \frac{N}{t}$. Split $\{0, 1, \ldots, r-1\}$ into $u$ residue classes modulo $u$, so each residue class is an arithmetic progression. For any set $P$ of at most $\frac{s}{N}$ consecutive elements of such an arithmetic progression, we easily have that $\phi(P)$ has diameter at most $s$. Since the congruence classes have sizes differing by at most 1, it’s easy to divide them into arithmetic progressions $P_j$ of sizes differing by at most 1 with sizes between $\frac{st}{2N}$ and $\frac{st}{N}$ each, hence with diameter at most $s$. This makes the number of $P_j$ at most $\frac{2rN}{st} \leq \sqrt{\frac{2rN}{s}}$, as desired. \qed

Proposition 6.3. There is an absolute constant $C$ such that the following holds. Let $\psi : \mathbb{Z}_N \rightarrow \mathbb{Z}_N$ be any quadratic polynomial (so $\psi(x) = ax^2 + bx + c$) and let $r \in \mathbb{N}$. For some $m \leq Cr^{-\frac{1}{12}}$, the set $\{0, 1, \ldots, r-1\}$ can be partitioned into arithmetic progressions $P_1, \ldots, P_m$ such that the diameter of $\psi(P_j)$ is at most $Cr^{-\frac{1}{12}} N$, for each $j$, and the lengths of any two $P_j$ differ by at most 1.

Proof. The idea will be to use proposition 6.1 to break $\mathbb{Z}_N$ into progressions with common difference $p$, where $|ap^2|$ is small, and then account for the linear terms of $\psi$ using lemma 6.2 separately on each progression.
By Theorem 6.1 there is \( p \leq \sqrt{r} \) such that \(|ap^2| \leq C_1 r^{-\frac{1}{128}} N\). Then for any \( s \) we have

\[
\psi(x + sp) = s^2 (ap^2) + \theta_x(p)
\]

where \( \theta_x(\cdot) \) is an affine function depending on \( x \). Take \( u = \frac{1}{32} \) and partition \( \{0, 1, \ldots, r-1\} \) into progressions

\[
Q_j = \{x_j, x_j + p, \ldots, x_j + (u_j - 1)p\}.
\]

For any \( P \subseteq Q_j \) we have

\[
diam(\psi(P)) \leq C_1 u^2 r^{-\frac{1}{128}} N + diam(\theta_{x_j}(P))
\]

Now applying lemma 6.2 to each \( Q_j \) separately with \( s = 2r^{-\frac{1}{128}} N \) gives the result.

\[\square\]

**Corollary 6.4.** Let \( \psi : \mathbb{Z}_N \rightarrow \mathbb{Z}_N \) be a quadratic polynomial (so \( \psi(x) = ax^2 + bx + c \)) and \( r \leq N \), and fix \( \alpha > 0 \). There exists \( m \leq Cr^{-\frac{1}{1024}} \) and a partition of \( \{0, 1, \ldots, r-1\} \) into arithmetic progressions \( P_1, \ldots, P_m \) such that the sizes of the \( P_j \) differ by at most 1, and if \( f : \mathbb{Z}_N \rightarrow D \) is any function such that

\[
|\sum_{x=0}^{r-1} f(x) \omega^{-\psi(x)}| \geq \alpha r
\]

then

\[
\sum_{j=1}^{m} \left| \sum_{x \in P_j} f(x) \right| \geq \frac{\alpha r}{2}.
\]

**Proof.** Pick \( (P_j) \) as in proposition 6.3. For large \( r \), we have \( CNr^{-\frac{1}{1024}} \leq \frac{\alpha N}{4\pi} \) and so if \( x \in P_j \) then \( |\omega^{-\psi(x)} - \omega^{-\psi(x_j)}| \leq \frac{\pi}{2} \). The triangle inequality shows

\[
\sum_{j=1}^{m} \left| \sum_{x \in P_j} f(x) \omega^{-\psi(x)} \right| \geq \alpha r.
\]

Therefore we have

\[
\sum_{j=1}^{m} \left| \sum_{x \in P_j} f(x) \right| = \sum_{j=1}^{m} \left| \sum_{x \in P_j} f(x) \omega^{\psi(x_j)} \right| \\
\geq \sum_{j=1}^{m} \left| \sum_{x \in P_j} f(x) \omega^{\psi(x)} \right| - \frac{\alpha}{2} \sum_{j=1}^{m} |P_j| \\
\geq \frac{\alpha r}{2}
\]

as claimed.

\[\square\]
7. Putting everything together

**Theorem 7.1.** There is an absolute constant $C$ such that if $A \subseteq \mathbb{Z}_N$ has cardinality $|A| \geq \delta N$ and if $N \geq \exp \exp(\delta^{-C})$, then $A$ contains an arithmetic progression of length 4.

**Proof.** We use the preceding work to obtain a density increment if $A$ has no arithmetic progression of length 4. For in such a case, by corollary 2.13 we see that $A$ is not quadratically $\alpha$-uniform for $\alpha = 2^{-208} \delta^{112}$. Letting $f = 1_A - |A|$ be the balanced function of $A$, we see from Proposition 3.1 that there exists $B \subseteq \mathbb{Z}_N$ of cardinality at least $\alpha N/2$, and $\phi : B \to \mathbb{Z}_N$ such that $|\Delta(f; k)^{-1}(\phi(k))| \geq (\alpha/2)^{1/2} N$ for all $k \in B$. In particular, $\sum_{k \in B} |\Delta(f; k)^{-1}(\phi(k))|^2 \geq (\alpha/2)^2 N^3$.

But then Proposition 3.3 says that $\phi$ has $(\alpha/2)^8 N^3 = \Omega(N^3)$ additive quadruples. By 4.9 there is an arithmetic progression $P$ with cardinality $T \geq N^{\gamma(\delta)}$ and $\beta(\delta) > 0$ and constants $\lambda, \mu$ such that

$$\sum_{k \in P} |\Delta(f; k)^{-1}(\lambda k + \mu)|^2 \geq \beta N^2 T.$$  

Applying 5.1 gives locally defined quadratic polynomials $\psi_0, \ldots, \psi_{N-1}$ such that

$$\sum_s |\sum_{z \in P+s} f(z)\omega^{-\psi_s(z)}| \geq \frac{\beta}{\sqrt{2}} NT.$$  

Now by corollary 6.4 we can partition each $P + s$ into progressions $P_{s1}, \ldots, P_{sm}$ modulo $\mathbb{Z}_N$ with cardinalities differing by at most 1 of size $\Omega(T^{1/2})$ such that

$$\sum_s \sum_{j=1}^m |\sum_{x \in P_{sj}} f(x)| \geq \frac{\beta}{2\sqrt{2}} NT.$$  

Now let $p_{sj} = \sum_{x \in P_{sj}} f(x)$ and $q_{sj} = \max(p_{sj}, 0)$. Since $\sum_{s,j} p_{sj} = 0$, we have $\sum_{s,j} q_{sj} \geq \frac{\beta}{4\sqrt{2}} NT$. Hence there exists some $s, j$ with $\sum_{x \in P_{sj}} f(x) \geq \frac{\beta}{4m\sqrt{2}} T \geq \Omega(\beta T^{1/2})$. Restricting to this $P_{sj}$ thus gives density $\delta + \varepsilon(\delta)$. Since $P_{sj}$ was large depending on $N, \delta$, iterating this density increment gives an eventual contradiction.

A small point is that for the induction, we have to treat $\mathbb{Z}_N$ as the set $(1, 2, \ldots, N)$ rather than as a cyclic group. But we can just break the progressions $P_{sj}$ up into smaller pieces when they wrap around, and since all common differences of progressions were fairly small, this doesn’t cause real issues. The theorem stated follows by chasing dependencies of constants through the proof. \qed

8. Remarks on Generalizations and Other Approaches

Gowers extended his proof above to show the existence of length $k$ progressions in dense subsets of $\mathbb{Z}_N$ for arbitrary $k$ (2). An improvement in the use of Freiman’s theorem results in the bound $N \geq \exp \exp(\delta^{-C})$, a bit better than the bound stated above, where $C(k) = 2^{2k+9}$. The full proof is more involved than the proof for length 4 progressions presented above, though it follows the same general steps. Szemerédi’s theorem also generalizes in much the same way to arbitrary finite abelian groups; in fact, to show the theorem in an arbitrary abelian group it suffices to establish it in $\mathbb{Z}_N$ as above, and for vector spaces over $\mathbb{F}_p$ for fixed $p$. 

The Fourier analytic argument above is one of 4 known approaches to Szemerédi’s theorem. Aside from Szemerédi’s original proof, there are proofs based on hypergraph regularity and ergodic theory, which also exploit a dichotomy between structure and randomness; see chapter 11 of [4] for proof outlines and more detailed sources, as well as extensions such as the Green-Tao theorem on the existence of arbitrarily long progressions in the primes.

**References**


