1 Introduction

Given a function $h : \mathbb{F}_p \times \mathbb{F}_p \times \cdots \times \mathbb{F}_p \to \mathbb{C}$, we define the usual expectation operator

$$E_{n_1, \ldots, n_k}(h) := p^{-k} \sum_{n_1, \ldots, n_k \in \mathbb{F}_p} h(n_1, \ldots, n_k).$$

We also define, for $f : \mathbb{F}_p \to \mathbb{C}$, the operator

$$\Lambda(f) := E_{n,d}(f(n)f(n+d)f(n+2d)).$$

If $f$ were an indicator function for some set $S \subseteq \mathbb{F}_p$, this would give a normalized count of the number of three-term progressions in $S$.

In the present paper we establish a new structure theorem for functions $f : \mathbb{F}_p \to [0, 1]$ that minimize the number of three-term progressions, subject to a density constraint; and, as a consequence of this result, we prove a further structural result, which can also be deduced from the work of Green [3], though only for high densities (Green’s result only works for densities exceeding $1/\log^*(p)$, though perhaps his method can be generalized for this particular problem to handle lower densities).

Before stating the theorem, it is worth mentioning that Green and Sisask [5] have shown that sets of high density (density close to 1) that minimize the number of three-term arithmetic progressions, are the complement of the union of two long arithmetic progressions (actually, their result is stated in terms of sets that maximize the number of three-term progressions, but there is a standard trick to relate this to the minimizing sets).

Our main theorem is now given as follows:

**Theorem 1** Suppose that

$$f : \mathbb{F}_p \to [0, 1]$$

...
minimizes \( \Lambda(f) \), subject to the constraint that

\[ \Lambda(f) \geq \theta \in (0, 1]. \]

Then,

- Let \( C(n) \) equal \( f(n) \) rounded to the nearest integer, which is therefore 0 or 1. Then,

\[ \sum_n |f(n) - C(n)| \ll p(\log p)^{-2/3}. \]

So, \( f \) must be approximately an indicator function. Furthermore, we get the same conclusion if \( f \) satisfies \( \mathbb{E}(f) \geq \theta \), and \( \Lambda(f) \) comes within \( O(1/p) \) of the minimal value for this density constraint.

- There exists a function \( r : \mathbb{F}_p \to [0, 1] \) such that \( \mathbb{E}(r) = \mathbb{E}(f) \), where \( \Lambda(r) \) is very close to the minimal \( \Lambda(f) \), specifically

\[ \Lambda(r) = \Lambda(f) + O(p^{-1}), \]

such that if we let, for some \( L \),

\[ S := \{ n \in \mathbb{F}_p : (r \ast r)(2n) + 2(r \ast g)(-n) \leq L \}, \]

where \( g(n) := r(-n/2) \), then

\[ \sum_n |r(n) - S(n)| \ll p|B|(\log \log p)^{-2/3}. \quad (1) \]

(Please see subsection 1.1 for an explanation of this part of the theorem.)

- We have that there exist sets \( A \) and \( B \) of \( \mathbb{F}_p \), with \( |A| > p^{1-o(1)} \) and \( |B| > p^{1/2} \), such that the set for which \( f \) is approximately an indicator function, is roughly the sumset \( A + B \). More precisely: If we let \( C(n) \) denote \( f \) rounded to the nearest integer, as in the first bullet above, then

\[ \sum_n |(A \ast B)(n) - |B|C(n)| \ll p|B|(\log \log p)^{-2/3}. \]

Furthermore, we may take \( A = C \) and take \( B \) to be a certain “Bohr neighborhood” \( B \), which is described in the proof of the theorem.

1.1 A remark on the second part of the theorem

By (1) we see that \( r \) is nearly an indicator function for the set \( S \). Let us suppose, for the purposes of discussion, that it is exactly an indicator function for some set, and let \( R \) be this set. Note that \( R \) and \( S \) must have small symmetric difference.

When does an \( n \in \mathbb{F}_p \) belong to the set \( S \)? To decide this, given \( n \), we let \( N_1 \) be the number of pairs \((x, y) \in R \times R \) such that \( n, x, y \) forms an arithmetic progression; we let \( N_2 \) be the number of pairs \((x, y) \in R \times R \) such that \( x, n, y \)
is an arithmetic progression; and, we let $N_3$ be the number of such ordered pairs where $x, y, n$ is an arithmetic progression. For $n$ to belong to $S$, we must have that

$$N_1 + N_2 + N_3 \leq L.$$  

Since $S$ and $R$ have small symmetric difference, we see that our $r$ can be thought of as enjoying a “local minimal” property: Not only does $r$ minimize $\Lambda(r)$ up to an error $O(1/p)$, subject to $E(r) \geq \theta$, but we can easily decide whether $n \in \mathbb{F}_p$ belongs to $R$ or not, simply by checking to see how many progressions pass through the point $n$, with the other two end-points in $R$. If this count is small enough, then $n$ likely belongs to $R$ (though it certainly belongs to $S$); but, if the count is large, $n$ likely does not belong to $R$.

The most difficult part of this proof that $R$ and $S$ are nearly the same, is handling those $n$ where $N_1 + N_2 + N_3$ exactly equals $L$.

1.2 Remarks on the third part of the theorem

One reason to believe the third bullet above is that from the second bullet we expect that $f$ is an indicator function for a level set of a “smooth function” $(r * r)(2n) + 2(r * g)(-n)$; and, as is well known, such level sets must be approximately the union of a bunch of translates of a Bohr neighborhood of the function, at least when their density is large enough.

It should be remarked that sumsets are quite special structures, as are smooth functions of the type $(r * r)(2n) + 2(r * g)(-n)$, and only a vanishingly small proportion of the subsets of $\mathbb{F}_p$ are sumsets or form the support of a smooth function; so, the third bullet is saying something fairly non-trivial about our minimal $f$.

Also, there are loads of other consequences that one can deduce from the third bullet. One of these is that, upon decomposing the Bohr neighborhood $\mathcal{B}$ into a union of arithmetic progressions, one can deduce that $C$ is essentially the union of a “small number” of somewhat “long” arithmetic progressions (“small number” can mean a power of $p$, say $p^c$, where $c < 1$), all having the same common difference.

2 Proof of Theorem 1

The proof of this structure theorem depends on a certain function $r_3$, which we presently define.

**Definition.** Given a subset $S$ of a group $G$, we let $r_3(S)$ denote the size of the largest subset of $S$ free of solutions to $x + y = 2z, x \neq y$. In all the uses
of $r_3$ in the present paper, $G = \mathbb{Z}$ and $S = [N] := \{1, 2, ..., N\}$, for various different values of $N$.

Bourgain [2] has recently shown that

$$r_3([N]) \ll N(\log N)^{-2/3},$$  

(2)

and from a result of Behrend [1], we know that for $N$ sufficiently large,

$$r_3([N]) > N \exp(-c\sqrt{\log N}),$$

for a certain constant $c > 0$.

2.1 Proof of the first part of Theorem [1]

For this part we will begin by assuming that $\mathbb{E}(f) > \kappa p(\log p)^{-2/3}$, for as large a $\kappa > 0$ as we might happen to need, since this part of the theorem is trivially true otherwise.

Here we will first show that the minimal $f$ is well-approximated by an indicator function; actually, we will prove even more – we will show that if $\Lambda(f)$ comes within $O(p^{-1})$ of this smallest value, subject to the density constraint $\mathbb{E}(f) > \theta$, then $f$ must be approximately an indicator function. To do this, we will require the following proposition, proved in subsection [2.3].

Proposition 1 Suppose that $A$ and $B$ are disjoint subsets of $\mathbb{F}_p$, such that $f : \mathbb{F}_p \rightarrow [0, 1]$ has the property

$$\text{for } n \in A, \ f(n) \leq 1 - \varepsilon, \ 0 < \varepsilon < 1/3,$$

and suppose that

$$\text{support}(f) = A \cup B.$$

Then, for $\beta > 0$ satisfying

$$\varepsilon\beta \geq p^{-1/2}\log p,$$

there exists a function $g : \mathbb{F}_p \rightarrow [0, 1]$ such that

$$\mathbb{E}(g) \geq \mathbb{E}(f),$$

and yet

$$\Lambda(g) < \Lambda(f) + 2\beta - \varepsilon^2 p^{-2}W_0/4 + O(p^{-1}),$$

where

$$W_0 := \sum_{a,a+d,a+2d \in A} f(a)f(a+d)f(a+2d).$$
We also will require the following quantitative version of Varnavides’s theorem [7].

**Lemma 1** If $S \subseteq \mathbb{F}_p$ satisfies $|S| \geq 2(r_3(N)/N)p$, we will have for any $2 \leq N \leq p$ that

$$\Lambda(S) \geq \frac{2r_3([N])}{N^3 + O(N^2)}.$$

**Proof of the Lemma.** The proof of this lemma is via some easy averaging: We let $A_N$ denote the set of all arithmetic progressions $A \subseteq \mathbb{F}_p$ having length $N$. These arithmetic progressions are to be identified by ordered pairs $(a, d)$, $d \neq 0$, where $a$ is the first term in the progression, and where $d$ is the common difference. Note that this means we “double count” arithmetic progressions in that the progression $a, a+d, a+2d, \ldots, a+kd$ is distinct from $a+kd, a+(k-1)d, \ldots, a$.

It is easy to check that each sequence $a, a+d, a+2d, d \neq 0$ is contained in exactly $N^2/2 + O(N)$ of these $A \in A_N$: We have that each three-term progression is contained in the same number of $A \in A_N$; and each $A \in A_N$ contains $N^2/2 + O(N)$ three-term progressions; hence, if $P$ denotes the number of $A \in A_N$ containing a particular sequence $a, a+d, a+2d$, we have since there are $p(p-1)$ non-trivial progressions in $\mathbb{F}_p$, that

$$p(p-1)P = |A_N|(N^2/2 + O(N)),$$

whence $P = N^2/2 + O(N)$.

So, if we let $T_3(X)$ denote the number of sequences $a, a+d, a+2d \in X$, $d \neq 0$, we have that

$$T_3(S) = (N^2/2 + O(N))^{-1} \sum_{A \in A_N} T_3(A \cap S). \tag{3}$$

Next, we need a lower bound on how many $A \in A_N$ satisfy $|A \cap S| \geq r_3(N)$: First, note that for each $d \in \mathbb{F}_p$, $d \neq 0$, there are exactly $N$ arithmetic progressions $A \in A_N$ having common difference $d$ that contain a particular point $a \in \mathbb{F}_p$. So,

$$\sum_{A \in A_N} |A \cap S| = \sum_{s \in S} \sum_{d \in \mathbb{F}_p \backslash \{0\}} N = (p-1)N|S|.$$

Let $Y$ be the number of $A \in A_N$ for which $|A \cap S| > r_3(N)$. Then, we have

$$(|A_N| - Y)r_3(N) + YN \geq (p-1)N|S|,$$

which implies

$$Y \geq \frac{(p-1)N|S| - |A_N|r_3(N)}{N - r_3(N)} \geq (p-1)|S| - |A_N|(r_3(N)/N).$$
For each of these \( Y \) progressions \( A \in \mathcal{A}_N \) we will have that \( T_3(A \cap S) \geq 1; \) and so, we deduce from (3) that
\[
T_3(S) \geq \frac{(p - 1)|S| - |\mathcal{A}_N|(r_3(N)/N)}{N^2/2 + O(N)}.
\]

Using the easy to see fact that \( |\mathcal{A}_N| = p(p - 1) \), we deduce that if
\[
|S| > 2(r_3(N)/N)p,
\]
then
\[
T_3(S) \geq \frac{2p^2(r_3(N)/N)}{N^2 + O(N)}.
\]

The lemma easily follows on rephrasing this in terms of \( \Lambda(S) \). ■

Now we let
\[
A := \{ n \in \mathbb{F}_p : f(n) \in [\varepsilon, 1 - \varepsilon]\},
\]
where \( \varepsilon > 0 \) will be determined later. In order for \( f \) to be minimal, from Proposition 1 we deduce that we must have that if \( \varepsilon \beta = p^{-1/2} \log p \), then
\[
\beta \geq \varepsilon^2 p^{-2} W_0/8 + O(1/p).
\]

So, since we trivially have that
\[
W_0 \geq \varepsilon^3 p^2 \Lambda(A),
\]
it follows that
\[
\Lambda(A) \leq 8\varepsilon^{-4} p^{-1/2} \log p. \tag{4}
\]

We would like to now apply Lemma 1 to this, but in order to do so, we must solve for \( N \) such that
\[
|A| > 2r_3(N)p/N.
\]

To this end, we require the bound (2) of Bourgain, which implies that if we let
\[
N = \exp(c(p/|A|)^{3/2}) < p, \text{ since } |A| > \kappa p (\log p)^{-2/3},
\]
then we will have that
\[
|A| > p(\log N)^{-2/3} > 2r_3(N)p/N,
\]
as we require.

From this it follows from Lemma 1 that
\[
\Lambda(A) > r_3(N)/N^3 > 1/N^3 > \exp(-3c(p/|A|)^{3/2}).
\]
It follows now from (4) that
\[
|A| \ll p \log^{-2/3}(\varepsilon^{12}p), \text{ for } \varepsilon > p^{-1/12} \log p.
\]

So, if we let \( C \) be the function \( f \) rounded to the nearest integer (which will be either 0 or 1), then for \( n \in A \) we will have \( |f(n) - C(n)| \leq 1 \), while for all other \( n \) we will have \( |f(n) - C(n)| \leq \varepsilon \). It follows that
\[
\sum_n |f(n) - C(n)| \ll (\varepsilon + (\log \varepsilon^{12}p)^{-2/3})p, \text{ for } \varepsilon > p^{-1/12} \log p.
\]
Choosing \( \varepsilon = (\log p)^{-2/3} \), we deduce that this sum is \( O(p(\log p)^{-2/3}) \), just as in Bourgain’s theorem (2). This completes the proof of the first part of our theorem.

### 2.2 Proof of the second part of Theorem 1

Given a function \( h : \mathbb{F}_p \to [0, 1] \), we let
\[
h_2(n) := h(-n/2),
\]
and then we define
\[
F_h(n) := (h * h)(2n) + (h * h_2)(-n).
\]

In order to proceed further, we will require the following proposition.

**Proposition 2** Fix \( A \subseteq \mathbb{F}_p \), and associate to each \( a \in A \) a real number \( w_a \in [0, 1] \). Among all functions \( h : \mathbb{F}_p \to [0, 1] \) satisfying
\[
\mathbb{E}(h) = \gamma > \sum_{a \in A} w_a,
\]
those which minimize \( \Lambda(h) \) have the property that there exists \( L > 0 \) such that
\[
\text{for } n \in \mathbb{F}_p \setminus A, \ h(n) = \begin{cases} 1, & \text{if } F_h(n) < L; \\ 0, & \text{if } F_h(n) > L. \end{cases}
\]

The proof of this proposition can be found in subsection 2.5.

We will use this proposition to construct a sequence of sets
\[
A_1, A_2, \ldots \subseteq \mathbb{F}_p,
\]
and a sequence of functions
\[
r_1, r_2, \ldots : \mathbb{F}_p \to [0, 1], \text{ and all } \mathbb{E}(r_i) = \mathbb{E}(f).
\]
such that the following all hold.
• First, $|A_1| = 2$ and $A_{i+1} = A_i \cup \{x_{i+1}, y_{i+1}\}$;
• second, $\Lambda(r_i) \leq \Lambda(f) + 5p^{-2}|A_i|$;
• third, given particular fixed values for $r_i(n)$ on $A_i$, we have that $r_i$ minimizes $\Lambda(r_i)$, subject to the density constraint $E(r_i) = E(f)$;
• and finally, for each $n \in A_i$, $r_i(n) \in [1/4, 3/4]$.

Clearly, this process cannot continue past the $\lfloor p/2 \rfloor$th iteration, as the sets $A_i$ grow by two elements after each iteration. Furthermore, we will show that whenever the process does terminate (which it will in either case 1 or case 2 below), we will be left with a function $r : \mathbb{F}_p \to [0, 1]$ satisfying the conclusion in the second bullet of Theorem 1.

For the time being, let us suppose that these sequences can be constructed as claimed: Suppose we have constructed $A_i$; we will now show how to construct $A_{i+1}$. To this end, we apply Proposition 2 with $A = A_i$ (in the case $i = 0$ we let $A$ be the empty set), and then we deduce that for some $L > 0$, $r_i(n) = 1$ (we use $r_0(n) := f(n)$) for $F_{r_i}(n) < L$ and $r_i(n) = 0$ for $F_{r_i}(n) > L$. We furthermore apply the already-proved first part of Theorem 1 from subsection 2.1 and deduce that since $\Lambda(r_i) \leq \Lambda(f) + 10ip^{-2}$,

$$\sum_n |r_i(n) - C(n)| \ll p(log p)^{-2/3},$$

where $C(n)$ is $r_i(n)$ rounded to the nearest integer.

Now, if there are two distinct places $x, y \in \mathbb{F}_p \setminus A_i$ for which

$$r_i(x), r_i(y) \in [1/4, 3/4],$$

then we just let

$$A_{i+1} := \{x, y\}, \text{ and } r_{i+1} := r_i.$$

So suppose that there are no such $x$ and $y$; there are three possibilities to consider.

2.2.1 Case 1: $r_i(n) \geq 1/2$ for all $n \in \mathbb{F}_p \setminus A_i$ where $F_{r_i}(n) = L$.

Note that we include in this case the possibility that there are no $n \in \mathbb{F}_p \setminus A_i$ such that $F_{r_i}(n) = L$.

If we are in this case, then it means that

$$\sum_{n \in \mathbb{F}_p \setminus A_i : F_{r_i}(n) = L} |r_i(n) - 1| \ll p(log p)^{-2/3};$$

and so, $r_i(n)$ is very close to 1 at most places $n \in \mathbb{F}_p \setminus A_i$ where $F_{r_i}(n) = L$. It follows that if we were to let $S$ be the set of all $n \in \mathbb{F}_p \setminus A_i$ with $F_{r_i}(n) \leq L$, then

$$\sum_{n \in S} |r_i(n) - 1| \ll p(log p)^{-2/3}.$$
In order to extend this sum to all $n \in \mathbb{F}_p$, we will need to show that $|A_i|$ cannot be too big. Basically, we will show that if it is, then $\Lambda(f)$ could not be minimal.

To see this last point, we apply Proposition 1 using $A := A_i$, $h := r_i$, $\varepsilon = 1/4$, and $\varepsilon\beta = p^{-1/2}\log p$, and we deduce that there exists

$$g : \mathbb{F}_p \to [0, 1], \quad E(g) \geq E(r_i),$$

and yet

$$\Lambda(g) \leq \Lambda(f) + 8p^{-1/2}\log p - p^{-2}W_0/64 + O(p^{-1}),$$

where

$$W_0 := \sum_{a,a+d,a+2d \in A} r_i(a)r_i(a+d)r_i(a+2d) \geq 4^{-3}p^2\Lambda(A).$$

We wish to apply Lemma 1: First, let

$$N = \exp(c(p/|A_i|)^{3/2}) < p,$$

such that from (2) we deduce that

$$|A_i| > 2p(\log N)^{-2/3} > 2r_3([N])p/N,$$

as we require.

From this it follows now from Lemma 1 that

$$\Lambda(A_i) \geq \frac{2r_3(N)}{N^3 + O(N^2)} > 1/N^3 \gg \exp(-3c(p/|A_i|)^{3/2}),$$

(5)

for $N$ sufficiently large.

In order for $f$ to minimize $\Lambda(f)$, we must have that

$$8p^{-1/2}\log p = 2\beta \geq p^{-2}W_0/64 + O(1/p);$$

so, ignoring the $O(1/p)$, we see that

$$\Lambda(A_i) \leq 4^3p^{-2}W_0 \leq 4^8p^{-1/2}\log p.$$

It follows from this and (5) that

$$|A_i| \ll p(\log p)^{-2/3},$$

as claimed. It follows that if we extend $S$ to be the set of all $n$ where $F_{r_i}(n) \leq L$, then

$$\sum_{n}|r_i(n) - S(n)| \ll p(\log p)^{-2/3},$$

and the second bullet of Theorem 1 is proved upon setting $r = r_i$. 

9
2.2.2 Case 2: \( r_i(n) < 1/2 \) for all \( n \in \mathbb{F}_p \setminus A_i \) where \( F_{r_i}(n) = L \).

If we are in this case, then it means that

\[
\sum_{n \in \mathbb{F}_p \setminus A_i} \text{ for } F_{r_i}(n) = L \sum_{n \in \mathbb{F}_p \setminus A_i} r_i(n) \ll p (\log p)^{-2/3},
\]

and so, if we let \( L' = L - \delta \), for small enough \( \delta > 0 \), then we will have that for \( n \in \mathbb{F}_p \setminus A_i \), \( r_i(n) = 1 \) for \( F_{r_i}(n) \leq L' \), while \( r_i(n) \) is usually near 0 when \( F_{r_i}(n) > L' \). It follows then that if we let \( S \) be the set of \( n \in \mathbb{F}_p \setminus A_i \) where \( F_{r_i}(n) \leq L' \), then

\[
\sum_{n \in S} |r_i(n) - 1| \ll p (\log p)^{-2/3},
\]

We wish to extend this to where \( S \) is the set of all \( n \) satisfying \( F_{r_i}(n) \leq L' \), by showing that \( |A_i| \) cannot be too big, and we proceed exactly the same way as in Case 1 above. We then deduce that, upon redefining \( S \) in this way, that

\[
\sum_{n} |r_i(n) - S(n)| \ll p (\log p)^{-2/3},
\]

and again this proves the second bullet of Theorem 1 upon setting \( r = r_i \).

2.2.3 Case 3: There exists \( x, y \in \mathbb{F}_p \setminus A_i \) where \( F_{r_i}(x) = F_{r_i}(y) = L \), and \( r_i(x) < 1/2 < r_i(y) \).

In order to decide what to do in this case, we will require the following basic fact, which is an immediate consequence of the formula for \( \Lambda(h_3) \) in the proof of Proposition 2 in section 2.5 in equation (10): We have that if we let

\[
r_{i+1}(n) := \begin{cases} \begin{array}{ll}
  r_i(n), & \text{if } n \neq x, y; \\
  \frac{(r_i(x) + r_i(y))}{2}, & \text{if } n = x \text{ or } y,
\end{array} \end{cases}
\]

then

\[
\Lambda(r_{i+1}) \leq \Lambda(r_i) + p^{-2} (r_{i+1}(x) - r_i(x)) F_{r_i}(x) + p^{-2} (r_{i+1}(y) - r_i(y)) F_{r_i}(y) + 10 p^{-2}
\]
\[
\leq \Lambda(r_i) + 10 p^{-2}.
\]

So, when we are in this case, we just let

\[
A_{i+1} := \{x, y\},
\]

and so the properties of \( A_{i+1}, r_{i+1} \) that we require all hold.

2.3 Proof of the third part of Theorem 1

We assume for this part of the proof of our theorem that \( \theta > (\log \log p)^{-2/3} \), since our problem is trivial otherwise.
We now prove the third bullet of Theorem \[ \text{III} \]. To this end, we let
\[
f_3(n) := (f * \mu)(n),
\]
where \( \mu \) is defined as follows: First, we locate the places \( b_1, ..., b_t \) where the Fourier transform
\[
|\hat{f}(b_i)| > \varepsilon_0 p,
\]
where \( \varepsilon_0 > 0 \) will be decided later, and then we define the Bohr neighborhood \( \mathcal{B} \) to be all those \( n \in \mathbb{F}_p \) where
\[
||b_in/p|| < \varepsilon_0, \text{ for all } i = 1, ..., t.
\]
Finally, we just let \( \mu(n) = 1/|\mathcal{B}| \) if \( n \in \mathcal{B} \), and \( \mu(n) = 0 \) otherwise.

Our goal now will be to show that
\[
\sum_n|f_3(n) - f(n)| \ll p(\log \log p)^{-2/3},
\]
for this will imply the third bullet of Theorem \[ \text{III} \] holds: To see this, note that from the already-proved first bullet, we know that if we let \( C(n) \) be \( f(n) \) rounded to the nearest integer, then
\[
\sum_n||\mathcal{B}||^{-1}(C * \mathcal{B})(n) - C(n)| = \sum_n||\mathcal{B}||^{-1}(f * \mathcal{B})(n) - f(n)| + O(p(\log p)^{-2/3})
\]
\[
= \sum_n|f_3(n) - f(n)| + O(p(\log p)^{-2/3})
\]
\[
\ll p(\log \log p)^{-2/3},
\]
which is just what the third bullet claims.

Now we show that (6) holds: First note that Parseval gives
\[
t \leq \theta \varepsilon_0^{-2};
\]
and the following standard lemma tells us that our Bohr neighborhood is "large".

**Lemma 2** We have that
\[
|\mathcal{B}| \geq (\varepsilon_0 + O(1/p))^t p.
\]

**Proof of the lemma.** For \( i = 1, 2, ..., t \), we let
\[
\alpha_i(x) := (\varepsilon_0 p + 1)^{-1} \left( \sum_{||b_in/p||<\varepsilon_0/2} e^{2\pi i nx/p} \right)^2
\]
We note that \( \alpha_i(x) \) is always a non-negative real for all real numbers \( x \), and \( \alpha_i \) is the Fourier transform of a function \( \beta_i : \mathbb{F}_p \to [0, 1] \). Furthermore,
\[
|\alpha_i(0)| = \varepsilon_0 p + O(1).
\]
Now letting
\[ \beta(n) := (\beta_1 \cdots \beta_t)(n), \]
we find that \( \beta : \mathbb{F}_p \to [0, 1] \), and has support contained within \( \mathcal{B} \). So,
\[
|\mathcal{B}| \geq \hat{\beta}(0) = p^{-t+1}(\hat{\beta}_1 * \cdots * \hat{\beta}_t)(0) = p^{-t+1}(\alpha_1 * \alpha_2 * \cdots * \alpha_t)(0) \geq p^{-t+1}\alpha_1(0) \cdots \alpha_t(0) \geq (\varepsilon_0 + O(1/p))^t p.
\]

Now, from the easy-to-check fact that
\[
||\hat{f}_3(a) - \hat{f}(a)||_\infty = ||\hat{f}(a)(1 - \hat{\mu}(a))||_\infty \leq \varepsilon_0 p,
\]
we easily deduce, via standard arguments (Parseval and Cauchy-Schwarz) that
\[
\Lambda(f_3) = p^{-3}\sum_a \hat{f}_3(a)^2 \hat{f}_3(-2a) = p^{-3}\sum_a \hat{f}(a)^2 \hat{f}(-2a) + E = \Lambda(f) + E,
\]
where the “error” \( E \) satisfies
\[
|E| \leq 10\varepsilon_0.
\]

Now let \( A \) be all those \( n \in \mathbb{F}_p \) for which
\[
f_3(n) \in [\varepsilon_1, 1 - \varepsilon_1].
\]
Then, we have that
\[
W_0 := \sum_{a,a+d,a+2d \in A} f_3(a)f_3(a+d)f_3(a+2d) \geq \varepsilon_3^3 p^2 \Lambda(A).
\]
In order to apply Lemma 1 to this, we let
\[
N = \exp(c(p/|A|)^{3/2}) < p,
\]
so that from (2) we deduce that
\[
|A| > p(\log N)^{-2/3} > 2r_3(N)p/N,
\]
as we require.

From this it follows now from Lemma 1 that
\[
\Lambda(A) \geq \frac{2r_3(N)}{N^3 + O(N^2)} > \frac{1}{N^3} \gg \exp(-3(2p/|A|)^{3/2}),
\]

12
for $N$ sufficiently large.

In order for $\Lambda(f)$ to be minimal, we must have that

$$\Lambda(f) \leq \Lambda(f_3) \leq \Lambda(f) + 2\beta + 10\varepsilon_0 - \varepsilon_1^2 p^{-2} W_0/4 + O(1/p).$$

Setting $\beta = 5\varepsilon_0$ we must have

$$20\varepsilon_0 \geq \varepsilon_1^2 p^{-2} W_0/2 + O(1/p) \geq \varepsilon_1^5 \Lambda(A)/2 + O(1/p);$$

and so,

$$\Lambda(A) \leq 80\varepsilon_0 \varepsilon_1^{-5} + O(1/p).$$

Combining this with our lower bound for $\Lambda(A)$ above, we deduce that

$$|A| \ll p(\log \varepsilon_1^5 \varepsilon_0^{-1})^{-2/3}.$$

It now follows that if $C(n)$ is $f_3(n)$ rounded to the nearest integer, then

$$\sum_n |f_3(n) - C(n)| \leq \sum_{n \in A} 1/2 + \sum_{n \in \mathbb{F}_p \setminus A} \varepsilon_1 \ll p(\log \varepsilon_1^5 \varepsilon_0^{-1})^{-2/3} + \varepsilon_1 p.$$

Now we will set

$$\varepsilon_0 := \sqrt{\theta \log \log p / \log p}, \text{ and } \varepsilon_1 := (\log \log p)^{-2/3},$$

which will give

$$|B| > p^{1/2},$$

and then our sum on $|f_3(n) - C(n)|$ will be at most

$$\sum_n |f_3(n) - C(n)| \ll p(\log \log p)^{-2/3},$$

which completes the proof of Theorem 1.

### 2.4 Proof of Proposition 1

#### 2.4.1 Technical lemmas needed for the proof of the Proposition

We will need to assemble some lemmas to prove this proposition. We begin with the following standard fact:

**Lemma 3** Suppose that $S \subseteq \mathbb{F}_p$ satisfies $|S| = \alpha p$. Let $T$ denote the complement of $S$. Then, we have that

$$\Lambda(S) + \Lambda(T) = 1 - 3\alpha + 3\alpha^2.$$
Proof of the lemma. One way to prove this is via Fourier analysis: We have that
\[ \Lambda(S) + \Lambda(T) = p^{-3} \sum_a \hat{S}(a)^2 \hat{S}(-2a) + \hat{T}(a)^2 \hat{T}(-2a). \]
Since \( \hat{S}(a) = -\hat{T}(a) \) for \( a \neq 0 \), we have that all the terms except for \( a = 0 \) vanish. So,
\[ \Lambda(S) + \Lambda(T) = p^{-3}(\hat{S}(0)^3 + \hat{T}(0)^3) = \alpha^3 + (1 - \alpha^3) = 1 - 3\alpha + 3\alpha^2. \]

From this lemma, one can deduce the following corollary, which we state as another lemma:

Lemma 4 For \( \alpha > 2/3 \) we have that there exists a set \( S \subseteq \mathbb{F}_p \) satisfying \( |S| = \lfloor \alpha p \rfloor \), and
\[ \Lambda(S) \leq \alpha^3 \left( 1 - (1 - \alpha)^2 / 2 \right) + O(1/p). \]

Proof of the Lemma. Let \( \beta = 1 - \alpha < 1/3 \), and then let \( S \) just be the arithmetic progression \( \{0, 1, \ldots, \lfloor \alpha p \rfloor - 1\} \), and then let \( T \) be the complement of \( S \), which is also just an arithmetic progression. It is easy to check that
\[ \Lambda(T) = |T|^2 / 2p^2 + O(|T|/p^2) = \beta^2 / 2 + O(1/p), \]
as the solutions to \( x + y = 2z \), \( x, y, z \in T \) are exactly those ordered pairs \((x, z) \in T \times T\) of the same parity.

Applying Lemma 3 to this set \( T \), we find that
\[ \Lambda(S) = (1 - 3\beta + 3\beta^2) - \beta^2 / 2 + O(1/p) \]
\[ = 1 - 3\beta + 5\beta^2 / 2 + O(1/p) \]
\[ < (1 - \beta)^3 (1 - \beta^2 / 2) + O(1/p), \]
as claimed.

2.4.2 Body of the proof of Proposition

We will define the function \( g : \mathbb{F}_p \rightarrow [0, 1] \) such that
\[ \text{support}(g) \subseteq A \cup B, \]
where
\[ \text{for } n \in B, \ g(n) = f(n), \]

14
but on the set $A$, the function $g$ will be different from $f$: Basically, we let $S$ be the set produced by Lemma 4 with $\alpha = 1 - \varepsilon$, then take $T$ to be a random translate and dilate of $S$, say

$$T := m.S + t = \{ms + t : s \in S\}.$$  

Then, we let

$$g(n) = (1 - \varepsilon)^{-1}f(n)T(n).$$

Note that this is $\leq 1$, because we know $f(n) \leq 1 - \varepsilon$ on $A$.

We will show that, so long as there are “enough” three-term progressions lying in $A$, this new function $g$ will have the property that $\Lambda(g)$ is much smaller than $\Lambda(f)$. To this end, we consider three types of arithmetic progressions that give rise to the counts $\Lambda(f)$ and $\Lambda(g)$: Those progressions that pass through both $A$ and $B$ (say one point in $A$ and two in $B$; or two in $A$ and one in $B$); those that lie entirely within $A$; and those that lie entirely within $B$.

The contribution to $\Lambda(g)$ of those arithmetic progressions lying entirely within $B$ is the same as the contribution to $\Lambda(f)$. So, we don’t need to account for these when trying to prove our upper bound on $\Lambda(g)$; and therefore there are only two non-trivial cases that we need to work out:

**Case 1 (all three points in $A$).**

Define the random variable

$$Z_0 := \sum_{a,a+d,a+2d \in A} a(a+d)(a+2d),$$

and let $W_0$ be the analogous sum but with $g$ replaced by $f$. We note that if we only consider those terms with $d \neq 0$, we lose at most $O(p)$ in estimating $Z_0$.

We have that

$$\mathbb{E}(Z_0) = \sum_{a,a+d,a+2d \in A} \mathbb{E}(a(a+d)(a+2d)) + O(p)$$

$$= p^{-2}(1 - \varepsilon)^{-3} \sum_{a,a+d,a+2d \in A} f(a)f(a+d)f(a+2d) \sum_{m.t \in T_p} 1 + O(p)$$

$$= p^{-2}(1 - \varepsilon)^{-3} \sum_{a,a+d,a+2d \in A} \sum_{b,b+d',b+2d'+S} f(a)f(a+d)f(a+2d) + O(p).$$

To estimate this inner sum, we note that the contribution of those terms with $d' = 0$ is 0; and, when $d' \neq 0$, we get a contribution of $f(a)f(a+d)f(a+2d)$ to just the inner sum, because there is only one pair $m,t$ which works. Thus, we deduce from this and Lemma 4 that

$$\mathbb{E}(Z_0) = p^{-2}(1 - \varepsilon)^{-3} \sum_{a,a+d,a+2d \in A} f(a)f(a+d)f(a+2d) + O(p)$$

$$= (1 - \varepsilon)^{-3} \Lambda(S)W_0 + O(p)$$

$$< (1 - \varepsilon^2/2)W_0 + O(p).$$
Case 2 (at least one point in $A$, and at least one in $B$).

Define the random variables

$$Z_1 := \sum_{a,a+2d \in A} g(a)g(a+d)g(a+2d)$$
$$Z_2 := \sum_{a,a+2d \in A} g(a)g(a+d)g(a+2d)$$
$$Z_3 := \sum_{a+d,a+2d \in A} g(a)g(a+d)g(a+2d)$$
$$Z_4 := \sum_{a,a+2d \in A} g(a)g(a+d)g(a+2d)$$
$$Z_5 := \sum_{a+d,a+2d \in B} g(a)g(a+d)g(a+2d)$$
$$Z_6 := \sum_{a+2d \in A} g(a)g(a+d)g(a+2d).$$

Also, let $W_1, \ldots, W_6$ be the analogous constants with $g$ replaced by $f$ (note that these are not random variables).

We will now compute the expectations of these random variables; though, we will not do all of these here, and instead will just work it out for $Z_1$, as showing it for all the others can be done in exactly the same way, and leads to the same bounds.

We have that

$$\mathbb{E}(Z_1) = \sum_{a+2d \in B} f(a+2d) \sum_{a,a+2d \in A} \mathbb{E}(g(a)g(a+d)).$$

To evaluate this last expectation, let us suppose that $a+2d \in B$ and $a,a+d \in A$, where $d \neq 0$ (if $d = 0$ then we would have that $a$ lies both in $A$ and $B$, which is impossible). Then, given any pair of distinct elements $x, y \in S$, there exists a unique pair $(m, t) \in \mathbb{F}_p \times \mathbb{F}_p$ such that

$$mx + t = a \quad \text{and} \quad my + t = b.$$

So, the probability that

$$g(a)g(a+d) = (1-\varepsilon)^{-2}f(a)f(a+d),$$

given $a+2d \in B$, $a,a+d \in A$, is $1/p^2$ times the number of ordered pairs $(x, y)$ of distinct elements of $S$, which is $|S|(|S|-1)$. Note that if $g(a)g(a+d)$ is not equal to this, then it must take the value 0. It follows that

$$\mathbb{E}(Z_1) = p^{-2}|S|(|S|-1)(1-\varepsilon)^{-2}W_1 = W_1 + O(p). \quad (7)$$

Likewise for the other $Z_i$, we will have that

$$\mathbb{E}(Z_i) = W_i + O(p).$$
Collecting the two cases together.

Let $Z_7$ denote the contribution of arithmetic progressions lying entirely in $B$; that is,

$$Z_7 = \sum_{b,b+d,b+2d \in B} f(b)f(b+d)f(b+2d) = \sum_{b,b+d,b+2d \in B} g(b)g(b+d)g(b+2d).$$

Note that in this case $W_7 = Z_7$.

Putting together our above estimates, and using the fact that $\Lambda(g) = \frac{p-2}{2} (Z_0 + \cdots + Z_7)$, we find that

$$E(\Lambda(g)) = p^{-2}(W_0 + \cdots + W_7 - \varepsilon^2 W_0/2 + O(p)) = \Lambda(f) - \varepsilon^2 p^{-2} W_0/2 + O(1/p).$$

Using Markov’s inequality we have

$$\text{Prob}(\Lambda(g) < \Lambda(f) - \varepsilon^2 p^{-2} W_0/4) \geq 1 - \frac{E(\Lambda(g))}{\Lambda(f) - \varepsilon^2 p^{-2} W_0/4} > \varepsilon^2/8,$$

since $\Lambda(f) \geq p^{-2} W_0$.

$E(g)$ is close to $E(f)$ with high probability.

Before we “derandomize” and pass to an instantiation of $g$, we will need to also show that $E(g)$ is close to $E(f)$ with high probability. This can be accomplished in several different ways, though here we will just use the second moment method: First, let

$$F := \sum_{a \in A} f(a), \text{ and } G := \sum_{a \in A} g(a).$$

Now, as is easy to show, $F + O(1/p) = E(G)$; and so, since $\varepsilon \beta > p^{-1/2} \log p$, we have that

$$\text{Prob}(|F - G| \geq 2 \beta p) \leq \text{Prob}(|G - E(G)| \geq \beta p). \quad (8)$$

It follows from Chebychev’s inequality that this last probability is at most

$$\frac{\text{Var}(G)}{\beta^2 p^2} = \frac{E(G^2) - E(G)^2}{\beta^2 p^2}.$$

To bound this from above we observe that

$$E(G^2) = \sum_{a,b \in A} E(g(a)g(b)).$$
Now, as a consequence of what we worked out just before (7), we have that $g(a)$ and $g(b)$ are independent whenever $a \neq b$. So,

$$E(G^2) = E(G^2) + O(p),$$

and it follows that the probability of the right-most event in (8) is at most $O(\beta^{-2}/p)$. It is easy to see that with probability $1 - O(\beta^{-2}/p)$ we will have

$$E(g) \geq E(f) - 2\beta. \quad (9)$$

**Conclusion of the proof.**

It follows that with probability at least

$$(1 - O(\beta^{-2}/p)) + \varepsilon^2/8 - 1$$

we will have that

$$E(g) \geq E(f) - 2\beta \quad \text{and} \quad \Lambda(g) \leq \Lambda(f) - \varepsilon^2p^{-2}W_0/4 + O(1/p).$$

Using our assumption that

$$\varepsilon\beta > p^{-1/2}\log p,$$

we have that this probability is positive. So, there exists an instantiation of $g$ such that both hold; henceforth, $g$ will no longer be random, but will instead be one of these instantiations.

By reassigning at most $2\beta p$ places $a \in A$ where $g(a) = 0$ to the value 1, we can guarantee that $E(g) \geq E(f)$, and one easily sees that

$$\Lambda(g) < \Lambda(f) + 2\beta - \varepsilon^2p^{-2}W_0/4 + O(1/p).$$

This completes the proof of our proposition. \[\square\]

### 2.5 Proof of Proposition 2

We have that if we define the new function $h_3(n) = h(n)$ at all $n \in \mathbb{F}_p$, except for $n = x$ and $n = y$, then

$$\Lambda(h_3) = \Lambda(h) + E_1 + \cdots + E_{13},$$
where if we let \( \omega = e^{2\pi i/p} \), then

\[
\begin{align*}
E_1 &= p^{-3} \sum_a \hat{h}(a)^2 (h_3(y) - f(y)) \omega^{-2ay} = p^{-2}(h \ast h)(2y)(h_3(y) - h(y)) \\
E_2 &= p^{-3} \sum_a \hat{h}(a)^2 (h_3(x) - h(x)) \omega^{-2ax} = p^{-2}(h \ast h)(2x)(h_3(x) - h(x)) \\
E_3 &= 2p^{-3} \sum_a \hat{h}(a) \hat{h}(-2a)(h_3(y) - h(y)) \omega^{ay} \\
&= 2p^{-2}(h \ast h)(-y)(h_3(y) - h(y)) \\
E_4 &= 2p^{-3} \sum_a \hat{h}(a) \hat{h}(-2a)(h_3(x) - h(x)) \omega^{ax} \\
&= 2p^{-2}(h \ast h)(-x)(h_3(x) - h(x)) \\
E_5 &= 2p^{-3} \sum_a \hat{h}(a)(h_3(y) - h(y))^2 \omega^{-ay} = 2p^{-2}(h_3(y) - h(y))^2 h(y) \\
E_6 &= 2p^{-3} \sum_a \hat{h}(a)(h_3(y) - h(y))(h_3(x) - h(x)) \omega^{a(y-2x)} \\
&= 2p^{-2}(h_3(y) - h(y))(h_3(x) - h(x)) h(2x - y) \\
E_7 &= 2p^{-3} \sum_a \hat{h}(a)(h_3(y) - h(y))(h_3(x) - f(x)) \omega^{a(x-2y)} \\
&= 2p^{-2}(h_3(y) - h(y))(h_3(x) - f(x)) h(2y - x) \\
E_8 &= 2p^{-3} \sum_a \hat{h}(a)(h_3(x) - h(x))^2 \omega^{-ax} = 2p^{-2}(h_3(x) - h(x))^2 h(x) \\
E_9 &= p^{-3} \sum_a \hat{h}(-2a)(h_3(y) - h(y))^2 \omega^{2ay} = p^{-2}(h_3(y) - h(y))^2 h(y) \\
E_{10} &= p^{-3} \sum_a \hat{h}(-2a)(h_3(x) - h(x))^2 \omega^{2ax} = p^{-2}(h_3(x) - h(x))^2 h(x) \\
E_{11} &= 2p^{-3} \sum_a \hat{h}(-2a)(h_3(x) - h(x))(h_3(y) - h(y)) \omega^{a(x+y)} \\
&= 2p^{-2}(h_3(x) - h(x))(h_3(y) - h(y)) h((x + y)/2) \\
E_{12} &= p^{-2}(h_3(y) - h(y))^3 \\
E_{13} &= p^{-2}(h_3(x) - h(x))^3.
\end{align*}
\]

There are actually 6 more terms that make up the above “error”; however, all of these give a contribution of 0, which is why they were not listed.

So, one sees that

\[
\Lambda(h_3) = \Lambda(h) + p^{-2}(h_3(x) - h(x)) F_h(x) + p^{-2}(h_3(y) - h(y)) F_h(y) \\
+ E_5 + \cdots + E_{13}.
\]  

To prove our proposition, all we need to show is that if there is a pair \( x, y \in \mathbb{F}_p \), \( x, y \not\in A \), with

\[
F_h(x) < F_h(y), \text{ and } h(y) > 0,
\]

then in fact

\[
h(x) = 1.
\]

Suppose there were such a pair \( x, y \) for which \( h(x) < 1 \). Then, we will show that \( h \) fails to minimize \( \Lambda(h) \) subject to the various constraints: Basically, we let

\[
0 < \varepsilon < \min(1 - h(x), h(y))
\]

19
(its exact value will be decided later) and then we consider the function $h_3$ given by

$$h_3(n) = \begin{cases} h(n) & \text{for } n \in \mathbb{F}_p, \ n \neq x, y, \\
h(x) + \varepsilon & h_3(x) = h(x) + \varepsilon, \\
h(y) - \varepsilon & h_3(y) = h(y) - \varepsilon. \\
\end{cases}$$

From our formula (10), we easily deduce that

$$\Lambda(h_3) \leq \Lambda(h) + \varepsilon p^{-2} (F_h(x) - F_h(y)) - O(\varepsilon^2 p^{-2}).$$

Clearly, if we take $\varepsilon > 0$ small enough, we will get

$$\Lambda(h_3) < \Lambda(h),$$

which contradicts the minimality of $h$. We conclude, therefore, that $h(x) = 1$, as claimed.

References

[1] F. A. Behrend, *On the Sets of Integers Which Contain No Three in Arithmetic Progression*, Proc. Nat. Acad. Sci. 23 (331-332), 1946.

[2] J. Bourgain, *Roth’s Theorem on Progressions Revisited*, preprint.

[3] B. Green, *A Szemerédi-Type Regularity Lemma in Abelian Groups*, Geom. and Funct. Anal. 15 (2005), 340-376.

[4] ————, *Roth’s Theorem in the Primes*, Annals of Math. 161 (2005), 1609-1636.

[5] B. Green and O. Sisask, *On the Maximal Number of Three-Term Arithmetic Progressions in Subsets of $\mathbb{Z}/p\mathbb{Z}$*, preprint on the ARXIVES.

[6] B. Green and I. Ruzsa, *Counting Sumsets and Sumfree Sets Modulo a Prime*, Studia Sci. Math. Hungar. 41 (2004), 285-293.

[7] P. Varnavides, *On Certain Sets of Positive Density*, J. London Math. Soc. 34 (1959), 358-360.