The structure of critical sets for $\mathbb{F}_p$ arithmetic progressions

Ernie Croot

October 1, 2009

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

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

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

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

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*

$$\mathbb{E}(f) \;\geq\; \theta \;\in\; (0,1].$$

Then,

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

$$\textstyle\sum_n |f(n) - C(n)| \;\ll\; p(\log p)^{-2/3}.$$

So, f *must be approximately an indicator function.*

- *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*

$$\textstyle\sum_n |(A * B)(n) - |B|C(n)| \;\ll\; p|B|(\log\log p)^{-2/3}.$$

Furthermore, we may take $A = C$ *and take* B *to be a certain "Bohr neighborhood"* $\mathcal{B}$, *which is described in the proof of the theorem.*

1.1 Remarks about the second bullet

What the second bullet is basically saying is the following: from the first bullet we know that $f(n)$ is usually very close to $C(n)$, which is an indicator function. So, $|B|f(n)$ is very close to $|B|C(n)$; and therefore, the conclusion of the second bullet is basically telling us that

$$f(n) \;\sim\; |B|^{-1}(A * B)(n)$$

for "most" elements $n \in \mathbb{F}_p$.

It should be remarked that sumsets are quite special structures, 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 second bullet is saying something fairly non-trivial about our minimal f.

Also, there are loads of other consequences that one can deduce from the second 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

$$r_3([N]) \ \ll \ N(\log N)^{-2/3}, \tag{1}$$

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 $f : \mathbb{F}_p \to [0,1]$ 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 $g : \mathbb{F}_p \to [0,1]$ 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 $\mathcal{A}_N$ 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 $a, a + d, a + 2d, ..., a + kd$ is distinct from $a + kd, a + (k - 1)d, ..., a$.

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 $A \in \mathcal{A}_N$: We have that each three-term progression is contained in the same number of $A \in \mathcal{A}_N$, and each $A \in \mathcal{A}_N$ contains $N^2/2 + O(N)$ three-term progressions; hence, if P denotes the number of $A \in \mathcal{A}_N$ 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

$$p(p-1)P \ = \ |\mathcal{A}_N|(N^2/2 + O(N)),$$

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

$$T_3(S) \ = \ \left(N^2/2 + O(N)\right)^{-1} \sum_{A \in \mathcal{A}_N} T_3(A \cap S). \qquad (2)$$

Next, we need a lower bound on how many $A \in \mathcal{A}_N$ 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 $A \in \mathcal{A}_N$ having common difference d that contain a particular point $a \in \mathbb{F}_p$. So,

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

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

$$(|\mathcal{A}_N| - Y) r_3(N) + YN \ \geq \ (p-1)N|S|,$$

which implies

$$Y \ \geq \ \frac{(p-1)N|S| - |\mathcal{A}_N| r_3(N)}{N - r_3(N)} \ \geq \ (p-1)|S| - |\mathcal{A}_N|(r_3(N)/N).$$

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 (2) 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

$$\Lambda(A) \ \leq \ 8\varepsilon^{-6} p^{-1/2} \log p. \tag{3}$$

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 (1) 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 (3) 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

$$\textstyle\sum_n |f(n) - C(n)| \ \ll \ (\varepsilon + (\log \varepsilon^{12} p)^{-2/3})p, \ \text{for } \varepsilon \ > \ p^{-1/12} \log p.$$

Choosing $\varepsilon = (\log p)^{-2/3}$, we deduce that this sum is $O(p(\log p)^{-2/3})$, just as in Bourgain's theorem (1). This completes the proof of the first part of our theorem.

2.2 Proof of the second 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 second bullet of Theorem 1. To this end, we let

$$f_3(n) \ := \ (f * \mu)(n),$$

where μ 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

$$||b_i n/p|| \ < \ \varepsilon_0, \ \text{for all } i = 1, ..., t.$$

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

$$\textstyle\sum_n |f_3(n) - f(n)| \ \ll \ p(\log \log p)^{-2/3}, \tag{4}$$

for this will imply the second bullet of Theorem 1 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

$$\begin{aligned} \textstyle\sum_n ||\mathcal{B}|^{-1}(C * \mathcal{B})(n) - C(n)| \ &= \ \textstyle\sum_n ||\mathcal{B}|^{-1}(f * \mathcal{B})(n) - f(n)| + O(p(\log p)^{-2/3}) \\ &= \ \textstyle\sum_n |f_3(n) - f(n)| + O(p(\log p)^{-2/3}) \\ &\ll \ p(\log \log p)^{-2/3}, \end{aligned}$$

which is just what the second bullet claims.

Now we show that (4) 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

$$\alpha_i(x) \ := \ (\varepsilon_0 p + 1)^{-1} \left(\sum_{||b_i n/p|| < \varepsilon_0/2} e^{2\pi i n x/p} \right)^2$$

We note that $\alpha_i(x)$ is always a non-negative real for all real numbers x, and α_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,

$$\begin{aligned} |\mathcal{B}| \ \geq \ \hat{\beta}(0) \ &= \ p^{-t+1}(\hat{\beta}_1 * \hat{\beta}_2 * \cdots * \hat{\beta}_t)(0) \\ &= \ p^{-t+1}(\alpha_1 * \alpha_2 * \cdots * \alpha_t)(0) \\ &\geq \ p^{-t+1}\alpha_1(0) \cdots \alpha_t(0) \\ &\geq \ (\varepsilon_0 + O(1/p))^t p. \end{aligned}$$

■

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

$$\begin{aligned} \Lambda(f_3) \ = \ p^{-3} \textstyle\sum_a \hat{f}_3(a)^2 \hat{f}_3(-2a) \ &= \ p^{-3} \textstyle\sum_a \hat{f}(a)^2 \hat{f}(-2a) \ + \ E \\ &= \ \Lambda(f) \ + \ E, \end{aligned}$$

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

$$W_0 \ := \ \textstyle\sum_{a,a+d,a+2d \in A} f_3(a) f_3(a+d) f_3(a+2d) \ \geq \ \varepsilon_1^3 p^2 \Lambda(A).$$

In order to apply Lemma 1 to this, we let

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

so that from (1) 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

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

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

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

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

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

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

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

which completes the proof of Theorem 1.

2.3 Proof of Proposition 1

2.3.1 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

$$\Lambda(S) + \Lambda(T) \ = \ p^{-3} \textstyle\sum_a (\hat{S}(a)^2 \hat{S}(-2a) + \hat{T}(a)^2 \hat{T}(-2a)).$$

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*

$$\Lambda(S) \ \leq \ \alpha^3(1 - (1 - \alpha)^2/2) + O(1/p).$$

Proof of the Lemma. Let $\beta = 1 - \alpha < 1/3$, and then let S just be the arithmetic progression $\{0, 1, ..., \lfloor \alpha p \rfloor - 1\}$, 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

$$\begin{aligned} \Lambda(S) \ &= \ (1 - 3\beta + 3\beta^2) - \beta^2/2 + O(1/p) \\ &= \ 1 - 3\beta + 5\beta^2/2 + O(1/p) \\ &< \ (1 - \beta)^3(1 - \beta^2/2) + O(1/p), \end{aligned}$$

as claimed. ■

2.3.2 Body of the proof of Proposition 1

We will define the function $g : \mathbb{F}_p \to [0,1]$ such that

$$\text{support}(g) \ \subseteq \ A \cup B,$$

where

$$\text{for } n \in B,\ g(n) \ = \ f(n),$$

but on the set A, the funciton 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

$$\text{for } n \in A,\ g(n) \ = \ (1-\varepsilon)^{-1} f(n) T(n).$$

Note that this is ≤ 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

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

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

$$\begin{aligned}
\mathbb{E}(Z_0) \quad &= \quad \textstyle\sum_{\substack{a,a+d,a+2d \in A \\ d \neq 0}} \mathbb{E}(g(a)g(a+d)g(a+2d)) + O(p) \\
&= \quad p^{-2}(1-\varepsilon)^{-3} \textstyle\sum_{\substack{a,a+d,a+2d \in A \\ d \neq 0}} f(a)f(a+d)f(a+2d) \sum_{\substack{m,t \in \mathbb{F}_p \\ a,a+d,a+2d \in m.S+t}} 1 + O(p) \\
&= \quad p^{-2}(1-\varepsilon)^{-3} \textstyle\sum_{\substack{a,a+d,a+2d \in A \\ d \neq 0}} \sum_{b,b+d',b+2d' \in S} \\
&\qquad\qquad \textstyle\sum_{\substack{m,t \in \mathbb{F}_p \\ mb+t=a,\ m(b+d')+t=a+d}} f(a)f(a+d)f(a+2d) + O(p).
\end{aligned}$$

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

$$\begin{aligned} \mathbb{E}(Z_0) &= p^{-2}(1-\varepsilon)^{-3}\sum_{\substack{b,b+d',b+2d'\in S\\ a,a+d,a+2d\in A}} f(a)f(a+d)f(a+2d) \ + \ O(p) \\ &= (1-\varepsilon)^{-3}\Lambda(S)W_0 \ + \ O(p) \\ &< (1-\varepsilon^2/2)W_0 \ + \ O(p). \end{aligned}$$

Case 2 (at least one point in A, and at least one in B).

Define the random variables

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

Also, let $W_1, ..., W_6$ 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

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

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

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

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

$$\mathbb{E}(Z_1) \ = \ p^{-2}|S|(|S|-1)(1-\varepsilon)^{-2}W_1 \ = \ W_1 + O(p). \tag{5}$$

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,

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

Note that in this case $W_7 = Z_7$.

Putting together our above estimates, and using the fact that

$$\Lambda(g) \ = \ p^{-2}(Z_0 + \cdots + Z_7),$$

we find that

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

Using Markov's inequality we have

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

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

$\mathbb{E}(g)$ is close to $\mathbb{E}(f)$ with high probability.

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

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

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

$$\mathrm{Prob}(|F-G| \geq 2\beta p) \ \leq \ \mathrm{Prob}(|G - \mathbb{E}(G)| \geq \beta p). \tag{6}$$

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

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

To bound this from above we observe that

$$\mathbb{E}(G^2) \ = \ \textstyle\sum_{a,b\in A}\mathbb{E}(g(a)g(b)).$$

Now, as a consequence of what we worked out just before (5), we have that $g(a)$ and $g(b)$ are independent whenever $a \neq b$. So,

$$\mathbb{E}(G^2) \ = \ \mathbb{E}(G^2) + O(p),$$

and it follows that the probability of the right-most event in (6) is at most $O(\beta^{-2}/p)$. It is easy to see that with probability $1 - O(\beta^{-2}/p)$ we will have

$$\mathbb{E}(g) \ \geq \ \mathbb{E}(f) - 2\beta. \tag{7}$$

Conclusion of the proof.

It follows that with probability at least

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

we will have that

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

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 $\mathbb{E}(g) \geq \mathbb{E}(f)$, and one easily sees that

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

This completes the proof of our proposition. ∎

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.