POLYNOMIAL CONFIGURATIONS IN DIFFERENCE SETS (REVISED VERSION)

NEIL LYALL ÁKOS MAGYAR

ABSTRACT. We prove a quantitative version of the Polynomial Szemerédi Theorem for *difference sets*. This result is achieved by first establishing a higher dimensional analogue of a theorem of Sárközy (the simplest non-trivial case of the Polynomial Szemerédi Theorem asserting that the difference set of any subset of the integers of positive upper density necessarily contains a perfect square) and then applying a simple lifting argument.

An earlier version of this article, in which the main results were restricted to the case of linearly independent polynomials, has already appeared in [8]. *The main substantive changes in this revision occur in the statements of Theorem 1 and Corollary 2 which are obtained through a simple modification of the original argument in Section 2.1.*

1. INTRODUCTION

1.1. **Background.** A striking and elegant result in density Ramsey theory states that in any subset of the integers of positive upper density there necessarily exist two distinct elements whose difference is a perfect square.

This result was originally conjectured by L. Lovász and eventually verified independently by Furstenberg [3], using techniques from ergodic theory, and Sárközy [13], using an approach similar in spirit to Roth's Fourier analytic (circle method inspired) proof of Szemerédi's theorem for arithmetic progressions of length three.

Sárközy actually obtained the following stronger quantitative result.

Theorem A (Sárközy [13])**.** *If $A \subseteq [1,N]$ and $d^2 \notin A-A$ for any $d \neq 0$, then there exists an absolute constant $C>0$ such that*

$$\frac{|A|}{N} \leq C\left(\frac{(\log\log N)^2}{\log N}\right)^{1/3}.$$

Notation: In the theorem above and in the sequel we will use N (and later also M) to denote an arbitrary positive integer, $[1,N]$ to denote $\{1,\dots,N\}$ as is customary, and $A \pm A$ to denote the usual difference and sum sets of A, namely $A \pm A = \{a \pm a' \,|\, a,a' \in A\}$.

The current best known quantitative bound of $(\log N)^{-c\log\log\log\log N}$ in Theorem A is due to Pintz, Steiger and Szemerédi [10]. These methods were later extended by Balog, Pelikán, Pintz and Szemerédi [1] to obtain the same bounds, with the implicit constant now depending on k, for sets with no kth power differences.

We note that it is conjectured that for any $\epsilon > 0$ and $N \geq N_0(\epsilon)$ sufficiently large there exists a set $A \subseteq [1,N]$ with $|A| \geq N^{1-\epsilon}$ that contains no square differences, see for example [5]. Ruzsa [11] has demonstrated that this is at least true for $\epsilon = 0.267$.

Bergelson and Leibman (extending on the ideas of Furstenberg) established a far reaching qualitative generalization of Sárközy's theorem, the so-called Polynomial Szemerédi Theorem.

Theorem B (Bergelson and Leibman [2])**.** *If A is a subset of the integers of positive upper density and $P_1(d),\dots,P_\ell(d)$ are polynomials in $\mathbb{Z}[d]$ with $P_i(0)=0$ for $i=1,\dots,\ell$, then there exists $a \in A$ and $d \neq 0$ such that*

$$a + \{P_1(d),\dots,P_\ell(d)\} \subseteq A.$$

Key words and phrases. Difference sets, Sárközy's theorem, polynomial configurations.
Both authors were partially supported by NSF grants.

We note that no quantitative version of this multiple recurrence result is known beyond the linear case of Szemerédi's theorem, see Gowers [4], and the general single recurrence case of Sárközy's theorem (the case $\ell = 2$ above).

The purpose of this paper is to establish a quantitative result on the existence of polynomial configurations of the type in Theorem B in the *difference set* of sparse subsets of the integers.

1.2. **Statement of Main Results.** We now fix a family of polynomials

$$P_1(d), \dots, P_\ell(d) \in \mathbb{Z}[d]$$

with $P_i(0) = 0$ for $i = 1, \dots, \ell$ and set $k = \max_i \deg P_i$. We will assume throughout this paper that $k \geq 2$.

Theorem 1. *If $A \subseteq [1, N]$ and $\{P_1(d), \dots, P_\ell(d)\} \nsubseteq A - A$ for any $d \neq 0$, then we necessarily have*

$$\frac{|A|}{N} \leq C \left(\frac{\log\log N}{\log N} \right)^{1/\ell(k-1)}$$

for some absolute constant $C = C(P_1, \dots, P_\ell)$.

In the case of a single polynomial ($\ell = 1$), this result has also recently been obtained by Lucier [7] and, to the best of our knowledge, constitutes the best bounds that are currently known for arbitrary polynomials with integer coefficients and zero constant term.

We remark that one can immediately deduce, from Theorem 1, the following result on the existence of polynomial configurations in (a shift of) the sumset $A + B$ of two given sets $A, B \subseteq [1, N]$.

Corollary 2. *If $A, B \subseteq [1, N]$ and $m + \{P_1(d), \dots, P_\ell(d)\} \nsubseteq A + B$ for any $d \neq 0$ and $m \in [2, 2N]$, then we necessarily have*

$$\frac{|A||B|}{N^2} \leq C' \left(\frac{\log\log N}{\log N} \right)^{1/\ell(k-1)}$$

for some absolute constant $C' = C'(P_1, \dots, P_\ell)$.

Proof. Since

$$\sum_{2 \leq m \leq 2N} |B \cap (m - A)| = |A||B|$$

it follows that there exists $m \in [2, 2N]$ such that if we set $D = B \cap (m - A)$, then

$$|D| \geq \frac{|A||B|}{2N - 1}.$$

The result then follows from Theorem 1 since

$$D - D \subseteq A + B - m. \qquad \square$$

The strategy we will employ to prove Theorem 1 is to *lift* the problem to $\mathbb{Z}^k$ in such a way that we may then apply the following higher dimensional analogue of Sárközy's theorem.

Theorem 3. *If $B \subseteq [1, N]^k$ and $(d, d^2, \dots, d^k) \notin B - B$ for any $d \neq 0$ then we necessarily have*

$$\frac{|B|}{N^k} \leq C \left(\frac{\log\log N}{\log N} \right)^{1/(k-1)}$$

for some absolute constant $C = C(k)$.

Since Theorem 3 is concerned with the intersection of a difference set with the *monomial* curve $(d, d^2, \dots, d^k)$ we speculate that the methodology of Balog et al. [1] may be applied in this higher dimensional situation to obtain far superior bounds in Theorem 3 and hence also in Theorem 1.

Further notational convention: Throughout this paper the letters c and C will denote absolute constants that will generally satisfy $0 < c \ll 1 \ll C$, whose values may change from line to line and even from step to step, and will unless otherwise specified depend only on the dimension k.

2. Reduction to the key dichotomy proposition

We first present the *lifting* argument that allows us to deduce Theorem 1 from Theorem 3.

2.1. **Proof that Theorem 3 implies Theorem 1.** Let $P_i(d) = c_{i1}d + \cdots + c_{ik}d^k$ for $1 \leq i \leq \ell$.

Suppose that the coefficient matrix $\mathcal{P} = \{c_{ij}\}$ has rank r with $1 \leq r \leq \ell$. Without loss in generality we will make the additional assumption that it is in fact the first r polynomials $P_1, \ldots, P_r$ that are linearly independent and use $\mathcal{R}$ to denote the $r \times k$ matrix corresponding to the first r rows of $\mathcal{P}$.

As a consequence of this assumption it follows that the remaining polynomials, P_{r+i} with $1 \leq i \leq \ell - r$, can be expressed as

$$P_{r+i} = d_{i1}P_1 + \cdots + d_{ir}P_r$$

where $\mathcal{D} = \{d_{ij}\}$ is some $(\ell - r) \times r$ matrix with rational coefficients.

Note that

$$\begin{aligned} \mathcal{P} &: \mathbb{Z}^k \to \mathbb{Z}^\ell \\ \mathcal{R} &: \mathbb{Z}^k \to \mathbb{Z}^r \\ \mathcal{D} &: \mathcal{R}(\mathbb{Z}^k) \to \mathbb{Z}^{\ell - r} \end{aligned}$$

and

$$\mathcal{P}(b) = \begin{pmatrix} \mathcal{R}(b) \\ \mathcal{D}(\mathcal{R}(b)) \end{pmatrix}.$$

Let $A^\ell = A \times \cdots \times A \subseteq [1, N]^\ell$ and $\delta = |A|/N$.

The full rank assumption on the matrix $\mathcal{R}$ ensures that there exists an absolute constant c, depending only on the coefficients of the matrix $\mathcal{R}$, such that

$$\left|\mathcal{R}(\mathbb{Z}^k) \cap (A^r - s)\right| \geq c\delta^r N^r$$

for some $s \in [1, c^{-1}]^r$. Thus, if we choose N' to be a large enough multiple of N (again depending only the coefficients of the matrix $\mathcal{R}$) and let

$$B' = \left\{b \in [-N', N']^k \,:\, \mathcal{R}(b) \in A^r - s\right\},$$

it follows that

$$|B'| \geq c\,\delta^r N^k.$$

Since

$$\sum_{t \in \mathbb{Z}^{\ell - r}} \sum_{b \in B'} 1_{A^{\ell - r}}(\mathcal{D}(\mathcal{R}(b)) + t) = |A|^{\ell - r}|B'|$$

it follows that there exists $c = c(\mathcal{P})$ and $t \in \mathbb{Z}^{\ell - r}$ such that

$$\left|\left\{b \in B' \,:\, \mathcal{D}(\mathcal{R}(b)) \in A^{\ell - r} - t\right\}\right| \geq c\delta^{\ell - r}|B'|.$$

Hence, if we let

$$B = \left\{b \in [-N', N']^k \,:\, \mathcal{P}(b) \in A^\ell - m\right\},$$

where $m = (s, t) \in \mathbb{Z}^\ell$, it follows that

$$|B| \geq c\,\delta^\ell N^k.$$

The result now follows from Theorem 3 since if there were to exist a $d \neq 0$ such that

$$(d, d^2, \ldots, d^k) \in B - B$$

this would immediately implies that

$$(P_1(d), \ldots, P_\ell(d)) \in A^\ell - A^\ell,$$

since $\mathcal{P}(B) \subseteq A^\ell - m$. □

Matters therefore reduce to proving Theorem 3.

2.2. **Dichotomy between randomness and arithmetic structure.** Our approach will be to establish a dichotomy between randomness and structure of the following form.

Let us fix the notation $Q_M = [1, M] \times \cdots \times [1, M^k]$ and $\varepsilon = (10k)^{-1}$.

Proposition 4. *Let $B \subseteq Q_M$, $\delta = |B|/|Q_M|$, and $\sigma = c_k \delta^{k-1}$. If $M \geq \delta^{-C}$, with $C > 0$ sufficiently large (depending only on k), then either B behaves as though it were a random set in the sense that*

$$\sum_{d=1}^{M} \left| B \cap \left(B + (d, d^2, \ldots, d^k) \right) \right| \geq \frac{\varepsilon}{4} \delta |B| M \tag{1}$$

or B has arithmetic structure in the sense that there exists a grid $\Lambda \subseteq Q_M$ of the form

$$\Lambda = \{ m + (\ell_1 q, \ldots, \ell_k q^k) \,|\, (\ell_1, \ldots, \ell_k) \in Q_L \} \tag{2}$$

with $L \geq \delta^{k+2} \sigma M$ such that

$$|B \cap \Lambda| > \delta(1 + \sigma)|\Lambda|.$$

In contrast with the standard L^∞ increment strategy of Roth, we will obtain the dichotomy in Proposition 4 by exploiting the concentration of the L^2 mass of the Fourier transform. Similar arguments of this type can be found in Heath-Brown [6] and Szemerédi [15], see also Ruzsa and Sanders [12]. The proof of Proposition 4 will be presented in Sections 3 and 4.

2.3. **Proof that Proposition 4 implies Theorem 3.** It is easy to see, by partitioning $[1, N]^k$ into boxes of size $M \times M^2 \times \cdots \times M^k$ with M essentially equal to $N^{1/k}$, that we may, with no loss in generality, assume that $B \subseteq Q_M$ with $\delta = |B|/|Q_M| \geq |B|/N^k$.

If $(d, d^2, \ldots, d^k) \notin B - B$ for any $d \neq 0$ (as is the assumption in Theorem 3), then Proposition 4 allows us to perform an iteration. At the nth step of this iteration we will have a set $B_n \subseteq Q_{M_n}$ of size $\delta_n |Q_{M_n}|$, this set will be an appropriately rescaled version of a subset of B itself and hence will also contain no non-trivial differences of the form $(d, d^2, \ldots, d^k)$.

Let $B_0 = B$, $M_0 = M$ and $\delta_0 = \delta$. Proposition 4 ensures that either

$$M_n \leq \delta_n^{-C} \tag{3}$$

or else the iteration proceeds allowing us to choose M_{n+1}, δ_{n+1} and B_{n+1} such that

$$M_{n+1} \geq c \delta_n^{(2k+1)} M_n$$

and

$$\delta_{n+1} \geq \delta_n + c\delta_n^k.$$

Now as long as the iteration continues we must have $\delta_n \leq 1$, and so after $O(\delta^{1-k})$ iterations condition (3) must be satisfied, giving

$$(\delta^{-(2k+1)})^{-C\delta^{1-k}} M \leq \delta^{-C}.$$

From this it follows that

$$\log M \leq C \delta^{-(k-1)} \log \delta^{-1}$$

and consequently (after a short calculation that we leave to the reader) that

$$\delta \leq C \left(\frac{\log \log M}{\log M} \right)^{1/(k-1)}.$$

This establishes Theorem 3. □

The rest of this article is devoted to the proof of Proposition 4.

3. Setting the stage for the proof of Proposition 4

We suppose that $B \subseteq Q_M$, $\delta = |B|/|Q_M|$, and $M \geq \delta^{-C}$. Our approach will be to assume that B exhibits neither of the two properties described in Proposition 4 and then seek a contradiction.

3.1. **A simple consequence of B being non-random.** If we were to suppose that B is non-random, in the sense that inequality (1) does not hold, then it would immediately follows that

$$\sum_{m,n\in\mathbb{Z}^k} 1_B(m)1_B(n)1_S(m-n) \leq \frac{1}{4}\,\delta|B||S| \tag{4}$$

where

$$S = \{(d, d^2, \ldots, d^k) \,:\, 1 \leq d \leq \varepsilon M\}.$$

3.2. **A simple consequences of B being non-structured.** If we were to assume that B is *regular*, in the sense that B in fact satisfies the inequality

$$|B \cap \Lambda| \leq \delta(1+\sigma)|\Lambda|$$

for all arithmetic grids $\Lambda \subseteq Q_M$ of the form (2) with $L \geq \delta^{k+2}\sigma M$, then the set

$$B' = B \cap \big((\varepsilon M, (1-\varepsilon)M] \times \cdots \times (\varepsilon M^k, (1-\varepsilon)M^k]\big)$$

must contain most of the elements of B. In particular we must have

$$|B'| \geq (3/4)|B| \tag{5}$$

since if this were not the case we would immediately obtain a grid $\Lambda \subseteq Q_M$ of the form (2) with $q = 1$ and $L \geq \varepsilon M$ such that

$$|B \cap \Lambda| \geq \delta(1 + 1/4)|\Lambda|.$$

3.3. **The balance function.** We define the *balance function of B* to be

$$f_B = 1_B - \delta 1_{Q_M},$$

and note that f_B has mean value zero, that is $\sum f_B(m) = 0$. This property of the balance function f_B will be critically important in our later arguments.

It easy to verify that if B satisfies inequalities (4) and (5), then

$$\sum_{m,n\in\mathbb{Z}^k} f_B(m)f_B(n)1_S(m-n) \leq -\frac{1}{4}\,\delta|B||S|. \tag{6}$$

One can see this by simply expanding the sum into a sum of four sums, one involving only the function 1_B on which we can apply (4), two involving the functions 1_B and $-\delta 1_{Q_M}$ on which we can apply (5), and one involving only the function $-\delta 1_{Q_M}$ which can be estimated trivially.

3.4. **Fourier analysis on $\mathbb{Z}^k$.** For $f : \mathbb{Z}^k \to \mathbb{C}$ with finite support we define the *Fourier transform of f* to be

$$\widehat{f}(\alpha) = \sum_{m\in\mathbb{Z}^k} f(m)e^{-2\pi i m\cdot\alpha}.$$

The finite support assumption on f ensures that $\widehat{f}$ is a continuous function on $\mathbb{T}^k$ and that orthogonality immediately gives both the Fourier inversion formula and Plancherel's identity, namely

$$f(m) = \int_{\mathbb{T}^k} \widehat{f}(\alpha)e^{2\pi i m\cdot\alpha}d\alpha \qquad \text{and} \qquad \int_{\mathbb{T}^k} |\widehat{f}(\alpha)|^2 d\alpha = \sum_{m\in\mathbb{Z}^k} |f(m)|^2.$$

It is then easy to verify that from inequality (6) we immediately obtain the estimate

$$\int_{\mathbb{T}^k} |\widehat{f_B}(\alpha)|^2 |\widehat{1_S}(\alpha)|\, d\alpha \geq \frac{1}{4}\,\delta|B||S| \tag{7}$$

where we recognize

$$\widehat{1_S}(\alpha) = \sum_{d=1}^{\varepsilon M} e^{-2\pi i(\alpha_1 d + \alpha_2 d^2 + \cdots + \alpha_k d^k)}, \tag{8}$$

as a classical Weyl sum.

3.5. **Estimates for Weyl sums.** Since $\varepsilon = (10k)^{-1}$ is fixed it is clear that whenever $|\alpha_j| \ll M^{-j}$ there can be no cancellation in the Weyl sum (8), in fact the same is also true when each α_j is close to a rational with *small* denominator (in other words there is no cancellation over sums in residue classes modulo q).

We now state a precise formulation of the well-known fact that this is indeed the only obstruction to cancellation. Let $\eta > 0$. We define

$$\mathbf{M}_q = \mathbf{M}_q(\eta) = \left\{ \alpha \in \mathbb{T}^k \,:\, \left|\alpha_j - \frac{a_j}{q}\right| \le \frac{1}{\eta^k M^j} \ (1 \le j \le k) \text{ for some } a \in [1,q]^k \right\}. \tag{9}$$

Lemma 5. *Let $\eta > 0$ and $M \ge \eta^{-C}$ (with C sufficiently large depending on k).*

(i) (Minor box estimate) *If $\alpha \notin \mathbf{M}_q$ for any $1 \le q \le \eta^{-k}$, then*
$$|\widehat{1_S}(\alpha)| \le C\eta|S|.$$
(ii) (Major box estimate) *If $\alpha \in \mathbf{M}_q$ for some $1 \le q \le \eta^{-k}$, then*
$$|\widehat{1_S}(\alpha)| \le Cq^{-1/k}|S|.$$

The proof of this result is a straightforward (and presumably well-known) consequence of the standard estimates for Weyl sums, for the sake of completeness we include these arguments in Appendix A.

4. The proof of Proposition 4

In the previous section we established that inequalities (4) and (5) would be immediate consequences of B not exhibiting either of the two properties described in Proposition 4. We now present the two lemmas from which we will obtain our desired contradiction.

In both lemmas below we set $\eta = \delta/8C$, where C is the large constant in Lemma 5.

Lemma 6. *Let $\eta = \delta/8C$. If B is neither random nor structured, in the sense outlined in Proposition 4, then there exists $1 \le q \le \eta^{-k}$ such that*

$$\frac{1}{\delta|B|}\int_{\mathbf{M}_q} |\widehat{f_B}(\alpha)|^2 d\alpha \ge c\delta^{k-1}. \tag{10}$$

The second lemma is a precise quantitative formulation, in our setting, of the standard L^2 density increment lemma.

Lemma 7. *Let $\eta = \delta/8C$ and $\sigma \le \eta^{k-2}/8\pi$. If B is* regular*, in the sense that*
$$|B \cap \Lambda| \le \delta(1+\sigma)|\Lambda|$$
for all arithmetic grids $\Lambda \subseteq Q_M$ of the form (2) with $qL \ge \eta^2\sigma M$, then

$$\frac{1}{\delta|B|}\int_{\mathbf{M}_q} |\widehat{f_B}(\alpha)|^2 d\alpha \le 12\sigma. \tag{11}$$

We therefore obtain a contradiction if $\sigma \le c\delta^{k-1}$, proving Proposition 4.

4.1. **Proof of Lemma 6.** It follows from the minor box estimate of Lemma 5 and Plancherel's identity that

$$\int_{\text{minor boxes}} |\widehat{f_B}(\alpha)|^2 |\widehat{1_S}(\alpha)|\, d\alpha \le C\eta|S||B|.$$

Therefore, if $\eta = \delta/8C$, it follows from estimate (7) and the major box estimate of Lemma 5 that

$$\sum_{q=1}^{\eta^{-k}} q^{-1/k} \int_{\mathbf{M}_q} |\widehat{f_B}(\alpha)|^2 d\alpha \ge \eta|B|.$$

It therefore follows that

$$\max_{1\le q\le \eta^{-k}} \int_{\mathbf{M}_q} |\widehat{f_B}(\alpha)|^2 d\alpha \ge \eta^k|B|$$

as required. □

4.2. **Proof of Lemma 7.** We fix q and L so that $qL = \eta^2\sigma M$ and define

$$\Lambda = \{-(\ell_1 q, \ell_2 q^2, \dots, \ell_k q^k) \,|\, 1 \le \ell_j \le L^j\}.$$

Claim 1. *If $\alpha \in \mathbf{M}_q$, then $|\widehat{1_\Lambda}(\alpha)| \ge |\Lambda|/2$.*

Proof of Claim 1. Since

$$\sum_{j=1}^{k} L^j \|q^j\alpha_j\| \le \sum_{j=1}^{k} (Lq)^j \eta^{-k} M^{-j} = \eta^{-k} \sum_{j=1}^{k} (\eta^2\sigma)^j \le 2\eta^{-(k-2)}\sigma,$$

for all $\alpha \in \mathbf{M}_q$, where $\|\cdot\|$ denotes the distance to the nearest integer, it follows that

$$|\widehat{1_\Lambda}(\alpha)| \ge |\Lambda| - \sum_{\ell_j=1}^{L^j} \left|e^{2\pi i(\ell_1 q\alpha_1 + \dots + \ell_k q^k \alpha_k)} - 1\right| \ge |\Lambda|\Big(1 - 2\pi \sum_{j=1}^{k} L^j \|q^j\alpha_j\|\Big) \ge |\Lambda|/2,$$

for all $\alpha \in \mathbf{M}_q$, provided $\sigma \le \eta^{k-2}/8\pi$. □

Plancherel's identity (applied to the function $f_B * 1_\Lambda$) and Claim 1 imply that

$$\frac{1}{\delta|B|} \int_{\mathbf{M}_q} |\widehat{f_B}(\alpha)|^2 d\alpha \le \frac{4}{\delta|B||\Lambda|^2} \sum_{m\in\mathbb{Z}^k} |f_B * 1_\Lambda(m)|^2. \tag{12}$$

The conclusion of Lemma 7 will therefore be an immediate consequence of the following.

Claim 2. *As a consequence of the assumptions in Lemma 7 if follows that*

$$\sum_{m\in\mathbb{Z}^k} |f_B * 1_\Lambda(m)|^2 \le 3\sigma\,\delta|B||\Lambda|^2.$$

Proof of Claim 2. We let

$$\mathcal{M} = \{m \in \mathbb{Z}^k \,|\, m - \Lambda \subseteq Q_M\}$$
$$\mathcal{E} = (Q_M + \Lambda) \setminus \mathcal{M}$$

and write

$$\sum_{m\in\mathbb{Z}^k} |f_B * 1_\Lambda(m)|^2 = \sum_{m\in\mathcal{M}} |f_B * 1_\Lambda(m)|^2 + \sum_{m\in\mathcal{E}} |f_B * 1_\Lambda(m)|^2.$$

We note that since

$$f_B * 1_\Lambda(m) = |B \cap (m - \Lambda)| - \delta|Q_M \cap (m - \Lambda)|$$

it follows from our *regularity* assumption on B that if $m \in \mathcal{M}$, then

$$-\delta|\Lambda| \le f_B * 1_\Lambda(m) \le \delta\sigma|\Lambda|,$$

while for $m \in \mathcal{E}$ we can only conclude that

$$|f_B * 1_\Lambda(m)| \le |\Lambda|.$$

Now since f_B has mean value zero the convolution

$$f_B * 1_\Lambda(m) = \sum_n f_B(n) 1_\Lambda(m - n)$$

also has mean value zero. Thus, using the fact that $|g| = 2g_+ - g$, where $g_+ = \max\{g, 0\}$ denotes the *positive-part* function, and the trivial size estimate $|\mathcal{M}| \le |Q_M|$, we can deduce that

$$\begin{aligned}\sum_{m\in\mathcal{M}} |f_B * 1_\Lambda(m)|^2 &\le 2\left(\sup_{m\in\mathcal{M}} |f_B * 1_\Lambda(m)|\right) \sum_{m\in\mathcal{M}} (f_B * 1_\Lambda)_+(m) \\ &\le 2(\delta|\Lambda|)(\delta\sigma|\Lambda|)|\mathcal{M}| \\ &\le 2\delta^2\sigma|\Lambda|^2|Q_M|.\end{aligned}$$

We leave it to the reader to verify that

$$|\mathcal{E}| \le \left((1 + 2\eta^2\sigma)^k - (1 - 2\eta^2\sigma)^k\right)|Q_M| \le 8k\eta^2\sigma|Q_M|,$$

and hence

$$\sum_{m\in\mathcal{E}}|f_B*1_\Lambda(m)|^2\le|\Lambda|^2|\mathcal{E}|\ll\frac{1}{2}\sum_{m\in\mathcal{M}}|f_B*1_\Lambda(m)|^2,$$

provided $8k\eta^2\ll\delta^2$.

This concludes the proof of Claim 2 and establishes Lemma 7. □

Appendix A. Weyl sum estimates

A.1. **Standard major and minor arc estimates.** Let $P(\alpha,d)=\alpha_1 d+\cdots+\alpha_k d^k$.

Lemma 8 (Weyl inequality)**.** *If $|\alpha_k-a_k/q|\le q^{-2}$ and $(a,q)=1$, then*

$$\left|\sum_{d=1}^{N}e^{2\pi iP(\alpha,d)}\right|\le C_{k,\epsilon}N^{1+\epsilon}\left(\frac{1}{q}+\frac{1}{N}+\frac{q}{N^k}\right)^{1/2^{k-1}}.$$

This result is completely standard, see for example [9]. We now fix a sufficiently small $\mu=\mu(k)>0$ and define

$$\mathbf{M}'_{a/q}=\left\{\alpha\in\mathbb{T}^k\,:\,\left|\alpha_j-\frac{a_j}{q}\right|\le\frac{1}{N^{j-\mu}}\ (1\le j\le k)\right\}. \tag{13}$$

Successive applications of Dirichlet's principle and the Weyl inequality, starting with the highest power k, gives the following qualitative estimate (a quantitative version of which can be found in Vinogradov [17]).

Proposition 9 (Minor arc estimate I)**.** *If $\alpha\notin\mathbf{M}'_{a/q}$ for any $(a,q)=1$ with $1\le q\le N^\mu$, then*

$$\left|\sum_{d=1}^{N}e^{2\pi iP(\alpha,d)}\right|\le CN^{1-\nu}. \tag{14}$$

for some $\nu=\nu(k,\mu)>0$.

Proposition 10 (Major arc estimate)**.** *If $\alpha\in\mathbf{M}'_{a/q}$ for some $(a,q)=1$ with $1\le q\le N^\mu$, then*

$$\left|\sum_{d=1}^{N}e^{2\pi iP(\alpha,d)}\right|\le CNq^{-1/k}\Big(1+\sum_{j=1}^{k}N^j|\alpha_j-a_j/q|\Big)^{-1/k}+O(N^{1/2}). \tag{15}$$

Proof. It is straightforward to write

$$\sum_{d=1}^{N}e^{2\pi iP(\alpha,d)}=q^{-1}S(a,q)v_N(\alpha-a/q)+O(N^{1/2})$$

where

$$S(a,q):=\sum_{r=0}^{q-1}e^{2\pi iP(a,r)/q}\quad\text{and}\quad v_N(\beta):=\int_0^N e^{2\pi iP(\beta,x)}dx.$$

The result then follows from the observation that

$$|S(a,q)|\le Cq^{1-1/k} \tag{16}$$

whenever $(a,q)=1$, which is a result of Hua (see for example [16]), and

$$|v_N(\beta)|\le CN\Big(1+\sum_{j=1}^{k}N^j|\beta_j|\Big)^{-1/k} \tag{17}$$

which follows from van der Corput's lemma for oscillatory integrals (see for example [14]) and rescaling. □

A.2. **Refinement of the major arcs.** Let $0 < \eta \leq 1$ and

$$\mathbf{M}_{a/q} = \mathbf{M}_{a/q}(\eta) = \left\{ \alpha \in \mathbb{T}^k \, : \, \left| \alpha_j - \frac{a_j}{q} \right| \leq \frac{1}{\eta^k N^j} \ (1 \leq j \leq k) \right\}. \tag{18}$$

Combining Propositions 9 and 10 we easily obtain the following result from which Lemma 5 is an immediate consequence.

Proposition 11 (Minor arc estimate II)**.** *If $\alpha \notin \mathbf{M}_{a/q}$ for any $(a,q) = 1$ with $1 \leq q \leq \eta^{-k}$, then*

$$\left| \sum_{d=1}^{N} e^{2\pi i P(\alpha,d)} \right| \leq C\eta N + O(N^{1-\nu}).$$

Proof. It follows from Proposition 10 that on $\mathbf{M}'_{a/q}$ we have

$$\left| \sum_{d=1}^{N} e^{2\pi i P(\alpha,d)} \right| \leq C\eta N$$

provided $(a,q) = 1$ and either

$$\eta^{-k} \leq q \leq N^{\mu}$$

or there exists j such that

$$\eta^{-k} N^{-j} \leq |\alpha_j - a_j/q| \leq N^{-j+\mu}.$$

□

References

[1] A. Balog, J. Pelikán, J. Pintz, E. Szemerédi, *Difference sets without κ-th powers*, Acta Math. Hungar. 65 (1994), 165-187.

[2] V. Bergelson and A. Leibman, *Polynomial extensions of van der Waerden's and Szemerédi's theorems*, J. Amer. Math. Soc., 9, No. 2 (1996), 725-753.

[3] H. Furstenberg, *Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions*, J. d'Analyse Math, 71 (1977), 204-256.

[4] W. T. Gowers, *A new proof of Szemerédi's theorem*, GAFA, 11 (2001), 465-588.

[5] B. Green, *On arithmetic structures in dense sets of integers*, Duke Math. Jour., 114, (2002) (2), 215-238.

[6] D. R. Heath-Brown, *Integer sets containing no arithmetic progressions*, J. London Math. Soc. (2) 35(3) (1987), 385-394.

[7] J. Lucier, *Intersective sets given by a polynomial*, Acta Arith. 123 (2006), no. 1, 57-95.

[8] N. Lyall and Á. Magyar, *Polynomial configurations in difference sets*, J. Num. Theory, v. 129/2, pp. 439-450, 2009.

[9] H. L. Montgomery, *Ten Lectures on the Interface Between Analytic Number Theory and Harmonic Analysis*, CBMS Regional Conference Series in Mathematics, 84.

[10] J. Pintz, W. L. Steiger, E. Szemerédi, *On sets of natural numbers whose difference set contains no squares*, J. London Math. Soc. 37 (1988), 219-231.

[11] I. Z. Ruzsa, *Difference sets without squares*, Period. Math. Hungar. 15 (1984), 205-209.

[12] I. Z. Ruzsa and T. Sanders, *Difference sets and the primes*, preprint.

[13] A. Sárközy, *On difference sets of sequences of integers III*, Acta Math. Acad. Sci. Hungar., 31 (1978), 355-386.

[14] E. M. Stein, *Harmonic Analysis: Real–Variable Methods, Orthogonality, and Oscillatory Integrals*, Princeton Univ. Press, Princeton, 1993.

[15] E. Szemerédi, *Integer sets containing no arithmetic progressions*, Acta Math. Hungar., 56(1-2) (1990), 155-158.

[16] R. C. Vaughan, *The Hardy–Littlewood Method*, 2nd ed., Cambridge Univ. Press, Cambridge, 1997.

[17] I. M. Vinogradov, *The Method of Trigonometrical Sums in the Theory of Numbers*, Interscience, New York, 1954.

Department of Mathematics, The University of Georgia, Boyd Graduate Studies Research Center, Athens, GA 30602, USA

E-mail address: lyall@math.uga.edu

Department of Mathematics, The University of Georgia, Boyd Graduate Studies Research Center, Athens, GA 30602, USA

E-mail address: magyar@math.uga.edu