SIMULTANEOUS POLYNOMIAL RECURRENCE

NEIL LYALL   ÁKOS MAGYAR

ABSTRACT. Let $A \subseteq \{1, \ldots, N\}$ and $P_1, \ldots, P_\ell \in \mathbb{Z}[n]$ with $P_i(0) = 0$ and $\deg P_i = k$ for every $1 \leq i \leq \ell$. We show, using Fourier analytic techniques, that for every $\varepsilon > 0$, there necessarily exists $n \in \mathbb{N}$ such that

$$\frac{|A \cap (A + P_i(n))|}{N} > \left(\frac{|A|}{N}\right)^2 - \varepsilon$$

holds simultaneously for $1 \leq i \leq \ell$ (in other words all of the polynomial shifts of the set A intersect A "ε-optimally"), as long as $N \geq N_1(\varepsilon, P_1, \ldots, P_\ell)$. The quantitative bounds obtained for N_1 are explicit but poor; we establish that N_1 may be taken to be a constant (depending only on $P_1, \ldots, P_\ell$) times a tower of 2's of height $C^*_{k,\ell} + C\varepsilon^{-2}$.

1. INTRODUCTION.

1.1. Background.

The study of recurrence properties of dynamical systems goes back to the beginnings of ergodic theory. If A is a measurable subset of a probability space $(X, \mathcal{M}, \mu)$ with $\mu(A) > 0$ and T is a measure preserving transformation, then it was already shown by Poincaré [23] that $\mu(A \cap T^{-n} A) > 0$ for some natural number n (and hence for infinitely many). This result was subsequently sharpened by Khintchine [15], who observed that for every $\varepsilon > 0$ there in fact exist $n \in \mathbb{N}$ such that $\mu(A \cap T^{-n} A) > \mu(A)^2 - \varepsilon$. Note that in general this lower bound is sharp, since $\mu(A \cap T^{-n} A) \to \mu(A)^2$ as $n \to \infty$ whenever T is a mixing transformation.

A polynomial version of Khintchine's result, where the set of natural numbers n is replaced by the values of an integral polynomial $P(n)$ that satisfies $P(0) = 0$, was established by Furstenberg [9], for a proof see also [22] or [3]. Recently, far reaching generalizations of Furstenberg's result have been obtained in the settings of multiple recurrence: let $(X, \mathcal{M}, \mu, T)$ be an invertible measure preserving system, $A \in \mathcal{M}$ and $P_1, \ldots, P_\ell$ be any linearly independent family of integral polynomials with $P_i(0) = 0$ for all $1 \leq i \leq \ell$, then Frantzikinakis and Kra [7] have shown that for any $\varepsilon > 0$, there necessarily exists $n \in \mathbb{N}$ such that

$$(1) \qquad \mu(A \cap T^{-P_1(n)} A \cap \ldots \cap T^{-P_\ell(n)} A) > \mu(A)^{\ell+1} - \varepsilon.$$

We note that it follows from an earlier counterexample of Ruzsa [4] that this result cannot hold in general for dependent polynomials when $\ell \geq 2$ nor even in the setting of ergodic systems when $\ell \geq 4$. Bergelson, Host and Kra established in [4] that (1) does hold, under this additional assumption that T is ergodic in the case of (dependent) linear polynomials when $\ell = 2, 3$. Frantzikinakis [6] has investigated the situation for higher degree polynomials.

This contrasts sharply with the situation when one drops the requirement that the measure of the intersections in (1) are "optimally" large: the Polynomial Szemerédi Theorem of Bergelson and Leibman [5] states that if $(X, \mathcal{M}, \mu, T)$ is an invertible measure preserving system and $P_1, \ldots, P_\ell \in \mathbb{Z}[n]$ with $P_i(0) = 0$ for all $1 \leq i \leq \ell$, then for any $A \in \mathcal{M}$ with $\mu(A) > 0$ there necessarily exists $n \in \mathbb{N}$ such that

$$(2) \qquad \mu(A \cap T^{-P_1(n)} A \cap \ldots \cap T^{-P_\ell(n)} A) > 0.$$

Note that the case when all the polynomials are linear corresponds to Furstenberg's Multiple Recurrence Theorem [9], which is, via Furstenberg's correspondence principle, equivalent to Szemerédi's Theorem on arithmetic progressions.

For a comprehensive survey of the impact of the Poincaré recurrence principle in ergodic theory, especially as pertains to the field of ergodic Ramsey theory/additive combinatorics, see [8] and [17].

1.2. **Statement of Main Results.** In this article we will concern ourselves with the study of simultaneous (single) polynomial recurrence. The following result gives a full generalization of Furstenberg's result in this direction and can be established, as we shall see below, using current and well-known methods in ergodic Ramsey theory.

Theorem 1.1. *Let $(X, \mathcal{M}, \mu, T)$ be an invertible measure preserving system, $P_1, \ldots, P_\ell \in \mathbb{Z}[n]$ with $P_i(0) = 0$ for all $1 \leq i \leq \ell$ and $A \in \mathcal{M}$. For every $\varepsilon > 0$, there exists $n \in \mathbb{N}$ such that*

$$(3) \qquad \mu(A \cap T^{-P_i(n)} A) > \mu(A)^2 - \varepsilon \quad \text{for all} \quad 1 \leq i \leq \ell.$$

Note that there are no assumptions that the polynomials in Theorem 1.1 are linearly independent.

In the special case when $k = 1$, that is when all of the polynomials are linear, this result can be established using only combinatorial methods and the following quantitative result can be obtained.

Theorem 1.2 (Griesmer [13]). *Let $(X, \mathcal{M}, \mu, T)$ be an invertible measure preserving system, $c_1, \ldots, c_\ell \in \mathbb{Z} \setminus \{0\}$ with $\ell \leq 2^m$ for some $m \in \mathbb{N}$ and $A \in \mathcal{M}$. For every $\varepsilon > 0$ and $B \subseteq \mathbb{N}$ with $\log^m |B| \geq C\varepsilon^{-1}$, there exists a non-zero $n \in B - B$ such that*

$$(4) \qquad \mu(A \cap T^{-c_i n} A) > \mu(A)^2 - \varepsilon \quad \text{for all} \quad 1 \leq i \leq \ell.$$

It follows from a variant of Furstenberg's correspondence principle, see Frantzikinakis and Kra [7] (in particular the proof of Theorem 2.2), that Theorem 1.1 has the following combinatorial consequence.

Corollary 1.3. *Let $P_1, \ldots, P_\ell \in \mathbb{Z}[n]$ with $P_i(0) = 0$ for all $1 \leq i \leq \ell$. For every $\varepsilon > 0$ there exists $N_0 = N_0(\varepsilon, P_1, \ldots, P_\ell)$ such that if $N \geq N_0$ and $A \subseteq [1, N]$, then there exists $n \in \mathbb{N}$ such that*

$$(5) \qquad \frac{|A \cap (A + P_i(n))|}{N} > \left(\frac{|A|}{N}\right)^2 - \varepsilon \quad \text{for all} \quad 1 \leq i \leq \ell.$$

We note that this correspondence give no quantitative bounds in the finite setting of Corollary 1.3 (other than the special case when all of the polynomials are linear). However, if we relax the requirement that the intersections are "optimally" large and ask merely that they are non-empty then one has the following result.

Theorem 1.4 (Lyall and Magyar [20]). *Let $0 < \delta < 1$ and $P_1, \ldots, P_\ell \in \mathbb{Z}[n]$ with $P_i(0) = 0$ and $\deg P_i \leq k$ for all $1 \leq i \leq \ell$. There exists a constant $C = C(P_1, \ldots, P_\ell)$ such that if $N \geq \exp(C\delta^{-\ell(k-1)} \log \delta^{-1})$ and $A \subseteq [1, N]$ with $|A| \geq \delta N$, then there exists $n \in \mathbb{N}$ for which*

$$(6) \qquad A \cap (A + P_i(n)) \neq \emptyset \quad \text{for all} \quad 1 \leq i \leq \ell.$$

While if we continue to insist on "optimally" large intersections, but restrict ourselves to the case $\ell = 1$, namely the case of a single polynomial, then we have the following result.

Theorem 1.5 (Lyall and Magyar [21]). *Let $A \subseteq [1, N]$, $P(n) \in \mathbb{Z}[n]$ with $P(0) = 0$ and $\varepsilon > 0$. There exists a constant $C = C(P)$ such that if $N \geq \exp\exp(C\varepsilon^{-1} \log \varepsilon^{-1})$, then there exists $n \in \mathbb{N}$ for which*

$$(7) \qquad \frac{|A \cap (A + P(n))|}{N} > \left(\frac{|A|}{N}\right)^2 - \varepsilon.$$

The main objective of the present paper is to present the proof of a (partial) common generalization of Theorems 1.4 and 1.5. To be more precise, our objective is to establish, using Fourier analytic methods, Corollary 1.3 with explicit quantitative bounds, in the special case when all of the polynomials are of the *same* degree. In particular we are able to establish the following.

Theorem 1.6. *Let $P_1, \ldots, P_\ell \in \mathbb{Z}[n]$ with $P_i(0) = 0$ and $\deg P_i = k$ for all $1 \leq i \leq \ell$. For every $\varepsilon > 0$ there exists $N_1 = N_1(\varepsilon, P_1, \ldots, P_\ell)$ such that if $N \geq N_1$ and $A \subseteq [1, N]$, then there exist $n \in \mathbb{N}$ such that*

$$(8) \qquad \frac{|A \cap (A + P_i(n))|}{N} > \left(\frac{|A|}{N}\right)^2 - \varepsilon \quad \text{for all} \quad 1 \leq i \leq \ell.$$

In particular, the number $N_1(\varepsilon, P_1, \ldots, P_\ell)$ may be taken to be a constant (depending only on $P_1, \ldots, P_\ell$) times a tower of 2's of height $C_{k,\ell}^* + C\varepsilon^{-2}$.

As the reader will no doubt have noticed, the bounds obtained for $N_1(\varepsilon, P_1, \ldots, P_\ell)$, while explicit, are rather poor. It is our belief that these are far from the best bounds possible whose dependence on ε we would expect to be at least of exponential type.

We also note that, because of the introduction of different scales (specifically in Lemma 3.1), our current Fourier analytic approach appears to be insufficient for the task establish a quantitative result along the lines of Theorem 1.6 for polynomials with different degrees. In particular, we are not aware of any quantitative result of this type even in the simplest case, namely $\ell = 2$ with $P_1(n) = n$ and $P_2(n) = n^2$.

1.3. **An outline of the paper.** As we have been unable to find a proof of Theorem 1.1 in the literature, we give a complete proof of this result in Section 2. We feel that the inclusion of this argument will also help illuminate for the reader the proof of our main result, namely Theorem 1.6, which we present in Section 3.

In Section 4 we communicate an elegant combinatorial proof of Theorem 1.2, using Ramsey's theorem, that was shown to us by John Griesmer. We are grateful to John for both showing us this argument and giving us his permission to include it here.

Theorem 1.4 was first established in [19], but only in the case of linearly independent polynomials. In Section 5 we include a simple modification of the *lifting* argument used in [19], to extend the original result in [19] to the case of linearly dependent polynomials, thus establishing Theorem 1.4 as stated above and in [20].

Finally, we also include a short appendix on counting solutions to systems of polynomial diophantine equations as well as a somewhat lengthier appendix on simultaneous polynomial diophantine approximation.

2. THE PROOF OF THEOREM 1.1

Let $\varepsilon > 0$ and Let $(X, \mathcal{M}, \mu, T)$ be an invertible measure preserving system.

We define $U_T f(x) := f(Tx)$ and note that U_T then defines a unitary operator on the Hilbert space of all square integrable function $L^2(X, \mu)$. If we define $f = 1_A$, then

$$(9) \qquad \mu(A \cap T^{-P_i(n)} A) = \langle f, U_T^{P_i(n)} f \rangle$$

for each $1 \le i \le \ell$.

2.1. **Decomposition.** We now proceed by decomposing f into an almost periodic (structured) component and a weakly-mixing (anti-structured) component, the so-called Koopman-von Neumann decomposition.

Proposition 2.1 (Koopman and von Neumann [16], see also [22]). *Let* $\mathcal{H} := L^2(X, \mu)$, *then*

$$(10) \qquad \mathcal{H} = \mathcal{H}_c \oplus \mathcal{H}_{wm}$$

where

$$(11) \qquad \mathcal{H}_c = \{f \in \mathcal{H} : \{U_T f : n \in \mathbb{Z}\} \text{ is pre-compact}\}$$

and

$$(12) \qquad \mathcal{H}_{wm} = \left\{f \in \mathcal{H} : \lim_{N \to \infty} \frac{1}{N} |\{1 \le n \le N : |\langle U_T^n f, g \rangle| \ge \varepsilon\}| = 0 \text{ for all } \varepsilon > 0 \text{ and } g \in \mathcal{H}\right\}.$$

Moreover,

$$\mathcal{H}_{wm} = \mathcal{H}_c^\perp.$$

Using Proposition 2.1 we can therefore uniquely decompose

$$(13) \qquad f = f_1 + f_2,$$

with $f_1 \in \mathcal{H}_c$ and $f_2 \in \mathcal{H}_{wm}$. Moreover, it is easy to see that these functions also enjoy the property that $0 \le f_1 \le 1$ and $\langle f_2, 1 \rangle = 0$. We note that it follows immediately from the Cauchy-Schwarz inequality that

$$(14) \qquad \mu(A)^2 = \langle f, 1 \rangle^2 = \langle f_1, 1 \rangle^2 \le \langle f_1, f_1 \rangle.$$

2.2. **Proof of Theorem 1.1.** Inserting our decomposition $f = f_1 + f_2$ into (9) we see that

$$(15) \qquad \mu(A \cap T^{-P_i(n)}A) = \langle f_1, U_T^{P_i(n)} f_1 \rangle + \langle f, U_T^{P_i(n)} f_2 \rangle + \langle U_T^{-P_i(n)} f_2, f_1 \rangle$$

for each $1 \leq i \leq \ell$.

Our stategy to prove Theorem 1.1 will be to show that $U_T^{P_i(n)} f_1 \approx f_1$ and hence

$$\langle U_T^{P_i(n)} f_1, f_1 \rangle \approx \langle f_1, f_1 \rangle \geq \mu(A)^2$$

simultaneously for all $1 \leq i \leq \ell$ for a positive proportion of $1 \leq n \leq N$, while for any given $g \in L^2(X, \mu)$ and $P \in \mathbb{Z}[n]$ the proportion of $1 \leq n \leq N$ for which $\langle g, U_T^{P(n)} f_2 \rangle \approx 0$ tends to 1 as $N \to \infty$.

More precisely, we will establish the following two lemmas from which Theorem 1.1 follows immediately.

Lemma 2.2 (Main term estimate). *There exists $c_0 = c_0(\varepsilon, k, \ell, f_1) > 0$ such that for all large N*

$$(16) \qquad \left| \left\{ 1 \leq n \leq N \, : \, \left\| U_T^{P_i(n)} f_1 - f_1 \right\| \leq \varepsilon/2 \text{ for all } 1 \leq i \leq \ell \right\} \right| \geq c_0 N.$$

Lemma 2.3 (Error term estimate). *Let $P \in \mathbb{Z}[n]$ with $P(0) = 0$ and $g \in L^2(X, \mu)$, then*

$$(17) \qquad \lim_{N \to \infty} \frac{1}{N} \left| \left\{ 1 \leq n \leq N \, : \, |\langle U_T^{P(n)} f_2, g \rangle| \geq \varepsilon/4 \right\} \right| = 0.$$

Indeed, if N is large enough then from Lemma 2.2 it follows that there must exist at least $c_0 N$ values of $n \in [1, N]$ for which $\| U_T^{P_i(n)} f_1 - f_1 \| \leq \varepsilon/2$ and hence

$$\langle f_1, U_T^{P_i(n)} f_1 \rangle \geq \langle f_1, f_1 \rangle - |\langle f_1, U_T^{P_i(n)} f_1 - f_1 \rangle| \geq \mu(A)^2 - \varepsilon/2$$

simultaneously for all $1 \leq i \leq \ell$. While from Lemma 2.3 it follows that if N is taken sufficiently large then the absolute value of the last two error terms in (15) can be made less than $\varepsilon/4$ for all but at most $c_0 N/2$ values of $n \in [1, N]$ simultaneously for $1 \leq i \leq \ell$. Thus for any $\varepsilon > 0$ there exists $n \in [1, N]$ (in fact a positive proportion) for which

$$\mu(A \cap T^{-P_i(n)}A) > \mu(A)^2 - \varepsilon$$

for all $1 \leq i \leq \ell$. This completes the proof of Theorem 1.1.   □

2.3. **Proof of Lemmas 2.2 and 2.3.** The proof of Lemma 2.2 is based on van der Waerden's theorem and the magical identity

$$(18) \qquad \sum_{t=0}^{j} (x + td)^j \binom{j}{t} (-1)^{j-t} \, = \, j! \, d^j$$

the validity of which can be easily verified for all $j \in \mathbb{N}$ by induction. Lemma 2.2 is of course in essence a result on simultaneous diophantine approximation and the proof we present below is essentially an adaptation of the proof of Proposition 1.5 (on quadratic recurrence) in [24].

The proof of Lemma 2.3 follows from the Hilbert space version of van der Corput's Lemma (for a statement of this version see either [3] or [22]) and is well-known, but for the sake of completeness we have chosen to also sketch its proof below.

Proof of Lemma 2.2. Let $P_i(x) = \sum_{j=1}^{k} c_{ij} x^j$, and let $\eta = \eta(\varepsilon, k) > 0$ be a small constant to be chosen later.

Cover the orbit $\{ T^n f \, : \, n \in \mathbb{Z} \}$ by balls $B_1, \ldots, B_M$ of diameter η and use this to define a (matrix-valued) coloring $\chi : \mathbb{Z} \to [1, M]^{k\ell}$ of the integers by setting, for each $1 \leq i \leq \ell$ and $1 \leq j \leq k$, $\chi_{ij}(n) = r$ if $U_T^{c_{ij} n^j} f_1 \in B_r$.

By (the averaged version of) van der Waerden's theorem there is a constant $c_1 = c_1(M, k, \ell)$ such that there will be at least $c_1 N$ monochromatic $(k+1)$-term arithmetic progressions in $[1, N]$ with *different* step sizes d, provided N is sufficiently large.

Let d be the step size of a monochromatic arithmetic progression $\{ x + td : \ 0 \leq t \leq k \}$ with respect to the coloring χ in $[1, N]$. Then for any fixed i, j one has

$$\| U_T^{c_{ij}(x+td)^j} f_1 - U_T^{c_{ij} x^j} f_1 \| \leq \eta$$

for all $0 \leq t \leq k$, thus by (18) it follows that

$$\left\| U_T^{c_{ij} j! d^j} f_1 - f_1 \right\| = \left\| U_T^{c_{ij} \sum_{t=0}^{j} (x+td)^j \binom{j}{t} (-1)^{j-t}} f_1 - f_1 \right\|$$

$$= \left\| U_T^{c_{ij} \sum_{t=0}^{j} (x+td)^j \binom{j}{t} (-1)^{j-t}} f_1 - U_T^{c_{ij} \sum_{t=0}^{j} x^j \binom{j}{t} (-1)^{j-t}} f_1 \right\|$$

$$\leq \sum_{t=0}^{j} \binom{j}{t} \left\| U_T^{c_{ij}(x+td)^j} f_1 - U_T^{c_{ij} x^j} f_1 \right\|$$

$$\leq 2^j \eta.$$

Here we have used the facts that $\|U_T^{m_1+m_2} f_1 - U_T^{n_1+n_2} f_1\| \leq \|U_T^{m_1} f_1 - U_T^{n_1} f_1\| + \|U_T^{m_2} f_1 - U_T^{n_2} f_1\|$ and $\|U_T^{bn} f_1 - U_T^{bm} f_1\| \leq |b| \|U_T^{n} f_1 - U_T^{m} f_1\|$ which follows from the triangle inequality and the fact that U_T is a unitary operator on $L^2(X,\mu)$. Thus

$$\left\| U_T^{c_{ij}(k!d)^j} f_1 - f_1 \right\| \leq (k!^j/j!) \left\| U_T^{j! c_{ij} d^j} f_1 - f_1 \right\| \leq (k!)^k 2^k \eta.$$

Letting $n = k!d$ it follows that for all $1 \leq i \leq \ell$ we have

(19)
$$\left\| U_T^{P_i(n)} f_1 - f_1 \right\| \leq k(k!)^k 2^k \eta \leq \varepsilon/2$$

provided η is chosen small enough. Since the number of such $d \in [1, N/k!]$ is at least $(c_1/k!)N$, the lemma follows. □

Proof of Lemma 2.3. We give a proof by induction, using the fact that (17) is equivalent to

(20)
$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \left| \langle U_T^{P(n)} f_2, g \rangle \right|^2 = 0$$

and that when $\deg P = 1$, that is $P(x) = mx$ for some $m \in \mathbb{Z}$, then the conclusion of the lemma is an immediately consequence of the weak-mixing properties of f_2.

Let $k \geq 1$, $P \in \mathbb{Z}[n]$ be a polynomial of degree $k+1$, and assume that (20) holds for all polynomials of degree at most k. We will show that (20) holds for all $g \in L^2(X,\mu)$. To this end we note that

$$\frac{1}{N} \sum_{n=1}^{N} \left| \langle U_T^{P(n)} f_2, g \rangle \right|^2 = \frac{1}{N} \sum_{n=1}^{N} \left\langle (U_T \times U_T)^{P(n)} (f_2 \times f_2), g \times g \right\rangle_{X \times X}$$

and hence that it suffices to show that

$$\lim_{N \to \infty} \left\| \frac{1}{N} \sum_{n=1}^{N} (U_T \times U_T)^{P(n)} (f_2 \times f_2) \right\|_{L^2(X \times X)} = 0.$$

Let $x_n := (U_T \times U_T)^{P(n)} (f_2 \times f_2)$ for $n \in \mathbb{Z}$, and let $h \in \mathbb{Z} \setminus \{0\}$. Since

$$\langle x_{n+h}, x_n \rangle_{X \times X} = \left\langle (U_T \times U_T)^{P(n+h)-P(n)-P(h)} (f_2 \times f_2), (U_T \times U_T)^{-P(h)} (f_2 \times f_2) \right\rangle_{X \times X}$$

$$= \left| \left\langle U_T^{P(n+h)-P(n)-P(h)} f_2, U_T^{-P(h)} f_2 \right\rangle \right|^2$$

it follows from the inductive hypothesis, since the polynomial $P(n+h) - P(n) - P(h)$ has degree at most k and $U_T^{-P(h)} f_2 \in L^2(X,\mu)$, that

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \langle x_{n+h}, x_n \rangle_{X \times X} = 0.$$

The claim now follows from the Hilbert space version of van der Corput's Lemma, see either [3] or [22]. □

3. THE PROOF OF THEOREM 1.6

Let $\varepsilon > 0$ and $P_1, \ldots, P_\ell \in \mathbb{Z}[n]$ with $P_i(0) = 0$ and $\deg P_i = k$ for all $1 \leq i \leq \ell$.

3.1. The Fourier transform, uniformity and polynomial shifts. Let $\mathbb{Z}_N$ denote the group $\mathbb{Z}/N\mathbb{Z}$.

3.1.1. *The Fourier transform.* Given $f : \mathbb{Z}_N \to \mathbb{C}$ we define its (discrete) Fourier transform, $\widehat{f} : \widehat{\mathbb{Z}_N} \to \mathbb{C}$, by

$$\widehat{f}(\xi) = \frac{1}{N} \sum_{x \in \mathbb{Z}_N} f(x) e(-x\xi/N)$$

where $e(x) = e^{2\pi i x}$ and $\widehat{\mathbb{Z}_N}$ denote the *dual group* of of all characters on $\mathbb{Z}_N$.

It is easy to see that we can identify $\mathbb{Z}_N$ with its dual. There are two natural measures that one can put on $\mathbb{Z}_N$, namely uniform probability measure and counting measure. As is customary, we shall use the uniform probability measure on $\mathbb{Z}_N$ and the counting measure on $\mathbb{Z}_N$ when it is being identified with its dual group. We then define L^p-norms and ℓ^p-norms as follows.

We define L^p to be the space of all functions from $\mathbb{Z}_N$ to $\mathbb{C}$, with the norm

$$\|f\|_p = \Big(\frac{1}{N} \sum_{x \in \mathbb{Z}_N} |f(x)|^p \Big)^{1/p},$$

where this is interpreted as $\max_{x \in \mathbb{Z}_N} |f(x)|$ when $p = \infty$. We define ℓ^p to be the space of all functions from $\widehat{\mathbb{Z}_N}$ to $\mathbb{C}$, with the norm

$$\|F\|_p = \Big(\sum_{x \in \mathbb{Z}_N} |F(x)|^p \Big)^{1/p},$$

where we again interpreted this as $\max_{x \in \mathbb{Z}_N} |F(x)|$ when $p = \infty$.

In contrast to the situation for the Fourier transform on $\mathbb{R}$, the Fourier inversion formula and Plancherel's identity, namely

$$f(x) = \sum_{\xi \in \mathbb{Z}_N} \widehat{f}(\xi) e(x\xi/N) \qquad \text{and} \qquad \|f\|_2 = \|\widehat{f}\|_2$$

are, in this setting, immediate and simple consequences of the familiar orthogonality relation

$$\frac{1}{N} \sum_{x \in \mathbb{Z}_N} e(x\xi/N) = \begin{cases} 1 & \text{if } \xi = 0 \\ 0 & \text{if } \xi \neq 0 \end{cases} .$$

3.1.2. *Uniformity and polynomial shifts.* We now fix a set $A \subseteq [1, N]$. In order to use Fourier analytic techniques we will, as is customary, identify $[1, N]$ with $\mathbb{Z}_N$ and consider A as a subset of $\mathbb{Z}_N$. In order to ensure that working in $\mathbb{Z}_N$ will not, in any essential way, affect the validity of (8), we will restrict our attention to those values of n for which

$$(21) \qquad\qquad 1 \leq n \leq M := c(\varepsilon N)^{1/k}$$

with $c = c(P_1, \ldots, P_\ell)$ chosen sufficiently small such that $|P_i(n)| \leq \varepsilon N$ for all $1 \leq i \leq \ell$. Note that doing this we will ensure that the size of $A \cap (A + P_i(n))$ will increase by at most εN, due to overlapping when the shifts $P_i(n)$ take place in $\mathbb{Z}_N$, and as such working in $\mathbb{Z}_N$ will not affect the validity of (8) (other than changing ε to 2ε).

Note that if we define $f = 1_A$, then

$$(22) \qquad\qquad \frac{|A \cap (A + P_i(n))|}{N} = \frac{1}{N} \sum_{x \in \mathbb{Z}_N} f(x) f(x - P_i(n))$$

for each $1 \leq i \leq \ell$.

It is well-known (and easy to verify, see Lemma 2.2 in [11]) that if $g, h : \mathbb{Z}_N \to \mathbb{C}$ with both $\|g\|_2$ and $\|h\|_2$ bounded by 1, then $\|\widehat{h}\|_\infty \leq \eta$ is equivalent to

$$\frac{1}{N} \sum_{n=1}^{N} \left| \frac{1}{N} \sum_{x \in \mathbb{Z}_N} h(x) g(x - n) \right|^2 \leq \eta$$

and consequently that

$$\left|\left\{1 \le n \le N \;:\; \left|\frac{1}{N}\sum_{x \in \mathbb{Z}_N} h(x)g(x-n)\right| \ge \eta^{1/3}\right\}\right| \le \eta^{1/3}N$$

whenever h satisfies the uniformity assumption that $\|\widehat{h}\|_\infty \le \eta$.

The main aim of this section is to show that this phenomenon continues to hold for each of the polynomial shifts $P_i(n)$ for almost all n satisfying (21). In particular we show the following.

Lemma 3.1 (Error term estimate). *Let $\varepsilon > 0$, $P \in \mathbb{Z}[n]$ with $\deg P = k \ge 2$ and $M = c(\varepsilon N)^{1/k}$.*

If $g, h : \mathbb{Z}_N \to \mathbb{C}$ with $\|g\|_2 \le 1$, $\|h\|_2 \le 1$ and $\|\widehat{h}\|_\infty \le \eta$ with $0 < \eta \le \varepsilon$, then

$$(23) \qquad \left|\left\{1 \le n \le M \;:\; \left|\frac{1}{N}\sum_{x \in \mathbb{Z}_N} h(x)g(x-P(n))\right| \ge \eta^{1/K}\right\}\right| \le C_1\, \eta^{1/K} M$$

for any positive integer $K \ge C\,k^2 \log k$, where then C_1 is a large constant depending only on P.

3.1.3. *Theorem 1.6 for uniform sets.* We quickly remark that from Lemma 3.1 we can immediately deduce Theorem 1.6 in the special case of (suitably) uniform sets A. Recall that a set A is said to be η-uniform if $|\widehat{1_A}(\xi)| \le \eta$ for all $\xi \in \mathbb{Z}_N \setminus \{0\}$, or equivalently $\|\widehat{f_A}\|_\infty \le \eta$, where $f_A = 1_A - |A|/N$ denotes the so-called balanced function of A. Inserting the decomposition $f = f_A + |A|/N$ into (22) and recalling that the function f_A has mean value zero, we obtain the following corollary to Lemma 3.1.

Corollary 3.2. *If $A \subseteq \mathbb{Z}_N$ is η-uniform with $0 < \eta \le \varepsilon^K$ and $K \ge C\,k^2 \log k$, then there necessarily exist at least $(1 - \ell C_1 \eta^{1/K})M$ values of $n \in [1,M]$ (and hence at least one if $\eta^{1/K} \le \min\{\varepsilon, 1/2\ell C_1\}$) for which*

$$\left|\frac{|A \cap (A + P_i(n))|}{N} - \left(\frac{|A|}{N}\right)^2\right| < \varepsilon$$

holds simultaneously for $1 \le i \le \ell$.

3.1.4. *Proof of Lemma 3.1.* We make use of the following well-known fact from number theory which we state here, and whose proof will be given in Appendix A for the sake of completeness.

Proposition 3.3. *Let $P \in \mathbb{Z}[n]$ with $\deg P = k$ and $K \ge C\,k^2 \log k$ be a natural number. Then for any $M \in \mathbb{N}$, the number of $2K$-tuples $(n_1, \ldots, n_K, m_1, \ldots, m_K) \in [1, M]^{2K}$ satisfying*

$$(24) \qquad P(n_1) + \cdots + P(n_K) = P(m_1) + \cdots + P(m_K)$$

is bounded by $C_0 M^{2K-k}$, where C_0 is a large constant depending only on P.

In order to prove Lemma 3.1, with $4K$ in place of K, it is suffices to show that

$$(25) \qquad \sum_{n=1}^{M} \left|\frac{1}{N}\sum_{x \in \mathbb{Z}_N} h(x)g(x - P(n))\right| \le C_1 \eta^{1/2K} M.$$

In order to verify (25) we introduce a "weight function" $w : \mathbb{Z}_N \to \{-1, 1\}$ such that the left side of (25) may be written (using Fourier inversion) as

$$(26) \qquad \frac{1}{N}\sum_{n=1}^{M}\sum_{x \in \mathbb{Z}_N} h(x)g(x - P(n))w(n) = \sum_{\xi \in \mathbb{Z}_N} \widehat{h}(\xi)\overline{\widehat{g}(\xi)} S_w(\xi)$$

where

$$(27) \qquad S_w(\xi) = \sum_{n=1}^{M} w(n)e(P(n)\xi/N).$$

Note that we have no control over the "weight function" $w(n)$ and hence cannot hope for any non-trivial pointwise bound on $|S_w(\xi)|$, nevertheless we can obtain sharp estimates for the higher moments of S_w using Proposition 3.3. Indeed for any positive integer $K \geq C\,k^2\log k$ one estimates

$$\|S_w\|_{2K}^{2K} = \sum_{\xi \in \mathbb{Z}_N} |S_w(\xi)|^{2K}$$

$$= \sum_{\substack{n_1,\dots,n_K=1 \\ m_1\dots,m_K=1}}^{M} w(n_1)\cdots w(m_K) \sum_{\xi \in \mathbb{Z}_N} e\left((P(n_1)+\cdots+P(n_K)-P(m_1)-\cdots-P(m_K))\,\xi/N\right)$$

$$\leq N\left|\left\{(n_1,\dots,n_K,m_1\dots,m_K)\in[1,M]^{2K}\,:\,P(n_1)+\cdots+P(n_K)=P(m_1)+\cdots+P(m_K)\right\}\right|.$$

Since $1\leq n_i,m_i\leq M$ and $M=c\,(\varepsilon N)^{1/k}$ with a sufficiently small constant c, the equality

$$P(n_1)+\cdots+P(n_K)=P(m_1)+\cdots+P(m_K)$$

holds in $\mathbb{Z}_N$ if and only it holds in $\mathbb{Z}$.

It therefore follows from Proposition 3.3 and Hölder's inequality that

$$\sum_{\xi\in\mathbb{Z}_N}|\widehat{h}(\xi)||\widehat{g}(\xi)||S_w(\xi)| \;\leq\; \|\widehat{g}\|_2\,\|\widehat{f}\|_{\frac{2K}{K-1}}\,\|S_w\|_{2K}\leq \eta^{\frac{1}{K}}(C_0 N M^{2K-k})^{1/2K}\leq C_1\eta^{1/2K}M$$

with $C_1=(C_0/c^k)^{1/2K}$ since $\|\widehat{h}\|_{\frac{2K}{K-1}}\leq\|\widehat{h}\|_\infty^{1/K}\|\widehat{h}\|_2^{(K-1)/K}\leq\eta^{\frac{1}{K}}$ and $M=c(\varepsilon N)^{1/k}$. $\qquad\square$

3.2. Decomposition.

In order to exploit the phenomenon exhibited in Lemma 3.1 in a proof of Theorem 1.6 one would naturally try, as we did in the proof of Theorem 1.1, to make use of a decomposition theorem that will allow use to decompose f into a structured component and a suitably uniform (anti-structuered) component.

It is easy to see that for any given $\eta>0$ and $f:\mathbb{Z}_N\to\mathbb{C}$ with $\|f\|_2\leq 1$, then the number of $\xi\in\mathbb{Z}_N$ such that $|\widehat{f}(\xi)|\geq\eta$ is at most η^{-2}, since $\|\widehat{f}\|_2\leq 1$. Using the Fourier inversion formula together with this fact we certainly split $f=g+h$, where

$$g(x)=\sum_{\xi\in\Gamma}\widehat{f}(\xi)e(\xi x/N)\quad\text{and}\quad h(x)=\sum_{\xi\notin\Gamma}\widehat{f}(\xi)e(\xi x/N)$$

with $\Gamma=\{\xi\in\mathbb{Z}_N\,:\,|\widehat{f}(\xi)|\geq\eta\}$. It is then immediate that $|\widehat{h}(\xi)|<\eta$, and that g is indeed "structured" in the sense that it involves only a bounded number of characters.

As we shall see below this simple decomposition will unfortunately be insufficient for our purpose. The problem being that we need a much stronger relationship between $\|\widehat{h}\|_\infty$ and the upper bound on the size of Γ. The following result, which shows that one can indeed obtain this, modulo a small L^2-error, is part of the standard folklore of additive combinatorics (see for example [12] Prop. 2.5).

Proposition 3.4. *Let $f:\mathbb{Z}_N\to\mathbb{C}$ with $\|f\|_2\leq 1$ and write $\mathbb{Z}_N=\{\xi_1,\dots,\xi_N\}$ so that $|\widehat{f}(\xi_1)|\geq\cdots\geq |\widehat{f}(\xi_N)|$.*

For every $\varepsilon>0$ and $\eta:\mathbb{N}\to\mathbb{R}_+$ a positive function that decreases to 0, there exists $m=m(\varepsilon,\eta)\in\mathbb{N}$ and a decomposition

$$f=f_1+f_2+f_3,$$

with

$$f_1(x)=\sum_{j=1}^{m}\widehat{f}(\xi_j)e(\xi_j x/N)$$

while $\|\widehat{f_2}\|_\infty\leq\eta(m)$ and $\|f_3\|_2\leq\varepsilon$.

We note that the proof of this result, we include below for completeness, gives us the desired decomposition for some natural number m that is bounded above by a number that results from starting with 1 and applying the function $t\mapsto\eta(t)^{-4}$ at most ε^{-2} times.

Proof of Proposition 3.4, [12]. Choose an increasing sequence of positive integers $m_1, m_2, \ldots$ with $m_1 = 1$ and $m_{r+1} \geq \eta(m_r)^{-4}$ for every r. We now choose r and attempt to prove the result using the decomposition $f = f_1 + f_2 + f_3$ with

$$f_1(x) = \sum_{j \leq m_r} \widehat{f}(\xi_j) e(\xi_j x/N), \quad f_2(x) = \sum_{j > m_{r+1}} \widehat{f}(\xi_j) e(\xi_j x/N), \quad f_3(x) = \sum_{m_r < j \leq m_{r+1}} \widehat{f}(\xi_j) e(\xi_j x/N).$$

Then f_1 is a linear combination of at most m_r characters. Also $\|\widehat{f_2}\|_\infty \leq m_{r+1}^{-1} \leq \eta(m_r)^4$, since, by Plancherel's theorem, there can be at most m_{r+1} Fourier coefficients of f whose magnitude is at least $m_{r+1}^{-1/2}$.

Therefore, we are done if $\|f_3\|_2 \leq \varepsilon$. But the possible functions f_3 (as r varies) are disjoint parts of the Fourier decomposition of f, so at most ε^{-2} of them can have norm greater then ε. Thus there exists $r \leq \varepsilon^{-2}$ such that the preposed decomposition works. $\square$

To find the appropriate function $\eta(m)$ we use the following result from the theory of diophantine approximation, the proof of this result is presented in Appendix B.

Lemma 3.5 (Main term estimate). *There exists a constant $C_{k,\ell}$ such that for all $0 < \varepsilon < 1/2$ we have*

$$(28) \qquad |\{1 \leq n < M \; : \; \|P_i(n)\xi_j/N\| < \varepsilon \text{ for all } 1 \leq i \leq \ell, \, 1 \leq j \leq m\}| \geq (\varepsilon/m)^{C_{k,\ell}m^2} M$$

provided $M \geq (\varepsilon/m)^{-C_{k,\ell}m^2}$, where $\|\alpha\|$ denotes, for each $\alpha \in \mathbb{R}$, the distance from α to the nearest integer.

3.3. **Proof of Theorem 1.6.** Inserting the decomposition from Proposition 3.4, for say $\varepsilon/8$ and a positive decreasing function $\eta : \mathbb{N} \mapsto \mathbb{R}_+$ to be chosen later (using Lemma 3.5), into (22) we obtain

$$\frac{|A \cap (A + P_i(n))|}{N} = \frac{1}{N} \sum_{x \in \mathbb{Z}_N} f_1(x) f_1(x - P_i(n)) + \frac{1}{N} \sum_{x \in \mathbb{Z}_N} (f_2 + f_3)(x) g(x - P_i(n))$$

$$\geq \left| \frac{1}{N} \sum_{x \in \mathbb{Z}_N} f_1(x) f_1(x - P_i(n)) \right| - \left| \frac{1}{N} \sum_{x \in \mathbb{Z}_N} (f_2 + f_3)(x) g(x - P_i(n)) \right|$$

for each $1 \leq i \leq \ell$, where $g : \mathbb{Z}_N \mapsto \mathbb{R}$ is some function that satisfies $\|g\|_2 \leq 2$.

By the Cauchy-Schwarz inequality it follows that

$$(29) \qquad \left| \frac{1}{N} \sum_{x \in \mathbb{Z}_N} f_3(x) g(x - P_i(n)) \right| \leq \|f_3\|_2 \|g\|_2 \leq \varepsilon/4$$

for all $n \in \mathbb{N}$, while from Lemma 3.1 it follows that for all but at most $\ell C_1 \eta(m)^{1/K} M$ values of $n \in [1, M]$ we also have

$$(30) \qquad \left| \frac{1}{N} \sum_{x \in \mathbb{Z}_N} f_2(x) g(x - P_i(n)) \right| < \eta(m)^{1/K}$$

simultaneously for all $1 \leq i \leq \ell$.

Using the fact that $\xi_1 = 0$ (a consequence of f is non-negative) one can conclude that for each $1 \leq i \leq \ell$ we have

$$\left| \frac{1}{N} \sum_{x \in \mathbb{Z}_N} f_1(x) f_1(x - P_i(n)) \right| = \left| \sum_{j=1}^{m} |\widehat{f}(\xi_j)|^2 e(\xi_j P_i(n)/N) \right|$$

$$\geq \sum_{j=1}^{m} |\widehat{f}(\xi_j)|^2 - \sum_{j=1}^{m} |\widehat{f}(\xi_j)|^2 \, |e(\xi_j P_i(n)/N) - 1|$$

$$\geq |\widehat{f}(\xi_1)|^2 - \varepsilon/2$$

$$= \left(\frac{|A|}{N} \right)^2 - \frac{\varepsilon}{2}$$

provided that $|e(\xi_j P_i(n)/N) - 1| \leq \varepsilon/2$ for all $1 \leq j \leq m$. It therefore follows from Lemma 3.5, since $|e(x) - 1| \leq 2\pi\|x\|$, that for at least $(\varepsilon/4\pi m)^{C_{k,\ell}m^2} M$ values of $n \in [1, M]$ the main terms will satisfy

$$\left| \frac{1}{N} \sum_{x \in \mathbb{Z}_N} f_1(x) f_1(x - P_i(n)) \right| \geq \left(\frac{|A|}{N} \right)^2 - \frac{\varepsilon}{2}$$

simultaneously for all $1 \leq i \leq \ell$, provided $M \geq (\varepsilon/4\pi m)^{-C_{k,\ell}m^2}$.

Therefore, if we choose $\eta(t) := (\ell C_1)^{-K}(\varepsilon/4\pi t)^{C_{k,\ell}Kt^2}/2$, it follows that $\ell C_1 \eta(m)^{1/K} < (\varepsilon/16m)^{C_{k,\ell}m^2}$, thus establishing Theorem 1.6 with

$$N_1(\varepsilon, P_1, \ldots, P_\ell) = C(\varepsilon/4\pi m)^{-kC_{k,\ell}m^2-1}$$

where the constant C depends only on the polynomials $P_1, \ldots, P_\ell$.

Since, as we remarked above, m is bounded above by a number that results from starting with 1 and applying the function $t \mapsto \eta(t)^{-4}$ at most ε^{-2} times, N_1 will clearly be a tower type bound with height proportional to ε^{-2}. To be more precise, it is not hard to verify that

$$(\varepsilon/4\pi m)^{-kC_{k,\ell}m^2-1} = T(C_{k,\ell}^* + C\varepsilon^{-2}),$$

where $T(1) = 1$, $T(j+1) = 2^{T(j)}$ is the tower function.

Finally we remark that if one uses explicit Vinogradov type bounds for the Weyl sums, then one obtains the constant $C_{k,\ell} = Ck^C \ell^C$ in Proposition 3.3 and hence $C_{k,\ell}^* \leq C \log^*(k) + C \log^*(\ell) + \log^* C$ where $\log^*$ denotes the inverse of the tower function $T(n)$. $\qquad \square$

4. Simultaneous Linear Recurrence

In this section we will give a combinatorial proof of Theorem 1.2, shown to us by John Griesmer [13].

4.1. A combinatorial proof of Khintchine's theorem.
We first establish the result for $\ell = 1$ (which of course corresponds to the case $m = 0$). This is a quantitative formulation of the ε-optimal extension of Poincaré's recurrence theorem, due to Khintchine, that was discussed in the introduction.

Lemma 4.1. *Let $(X, \mathcal{M}, \mu, T)$ be an invertible measure preserving system and $A \in \mathcal{M}$. For every $\varepsilon > 0$ and $B \subseteq \mathbb{N}$ with $|B| \geq \varepsilon^{-1}$, there exists a non-zero $n \in B - B$ such that*

$$(31) \qquad \mu(A \cap T^{-n}A) > \mu(A)^2 - \varepsilon.$$

Proof. Let $\varepsilon > 0$ and $v_1, \ldots, v_N$ be distinct natural numbers with $N \geq \varepsilon^{-1}$. Since

$$N\mu(A) = \int_X \sum_{j=1}^N 1_A(T^{v_j}x) \, d\mu$$

it follows from the Cauchy-Schwarz inequality that

$$N^2 \mu(A)^2 = \int_X \left(\sum_{j=1}^N 1_A(T^{v_j}x) \right)^2 d\mu$$

and hence that

$$\frac{1}{N^2} \sum_{1 \leq j,k \leq N} \mu(T^{-v_j}A \cap T^{-v_k}A) \geq \mu(A)^2.$$

It then follows from the fact that there are only N diagonal terms that there must exist a pair $1 \leq j < k \leq N$ for which

$$\mu(A \cap T^{-(v_k - v_j)}A) \geq \mu(A)^2 - \varepsilon. \qquad \square$$

4.2. **Proof of Theorem 1.2.** Let m be a non-negative integer. It clearly suffices to prove the Theorem for $\ell = 2^m$. We proceed by induction on m, noting that Lemma 4.1 above covers the base case, namely when $m = 0$, since we can simply apply that result to the transformation T^{c_1} in place of T for any given $c_1 \in \mathbb{Z} \setminus \{0\}$.

We now assume that the result holds for a given non-negative integer m, fix $c_1, \ldots, c_{2^m+1} \in \mathbb{Z} \setminus \{0\}$ and distinct natural numbers $v_1, \ldots, v_N$ with $N \geq 2^{2L}$ where $\log_2^m L \geq C\varepsilon^{-1}$.

We now define a coloring of the edges of the complete graph K_N on the vertices $v_1, \ldots, v_N$. Color the edge between vertices v_j and v_k *red* if

$$\mu(A \cap T^{-c_i(v_j - v_k)}A) > \mu(A)^2 - \varepsilon$$

for all $1 \leq i \leq 2^m$, and color it *blue* if this is not the case.

Since $N \geq 2^{2L}$ and $\log_2^m L \geq C\varepsilon^{-1}$, it follows from Ramsey's theorem and the inductive hypothesis that K_N must contain a complete *red* subgraph with at least L vertices. In other words, there exists a sub-collection

$$\{w_1, \ldots, w_L\} \subseteq \{v_1, \ldots, v_N\}$$

such that *for every* $1 \leq j, k \leq L$ we have

$$\mu(A \cap T^{-c_i(w_j - w_k)}A) > \mu(A)^2 - \varepsilon$$

for all $1 \leq i \leq 2^m$. Applying the inductive hypothesis once more, this time with $c_{2^m+1}, \ldots, c_{2^{2m}}$ to the collection $w_1, \ldots, w_L$ of distinct natural numbers obtained above, it follows that there necessarily exists a pair $1 \leq j < k \leq L$ such that

$$\mu(A \cap T^{-c_i(w_j - w_k)}A) > \mu(A)^2 - \varepsilon$$

for all $2^m + 1 \leq i \leq 2^{2m}$. This completes the proof. $\qquad\square$

5. The proof of Theorem 1.4

It is easy to see that in order to prove Theorem 1.4 it suffices to establish the following reformulation.

Theorem 5.1. *Let $k \geq 2$ and $P_1, \ldots, P_\ell \in \mathbb{Z}[n]$ with $P_i(0) = 0$ and $\deg P_i \leq k$ for all $1 \leq i \leq \ell$.*

If $A \subseteq [1, N]$ and $\{P_1(n), \ldots, P_\ell(n)\} \nsubseteq A - A$ for any $n \neq 0$, 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 $C = C(P_1, \ldots, P_\ell)$.

As remarked in Section 1.3, Theorem 5.1 was established for families of linearly independent polynomials in [19]. In the case of a single polynomial ($\ell = 1$), this result has originally obtained by Lucier [18] (with slightly weaker bounds) 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.

A simple modification of the *lifting* argument utilized in [19] will enable us to deduce Theorem 5.1 as a corollary of the following higher dimensional result, which was the main result in [19] (see also [20]).

Theorem 5.2 (Lyall and Magyar [19], [20])**.** *If $B \subseteq [1, N]^k$ and $(n, n^2, \ldots, n^k) \notin B - B$ for any $n \neq 0$ 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 $C = C(k)$.

As in [19], we again speculate that the methodology of Balog et al. [2] may be applied in this higher dimensional situation to obtain far superior bounds in Theorem 5.2 and hence also in Theorem 5.1.

5.1. **Proof that Theorem 5.2 implies Theorem 5.1.** Let $P_i(n) = c_{i1}n + \cdots + c_{ik}n^k$ 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 of 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

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

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 set $\delta = |A|/N$. 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

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

Theorem 5.1 now follows from Theorem 5.2 since if there were to exist an $n \neq 0$ such that

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

this would immediately implies that

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

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

APPENDIX A. COUNTING SOLUTIONS TO SYSTEMS OF POLYNOMIAL DIOPHANTINE EQUATIONS

The aim of this section is to prove Proposition 3.3. We do this by showing that it follows easily from counting the integer solutions $1 \leq x_1, \ldots, x_K, y_1, \ldots, y_K \leq M$ of the system of equations

$$(32) \qquad x_1^i + \cdots + x_K^i = y_1^i + \cdots + y_K^i$$

where the exponent i ranges from 1 to k, known as Tarry's problem. An asymptotic formula for the number of solutions $J_{K,k}(M)$ (as $M \to \infty$) was originally obtained by Hua [14], see also Wooley [25]. In particular, it follows from these results that

$$(33) \qquad J_{K,k}(M) \leq C_k M^{2K - k(k+1)/2}$$

as long as $K > Ck^2 \log k$. For additional discussion of Tarry's problem, see [1] and [25].

Proof of Proposition 3.3. Let $P(x) = c_k x^k + \cdots + c_1 x$ be an integral polynomial. For given

$$1 \leq x_1, \ldots, x_K, y_1, \ldots, y_K \leq M$$

and for $1 \leq i \leq k$, let

$$(34) \qquad s_i = x_1^i + \cdots + x_K^i \, - \, y_1^i - \cdots - y_K^i$$

Then $x_1, \ldots, x_K, y_1, \ldots, y_K$ is a solution of equation (24) if and only if $c_1 s_1 + \cdots + c_k s_k = 0$.

For given $s_1, \ldots, s_k$, let $J_{K,k}(M; s)$ denote the number of integer solutions of the system (34). Then, as usual, one can express the number of solutions as a multiple integral of the form

$$J_{K,k}(M; s) = \int_0^1 \cdots \int_0^1 |S(\theta_1, \ldots, \theta_k)|^{2K} e^{-2\pi i (s_1 \theta_1 + \cdots + s_k \theta_k)} \, d\theta_1 \ldots d\theta_k$$

where

$$S(\theta_1, \ldots, \theta_k) = \sum_{m=1}^M e^{2\pi i (m\theta_1 + \cdots + m^k \theta_k)}.$$

Therefore, we have that $J_{K,k}(M; s) \leq J_{K,k}(M)$ uniformly in $s_1, \ldots, s_k$. For a solution of (24) the values $s_1, \ldots, s_{k-1}$ determine s_k, and since $|s_i| \leq kN^i$ for $1 \leq i \leq k-1$, one estimates the number of solutions of (24) from above by $C_k M^{(k-1)k/2} M^{2K - k(k+1)/2} = C_k M^{2K-k}$. This proves Proposition 3.3. $\qquad\square$

APPENDIX B. SIMULTANEOUS POLYNOMIAL DIOPHANTINE APPROXIMATION

The purpose of this section is to supply a proof of Lemma 3.5. We in fact establish the following more general result.

Proposition B.1. *Let* $P_1, \ldots, P_\ell \in \mathbb{Z}[n]$ *with* $P_i(0) = 0$ *and* $\deg P_i \leq k$ *for all* $1 \leq i \leq \ell$ *and* $\theta_1, \ldots, \theta_m \in \mathbb{R}$. *Then for any* $0 < \varepsilon \leq 1/2$ *and* $N \in \mathbb{N}$ *we have that*

$$(35) \qquad |\{1 \leq n \leq N : \|P_i(n)\theta_{i'}\| < \varepsilon \text{ for all } 1 \leq i \leq \ell, \, 1 \leq i' \leq m\}| \geq (\varepsilon/d)^{C_k d^2}$$

where $d = k\ell m$ *and* $C_k > 0$ *is a constant depending only on* k.

It is easy to see that Proposition B.1 is an immediate corollary of the following result.

Proposition B.2. *Let* $d_1, \ldots, d_k \in \mathbb{N}$ *and* $\alpha_j \in \mathbb{R}^{d_j}$ *for* $1 \leq j \leq k$. *There exists* $C_k > 0$ *such that for any* $0 < \varepsilon \leq 1/2$ *and* $N \in \mathbb{N}$ *we have*

$$(36) \qquad |\{1 \leq n \leq N : \|n^j \alpha_j\| < \varepsilon \text{ for all } 1 \leq j \leq k\}| \geq (\varepsilon/d)^{C_k d^2}$$

where $d = d_1 + \cdots + d_k$ *and* $\|\alpha_j\|$ *denotes, for all* $\alpha_j \in \mathbb{R}^{d_j}$, *the distance from* α_j *to the nearest integer point in* $\mathbb{Z}_{d_j}$.

Proof of Proposition B.1. Let $P_i(n) = \sum_{j=1}^k c_{ij} n^j$ for $1 \leq i \leq k$. For any given $1 \leq i \leq \ell$ and $1 \leq i' \leq m$ we of course have $\|P_i(n)\theta_{i'}\| < \varepsilon$ whenever $\|n^j c_{ij} \theta_{i'}\| < \varepsilon/k$ for all $1 \leq j \leq k$. Thus, if we apply Proposition B.2 to the vectors $\alpha_1, \ldots \alpha_k \in \mathbb{R}^{\ell m}$ where $\alpha_j = (c_{ij} \theta_{i'})_{1 \leq i \leq \ell, \, 1 \leq i' \leq m}$ for $1 \leq j \leq k$, then Proposition B.1 follows. $\qquad\square$

We are therefore reduced to the task of proving Proposition B.2. The special case $k = 2$ and $\alpha_1 = 0$ is precisely Proposition A.2. in [10]. In fact, their argument generalizes to our case in a straightforward manner and as such we will sketch only the main steps and refer to the proofs in [10].

B.1. **The proof of Proposition B.2.** Let $\Lambda = \Lambda_1 \times \cdots \times \Lambda_k$ where each $\Lambda_i \subseteq \mathbb{R}^{d_i}$ is a full rank lattice. Recall, from [10], that the theta function

$$(37) \qquad \Theta_\Lambda(t, x) := \sum_{m \in \Lambda} e^{-\pi t |x - m|^2} = \frac{1}{t^{D/2} \det(\Lambda)} \sum_{\xi \in \Lambda^*} e^{-\pi |\xi|^2 / t} e(\xi \cdot x)$$

where $\Lambda^* := \{\xi \in \mathbb{R}^d : \xi \cdot m \in \mathbb{Z} \text{ for all } m \in \Lambda\}$ is the dual lattice of Λ, from which it follows that

$$(38) \qquad F_{\Lambda, \alpha}(N) := \det(\Lambda) \frac{1}{N} \sum_{n=1}^{N} \Theta_\Lambda(1, n \circ \alpha) = \sum_{\xi \in \Lambda^*} e^{-\pi |\xi|^2} \frac{1}{N} \sum_{n=1}^{N} e(\xi \cdot (n \circ \alpha))$$

where $\alpha = (\alpha_1, \ldots, \alpha_k)$ and for each $n \in \mathbb{N}$ we define $n \circ \alpha := (n\alpha_1, n^2\alpha_2, \ldots, n^k\alpha_k)$. As in [10] we also define the quantity

$$(39) \qquad A_\Lambda := \det(\Lambda) \sum_{m \in \Lambda} e^{-\pi |m|^2} = \sum_{\xi \in \Lambda^*} e^{-\pi |\xi|^2}.$$

The crucial ingredient in the proof of Proposition B.2 is the following lower bound on $F_{\Lambda, \alpha}(N)$, whose proof we outline in Section B.3.

Proposition B.3. *Let* $\Lambda = \Lambda_1 \times \cdots \times \Lambda_k$ *where* $\Lambda_i \subseteq \mathbb{R}^{d_i}$ *($d_i \geq 0$) is a full rank lattice, such that* $\det(\Lambda) \geq 1$. *Then for all* $\alpha = (\alpha_1, \ldots, \alpha_k)$, *with* $\alpha_i \in \mathbb{R}^{d_i}$ *and* $N \in \mathbb{N}$ *one has*

$$(40) \qquad F_{\Lambda, \alpha}(N) \geq (Cd)^{-C_k d^2} A_\Lambda^{-C_k d}$$

where $d = d_1 + \cdots + d_k \geq 1$.

Note that in the statement we allow the degenerate case $d_i = 0$, when $\Lambda_i = \mathbb{R}^{d_i} = \{0\}$.

Assuming Proposition B.3 we can now establish Proposition B.2 as in [10].

Proof of Proposition B.2. For simplicity of notation we set $c_d := (Cd)^{-C_k d^2}$. Let $\varepsilon > 0$ and $\Lambda := (R\mathbb{Z})^d$ where $R := C\, C_k d^2 \varepsilon^{-2}$ (with a suitable large constant C). Note that

$$(41) \qquad A_\Lambda = R^d \sum_{m \in \mathbb{Z}^d} e^{-\pi R^2 |m|^2} \leq (10R)^d.$$

If $\|n \circ \alpha\| \geq \varepsilon$ that is $|n \circ \alpha - m| \geq \varepsilon$ for all $m \in \mathbb{Z}^d$, then for $\beta = R\alpha$ one has that

$$|n \circ \beta - Rm|^2 \geq R^2 \varepsilon^2 / 2 + |n \circ \beta - Rm|^2 / 2$$

for all $m \in \mathbb{Z}^d$. Thus by (37) and the choice of R it follows that

$$\det(\Lambda) \Theta_\Lambda(1, n \circ \beta) \leq e^{-\pi R^2 \varepsilon^2 / 2} \det(\Lambda) \Theta_\Lambda(1/2, n \circ \beta)$$
$$\leq e^{-\pi R^2 \varepsilon^2 / 2} 2^{d/2} \sum_{\xi \in \Lambda^*} e^{-\pi |\xi|^2}$$
$$\leq \frac{c_d}{2} A_\Lambda^{-C_k d}.$$

If one defines the set $G := \{n \in [1, N] : \|n \circ \alpha\| \leq \varepsilon\}$, then by (38) and (41) it follows that

$$\frac{1}{N} \sum_{n \in G} \det(\Lambda) \Theta_\Lambda(1, n \circ \beta) \geq \frac{c_d}{2} A_\Lambda^{-C_k d}.$$

Also

$$\det(\Lambda) \Theta_\Lambda(1, n \circ \beta) \leq A_\Lambda \leq (C_k d^2 \varepsilon^{-2})^d$$

for all n. This implies (36) since

$$\frac{|G|}{N} \geq \frac{c_d}{2} A_\Lambda^{-C_d - 1} \geq (Cd)^{-C_k d^2} (C_k d^2 \varepsilon^{-2})^{-C_k d^2} \geq (\varepsilon/d)^{C_k' d^2}. \qquad \square$$

B.2. **Three Lemmas.** In order to prove Proposition B.3 we need the the following three lemmas, which we present without proof. These lemmas correspond to Lemmas A.5, A.6 and A.7 in [10] and are proved in essentially the same way. We leave this for the interested reader to verify.

Lemma B.4 (Properties of $F_{\Lambda,\alpha}$). *Let $\Lambda = \Lambda_1 \times \cdots \times \Lambda_k$, where each $\Lambda_i \subseteq \mathbb{R}^{d_i}$ is a full rank lattice. Let $\alpha = (\alpha_1, \ldots, \alpha_k)$, with $\alpha_i \in \mathbb{R}^{d_i}$ and $N > 20$.*

   (i) *For any $c \in (\frac{1}{10}, 1)$, we have $F_{\Lambda,\alpha}(N) \geq \frac{c}{2} F_{\Lambda,\alpha}(cN)$.*
   (ii) *For any integer $1 \leq q \leq \frac{N}{2}$, we have $F_{\Lambda,\alpha}(N) \geq \frac{1}{2q} F_{\Lambda,\alpha}(\frac{N}{q})$.*
   (iii) *Let $0 < \varepsilon < \frac{1}{d}$. If $\beta = (\beta_1, \ldots, \beta_k)$ such that $|\beta_i - \alpha_i| \leq \varepsilon N^{-i}$ for $1 \leq i \leq k$, then for all $1 \leq n \leq N$ we have*

$$\Theta_{\Lambda}(1, n \circ \alpha) \geq c^k \, \Theta_{(1+\varepsilon)\Lambda}(1, n \circ (1+\varepsilon)\beta)$$

   and hence

$$F_{\Lambda,\alpha}(N) \geq c^k \, F_{(1+\varepsilon)\Lambda,(1+\varepsilon)\beta}(N).$$

Lemma B.5 (Schmidt's alternative). *Let $\Lambda = \Lambda_1 \times \cdots \times \Lambda_k$, where each $\Lambda_i \subseteq \mathbb{R}^{d_i}$ is a full rank lattice and let $N > (4A_\Lambda)^{C_k}$. One of the following two alternatives holds:*

   (i) *$F_{\Lambda,\alpha}(N) \geq 1/2$;*
   (ii) *There is a positive integer $q \leq d(4A_\Lambda)^{C_k}$, and a primitive $\xi_i \in \Lambda_i^* \backslash \{0\}$, such that*

$$(42) \qquad |\xi_i| \leq C(\sqrt{d} + \sqrt{\log A_\Lambda})$$

   and

$$(43) \qquad \|q\xi_i \cdot \alpha_i\|_{\mathbb{R}/\mathbb{Z}} \leq (4A_\Lambda)^{C_k} N^{-i}$$

Recall that $\xi \in \Lambda_i^*$ is *primitive* if $\xi/n \notin \Lambda_i^*$ for any integer $n \geq 2$.

Lemma B.6 (Descent). *Suppose that $\Lambda' \subseteq \mathbb{R}^{d-1}$ and $\Lambda \subseteq \mathbb{R}^d$ are full rank lattices with $\Lambda' \subseteq \Lambda$, where $\mathbb{R}^{d-1}$ is regarded as a subset of $\mathbb{R}^d$. Suppose that $\alpha' \in \mathbb{R}^{d-1}$, that $\alpha \in \mathbb{R}^d$ and that $\alpha - \alpha' \in \Lambda$. Then*

$$(44) \qquad F_{\Lambda,\alpha}(N) \geq \frac{\det(\Lambda)}{\det(\Lambda')} F_{\Lambda',\alpha'}(N)$$

The only substantial difference from [10] in this section is in the proof of Lemma B.5, where we need estimates for the exponential sums, defined for $\theta = (\theta_1, \ldots, \theta_k) \in \mathbb{R}^k$ by

$$(45) \qquad S_N(\theta) = \frac{1}{N} \sum_{n=1}^{N} e(n\theta_1 + \cdots + n^k \theta_k).$$

The following is precisely what is required.

Lemma B.7 (Weyl Inequality). *Let $0 < \delta \leq 1/2$. There exist a positive constant $C_k > 0$, such that if $N \geq \delta^{-C_k}$ and*

$$(46) \qquad |S_N(\theta)| \geq \delta$$

then there exists a positive integer $q \leq \delta^{-C_k}$ such that

$$(47) \qquad \|q\theta_i\| < \delta^{-C_k} N^{-i}$$

for all $1 \leq i \leq k$.

This formulation follows easily form standard estimates on Weyl sums, see for example [19], Lemma 5. In fact using the sophisticated estimates of Vinogradov, one may take $C_k = C\, k^2 \log k$, however for simplicity we do not develop such bounds here.

B.3. **The proof of Proposition B.3.** As in [10], the proof of Proposition B.3 will follow, via an iteration, from the following result.

Proposition B.8 (Inductive lower bound on $F_{\Lambda,\alpha}$). *Suppose $\alpha = (\alpha_1, \ldots, \alpha_k)$, $\Lambda = \Lambda_1 \times \cdots \times \Lambda_k$ such that $\alpha_i \in \mathbb{R}^{d_i}$ and $\Lambda_i \in \mathbb{R}^{d_i}$ is a full rank lattice. Let $N > (4A_\Lambda)^{C_k}$ be an integer. Then either $F_{\Lambda,\alpha} \geq 1/2$ or there is an $\alpha' \in \mathbb{R}^{d-1}$ and a full rank lattice $\Lambda' = \Lambda_1' \times \cdots \Lambda_k' \subseteq \mathbb{R}^{d-1}$ with*

$$(48) \qquad\qquad A_{\Lambda'} \leq C(\sqrt{d} + \sqrt{\log A_\Lambda}) A_\Lambda$$

and an $N' \geq d^{-C}(4A_\Lambda)^{-C_k} N$ such that

$$(49) \qquad\qquad F_{\Lambda,\alpha} \geq D^{-C}(4A_\Lambda)^{-C_k} F_{\Lambda',\alpha'}(N')$$

Proof of Proposition B.8. Assuming $F_{\Lambda,\alpha} < 1/2$ and applying Lemma B.5, there exists an $1 \leq i \leq k$, a primitive $\xi_i \Lambda_i^* \backslash \{0\}$ and a positive integer $q \leq (4A_\Lambda)^{C_k}$, such that

$$\|\xi_i \cdot q^i \alpha_i\| \leq (4A_\Lambda)^{C_k} N^{-i}$$

(by changing the value of the constant C_k). Fixing the lattice Λ_i, and arguing as in [10], Proposition A.8, it follows that there is a $\beta_i \in \mathbb{R}^{d_i}$ such that

$$(50) \qquad\qquad |\beta_i - q^i \alpha_i| \leq (4A_\Lambda)^{C_k} N^{-i}$$

and an $m_i \in \Lambda_i$ such that $\beta_i' = \beta_i - m_i \in (\mathbb{R}\xi_i)^\perp \simeq \mathbb{R}^{d_i-1}$. Let $N_* = cd^{-C}(4A_\Lambda)^{-C_k} N$, then by the choice of N, we have

$$(51) \qquad\qquad |\beta_i - q^i \alpha_i| \leq d^{-1} N_*^{-i}$$

and moreover by Lemma B.4 (i) and (ii) it follows that

$$F_{\Lambda,\alpha}(N) \geq d^{-C}(4A_\Lambda)^{-C_k} F_{\Lambda,\alpha}(N_*) \geq d^{-C}(4A_\Lambda)^{-C_k} F_{\Lambda,q\circ\alpha}(N_*/q)$$

If $\beta = (\beta_1, \ldots, \beta_k)$ with β_i satisfying (50), and $\beta_j := q^j \alpha_j$ for each $j \neq i$, then by Lemma B.4 (iii) with $\varepsilon = 1/d$ we have

$$(52) \qquad F_{\Lambda,\alpha}(N) \geq d^{-C}(4A_\Lambda)^{-C_k} \geq d^{-C}(4A_\Lambda)^{-C_k} F_{(1+\varepsilon)\Lambda,(1+\varepsilon)\beta}(N')$$

where $N' = N_*/q$. Note that by the choice of N and the upper bound on q, N' satisfies the claimed lower bound. Finally let $\alpha' = (1+\varepsilon)\beta'$, where $\beta_j' = \beta_j$ for $j \neq i$, and let $\Lambda' = \Lambda_1' \times \cdots \times \Lambda_k'$ such that $\Lambda_i' = (1+\varepsilon)\Lambda_i \cap (\mathbb{R}\xi_i)^\perp$ and $\Lambda_j' := (1+\varepsilon)\Lambda_j$ for $j \neq i$. From Lemma B.6 using that $\det(\Lambda) = \prod_i \det(\Lambda_i)$ and $(1+1/d)^d \leq e$, one obtains

$$F_{\Lambda,\alpha}(N) \geq d^{-C}(4A_\Lambda)^{C_k} \frac{\det(\Lambda_i)}{\det(\Lambda_i \cap (\mathbb{R}\xi_i)^\perp)} F_{\Lambda',\alpha'}(N').$$

The rest of the argument goes exactly as in [10], one estimates

$$\frac{\det(\Lambda_i)}{\det(\Lambda_i \cap (\mathbb{R}\xi_i)^\perp)} = \frac{A_{\Lambda_i}}{A_{\Lambda_i \cap (\mathbb{R}\xi_i)^\perp}} \leq |\xi_i|^{-1} \leq C(\sqrt{d} + \sqrt{\log A_{\Lambda_i}}).$$

Since $A_\Lambda = \prod_i A_{\Lambda_i}$ and in particular $A_{\Lambda_i} \leq A_\Lambda$, the claimed bounds (48) and (49) follow. □

Note that if $d_i = 1$ the $(\mathbb{R}\xi_i)^\perp = \{0\}$, thus $\Lambda_i = \{0\}$, $d_i' = 0$ and $\Theta_{\Lambda_i}(t,x) \equiv 1$. We allow this to avoid the need to discuss separate cases. Iterating this proposition leads to the claimed lower bound on $F_{\Lambda,\alpha}(N)$ in Proposition B.3, in a straightforward manner, as in the proof of Proposition A.9 in [10].

Proof of Proposition B.3. By the trivial lower bound

$$F_{\Lambda,\alpha}(N) \geq \det(\Lambda)/N \geq 1/N$$

one may assume $N > d^{Cd^2}(4A_\Lambda)^{C_k D}$ for some suitably large constants C and C_k. Set $\Lambda_0 = \Lambda$, $\alpha_0 = \alpha$, $N_0 = N$. Applying Proposition B.8 repeatedly one obtains vectors $\alpha^{(j)} \in \mathbb{R}^{d-j}$, lattices $\Lambda^{(j)} \subseteq \mathbb{R}^{d-j}$ and integers $N^{(j)}$. Thus there must exist a $j \leq D$ such that $F_j := F_{\Lambda^{(j)},\alpha^{(j)}}(N^{(j)}) \geq 1/2$ (if $j = d$, then $\Lambda^{(j)} = \{0\}$ hence $F_j = 1$).

To check the validity of the iteration, since $A_\Lambda \geq 1$, one may use the crude bound

$$\sqrt{d} + \sqrt{\log X} \leq CdA_\Lambda^{1/d}$$

for $X \geq 1$, thus by (48) one has

$$A_{\Lambda^{(j)}} \leq (Cd)^d A_\Lambda^C$$

for all $1 \leq j \leq d$. This implies

$$N^{(j)} \geq d^{-C}(4A_{\Lambda^{(j)}})^{-C_k} N^{(j-1)} \geq (Cd)^{-C_k d} A_\Lambda^{-C_k}$$

thus by the choice of N, we have that $N^{(j)} \geq (4A_{\Lambda^{(j)}})^{C_k}$ for all $1 \leq j \leq d$, and hence by (49)

$$F_{j+1} \geq d^{-C}(4A_{\Lambda^{(j)}})^{-C_k} F_j \geq (Cd)^{-C_k d} A_\Lambda^{-C_k} F_j$$

which gives the desired lower bound for $F_{\Lambda,\alpha}(N)$. $\qquad \square$

References

[1] G. I. Arkhipov, V. N. Chubarikov and A. A. Karatsuba, *Trigonometric sums in number theory and analysis*, de Gruyter Expositions in Mathematics, 39. Walter de Gruyter GmbH & Co. KG, Berlin

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

[3] V. Bergelson, *Combinatorial and Diophantine applications of ergodic theory*, Handbook of dynamical systems. Vol. 1B, 745–869, Elsevier B. V., Amsterdam, 2006.

[4] V. Bergelson, B. Host and B. Kra, *Multiple recurrence and nilsequences*, with an appendix by Imre Ruzsa, Invent. Math. 160 (2005), no. 2, 261–303.

[5] 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.

[6] N. Frantzikinakis, *Multiple ergodic averages for three polynomials and applications*, Trans. Amer. Math. Soc. 360 (2008), no. 10, 5435–5475.

[7] N. Frantzikinakis and B. Kra, *Ergodic averages for independent polynomials and applications*, J. London Math. Soc. (2) 74 (2006), no. 1, 131–142.

[8] N. Frantzikinakis and R. McCutcheon, *Ergodic Theory: Recurrence*, Encyclopedia of Complexity and System Science, Springer, (2009), Part 5, 3083-3095

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

[10] B. Green and T. Tao, *New bounds for Szemerdi's theorem. II. A new bound for $r_4(N)$*, Analytic number theory, 180–204, Cambridge Univ. Press, Cambridge, 2009.

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

[12] W. T. Gowers, *Decompositions, approximate structure, transference, and the Hahn-Banach theorem*, arxiv.org/abs/0811.3103.

[13] J. Griesmer, Personal communication

[14] L. K, Hua, *Additive theory of prime numbers*, American Mathematical Society, Providence, R.I. 1965

[15] A. Y. Khintchine, *Eine Verschärfung des Poincaréscen "Wiederkehrsatzes"*, Composio Math. 1 (1934), 177-179.

[16] B. O. Koopman and J. von Neumann, *Dynamical systems of continuous spectra*, Proc. Nat. Acad. Sci. U.S.A. 18 (1932), 255-263

[17] B. Kra, *Ergodic methods in additive combinatorics*, Additive combinatorics, 103–144, CRM Proc. Lecture Notes, 43, Amer. Math. Soc., Providence, RI, 2007.

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

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

[20] N. Lyall and Á. Magyar, *Polynomial configurations in difference sets (Revised version)*, arxiv.org/abs/0903.4504.

[21] N. Lyall and Á. Magyar, *Optimal polynomial return times*, preprint

[22] R. McCutcheon, *Elemental methods in ergodic Ramsey theory*, Lecture Notes in Mathematics, 1722. Springer-Verlag, Berlin, 1999.

[23] H. Poincaré, *Les méthodes nouvelles de la mécanique céleste. I*, Gathiers-Villars, Paris 1892; II, 1893; III, 1899.

[24] T. Tao, *The ergodic and combinatorial approaches to Szemerédi's Theorem*, Additive combinatorics, 145–193, CRM Proc. Lecture Notes, 43, Amer. Math. Soc., Providence, RI, 2007.

[25] T. Wooley, *Some remarks on Vinogradov's mean value theorem and Tarry's problem*, Monatsh. Math. 122 (1996), no. 3, 265–273.

Department of Mathematics, The University of Georgia, Athens, GA 30602, USA

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

Department of Mathematics, University of British Columbia, Vancouver, B.C. V6T 1Z2, Canada

E-mail address: magyar@math.ubc.ca